Universidad de Valencia

Facultad de Derecho

Departamento de Derecho Civil

Programa de Doctorado en Derecho, Ciencia Política y Criminología

Tesis Doctoral

Derecho de la Inteligencia Artificial

Un enfoque global de responsabilidad desde la ética, la seguridad y las nuevas propuestas reguladoras europeas

Autor

José Manuel Muñoz Vela

Dirección de la Tesis

Prof. Dr. D. Javier Plaza Penadés

Prof. Dra. Dña. Raquel Guillén Catalán

Octubre 2021

Derecho de la Inteligencia Artificial

Un enfoque global de responsabilidad desde la ética, la seguridad y las nuevas propuestas reguladoras europeas

Mi agradecimiento a aquellas personas que siempre han estado a mi lado, que siempre están y siempre estarán de manera incondicional.

A mis padres, a mi mujer y a toda mi familia por su apoyo.

A todos los magníficos profesionales, investigadores y docentes con los que he tenido la oportunidad de colaborar y trabajar durante todos estos años y de cuyo talento y conocimiento tanto me he enriquecido.

Y, por supuesto, mi más profundo agradecimiento a Javier, director de esta investigación, quien siempre me ha proporcionado el aliento necesario durante este largo camino para proseguir con los esfuerzos y llegar hasta el final.

Gracias a tod@s.

Índice General

Abreviaturas

CC	Código Civil
CP	Código Penal
IA	Inteligencia Artificial
IoT	Internet of Things
LOPDGDD	Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales
LPI	Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia
RGPD/GDPR	Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos)
TRLGDCU	Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias
UE	Unión Europea

Introducción

La inteligencia artificial integra un conjunto de tecnologías que está cambiando y cambiará el mundo que conocemos. El potencial de la inteligencia artificial es inmenso.

La inteligencia artificial se está desarrollando y aplicando a un ritmo exponencial gracias, entre otros factores, a la mayor capacidad de computación y almacenamiento, la mayor disponibilidad de datos, el aumento de sus capacidades, mejora de la conectividad y el incremento de su interacción con otras tecnologías y elementos, y considero que se convertirá en un recurso o medio esencial para el futuro de la humanidad, en principio, orientado a lo que deberían ser sus principales finalidades, esto es, resolver problemas, mejorar nuestra vida y nuestro mundo.

La inteligencia artificial forma parte de nuestra vida diaria, si bien, en ocasiones, puede resultar confuso identificar qué es y dónde está presente, pero la sociedad en general la viene asociando casi sistemáticamente a cualquier sistema, dispositivo u objeto que pueda hacer cosas por sí mismo de manera automatizada.

Diariamente tenemos conocimiento de nuevos avances o aplicaciones basados en inteligencia artificial en todos los sectores, medicina, teleasistencia, biotecnología, conducción autónoma, agricultura, etc.

Los sistemas inteligentes ya emulan el comportamiento humano y pueden realizar actividades hasta ahora solo atribuibles a los seres humanos, sin embargo, muchas de las capacidades generalmente atribuidas a la inteligencia artificial hoy, distan mucho de la realidad tecnológica actual y del estado de desarrollo y aplicación de la misma.

La inteligencia artificial actual no dispone de un cerebro, no piensa como un humano, no tiene consciencia, no siente, pero sí imita.

Sin embargo, muchas decisiones que hasta ahora eran tomadas por humanos están siendo tomadas por sistemas inteligentes y pueden afectar a una o a muchas personas, empresas, Administraciones públicas o gobiernos, y en todo tipo cuestiones como la contratación

laboral, la concesión de un crédito o préstamo, la elaboración de una sentencia, la realización de un diagnóstico, tratamiento o intervención médica o la compraventa de valores.

Pero ni las decisiones humanas son perfectas ni, de manera consecuente, las artificiales o algorítmicas lo son, si bien, debe asegurarse la máxima legitimidad, licitud, objetividad, ética, seguridad, respeto, justicia y transparencia de las mismas, del mismo modo en el que pretendemos que lo sean las humanas.

En consecuencia, el sesgo inherente o la malsonante "estupidez", en ocasiones, son factores que pueden integrar el comportamiento del ser humano, por lo que esperemos que el mismo no lo proyecte a aquello que crea y en lo que pretende depositar su confianza y ceder la gestión de muchas de las parcelas de su vida, e incluso su dependencia y existencia.

Platón decía que "una buena decisión está basada en conocimiento, no en números (datos)". Arthur Sulzberger Hays, editor del New York Times durante más de 25 años, decía que "el juicio de un hombre no puede ser mejor que la información en la cual ha basado dicho juicio", por lo que, del mismo modo, tampoco podrá serlo el de la máquina respecto de la información en la que ha basado el mismo.

Algunos expertos afirman que la peor inteligencia artificial parece mejor que la inteligencia natural que nos gobierna, el problema que podrá suscitarse en tal caso es si el ser humano la hace a su imagen y semejanza.

La inteligencia artificial es una herramienta con un potencial y poder inmenso que previsiblemente se va a convertir en un medio cada vez más imprescindible para resolver algunos de los grandes problemas del ser humano.

Sin embargo, paralelamente, sus retos y riesgos pueden tener un enorme impacto, especialmente en el ámbito ético, jurídico, social, económico y de la seguridad, por lo que ni política ni jurídicamente pueden obviarse y deben ser conocidos, acometidos y gestionados adecuadamente, dado que, de otro modo, la inteligencia artificial se podría convertir en un medio para la desigualdad, el control y la destrucción.

Las respuestas a los nuevos retos y riesgos que plantea exigen profundos análisis y reflexiones desde todas las dimensiones a un ritmo síncrono, y por supuesto, desde una perspectiva de seguridad, que exige su identificación y adecuada gestión.

El aumento de las capacidades de los sistemas inteligentes, especialmente autoaprendizaje y autonomía, comportarán asociadamente una relativa impredecibilidad que comportará riesgos adicionales.

Las decisiones algorítmicas se basan en datos y pueden no ser perfectas. Las decisiones humanas tampoco lo son.

El funcionamiento de los sistemas inteligentes se halla sujeto, como cualquier sistema informático, a los riesgos precitados que incluyen el error en su programación, defectos en su funcionamiento o pérdida de conectividad y comunicación en muchos supuestos con redes, componentes y otros sistemas. Y el uso de los mismos también comporta riesgos inherentes, especialmente ante su uso indebido, ilegítimo o ilícito por parte del propio operador o usuario, o por parte de terceros.

Las decisiones y conductas de o mediante sistemas inteligentes pueden causar daños a los sujetos afectados por las mismas, lo que genera inevitablemente la necesaria reflexión sobre el régimen de responsabilidad aplicable en tales supuestos y si los marcos vigentes, hasta que dispongamos de otros, son suficientes para dar una respuesta adecuada a los daños causados por estos sistemas en los diferentes contextos que se pueden plantear, así como para garantizar el derecho a un resarcimiento efectivo a las personas afectadas.

La inteligencia artificial no se haya todavía regulada de manera global a nivel mundial y los marcos jurídicos vigentes pueden dar respuesta adecuada a algunos de los supuestos que puede plantear, especialmente en el ámbito de la responsabilidad por daños, pero no a todos, especialmente ante factores como la variedad de sistemas inteligentes, elementos que los integran, capacidades, características, sector, uso, contexto, finalidad, etapa de su ciclo de vida en el que se produce o sujetos intervinientes.

En definitiva, en la actualidad se están diseñando, desarrollando, comercializando, aplicando y utilizando sistemas inteligentes sin marcos éticos, jurídicos, de seguridad y

de responsabilidad claramente definidos y de carácter vinculante, a pesar del alto riesgo incuestionable de muchos de ellos.

La UE está liderando los esfuerzos por consensuar determinados principios y normas éticas de la inteligencia artificial y está siendo pionera en su regulación mediante la elaboración de distintas propuestas que son objeto de análisis en esta investigación, tanto en materia ética, jurídica como de responsabilidad civil, las cuales se han constituido ya en referente internacional en distintos países.

La necesidad de adecuación de los marcos regulativos actuales para dar respuestas a los distintos supuestos que se pueden plantear en materia de responsabilidad puede requerir su urgente revisión, previsiblemente profunda en algunos aspectos que constituyen sus pilares esenciales si no se crean marcos específicos, y podría exigir, además, nuevos enfoques y técnicas legislativas para crear nuevos marcos jurídicos para la inteligencia artificial y para adaptar, modificar o complementar los marcos actuales que permitan proporcionar soluciones adecuadas a las cuestiones que puede plantear la inteligencia artificial de hoy y la de mañana, especialmente la dotada de mayores capacidades y autonomía o más “avanzada”.

El objeto y alcance de esta investigación pretende focalizarse en todo ello y, en especial, en la responsabilidad por daños dimanante de su funcionamiento y/o uso desde una perspectiva global, esto es, ética, jurídica y de seguridad, para analizar desde la misma los marcos éticos y jurídicos actuales como los distintos marcos reguladores propuestos en el seno de la UE.

No concibo realizar el análisis de todo ello, ni de los marcos reguladores vigentes y de los propuestos que la pretenden regular, sin definir y comprender de manera previa esta realidad tan compleja desde una perspectiva integral, tanto técnica, ética, jurídica como de seguridad, al objeto de identificar los múltiples retos y riesgos, su ciclo de vida o los posibles sujetos participantes en el mismo.

Por ello, de manera previa al análisis de los distintos aspectos propuestos, consideré necesario abordar el concepto evolutivo de inteligencia artificial, así como otros conceptos esenciales relacionados, sus características y sus clases, y todo ello desde la

perspectiva global precitada, al objeto de sentar con claridad las bases sobre las que sustento mi posterior análisis, reflexiones, posicionamiento y conclusiones desde las dimensiones indicadas y su interrelación con la responsabilidad jurídica, especialmente por daños.

De antemano, se debe anticipar que no existe un consenso internacional ni sobre el concepto inteligencia artificial, ni sobre los principios y normas éticas exigibles a la misma, ni sobre la seguridad que debe integrar, ni sobre las técnicas e instrumentos normativos para su regulación.

No obstante, iniciar de este modo mi investigación me permitió poder delimitar con mayor claridad y precisión el objeto sobre el que se sustentan mis reflexiones y las nuevas propuestas regulatorias y, de otro, para disponer de una perspectiva más amplia para el posterior análisis de sus principales riesgos y retos en materia ética, jurídica y de seguridad, y la relación de todo ello con la responsabilidad jurídica derivada especialmente de los daños causados derivados de su uso y funcionamiento.

Consideré imprescindible partir de este enfoque global para abordar adecuadamente la inteligencia artificial, para poder comprenderla desde dicha perspectiva integral, para identificar sus principales características, sus capacidades, su tipología, su ciclo de vida, los posibles sujetos participantes, sus retos y sus riesgos, lo que a su vez, me permitió poder abordar con más profundidad los aspectos de responsabilidad, que constituyen el eje de esta investigación, y ello valorando adecuadamente con espíritu crítico las posibles soluciones que puedan contemplarse por los marcos jurídicos actuales, así como las cuestiones que considero deberán abordarse por los futuros marcos reguladores.

Los mayores dificultades que plantea la inteligencia artificial en relación con la depuración de responsabilidades sobre los daños causados por su funcionamiento y/o uso son sus propias características y capacidades de las que pueda estar dotada y correlativa gobernabilidad, especialmente su autonomía creciente, el autoaprendizaje, la impredecibilidad, la opacidad o la falta de explicabilidad, junto con la pluralidad de sujetos que pueden intervenir a lo largo de su ciclo de vida, máxime ante la ausencia de personalidad jurídica y capacidad de obrar de la misma o de las máquinas, robots o dispositivos que la integren.

Su análisis comporta la necesidad de revisar los marcos actuales y valorar futuros marcos de responsabilidad, especialmente los construidos sobre una posible responsabilidad específica basada en el riesgo y desde la mencionada perspectiva global, identificando quién de los sujetos participantes en su ciclo de vida tiene mayor control y supervisión sobre sus riesgos en cada contexto particular, determinado por el sector, uso, características, capacidades, entrenamiento e interacciones del sistema inteligente, entre otros factores.

Del mismo modo, dicha perspectiva integral también exige un enfoque global en cuanto a la multiplicidad de riesgos que pueden materializarse y causar daños, incluyendo los de cumplimiento regulatorio, evitando focalizarme en algunos de ellos como los de privacidad o de discriminación, sin perjuicio de poner en valor la madurez de la regulación europea de estos aspectos, que constituye una de las referencias a considerar para la regulación de distintos aspectos de la inteligencia artificial, especialmente los relativos a seguridad, transparencia, explicabilidad, rendición de cuentas -*accountability*- y responsabilidad.

Dejando a un lado todas estas cuestiones previas, es necesario significar que cuando me planteé definir mi investigación, mi primera reflexión fue concretar su objeto y alcance, así como el enfoque más adecuado para su desarrollo.

Desde mis primeras reflexiones decidí focalizarme inicialmente en sus principales retos y riesgos de cuya materialización puede derivarse responsabilidad por daños, por lo que consideré necesario abordarlos de manera previa desde la perspectiva global precitada, para abordar posteriormente las propuestas reguladoras acometidas hasta la fecha por la UE, los regímenes actuales de responsabilidad y los posibles planteamientos de futuro.

La vinculación entre ética, seguridad y responsabilidad en el marco de la inteligencia artificial es incuestionable. La necesidad de que el Derecho regule su exigencia también, en mi opinión.

La responsabilidad es una exigencia ética, como la seguridad, pero ambas también constituyen una exigencia jurídica. Y el apartamiento de estas exigencias conforman algunos de los marcos vigentes y propuestos de exigencia de responsabilidades de distinta

naturaleza, también civil, y tanto de naturaleza general -contractual o extracontractual-, como especial, por productos defectuosos.

Reconozco que la fuerte atracción que tengo hacia los avances tecnológicos, me impulsaba a tomar como referencia los últimos avances y aplicaciones más significativas en inteligencia artificial, para focalizarme en una aplicación específica. Solamente el análisis vertical de cualquiera de ellas, sus implicaciones y retos en el ámbito ético, jurídico y de seguridad, así como las cuestiones asociadas de responsabilidad, podría haber sido el objeto de esta investigación.

La tentación por hacer una inmersión global desde un punto de vista ético, jurídico y de seguridad en cualquier desarrollo o aplicación específica de la inteligencia artificial era muy grande, si bien, me hubiera exigido verticalizar mi enfoque y centrar mi análisis exclusivamente en una de ellas, conformando en su integridad el objeto y alcance de mi investigación.

La idea era atractiva y los *inputs* sobre nuevos desarrollos y aplicaciones, diarios y abrumadores.

En definitiva, la posibilidad de abordar su aplicación para el diseño y fabricación de coches autónomos, la gestión de sistemas inteligentes para la detección precoz de cáncer de piel, para la prevención y detección de ciberataques o para la dirección e incluso ejecución de intervenciones quirúrgicas en remoto era verdaderamente irresistible, máxime cuando he tenido la enorme suerte de participar profesionalmente en algunos proyectos desde mi óptica jurídico-tecnológica, lo que me ha permitido investigar "tocando la tecnología" y con la presión de tener que dar soluciones globales como profesional a aspectos no previstos en los distintos ordenamientos jurídicos vigentes a fecha actual.

No obstante, esos momentos iniciales de reflexión sobre la perspectiva más adecuada para una investigación de este tipo y con una pretensión global desde el punto de vista ético, jurídico y de seguridad, me llevaron a adoptar un enfoque horizontal y transversal de la inteligencia artificial desde todas estas dimensiones, sin perjuicio de abordar proyectos y aplicaciones reales específicas, como ejemplos de los distintos aspectos abordados para

ilustrar los mismos, a los que aludo en relación con los distintos aspectos tratados en mi investigación.

Tome la decisión de llevar a cabo una investigación horizontal desde un enfoque global, ante la necesidad de reflexionar desde estas perspectivas para acometer cualquier proyecto futuro, así como para el análisis y revisión de los marcos reguladores actuales, con el objetivo de determinar su posible adecuación y aplicación a los distintos retos que plantea la inteligencia artificial en estas dimensiones, así como sus posibles carencias y posibles soluciones a través de las presentes y futuras propuestas reguladoras, tanto para complementar los marcos actuales como, en su caso, para adaptarlos, complementarlos, modificar su estructura y fundamentos y/o proceder a la creación de marcos específicos sobre determinados aspectos.

La metodología utilizada en esta investigación ha combinado distintas técnicas de investigación, tanto la histórica, la descriptiva, la causal como la exploratoria, y desde un enfoque global, técnico-científico, ético, jurídico y de seguridad, para interrelacionar conceptos y marcos jurídicos con la realidad y contexto en el que fueron concebidos y su adecuación actual a las nuevas realidades y cuestiones complejas que suscitan a las que debe dar respuestas adecuadas.

He incorporado metodologías comparativas, dogmáticas y propositivas para poder abordar con profundidad desde una perspectiva global y completa una realidad no contemplada por el Derecho y que plantea conflictos de compleja solución en base a los marcos jurídicos vigentes, especialmente en materia de daños derivados de su funcionamiento y/o uso.

La investigación llevada a cabo se formalizó en una primera versión de este documento cerrada a 30.03.2021, si bien, la publicación de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, el 21 de abril de 2021, me obligó a analizar el contenido de la misma y revisar íntegramente la mayor parte de los capítulos de esta investigación conforme a su objeto y alcance, lo que me llevó a finalizar y cerrar la presente versión revisada el 30.06.2021.

En este sentido, significar que he partido de algunas obras y propuestas reguladoras publicadas en el momento de su análisis únicamente en inglés, por lo que cualquier mención o inserción expresa a su contenido en este documento pueden responder a una traducción libre del autor de esta investigación, sujeta a error o giro lingüístico diferencial que pueda tener cualquier traducción o versión posterior publicada en castellano o con corrección de errores.

Durante los próximos meses se producirán con toda seguridad novedades legislativas múltiples a nivel europeo e internacional en relación con los múltiples aspectos abordados, especialmente ante el estado de la tramitación de diversas iniciativas reguladoras citadas y tratadas a lo largo de esta investigación, de modo que en el momento en el que ésta investigación se deposite y publique, dichas novedades publicadas con posterioridad a su finalización no podrán haber sido consideradas en la misma, sin perjuicio de que los borradores y propuestas actuales hayan sido ya contempladas y, en su caso, analizadas en relación con los aspectos abordados a lo largo de la misma.

La investigación realizada ha tomado como referencia los marcos reguladores vigentes como las propuestas actualmente en tramitación, y no sólo en España y la UE, sino incluso, algunas referencias a nivel internacional en relación con concretos aspectos tratados durante mi investigación, considerando igualmente la doctrina y la jurisprudencia nacional e internacional sobre las cuestiones concretas objeto de análisis.

En base a todo ello, he llevado a cabo mi investigación sobre los distintos aspectos tratados en la misma, con el resultado y consideraciones finales que acompañan a cada capítulo y las conclusiones finales que se incluyen, que incorporan mi visión personal sobre las distintas cuestiones tratadas basada en la investigación llevada a cabo.

Los principales objetivos de la investigación han sido: a) La identificación de los principales retos y riesgos asociados a la inteligencia artificial -actuales y futuros-, b) La identificación y análisis de los marcos éticos, regulatorios, de seguridad y de responsabilidad vigentes aplicables a la misma y a los daños causados por su funcionamiento y/o uso; c) La determinación del grado de adecuación de los marcos jurídicos vigentes para resolver las cuestiones que plantea la inteligencia artificial en materia de responsabilidad y las que plantearán en lo sucesivo los sistemas inteligentes

más avanzados o con más capacidades; d) La necesidad y viabilidad jurídica de la atribución de una personalidad jurídica a los sistemas inteligentes más avanzados y con mayores capacidades; e) El análisis de las distintas propuestas reguladoras europeas; f) La identificación de las características y aspectos adicionales que deberían ser considerados para la posible construcción de los futuros marcos normativos que regulen la inteligencia artificial y que actualicen, adecuen, complementen o sustituyan a los existentes.

En consecuencia, la contribución al avance científico en estas materias se basa en los resultados obtenidos que pueden servir para, de un lado, identificar los marcos y requerimientos éticos, jurídicos y de seguridad de la inteligencia artificial, identificar los marcos de responsabilidad aplicables a la misma y los posibles sujetos relacionados durante su ciclo de vida, definir los distintos marcos y vías de actuación para la resolución de conflictos en materia de responsabilidad por daños derivados de su funcionamiento y/o uso y, por último, ofrecer algunos aspectos clave que deberían ser considerados por los futuros marcos normativos reguladores de la inteligencia artificial.

El enfoque adoptado por mi parte ha buscado la sencillez en los planteamientos para posteriormente analizar sus complejidades jurídicas y las posibles soluciones a las cuestiones detectadas, y en todo momento huyendo, en la medida de lo posible, de los tecnicismos propios del lenguaje científico, tecnológico y jurídico, para facilitar la comprensión de las reflexiones llevadas a cabo que conduzcan a las conclusiones consecuentes. En congruencia con ello, he intentado no usar ni abusar de abreviaturas, empezando por la de inteligencia artificial o "IA".

Las fuentes para llevar a cabo esa reflexión global han sido todas a las que me ha sido posible acceder, tanto científicas, éticas, económicas, filosóficas como jurídicas, con el objetivo de sentar y partir desde el enfoque global precitado y desde distintas opiniones y posicionamientos para disponer de una visión holística de una realidad, analizarla desde la óptica planteada, reflexionar sobre la misma y llegar a mis propias conclusiones y opiniones, que he pretendido cualificar mediante su razonamiento y desde mi formación híbrida y experiencia profesional, y obviamente con la necesaria prudencia, modestia y humildad en aspectos técnicos y científicos, como no puede ser de otra forma, dado mi perfil.

La investigación incorpora las distintas fuentes y referencias bibliográficas consultadas oportunamente citadas al pie de cada una de sus páginas, sin perjuicio de la bibliografía general que se relaciona al final más acotada y circunscrita exclusivamente a los autores referenciados a lo largo de la misma, habida cuenta de la extensión de las referencias consultadas y de la propia investigación.

Capítulo I

Inteligencia artificial

1. Introducción

La inteligencia artificial se ha convertido en un concepto de absoluta actualidad y cotidianidad. Compartimos nuestra vida diaria con elementos, sistemas, tecnologías, productos y servicios dotados de inteligencia artificial, y su desarrollo, despliegue y aplicación en todo tipo de ámbitos y sectores es incesante.

La inteligencia artificial es una expresión de moda y actualmente de obligada asociación a cualquier producto o servicio basado en tecnología y que pretenda presentarse al mercado como de vanguardia a efectos de marketing.

La palabra "inteligente" o "*smart*" forma parte de mensajes comerciales y publicitarios de cualquier producto o servicio vinculado a la tecnología como sinónimo de innovación y vanguardia. Cada vez es más habitual su utilización por empresas y entidades públicas para la comercialización de productos, venta o financiación de proyectos o captación de reconocimiento.

Este atributo forma parte imprescindible del marketing y la publicidad para posicionar productos y servicios en la mente de consumidor que motiven su adquisición por el mismo.

Incluso a nivel social se asocia este atributo a casi cualquier cosa que haga algo de manera automatizada y por sí misma.

Asistentes personales digitales, smartphones, relojes inteligentes, soluciones de domótica, aspiradoras, frigoríficos, smartTVs, buscadores, herramientas de traducción automática, compras en línea, banca electrónica, servicios de inversión, ciberseguridad,

robotización de procesos industriales, infraestructuras inteligentes, vehículos autónomos, *smartcities*, agricultura inteligente, drones o robots poseedores de dicha supuesta inteligencia artificial, forman parte ya de nuestra vida.

No obstante, dejando a un lado percepciones, constituye una tecnología que, en modo alguno, es nueva, pero cuyo creciente desarrollo y aplicación exponencial gracias a la mayor capacidad de computación y almacenamiento, mejora de sus capacidades y de la emulación del cerebro humano, mayor disponibilidad de datos y mayor interacción con otras tecnologías, está revolucionando y revolucionará el mundo que conocemos.

La inteligencia artificial está presente en el ámbito personal, profesional, empresarial, público y político, y en cualquier tipo de actividad, producto o servicio.

La inteligencia artificial se está asociando como capacidad a todo tipo de sistemas, máquinas, robots, vehículos y otros productos, así en todo tipo de sectores como salud, educación, alimentación, transporte, industria manufactura, financiero, agricultura, ganadería o servicios públicos.

En el ámbito de la salud, se utiliza en robots quirúrgicos, atención socio-sanitaria para el cuidado de personas mayores o grandes dependientes, *chatbots* de respuesta a llamadas de emergencia o sistemas de detección de enfermedades por la voz o la tos.

En el ámbito del transporte, se utiliza en sistemas de mejora de la seguridad, velocidad y eficiencia del tráfico ferroviario o en la gestión del tráfico rodado en ciudades.

En la agricultura, para la gestión responsable de fertilizantes, pesticidas o riego, mejora de la productividad o la reducción de impacto ambiental.

En la ganadería, para la trazabilidad de los productos o el control de movimiento, temperatura o consumo de alimentos por el ganado.

En el ámbito industrial, se aplica en las denominadas “fábricas inteligentes”, aportando principalmente precisión, eficiencia y reducción de costes, lo que a su vez permite mayor productividad. Permite mejorar los productos, los procesos y los modelos de negocio en todos los sectores económicos.

O, en el sector público, constituye una ayuda para la gestión de procedimientos administrativos o la prevención de desastres naturales.

La inteligencia artificial, como luego expondré, también comporta importantes retos éticos, técnicos (especialmente conectividad) y jurídicos, así como riesgos potenciales, como su falta de transparencia y opacidad en la toma de decisiones, la discriminación, el sesgo, la intromisión en la privacidad o su uso con fines maliciosos.

La complejidad, las capacidades, el nivel de desarrollo y de innovación y los retos y riesgos que plantea cada sistema inteligente son distintos, y no sólo por sus características propias, sino también por sus aplicaciones y usos potenciales, especialmente mediante su interacción con otras tecnologías y elementos -especialmente de *hardware*-, y otros sistemas.

Los avances en capacidad de computación, disponibilidad de grandes volúmenes de datos, su tratamiento masivo y selectivo, y la irrupción de nuevos algoritmos ha permitido un paralelo avance y eclosión exponencial del desarrollo y aplicación de la inteligencia artificial a nivel mundial.

Durante los últimos años, la inteligencia artificial ha pasado de ser considerada una capacidad específica asociada a un sistema de información, a adquirir "autonomía propia" y convertirse en una de las tecnologías de mayor impacto en el mundo, tanto a nivel político, económico como social. Los últimos informes y estudios publicados a nivel mundial así la posicionan[1].

El informe *Artificial Intelligence and life in 2030: The one hundred year study on artificial intelligence*[2] de la Universidad de Stanford identifica los principales ámbitos

[1] El informe de ICEMD-ESIC 2019 identifica la inteligencia artificial como una de las tecnologías más disruptivas, de mayor impacto y con mayor tendencia de desarrollo en lo sucesivo (https://www.icemd.com/digital-knowledge/infografias/tecnologias-disruptivas/). El informe *MIT Technology Review 2018*, identificaba la inteligencia artificial en la nube y libre, las redes generativas antagónicas (GAN) y las ciudades sensibles (además de inteligentes) dentro de las diez tecnologías emergentes de referencia mundial. El informe *Technology Industry Innovation Survey 2019* de *KPMG*, al igual que el *McKinsey Global Institute* de 2019, sitúa a la inteligencia artificial entre las tecnologías que mayor impacto están teniendo en la transformación del mundo de los negocios.

[2] STONE, P. Y OTROS (2016). *Artificial Intelligence and life in 2030: The one hundred year study on artificial intelligence*. 6 de septiembre de 2016. Universidad de Stanford 2016.

cotidianos en los que más se está desarrollando la inteligencia artificial y en los que tendrá mayor impacto, como el transporte, la salud, la educación, la seguridad pública, el empleo y el entretenimiento.

La inteligencia artificial cambiará nuestras vidas, mejorará la atención médica y social-sanitaria, la eficiencia de la agricultura, de la ganadería y de los procesos de producción, contribuirá a la mitigación del cambio climático, aumentará nuestra seguridad y nuestra calidad de vida y mejorará nuestro mundo, eso sí, si el ser humano lo permite.

La inteligencia artificial se ha convertido en una obligación para Estados, empresas y ciudadanos, al menos esta parece la percepción mayoritaria en los agentes que integran y operan en la denominada Sociedad Digital, y se ha iniciado una carrera desenfrenada para liderar la misma a nivel mundial en distintos sectores y aplicaciones por parte de las grandes tecnológicas -los llamados "Siete Gigantes de la era de la IA"[3], entre los que se incluyen Google, Facebook, Amazon, Microsoft, Baidu, Alibaba y Tencent-, así como por distintos gobiernos, especialmente China y EEUU, a las que se han unido países como Emiratos Árabes[4] y Europa.

EE.UU. ha venido evidenciando una superioridad en investigación y desarrollo de la inteligencia artificial que tradicionalmente la había dejado en manos del sector privado hasta la irrupción de China y por detrás Rusia, con sus estrategias e inversiones públicas multimillonarias.

A esta carrera se han unido países como Francia, Reino Unido, Alemania y la propia Unión Europea.

La UE ha apostado por un enfoque sólido en su *Estrategia Europea para la IA*[5] presentada en abril de 2018, con el objetivo de aprovechar sus oportunidades y abordar sus desafíos.

[3] LEE, KAI-FU (2018). *Superpotencias de la inteligencia artificial*". Versión Kindle. P. 106. Deusto 2018.
[4] Según el informe *Respuesta Política Global a la IA* de *FTI Consulting*, los países que pretenden liderar a nivel mundial la inteligencia artificial (IA) son China y Emiratos Árabes.
[5] *Inteligencia artificial para Europa*. COM (2018) 237 final.

Israel, por su parte, ha desarrollado su inteligencia artificial bajo un enfoque estratégico más local y orientado a mejorar su economía y lograr su supremacía regional.

La inteligencia artificial ha pasado a ser el foco de políticas y estrategias gubernamentales y, en algunos países como China, ha llegado a las aulas y en niveles de educación infantil. En España, a nivel autonómico, empezará posiblemente a incluirse como asignatura en los planes de estudios de Bachillerato para el próximo año académico.

El sector empresarial usuario de estos sistemas está identificando la inteligencia artificial como una tecnología "disruptiva" e "innovadora" dentro de sus planes y estrategias de transformación digital, a mi juicio, en ocasiones, desacertadamente, dado que lo realmente disruptivo no es la tecnología en sí misma, sino los modelos y procesos de negocio que puedan, o no, desarrollarse y sustentarse sobre la misma u otras tecnologías conexas para la consecución de los objetivos empresariales, lo que a veces se pierde vista para focalizarse en la tecnología como un fin, cuando realmente es un medio para la alcanzar aquellos, mejorar procesos y la eficiencia productiva.

La inteligencia artificial es una realidad y una tendencia ineludible, pero, sin embargo, no es tan nueva. El concepto fue ya acuñado en 1956 por uno de los considerados "padres" de la inteligencia artificial, John McCarthy, como expondré más adelante.

Desde hace ocho décadas se está trabajando en inteligencia artificial, pero la realidad que pretendía definir ha evolucionado y cambiado sustancialmente.

El crecimiento exponencial de sus usos y aplicaciones en todo tipo de ámbitos y sectores gracias a la mayor capacidad de computación actual, el aumento de su automatismo y su supuesta autonomía, la mayor disponibilidad de datos y su asociación e interacción con otras tecnologías actualmente disponibles, la están convirtiendo en un recurso esencial para el futuro de la humanidad, pero también podría ser un "arma" de potencial lesivo dantesco en manos equivocadas.

La realidad que pretende definir el concepto de inteligencia artificial debe ser revisado.

Como he anticipado anteriormente, el concepto "inteligente", en sus distintas acepciones, se halla asociado y vinculado desde hace muchos años a sistemas, equipos, aplicaciones, dispositivos, componentes y elementos de nuestra vida diaria como hogares, vehículos o robots, desde el teléfono o la televisión, hasta la aspiradora.

Sin embargo, durante estos últimos años, conceptos como "smart" asociado a nivel de marketing y publicidad a cualquier producto o servicio que quisiera posicionarse en la mente del consumidor como vanguardista, ha ido perdiendo ya su carácter diferenciador y empieza a disiparse en la medida que se le presupone ya esta característica a los productos o servicios sustentados en tecnología.

El atributo ha sido progresivamente sustituido a nivel de marketing por "Inteligencia artificial".

En este sentido conviene recordar que el marketing es una batalla de percepciones, no de productos[6], y que se libra en la mente del consumidor y no en el punto de venta.

La inteligencia artificial se ha instaurado en primera línea en nuestra mente, protagoniza noticias en medios de comunicación de manera incesante, centra el discurso público y es objeto de estrategias y planes estatales y regionales para su desarrollo y aplicación.

De hecho, en mi opinión, considero que se está haciendo un uso desmedido del concepto "Inteligencia artificial" a nivel de marketing y comunicación, y no sólo por parte de empresas y profesionales, sino de algunos gobiernos, que persiguen transmitir un mensaje de máxima innovación, vanguardia, etc.

En ocasiones, se transmiten y se asocian capacidades y atributos a los sistemas que integran inteligencia artificial que se encuentran muy alejados de lo que la realidad tecnológica actual permite y lejos todavía de alcanzarlas.

[6] RIES, A. Y TROUT, J. (1993). *Las 22 leyes inmutables del marketing*. McGraw-Hill, 1993, P. 9.

Y, es más, el lanzamiento de nuevos sistemas que no hagan referencia o uso de la inteligencia artificial en alguna de sus modalidades y técnicas, como *Machine Learning* o *Deep Learning*, puede privarlos de relevancia inicial para el mercado[7].

En paralelo, el desarrollo de los sistemas y aplicaciones dotados de inteligencia artificial están generando usos y aplicaciones que únicamente habíamos ideado en el cine o la literatura, pero una vez más, y como tantas otras, el ser humano ha sido capaz de mostrar, de un lado, su absoluta grandeza al convertir la ficción soñada en realidad y, de otro, su deseo social incesante por innovar.

Si viajar a la luna o navegar en un submarino hace dos siglos solo existía en la mente de soñadores y obras literarias, un siglo después se convirtió en realidad. Si viajar en vehículos sin conductor o taxis aéreos hace treinta años sólo existía igualmente en la mente de soñadores y obras cinematográficas, apenas treinta años después ya es una realidad.

Sin duda, predicciones inconscientes que se han cumplido como las de Julio Verne, que incluso llegó a asegurar que el mundo estaría conectado de algún modo a mediados del siglo XX, o más científicas como las de Arthur C. Clarke que, a finales de los años sesenta, hablaba de la conexión de las computadoras nivel mundial y de la que surgiría una nueva manera de entender la sociedad, el empleo o la vida. Y otras posteriores que ya nos mostraban situaciones de convivencia entre humanos y robots, como por ejemplo Asimov con su obra "*Yo, robot*" de 1950.

Durante los últimos años, la literatura y, especialmente, el cine, se han convertido también en fuentes de estímulo y motivación incesantes para el ser humano y para su capacidad inventiva y creativa, y que, a su vez, han trasladado anticipadamente a la sociedad el debate sobre los retos éticos y de seguridad que el desarrollo, aplicación y uso de la inteligencia artificial plantea.

[7] GONZÁLEZ, S.E. (2020). "La inteligencia de las sumas y las restas". *Inspiring Blog Tecnalia*. Publicado el 23 de julio de 2020. Disponible en: http://blogs.tecnalia.com/inspiring-blog/2020/07/23/la-inteligencia-las-sumas-restas. Consultado el 02.01.2021.

Entre otras, destacar algunas obras cinematográficas como "*Inteligencia Artificial*" (2001), que plantea el desafío ético de implantar emociones humanas en robots, "*2001: Una odisea del espacio*" (1968), "*Yo Robot*" (2004) o "*Blade Runner*" (1982) que abordan, entre otros aspectos, la autonomía de un sistema de inteligencia artificial y su control/descontrol, "*Her*" (2013) que plantea el desafío ético que supone la relación emocional y amorosa entre el ser humano y un sistema de inteligencia artificial, o películas como "*Terminator*" (1984) y "*The Machine*" (2013), donde los sistemas de inteligencia artificial toman conciencia de sí mismos, se autoprotegen y se revelan contra el ser humano, con mensajes transhumanistas[8].

Estamos asistiendo a una absoluta revolución incesante de nuestro mundo gracias a la tecnología.

La inteligencia artificial, como capacidad integrada en los sistemas que gobiernan sus aplicaciones y uso, como robots, máquinas y otros productos, está en constante evolución, desarrollo, despliegue y aplicación en todo tipo de sectores y ámbitos, y los cambios todavía serán más profundos y exponenciales conforme al incremento de su grado de automatización y supuesta "autonomía", el incremento constante de la capacidad de almacenamiento y de computación, la mayor disponibilidad de datos y su interacción y asociación con otras tecnologías.

La inteligencia artificial ha ido cualificándose, superando conceptos asociados a los meros automatismos e incorporando nuevos rasgos adicionales a su definición inicial y nuevas capacidades que, hasta ahora, eran propias del ser humano, en particular, su supuesta autonomía -total o parcial- y relativa independencia en la toma de decisiones o ejecución de acciones, y el aprendizaje profundo, cuya aplicación suponen una revolución para todo tipo de actividades y sectores.

[8] El transhumanismo puede definirse como la corriente cultural e intelectual internacional de corte filosófico que brinda al ser humano un plan de transformación de su condición natural o biológica limitada a otra ilimitada y posthumana, mejor que humana, a través del desarrollo, la fabricación y la aplicación de las nuevas tecnologías robóticas. BARRIO ANDRÉS, M. (2018). *Derecho de los Robots*. Wolters Kluwer España, S.A. Madrid. 2018.

En los próximos años los sistemas inteligentes conducirán nuestros coches, nos explorarán médicamente o nos gestionarán nuestros ahorros y valores, en la medida que la tecnología ya está y falta su aceptación y despliegue a gran escala.

No obstante, como ya he apuntado al inicio, todas estas "bondades" comportan inherentemente una serie de implicaciones tecnológicas, políticas, económicas, sociales, éticas y jurídicas, y también retos y de riesgos asociados a su uso y aplicación que requieren su identificación, evaluación y adecuada gestión, dado su potencial lesivo para los derechos, bienes e intereses en juego, y exige no perder de vista lo que debería ser su objetivo esencial, esto es, resolver problemas, atender necesidades y mejorar la vida del ser humano, mediante la innovación y el desarrollo tecnológico.

Los retos que plantea la inteligencia artificial a nivel social, político, económico, filosófico y ético, así como jurídico, son claves para el desarrollo de nuestra sociedad.

Sobre estas cuestiones han proliferado las predicciones y vaticinios más variados desde todo tipo de ópticas, y se han generado algunos mitos que difieren con cualquier realidad presente o futura posible conforme a la tecnología disponible, que algunos expertos y divulgadores se han encargado de analizar y desmentir como Tegmark[9].

La inteligencia artificial, como el resto de tecnologías creadas por el ser humano, puede derivar en aplicaciones o usos beneficiosos y también perjudiciales para el ser humano como, por ejemplo, lo fue en su día la pólvora o la energía nuclear.

La aplicación y uso de la tecnología en cualquier sentido no la califica como neutral, tampoco su diseño.

[9] Entre otros expertos, destacar a Max Tegmark, profesor de física del Massachussets Institute of Technology (MIT), director científico del Foundational Questions Institute y cofundador del Future of Life Institute: TEGMARK, M. (2018). *Vida 3.0: Qué significa ser humano en la era de la inteligencia artificial*. Penguin Random House Grupo Editorial. 2018. Pp. 57-63. Asimismo, destacar distintas reflexiones en este sentido de Richard Benjamins e Idoia Salazar en su libro "*El mito del algoritmo"*, que constituyen el hilo conductor del mismo y que cito en diversas ocasiones a lo largo de esta investigación.

El desarrollo tecnológico no es algo inocuo, sino que responde a unos objetivos previos, responde a un proceso de toma de decisiones y es diseñado por el ser humano, por lo que simplemente por este motivo puede comportar riesgos o resultar perjudicial.

La creación de sistemas inteligentes capaces de tomar decisiones y realizar acciones de forma supuesta y aparentemente "autónoma" exige, en mi opinión, integrar en los mismos referencias éticas así como legales y de seguridad, es decir, "líneas rojas" o marcos de conducta esperada, con el objetivo de que las mismas sean llevadas a cabo y respetadas mediante la inteligencia pretendida, como lo haría el ser humano, previendo las consecuencias asociadas a su infracción, incluyendo la reparación íntegra y efectiva de los daños y perjuicios derivados de la misma.

Durante los últimos años han sido múltiples las iniciativas a nivel internacional dirigidas a crear estos marcos éticos de referencia, al objeto de que cualquier sistema de inteligencia artificial los integre en su diseño y concepción, especialmente para garantizar la confiabilidad en los mismos y la seguridad jurídica para todas las partes involucradas.

No obstante, como abordaré a lo largo de esta investigación, los meros principios éticos y estándares de buenas prácticas se han mostrado insuficientes hasta la fecha, especialmente ante los diferentes objetivos inmediatos y mediatos de los distintos agentes intervinientes en la creación, explotación, aplicación y uso de la inteligencia artificial.

Por ello, anticipando ya mi opinión y algunas conclusiones de esta investigación, a mi juicio, sólo por este motivo, precisamos necesariamente nuevos marcos reguladores que establezcan el carácter vinculante y obligatorio de principios y normas éticas esenciales para los sistemas inteligentes, sin perjuicio de las normas y obligaciones jurídicas particulares que deban regir sistemas inteligentes considerados de riesgo significativo o alto riesgo pero para todos los agentes susceptibles de verse afectados por los mismo, esto es, personas, entidades públicas o privadas o incluso gobiernos y Estados.

Estos marcos regulativos deberán contemplar, adaptarse y proporcionar soluciones a los desafíos que plantea la inteligencia artificial, y considero que deben erigirse como marcos equilibrados, flexibles, dinámicos, adaptativos, evolutivos y *responsive*, que regulen los principios, requisitos, derechos y obligaciones jurídicas relacionadas con el diseño,

desarrollo, explotación, despliegue, funcionamiento, aplicación y uso de la inteligencia artificial y su asociación con otras tecnologías, si bien sólo en aquello que sea necesario, que complementen, modifiquen y/o sustituyan los aspectos necesarios de los marcos vigentes, y que establezcan el carácter vinculante de los marcos, principios y normas éticas que se consideren esenciales, sin perjuicio de promover la adhesión a códigos de conducta y buenas prácticas que regulen aspectos menos sustanciales cómo el modo de cumplir aquellas normas jurídicas y éticas vinculantes, así como otros aspectos de adscripción voluntaria y de menor criticidad para los derechos, bienes e intereses en juego. Es decir, un marco que convine equilibrada y adecuadamente el denominado *hard* y *soft law*.

Las decisiones de negocio a nivel empresarial se rigen principalmente por aspectos económicos, pero si realmente queremos salvaguardar todos los intereses en juego, deben establecerse normas y pautas claras, concretas, precisas, imperativas y no dispositivas a nivel ético, dado que su establecimiento como meros compromisos voluntarios se han mostrado insuficientes en otras materias, como ejemplificaré a lo largo de esta investigación.

Estos nuevos marcos deberían concretar los aspectos imperativos y dispositivos, así como aquéllos que puedan dejarse a la industria para su autorregulación, la creación de autoridades técnicas para crear estándares y guía más que meramente supervisoras, la creación de mecanismos de evaluación y control públicos y/o privados, un régimen sancionador en caso de incumplimiento, así como la definición del marco regulador de la responsabilidad por daños derivados de la inteligencia artificial que, en función del país, podrá ser sustitutivo, modificativo o complementario al existente, que se acompañe de mecanismos complementarios que garanticen el resarcimiento efectivo de la persona afectada, como seguros obligatorios y fondos de compensación.

Los futuros marcos regulatorios deberán ser construidos teniendo en cuenta los intereses y la protección tanto individual como la colectiva.

La UE ha apostado por regular todo ello desde una perspectiva proteccionista de sus intereses comerciales y de los derechos de los ciudadanos.

Esta investigación pretende profundizar sobre la evolución y estado actual de la inteligencia artificial, los retos y riesgos de seguridad, éticos y jurídicos que plantea, el marco legal aplicable a la misma, su adecuación ante una realidad tan compleja -especialmente en materia de responsabilidad civil por daños- y las propuestas actuales de regulación a nivel europeo, para llegar al análisis de las posibles claves de los futuros marcos reguladores.

El resultado de este análisis pretendo reflejarlo en este documento, huyendo de excesivos tecnicismos propios de la filosofía, la ciencia, la tecnología, el derecho y la ciberseguridad como he indicado en su prólogo, para centrarlo principalmente en su dimensión ética, jurídica y de seguridad.

Del mismo modo, pretendo evitar hacer referencia, en la medida de lo posible, a marcas, productos y empresas, sin perjuicio de que haga referencia o pueda aludir a algunas de ellas a modo ilustrativo o como ejemplo de cuestiones específicas abordadas, así como a aplicaciones concretas de sistemas específicos en contextos determinados, dado que el análisis de los retos y riesgos de sólo una de ellas podría ser el objeto de una investigación de esta naturaleza.

2. Origen y evolución histórica del concepto y de la tecnología subyacente

La inteligencia artificial es uno de esos conceptos del que todo el mundo ha oído hablar y buena parte utiliza, pero del que pocos saben exactamente en qué consiste.

La historia recoge la constante inquietud del ser humano por reproducir comportamientos inteligentes en máquinas, desde los primeros juegos matemáticos, como el de las Torres de Hanói (hacia el 3.000 a.C.), los planteamientos sobre máquinas inteligentes de Aristóteles sobre el año 322 a.C., o Herón de Alejandría con sus conocimientos sobre

robots o su invención de la Eolípila (precedente histórico de lo que luego sería la máquina de vapor) en el siglo I d.C[10].

El mundo científico sitúa mayoritariamente el nacimiento de la inteligencia artificial como área de investigación en 1.956, en la llamada Conferencia de Dartmouth[11], en la que se reunieron los principales investigadores estadounidenses del campo de la informática y de la psicología cognitiva para trabajar durante los meses de verano.

Durante la misma, John McCarthy, coorganizador, destacado profesor e investigador a nivel mundial en materia de inteligencia artificial, fundador del *Laboratorio de inteligencia artificial del MIT* junto con Marvin Minsky, y al que se atribuye también el concepto de "*cloud computing*"[12], acuñó por primera vez el término "Inteligencia artificial" en dicho evento.

La conferencia fue organizada por John McCarthy (Dartmouth College, New Hampshire), Marvin L. Minsky (Harvard University), Nathaniel Rochester (IBM Corporation) y Claude E. Shannon (Bell Telephone Laboratories) y propusieron unirse a la misma a un grupo de investigadores que quisieran trabajar sobre la hipótesis de que cada aspecto del aprendizaje y cada característica de la inteligencia podían ser tan precisamente descritos que se podían crear máquinas que las simularan, es decir, en su génesis, el objetivo de la investigación era crear sistemas que simularan la inteligencia humana, y 65 años después seguimos con el mismo anhelo.

La convocatoria del encuentro se encabezó del siguiente modo: "Proponemos que se lleve a cabo un estudio de inteligencia artificial de 2 meses y 10 hombres durante el verano de 1956 en el Dartmouth College de Hanover, New Hampshire. El estudio se basará en la conjetura de que todos los aspectos del aprendizaje o cualquier otra característica de la inteligencia pueden, en principio, describirse con tanta precisión que se puede hacer que

[10] RAINER, J.J. Y RODRÍGUEZ, L. (2019), "Perspectiva histórica y evolución de la inteligencia artificial". En *Documento de Seguridad y Defensa 79: La inteligencia artificial aplicada a la defensa*. Instituto Español de Estudios Estratégicos. Marzo 2019. Pp. 19-20.

[11] MCCARTHY, J.; MINSKY, M.; ROCHESTER, N.; SHANNON, C. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. 31.08.1955. Recuperado de: http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html. Consultado el 24.11.2020.

[12] *El origen de: El Cómputo en la Nube*. Recuperado de https://www.fayerwayer.com/2012/01/el-origen-de-el-computo-en-la-nube/. Consultado el 24.11.2020. *John McCarthy Biography*. Recuperado de http://www.bookrags.com/biography/john-mccarthy-wcs/#gsc.tab=0. Consultado el 24.11.2020.

una máquina los simule. Se intentará encontrar cómo hacer que las máquinas utilicen el lenguaje, formen abstracciones y conceptos, resuelvan tipos de problemas ahora reservados para los humanos y se mejoren. Creemos que se puede lograr un avance significativo en uno o más de estos problemas si un grupo de científicos cuidadosamente seleccionados trabajan juntos durante un verano"[13].

La convocatoria fue aceptada y se unieron a la misma Ray Solomonoff (fundador de la rama de la inteligencia artificial basada en el aprendizaje automático, la predicción y la probabilidad), Oliver Selfridge, Trenchard More, Arthur Samuel (pionero en el campo de los juegos informáticos y la inteligencia artificial), Herbert Simon y Allen Newell.

El presupuesto de base de aquel encuentro científico se basaba en la hipótesis anteriormente citada, la posibilidad de que "cualquier aspecto del aprendizaje y elemento de inteligencia puede, en principio, ser descrito de manera tan precisa que sea posible crear una máquina que lo emule".

Y conforme constaba ya en su convocatoria, se evidenciaba el nacimiento de una nueva disciplina orientada a la creación de máquinas capaces de replicar en las mismas la capacidad humana de emplear el lenguaje, de aprender y razonar creativamente. Un objetivo claramente definido, aunque de absoluta complejidad.

No obstante, resultaría imperdonable no hacer referencia obligada a Marian Rejeswky, Jerzy Eozyeahi y Henryk Zygalski, cuyos trabajos en las décadas previas fueron continuados por el matemático e informático Alan M. Turing, el considerado padre de la computación -con el permiso de la matemática Ada Byron, creadora en el siglo XIX del primer algoritmo destinado a ser procesado por una máquina que no llegó a convertirse en realidad-, y también considerado por otros el padre de la inteligencia artificial.

Turing hablaba ya de inteligencia artificial en su artículo *Computing Machinery and Intelligence* en 1.950 y su *The Imitation Game*, en el que ponía a prueba la capacidad de

[13] MCCARTHY, J.; MINSKY, M.; ROCHESTER, N.; SHANNON, C. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. 31.08.1955. Recuperado de: http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html. Consultado el 24.11.2020.

una máquina para exhibir un comportamiento inteligente similar al de un ser humano o *indistingible* de éste mediante la denominada *Prueba de Turing*.

Según el mismo, se consideraría "inteligente" a un programa cuando en interacción con un humano, no fuésemos capaces de distinguir entre las respuestas de una máquina y de un humano. La primera pregunta de este juego nos sitúa ante la primera reflexión científica y en los prolegómenos de toda una seria de cuestiones éticas y filosóficas: ¿Puede una máquina pensar?

Si situamos cronológicamente las bases de la inteligencia artificial en los años 50, las dos décadas posteriores, sin embargo, no tuvieron avances especialmente significativos, a pesar de las ambiciosas predicciones de algunos de los participantes en aquella conferencia, como Herbert Simon, que afirmaba que en veinte años las máquinas serían capaces de llevar a cabo cualquier tipo de trabajo que un hombre pudiera hacer. Estas predicciones las he escuchado personalmente en diversas ocasiones muchas décadas después y ni se cumplieron ni se han cumplido. Así mismo, expondré muchas más que tampoco lo han hecho hasta la fecha, entre las que destacan algunas que quizás, en un futuro no tan lejano, podrían cumplirse.

Durante este período aparecieron sistemas novedosos como "Eliza"[14], un sistema de supuesta inteligencia artificial que simulaba ser una especie de *coach*, permitiendo un diálogo con una persona y realizando preguntas para llevar a la misma a la solución. Se percibía entonces como un sistema inteligente partiendo de su resultado, pero no dejaba de ser de un conjunto de reglas y respuestas predefinidas. Seguramente hoy, la percepción social de esta herramienta como inteligente sería muy diferente.

Durante los años ochenta, la inteligencia artificial volvió a tener protagonismo gracias al éxito comercial de los denominados "sistemas expertos", que emulaban las capacidades analíticas y la toma de decisiones humanas. Algunos de estos sistemas empezaron a utilizarse en el ámbito de la gestión empresarial, especialmente para procesar órdenes y

[14] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Editorial Anaya Multimedia. Madrid 2020. Pos. 3245 (Edición Kindle).

pedidos, con el consiguiente ahorro. Un ejemplo de ello fue el programa "R1"[15] desarrollado por John P. McDermott.

Durante todos estos años se apreciaba un progreso constante, pero lento de la inteligencia artificial, lo cual comenzó a cambiar drásticamente en los años posteriores.

A finales de la década de los noventa y primera década del siglo XXI, se inició el despliegue y aplicación de la inteligencia artificial en sectores como la logística, la minería de datos o el diagnóstico médico.

La aceleración en su desarrollo y aplicación se produjo principalmente gracias al aumento de la potencia de cálculo y de los datos.

Uno de los puntos de inflexión en la inteligencia artificial y del aprendizaje profundo asociado a la misma se produjo gracias a los nuevos modos de entrenar de manera eficaz las nuevas capas de redes neuronales promovidos por el investigador Geoffrey Hinton.

Y todo ello conviviendo con un paralelo y ansioso afán de muchos por crear sistemas que superarán al ser humano, más inteligentes, más rápidos y más fuertes que el mismo, generando una especie de competición continua entre hombre-máquina y que se proyectaba en exhibiciones públicas como las partidas de ajedrez entre Garry Kasparov y "Deep Blue" de IBM durante 1996 y 1997, escenificadas como un enfrentamiento entre los EE.UU. y Rusia y que se disputaron hasta que el sistema inteligente ganó, interpretándose por algunos como el comienzo del fin de la supremacía de la raza humana.

A los logros de "Deep Blue" en el ajedrez, sucedieron en 2011 los de "Watson" -sistema de inteligencia artificial de IBM- en el concurso televisivo Jeopardy!. En 2016 los de "AlphaGo" -el sistema de inteligencia artificial de Google, "Deepmind"-, enfrentándose en este caso contra Lee Sedol -campeón del mundo del "Go"-[16], y posteriormente las de

[15] MCDERMOTT, J. P. (1982). "R1: A rule-based configurer of computer systems". *Artificial Intelligence*. Vol. 19, Issue 1. Elsevier. September 1982. Pp. 39-88.
[16] SILVER, D., HUANG, A., MADDISON, C. ET AL. (2016). "Mastering the game of Go with deep neural networks and tree search*". Nature* 529. Pp. 484-489. 2016. https://doi.org/10.1038/nature16961.

su sucesor "AlphaZero", sistema sin capacidad de proporcionar explicaciones de sus decisiones y sin supervisión humana.

Durante estos años y ante algunas de estas exhibiciones, me planteaba la esencia y objetivos primigenios de ésta y otras tecnologías, y si nos estábamos alejando de los mismos en nuestro afán por superar al ser humano. Nunca he considerado que su objetivo deba ser la generación de sistemas más inteligentes que puedan sustituir o superar al ser humano, sino sistemas que aumenten nuestras capacidades, que nos mejoren en aquello que sea necesario, que nos proporcionen valor, que nos complementen, que solucionen o nos ayudan a solucionar nuestros problemas, que satisfagan o nos ayuden a satisfacer nuestras necesidades y que mejoren nuestra vida y nuestro mundo, con independencia del grado de supuesta "Inteligencia" integrada en el sistema, comparable o no a la del ser humano.

Por el camino se proseguía consolidando ese halo absurdo de competición continua hombre-máquina en distintas esferas, que conformaba una especie de densa niebla que, en cierto modo, ocultaba a la sociedad en general las aportaciones potenciales de la inteligencia artificial a la humanidad, en un matrimonio indisoluble hombre-máquina, y que ya dejaba vislumbrar algunas visiones apocalípticas para cuando las máquinas superaran al hombre y fuesen capaces de mejorarse y crearse a sí mismas, es decir, al alcanzar la denominada "singularidad tecnológica[17]".

[17] La "singularidad" en el ámbito de la IA se concibe como el momento en que la IA superará a la inteligencia humana, cuyos defensores se basan en las especulaciones del matemático Irving J. Good que en 1965 definió una máquina ultrainteligente como aquella capaz de superar con creces a cualquier ser humano, por inteligente que sean llevando a cabo cualquier actividad intelectual. Recuperado DE BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Editorial Anaya Multimedia. Madrid 2020. Pos. 3343 (Edición Kindle). No obstante, el origen del término no es pacífico. La expresión "singularidad tecnológica se atribuye al matemático y físico húngaro John von Neumann, que en 1958 ya hablaba de un momento en el que las máquinas serían autosuficientes y de un mundo automático, si bien, no se popularizó hasta 1984 cuando el matemático y literato Vernor Vinge, lo empleó en su novela "*La guerra de la paz".* Recuperado de VIDAL, M. (2019). *La era de la humanidad*. Pos. 6005. Versión Kindle. Ediciones Deusto. Barcelona. 2019.

Y en este contexto, se producían nuevas aplicaciones especialmente significativas, como la "supuesta" primera operación quirúrgica completa de un robot cirujano y anestesiólogo robot "McSleepy"[18] en Montreal en 2010.

En los últimos años se ha producido la eclosión de la inteligencia artificial y de su implementación[19], con su aplicación en todo tipo de sectores, gracias la mayor capacidad de computación y de almacenamiento, la irrupción e interrelación con otras tecnologías, el aprendizaje profundo, la cualificación de ingenieros en inteligencia artificial, su asociación con otras disciplinas como las matemáticas, la economía y la estadística y, sobre todo, el Big data.

Es el momento de implementar todo lo investigado y desarrollado hasta la fecha, lo que cambiará el mundo que conocemos.

Andrew Ng[20], experto de referencia mundial en inteligencia artificial, comparó la inteligencia artificial con la electricidad y de este modo revolucionará todos los sectores: "Es la mayor revolución desde la introducción de la electricidad hace 100 años. No veo ningún sector que no vaya a transformar a medio plazo".

Durante el Foro Económico Mundial en Davos (Suiza), celebrado en enero de 2018, Sundar Pichai, CEO de Google, aseguró que "el avance de la inteligencia artificial (IA) hasta el momento, es lo más importante con lo que la humanidad ha trabajado alguna vez, incluso por arriba del descubrimiento del fuego o la electricidad". Esta afirmación no sólo la ha mantenido, sino que la ha ampliado en una reciente entrevista, considerando que supondrá un cambio "más profundo que el fuego, la electricidad o internet"[21].

---

[18] VIDAL, M. (2019). *La era de la humanidad*. Posición 2736. Versión Kindle. Ediciones Deusto. Barcelona. 2019.

[19] Kai-Fu Lee denomina a este período "la era de la implementación" a la que está sucediendo "la era de los datos". Recuperado de LEE, KAI-FU (2018). *Superpotencias de la inteligencia artificial*. Op.cit. Pos. 24. Deusto 2018.

[20] JEWELL, C. (2019). "Artificial intelligence: The new electricity". Publicado en *WIPO Magazine*. Junio 2019.

[21] RAJAN, A. (2021). "La inteligencia artificial supondrá un cambio "más profundo que el fuego, la electricidad o internet": Sundar Pichai, líder de Google". Publicado por *BBC News* el 13.07.2021. Recuperado de: https://www.bbc.com/mundo/noticias-57809469. Comsultado el 15.07.2021.

Y durante estos últimos años y hasta la fecha han seguido creciendo las "competiciones" hombre-máquina, pero con otras finalidades más relevantes para el ser humano, entre otras, la de salvar vidas.

A modo de ejemplo, en 2018, el hospital de Beijing Tiantan en China llevó a cabo un estudio -sin ánimo por mi parte de frivolizar sobre estos temas, sino de evidenciar su relevancia científica y valor para la humanidad-, que enfrentaba a su sistema de inteligencia artificial "Biomind" contra 15 de los mejores especialistas en cáncer del país[22], mediante la visualización de imágenes, debían diagnosticar correctamente tumores cerebrales.

En la primera prueba, la tasa de acierto del sistema fue del 87% con un mínimo tiempo de respuesta de minutos, el de los médicos del 66%. En la segunda prueba para diagnosticar casos de expansión de hematoma cerebral, la tasa del sistema fue del 83% y la de los médicos del 63%.

Simultáneamente los avances han sido incesantes.

El sistema denominado "Watson" de IBM evidenció su eficiencia en análisis de grandes volúmenes de datos de pacientes con cáncer, permitiendo predecir y personalizar tratamientos individualmente de manera inmediata y reduciendo su toxicidad[23]. De hecho, una evolución algorítmica de Watson permitió tratar de forma efectiva una leucemia mieloide en una paciente sobre la que hasta entonces ningún tratamiento estándar había dado resultado.

Google también se ha unido a la investigación de la aplicación de la inteligencia artificial en medicina a través de Google Health.

Pero no siempre el sistema inteligente es mejor.

[22] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Op.cit. Pos. 1526.
[23] MATHESON, R. (2018). "Artificial intelligence model 'learns' from patient data to make cancer treatment less toxic". Publicado en *MIT News*, 09.08.2018. Disponible en: https://news.mit.edu/2018/artificial-intelligence-model-learns-patient-data-cancer-treatment-less-toxic-0810. Consultado el 18.02.2021.

El prestigioso *International Symposium on Biomedical Imaging* -ISBI por sus siglas en inglés-[24] evaluó las aptitudes de máquinas y humanos a la hora de diagnosticar cáncer de pecho, concluyendo que los patólogos humanos lograron una diagnosis correcta un 3% superior a la obtenida por las máquinas. No obstante, el resultado más sorprendente fue observar que, al permitir a los humanos recibir asistencia de las máquinas, el error de los patólogos se reducía en un 85%, llegando a alcanzar una precisión final en la diagnosis superior a1 99%.

No sería "inteligente" sustituir a médicos humanos por máquinas, ni tampoco lo sería como humanos ignorar las aportaciones con las que la inteligencia artificial puede contribuir a anticipar diagnosis y prevenir enfermedades.

En mi opinión, el futuro exige pasar de la competición a la simbiosis, a una relación estrecha y permanente entre hombre y máquina que comparten un espacio común, que requieren de una colaboración continua y que siguen objetivos fijados por el ser humano, no por la máquina, de modo que la inteligencia artificial y las tecnologías asociadas estén alineadas con la consecución de los objetivos predefinidos por nosotros.

La consecución de los mismos exige el trabajo conjunto de gobiernos, corporaciones, universidades y centros de investigación para garantizar la innovación, la protección, la confianza, la seguridad y la educación sobre la inteligencia artificial.

Las predicciones indican un desarrollo exponencial de la inteligencia artificial en los próximos años gracias a las inversiones públicas y privadas en la misma, las expectativas de rentabilidad depositadas y la capacidad transformadora que se le atribuye.

La evolución del *Machine Learning* (aprendizaje automático), el *Deep Learning* (aprendizaje profundo), la capacidad de procesamiento de lenguaje natural y la robótica marcarán el desarrollo y aplicación de la inteligencia artificial durante los próximos años.

La inteligencia artificial está cada vez más presente en todos los ámbitos de nuestra vida y en productos y servicios más cotidianos, aportando mayor eficiencia en todos los

[24] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Op.cit. Pos. 1720.

procesos que la integren, permitiendo el surgimiento y evolución de una medicina predictiva ágil y eficaz, tratamientos médicos precoces personalizados, procedimientos judiciales más agiles y sentencias judiciales más rápidas.

Pero todo ello no es "gratuito", comporta riesgos que deberemos ser capaces de identificar y gestionar adecuadamente.

Hoy, la inteligencia artificial es identificada como una de las tecnologías de mayor impacto con un potencial abrumador para la mejora de la vida humana y del mundo en el que vivimos.

Es considerada una de las denominadas "DARQ"[25], junto a la Realidad Virtual/Extendida, Blockchain y la Computación Cuántica. De hecho, se considera uno de los pilares básicos de la cuarta revolución industrial junto con las tecnologías de mayor impacto como las precitadas, que permiten generar nuevos modelos de organización, interacción y de negocio totalmente disruptivos, y que será la protagonista de la quinta revolución industrial, si es que consideramos que no estamos ya en ella, dada la velocidad y profundidad del cambio constante en el que vivimos.

Actualmente, como he referido anteriormente, se ha iniciado una carrera mundial por liderar la inteligencia artificial, el gobierno de los datos y tecnologías como el Blockchain y la Computación Cuántica.

En materia de inteligencia artificial, las grandes tecnológicas y superpotencias internacionales como EE.UU. y China se sitúan a la vanguardia, si bien, desde enfoques y objetivos distintos.

EE.UU. ha apostado más por el desarrollo y la innovación y China por su despliegue, implantación y aplicación masiva, sustentándose en su mayor disponibilidad de su principal "fuente de energía": los datos. Además, la ventaja en este sentido de China no sólo es cuantitativa -por ejemplo, el número de usuarios de internet del país es mayor que

[25] Acrónimo de *Distributed ledgers* (Blockchain), *Artificial Intelligence* (inteligencia artificial), *Extended Reality* (Realidad Extendida) y *Quantum Computing* (Computación Cuántica), identificadas como los grandes catalizadores de la tecnología para el cambio.

el de EE.UU. y Europa juntos- sino cualitativa -tanto datos de la dimensión física como de la virtual-. La sociedad china acepta más fácilmente el acceso y tratamiento de sus datos y el marco regulador facilita igualmente todo ello.

Algunos expertos, como Ricardo Moya, considera que EE.UU. sigue liderando mundialmente la inteligencia artificial, si bien, vaticina que China podría adelantarle en los próximos años. Según estudios como *The Global AI Index*[26] el liderazgo de EE.UU. se sitúa en ámbitos como el talento, la infraestructura, la investigación e inversiones empresariales, si bien China empieza a liderar el desarrollo y la aplicación.

Además de todo ello, China estableció como objetivo convertirse en una superpotencia de la inteligencia artificial y promovió una fuerte inversión, apoyo político y coordinación nacional para la investigación, desarrollo y avance de la inteligencia artificial, especialmente intenso desde 2017, previendo que para 2030 se convertiría en el centro de la innovación global en inteligencia artificial, liderando tanto la tecnología como su aplicación.

De hecho, en 2017, los inversores de capital riesgo chino invirtieron en nuevas empresas de inteligencia artificial una cantidad equivalente al 48% de toda la financiación de capital riesgo en inteligencia artificial en todo el mundo, superando a EE.UU.[27].

Las claves del posicionamiento de China se hallan principalmente en la enorme disponibilidad de datos, la mentalidad, perseverancia y voracidad empresarial, el conocimiento, el número de investigadores cualificados en inteligencia artificial, un entorno normativo propicio y el entorno político totalmente volcado en apoyar la misma -ayudas, subvenciones, campañas para el fomento del espíritu emprendedor, empresarial y de innovación masiva, creación de incubadoras, zonas de emprendimiento, fondos de orientación, incentivos a la inversión, beneficios fiscales y simplificación de trámites para iniciar un negocio-[28].

[26] Recuperado de: https://www.tortoisemedia.com/intelligence/global-ai/. Consultado el 14.12.2020.
[27] LEE, KAI-FU (2018). *Superpotencias de la inteligencia artificial*. Op.cit. Pos. 13.
[28] LEE, KAI-FU (2018). *Superpotencias de la inteligencia artificial*. Op.cit. P. 26.

Todas estas claves parecen colocar a China en este momento en un plano aventajado respecto al resto de países. De hecho, China ha superado a EE.UU. en volumen de producción de datos, lo que proporciona a las empresas chinas una importante ventaja en el desarrollo de servicios basados en inteligencia artificial.

La UE está bien posicionada, tanto como usuaria como creadora y productora de sistemas de inteligencia artificial.

Según el *Libro blanco sobre la inteligencia artificial*[29] elaborado por la Comisión Europea, de 19.02.2020, considera que Europa es líder mundial en los sectores de la robótica, la fabricación y los servicios competitivos en inteligencia artificial, dispone de una infraestructura de computación sólida y posee un volumen de datos públicos y de la industria cuyo potencial está siendo infrautilizado en la actualidad.

Quizás algunas de las potencias mundiales en inteligencia artificial no están del todo de acuerdo con la primera aseveración en relación con su liderazgo mundial, pero sí en lo segundo, de ahí las estrategias de accesibilidad y gobierno de los datos de la UE, lo que, a mi juicio, constituye posiblemente una de las principales debilidades de la UE en la carrera por liderar la inteligencia artificial a nivel mundial.

El *Libro blanco sobre la inteligencia artificial* precitado hace una lectura muy clara de la realidad actual en relación con el acceso y tratamiento de los grandes volúmenes de datos, que, como he manifestado, constituye una debilidad para la UE frente a otras potencias y una desventaja competitiva para la misma y las empresas europeas. Según recoge el mismo, el volumen de datos producido en el mundo va en aumento rápidamente de modo que de los 33 zetabytes en 2018 pasaremos, según las últimas previsiones que cita el mismo[30], a 175 zetabytes en 2025.

Este documento también afirma que Europa produce más de un cuarto de todos los robots de servicios industriales y profesionales que se utilizan, situándose también a la vanguardia de la utilización de la inteligencia artificial en la fabricación, en concreto, "más de la mitad de los mayores fabricantes aplican al menos un elemento de inteligencia

[29] COM (2020) 65 final.
[30] IDC 2019

artificial en sus operaciones de fabricación" por delante de Japón (30%) y los EE.UU. (28%), según el informe tomado como referencia por el mismo de Capgemini de 2019.

Sin embargo, el precitado documento evidencia también la que considero, la segunda gran debilidad europea, que es la inversión actual en investigación e innovación en Europa, mucho menor que en otras regiones del mundo, destacando que, en 2016 se invirtieron unos 3.200 millones de euros en inteligencia artificial en Europa frente a los 12.100 millones en América del Norte y 6.500 millones de euros en Asia[31].

Las últimas previsiones de la UE sobre el impacto económico de la automatización de los conocimientos, los robots y los vehículos autónomos de aquí a 2025 se halla entre los 6.500 y los 12.000 millones de euros al año[32].

Si realizamos un análisis DAFO[33] de la situación de la inteligencia artificial que refleja el precitado documento, entre las principales debilidades estaría el menor acceso y disponibilidad de datos, la menor inversión, la falta de confianza y la falta de habilidades y competencias.

Entre sus principales fortalezas, su cultura por la seguridad, valores y la protección de los derechos fundamentales, su posicionamiento en la producción y aplicación de la tecnología precitado, su liderazgo en soluciones neuromórficas[34] y lo que creo que será un factor diferenciador y una ventaja competitiva en lo sucesivo, esto es, su apuesta clara por la excelencia científica en el ámbito de la inteligencia artificial, su apuesta por la ética y la regulación mediante la creación de marcos éticos vinculantes y marcos jurídicos específicos, especialmente sobre aspectos como la responsabilidad civil por daños causados por sistemas inteligentes, así como su explicación mediante mecanismos como la combinación del razonamiento simbólico con redes neurales profundas.

[31] Informe *10 imperatives for Europe in the age of AI and automation*. McKinsey. 2017.
[32] Comisión Europea (2019), IPOL (2020)
[33] Análisis de debilidades, amenazas, fortalezas y oportunidades.
[34] Sistemas a gran escala compuesto por circuitos integrados que imitan la arquitectura neuronal biológica.

La consultora PriceWaterhouseCoopers[35] estimó que el despliegue de la inteligencia artificial aportará 15,7 trillones de dólares a la economía mundial para 2030, de los que 7 trillones de ese total, casi la mitad, será para China.

Por lo que se refiere a solicitudes de patentes de inteligencia artificial en el período comprendido entre 1960 y 2019, EE.UU. encabeza la lista con 1.863 solicitudes, seguida de China con 1.085 y por la UE con 1.074, lo que es un elemento indicador, con algunos matices, del grado de protección de la innovación en materia de inteligencia artificial en estos países, que no tiene que ser equivalente necesariamente con su desarrollo, despliegue o aplicación. Durante la última década se ha producido un incremento exponencial de las solicitudes, en particular, más de un 400%.[36]

La UE ha impulsado un modelo de desarrollo de la inteligencia artificial centrado en el ser humano y ha definido recientemente su estrategia en materia de inteligencia artificial, que combina un conjunto de medidas para abordar sus oportunidades y desafíos, enfocadas a generar la necesaria confianza en la tecnología y en su potencial impacto tanto en los ciudadanos como en la sociedad y la economía.

Estas medidas se acompañan de propuestas sobre nuevos marcos normativos en materia ética, responsabilidad civil y propiedad intelectual, orientadas a proporcionar seguridad jurídica a todas las partes implicadas, generar confianza en su uso y promover la innovación, a la vez que se propicie un entorno adecuado para fomentar la investigación, el desarrollo de la tecnología y la creación y consolidación de empresas del sector, especialmente *startups*. También incluye medidas para el fomento de la inversión e inyecciones directas. En definitiva, la UE pretende fomentar una inteligencia artificial centrada en el ser humano lícita, ética, confiable y segura.

La Comisión pretende destinar 20.000 millones de euros al año en este momento para potenciar las inversiones privadas y públicas en tecnología de inteligencia artificial.

---

[35] RAO, A.S. Y VERMEIJ, G. (2017). *Sizing the Prize*. PwC- 27 de junio de 2017. Disponible en: https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf. Consultado el 02.03.2021.

[36] Comisión Europea (2019), IPOL (2020)

3. Definición

3.1. Una aproximación al concepto

De inicio, como he anticipado, considero necesario realizar una aproximación al concepto de "Inteligencia artificial" desde distintos enfoques, dada la falta de consenso en su definición incluso a nivel científico para, posteriormente, abordarlo desde un enfoque estrictamente jurídico.

En primer lugar, debo indicar que no existe una definición consensuada, ni en la literatura científica y ni mucho menos en la escasa literatura jurídica. Y, utilizando las palabras de la filósofa Adela Cortina[37], "definir conceptos que tienen una larga historia no es tarea fácil", y en tecnología y en mi opinión, 65 años pueden ser "una más que larga historia".

En segundo lugar, la inteligencia artificial no hace referencia a una mera tecnología, sino a conjunto heterogéneo de las mismas y a distintos campos de investigación, los cuales varían en función de la amplitud y alcance de la definición desde la que partamos.

Quizás, la forma más sencilla de definir el concepto de "Inteligencia artificial", partiendo exclusivamente de las palabras que componen este concepto, sería la inteligencia llevada a cabo por máquinas o la inteligencia propia de una máquina.

Habitualmente se asocia la capacidad y función de pensar a la inteligencia artificial, pero emular o aparentar hacerlo (pensar) no implica que lo haga o tenga la capacidad de hacerlo. Conforme expondré posteriormente, una definición igualmente sencilla, pero quizás más próxima a lo que pretende definir, sería la inteligencia emulada por máquinas.

No obstante, esta definición básica ya provoca en sí misma una discusión inicial, como lo es, si una máquina puede ser o no "inteligente" o si, a mi juicio, esta conceptualización pretende simplemente referirse a una inteligencia "gestionada por máquinas", pero no a un atributo o característica propia de las mismas.

[37] CORTINA, A. (2015). *Ética mínima. Introducción a la filosofía práctica*. Editorial Tecnos, Madrid 2015, 17ª Edición.

Otros debates paralelos que se podrían suscitar alrededor de este concepto es si todo aquello que comercialmente se presenta al mercado como dotado de inteligencia artificial es "inteligente", y si deberíamos hablar, en lugar de "máquinas", de "sistemas de información" (que pueden estar integrados por *hardware* -servidores, equipos, dispositivos, máquinas, vehículos, robots o aparatos-, *software* -programas informáticos-, algoritmos, información -datos-, entre otros elementos), lo que a mi juicio es más adecuado y ajustado a la realidad tecnológica actual que pretende definir este concepto, dado que puede incluir sistemas inteligentes exclusivamente basados en *software*.

Si consultamos la definición de "Inteligencia" recogida en el Diccionario de la Real Academia de la Lengua[38] nos encontramos con varias acepciones: Capacidad de resolver problemas, capacidad de entender o comprender o, por ejemplo, conocimiento, comprensión o acto de entender. De todas estas capacidades, a excepción de la primera, parece que inicialmente ninguna de las mismas puede ser propia de una "máquina" o "sistema de información", como entender o comprender.

Si partimos de una definición más científica de "Inteligencia", la declaración publicada por el Wall Street Journal firmada por cincuenta y dos investigadores bajo el título *Mainstream Science on Intelligence*[39], la define como una capacidad mental muy general que, entre otras cosas, implica la capacidad de razonar, planificar, resolver problemas, pensar de manera abstracta, comprender ideas complejas, aprender rápidamente y aprender de la experiencia. No es simplemente el aprendizaje de libros, una habilidad académica limitada o la inteligencia para tomar exámenes. "Más bien, refleja una capacidad más amplia y profunda para comprender nuestro entorno: 'comprender', 'dar sentido' a las cosas o 'descubrir' qué hacer".

Si analizamos algunas de los sistemas presentados y concebidos como "inteligentes" durante la historia reciente de la inteligencia artificial, podremos observar como este atributo ha cambiado a lo largo del tiempo, así como nuestra percepción.

[38] https://dle.rae.es/inteligencia
[39] GOTTFREDSON, L.S. (1994). "Mainstream Science on Intelligence". *Wall Street Journal*. 13.12.1994. Disponible en: http://www1.udel.edu/educ/gottfredson/reprints/1997mainstream.pdf. Consultado el 05.01.2021.

El precitado sistema "Eliza" no parece que encaje en el concepto actual de "inteligente".

Otros sistemas más recientes como "Google Duplex" parecen evidenciar un comportamiento inteligente. Se trata de un asistente de voz basado en aprendizaje automático supervisado que gestiona reservas en sitios como restaurantes y peluquerías y que es capaz de llamar telefónicamente y conversar de forma normal con seres humanos sin que estos puedan diferenciar que se trata de una máquina, salvo por el hecho de que se les informe previamente de ello.

Sin embargo, el sistema no sabe qué es un restaurante ni una mesa, ni sabe de lo que está hablando, por lo que respecto del resultado podría considerarse inteligente, pero quizás no tanto por el proceso. Es más, Google confirmó en The New York Times[40] que alrededor del 25% de las llamadas se iniciaban por parte de operadores humanos y el 15% de todas las llamadas acababan necesitando intervención humana.

De una manera más extendida, siguiendo la aproximación al concepto efectuada por Orozco González[41], la inteligencia se concibe como una capacidad de la mente para aprender, entender, valorar y adquirir una concepción de la realidad relativa a una persona, grupo o situación, y tomar una o varias decisiones consecuentes. De este modo permite recoger, analizar y procesa información para elegir la mejor opción frente a un problema o situación compleja. En definitiva, la inteligencia permite ejecutar procesos complejos de captación de datos, análisis, valoración y decisión sobre la base de la información procesada.

De este modo, la inteligencia se conforma por un conjunto de procesos complejos de recogida, tratamiento, análisis, valoración y decisión guiados por una serie de principios,

[40] CHEN, B.X. Y METZ, C. (2019). "Google's Duplex Uses A.I. to Mimic Humans (Sometimes)". Publicado en *The New York Times* el 22.05.2019. Disponible en: https://www.nytimes.com/2019/05/22/technology/personaltech/ai-google-duplex.html. Consultado el 23.02.2021.

[41] OROZCO GONZÁLEZ, M. (2021) "Reflexiones acerca de la relación entre inteligencia artificial y robótica", en ATAZ LÓPEZ, J. Y COBACHO GÓMEZ, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón.* Tomo III. Ed. Aranzadi-Thomson Reuters. 2021. P. 863.

valores, criterios y mandatos ética y legalmente establecidos y asumidos, con unos objetivos específicos.

La inteligencia fue definida igualmente por Stern[42] como la capacidad de resolver los problemas de la vida de forma adecuada, productiva e independiente, considerando que para que un comportamiento pueda ser considerado inteligente debe tener, como mínimo, las siguientes capacidades y competencias: autoconsciencia, razonar y/o inferir correctamente, aprender, responder adecuadamente a los cambios (adaptable), actuar teleológicamente y/o proactivamente, tener suficiente capacidad predictiva, competencia suficiente para resolver problemas de todo tipo y no sólo los algorítmicos, adquirir y aplicar conocimientos y habilidades.

En consecuencia, nos encontramos ante la inexistencia de un consenso en relación con el concepto "Inteligencia", si bien, según Ramalho[43], existe un consenso a nivel filosófico que la inteligencia humana integra un conjunto de elementos entre los que se encuentra la creatividad y, cuando se asocia al software, los expertos hablan de capacidad de juicio, aleatoriedad y autocrítica que deben estar presentes en los sistemas inteligentes, que actualmente disponen de las mismas, pero de manera limitada.

Por lo que se refiere a la relación entre la inteligencia y la creatividad a la que alude este autor, me permito remitirme al análisis de estas cuestiones en el último capítulo de esta investigación, relativo a la protección de las obras e invenciones generadas por sistemas inteligentes.

Si seguimos profundizando en el análisis lingüístico de este concepto y asociamos los conceptos "Inteligencia" y "artificial", podemos partir en una primera aproximación del concepto de "Inteligencia artificial" recogido por el Diccionario de la Real Academia de la Lengua[44], que lo define como la "disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana,

[42] MARTÍNEZ REY, M.A. Y PAZOS SIERRA, J. (2019). "La inteligencia artificial y el derecho: Pasado, presente y futuro", en Monterrosso Casado, E. (Dir.), *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch 2019. Pp. 549.
[43] RAMALHO, A. (2017). "Will Robots Rule the (Artistic) World? A Proposed Model for the Legal Status of Creations by Artificial Intelligence Systems". *Journal of Internet Law.* Julio 2017. SSRN: https://ssrn.com/abstract=2987757. Pp. 3-5.
[44] https://dle.rae.es/inteligencia?m=form#2DxmhCT

como el aprendizaje o el razonamiento lógico", pero científicamente y en la actualidad, podrían ser algunas, nunca todas.

Plaza Penadés[45] define la inteligencia desde el análisis etimológico del concepto proveniente del latín *intellegere*, "compuesto de *inter* 'entre' y *legere* 'leer, escoger', y por tanto hace referencia a esa habilidad humana de analizar todas las posibilidades y escoger la que se cree, en ese momento y según las circunstancias, que es más adecuada", y considera que este proceso es el que pueda ser ahora auxiliado por la inteligencia artificial.

Siguiendo a este autor, efectivamente la capacidad de analizar las distintas posibilidades y elegir la más correcta puede ser programada y dirigida por aplicaciones que permiten el desarrollo de la inteligencia artificial, incluso mediante autoaprendizaje de la mejor de las elecciones posibles. Adicionalmente a todo ello, en mi opinión, se trataría de una capacidad de elección conferida por el ser humano, que no estaría sujeta a determinados condicionantes o limitaciones del mismo, aunque debería estarlo a otras condiciones y sujeto a riesgos inherentes o adquiridos asociados a la inteligencia artificial como, por ejemplo, su falta de conectividad, error, entendimiento o relativa impredecibilidad en función de su autonomía.

Inicialmente, como he referido, desde un enfoque lógico, una máquina no siente, no puede realizar operaciones afectivas, no tiene consciencia. Quizás uno de los temores más arraigados sobre el futuro de la inteligencia artificial sea precisamente ese concepto de "inicialmente" y si realmente todas estas limitaciones podrían ser superadas y ser una realidad en el futuro.

La inteligencia artificial actual es principalmente una inteligencia artificial "estrecha" o "débil" como expondré más adelante, que puede llevar a cabo una tarea muy concreta con un excelente resultado, pero no tiene consciencia de la tarea que está ejecutando. Otra cosa es la denominada inteligencia artificial "fuerte" o "general", similar o superior a la

[45] PLAZA PENADÉS, J. (2018). "Primeras reflexiones desde el Derecho sobre la inteligencia artificial". *Revista Aranzadi de Derecho y Nuevas Tecnologías*. N. 47, Editorial. Thomson Reuters (Legal) Limited.

del ser humano, que integre sentido común, capacidad de inferencia lógica o de entender relaciones causales.

En definitiva, considero más acertado hablar en la actualidad de inteligencia emulada por máquinas, por lo que una definición más próxima a ello, sería la capacidad de ciertos sistemas de procesar información y producir un resultado mediante un razonamiento que emula en cierto modo o parcialmente la actividad inteligente de los seres humanos.

Prosiguiendo con mi análisis del concepto y dejando a un lado su dimensión lingüística, a nivel más coloquial y cotidiano, este concepto se asocia a las máquinas que emulan el pensamiento o actuación humana, es decir, máquinas capaces de imitar funciones cognitivas del ser humano como percibir, razonar, aprender o resolver problemas, así como incluso tomar decisiones "autónomas" y/o ejecutar acciones físicas y lógicas.

De hecho, la inteligencia artificial se está asociando a máquinas que imitan las funciones "cognitivas" que los humanos asocian con otras mentes humanas, por ejemplo, aprender y resolver problemas[46].

A nivel técnico-científico, como he indicado anteriormente, la inteligencia artificial ha sido definida de muchas maneras, que parecen partir de objetivos y concepciones distintas de "Inteligencia", de las que me permito destacar algunas de ellas para, posteriormente, abordar su concepto jurídico.

McCarthy definió en 1956 la "Inteligencia artificial" como "*la ciencia e ingenio de hacer máquinas inteligentes, especialmente programas de cómputo inteligentes*"[47]. Según los principales estudios científicos en esta materia, como he referido en el epígrafe precedente, la primera vez que se acuñó y utilizó este concepto públicamente fue en la denominada "Conferencia de Dartmouth"[48] por parte de McCarthy.

[46] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Op.cit. Pos. 3836.
[47] MCCARTHY, J. (2007). "What Is Artificial Intelligence". 11.11.2007. Consultado el 23 de noviembre de 2020. http://www-formal.stanford.edu/jmc/whatisai/node1.html
[48] MCCARTHY, J.; MINSKY, M.; ROCHESTER, N.; SHANNON, C. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. 31.08.1955. Recuperado de: http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html. Consultado el 24.11.2020.

Winston y Brown[49] la definieron como el estudio de la inteligencia por medio de las ideas y métodos de la computación.

Charniak y McDermott[50] la definieron en 1985 como una cierta capacidad o potencia de computación que permitiría crear sistemas y dispositivos dotados de las mismas capacidades cognitivas que los seres humanos, lo que actualmente y conforme a la tecnología actual no es posible.

Por su parte, Kaplan y Haenlein[51] definieron la inteligencia artificial como la capacidad de un sistema para interpretar correctamente datos externos, para aprender de dichos datos y emplear esos conocimientos para lograr tareas y metas concretas a través de la adaptación flexible.

David Poole utiliza más adelante el concepto *Computational Intelligence*[52] y, como disciplina, la define como "el estudio del diseño de agentes inteligentes", y considera "agente inteligente" aquel sistema que actúa de forma inteligente, esto es, lo que hace es apropiado para sus circunstancias y su objetivo, es flexible a los entornos cambiantes y a los objetivos cambiantes, aprende de la experiencia, y toma las decisiones apropiadas dadas las limitaciones perceptivas y la computación finita. De hecho, bajo su criterio, sería más apropiado denominarla "Inteligencia sintética" en lugar de inteligencia artificial, concepto con el que no puedo estar más de acuerdo.

Otras definiciones se centran en el concepto de "máquina inteligente" como un agente racional flexible que percibe su entorno y lleva a cabo acciones que maximizan sus posibilidades de éxito en algún objetivo o tarea. Otros expertos incluso no están de acuerdo con el término "Inteligencia artificial", decantándose por el concepto "Inteligencia aumentada"[53].

---

[49] YDEWALLE, G. Y DELHAYE, P. "La inteligencia artificial, la obtención del conocimiento y el estudio de la inteligencia humana". *Revista internacional de ciencias sociales*. Vol. XL, nº 1 1988. Pp. 63-72

[50] CHARNIAK, E- Y MCDERMOTT, D. (1985). *Introduction to Artificial Intelligence*. Addison Wesley, 1985.

[51] KAPLAN, A. Y HAENLEIN M. (2018). *Siri, Siri, in my Hand, who's the Fairest in the Land? On the Interpretations, Illustrations and Implications of Artificial Intelligence.* Business Horizons, 62 (1), 15-25. 04.08.2018.

[52] POOLE, D., GOEBEL, R.G. Y MACKWORTH, A.K. (1998). *Computational Intelligence: A Logical Approach*. Oxford University Press. New York 1998. P.1.

[53] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Op.cit. Pos. 1503.

Una definición mucho más simplificada de sistema de inteligencia artificial desde un enfoque científico sería el software o programa informático que procesa datos y genera decisiones en base a los datos analizados[54].

Para algunos expertos como García-Prieto, la inteligencia artificial es simplemente lo que aún no se ha logrado.

No obstante, a pesar de la inexistencia de una definición consensuada en el ámbito científico y de la investigación, como se he puesto de relieve, si existe un consenso al afirmar que la inteligencia artificial busca replicar la inteligencia humana en sistemas de información, como abordaré con mayor profundidad al analizar las clases de inteligencia artificial en el siguiente apartado. Y esta, pretendo que sea la base para mi posterior análisis, partiendo de que actualmente nos hallamos en el ámbito de la emulación no de su replicación global.

Las principales discrepancias a nivel técnico en cuanto al concepto se producen respecto de las facultades cognitivas que se pretenden emular y programar, y si todas o algunas específicas.

Si nos centramos en sus premisas desde un punto de vista tecnológico, la inteligencia artificial pretende descomponer la inteligencia humana en sus procesos más simples para poder describirlos y expresarlos lógicamente, es decir, hacerla computable, y transferirlos mediante/a algoritmos y programación para reproducir o emular la inteligencia humana en sistemas, programas informáticos y máquinas, es decir, en *software* y *hardware*.

Según Leibniz[55] "todo lo que sepamos describir de forma clara, completa, precisa e inequívoca es computable".

No puedo compartir esta afirmación en relación con la inteligencia artificial en función del concepto de "Inteligencia humana" del que partamos, especialmente por algunos de sus rasgos que forman parte de la misma, como la percepción o la comprensión, por lo

[54] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Op.cit. Pos. 380.
[55] LEIBNIZ, G.W. (1923). *Sämtliche Schriften und Briefe*. Deutsche Akademie der Wissenschaften zu Berlin. O. Reichl .1923. P. 174.

que de manera previa deberíamos plantearnos qué es y en qué consiste la "Inteligencia humana" que la inteligencia artificial pretende emular y replicar, y cuáles son sus principales atributos.

Abordar con la profundidad necesaria estos aspectos en el marco del objeto y alcance de esta investigación resulta obviamente imposible, pero considero necesario, cuanto menos, significar algunos de ellos, en particular, los atributos esenciales de la inteligencia humana de los que carece la actual inteligencia artificial.

De inicio, la inteligencia humana, como ya anticipé, puede definirse de maneras muy distintas en función del campo o área de conocimiento desde los que se aborde, como la filosofía, sociología, psicología o neurofisiología. Incluso desde un punto de vista psicológico podríamos hablar, no sólo de una inteligencia, sino de "Inteligencias múltiples"[56].

De hecho, muchas de las definiciones expuestas de inteligencia artificial en relación con los rasgos propios de la inteligencia humana, parten de expectativas, todavía hoy, muy alejadas de la realidad y de la probabilidad con la tecnología actual, especialmente respecto de la posibilidad de reproducir plenamente la inteligencia humana en sistemas y máquinas, especialmente aspectos como el pensamiento creativo o crítico, la percepción o la comprensión. Hasta qué punto será posible replicarla o emularla, constituye actualmente un debate abierto en la comunidad científica y filosófica.

Como he comentado anteriormente, cuando hablamos de inteligencia artificial no estamos ante un nuevo concepto o tecnología, sino que ya desde los años cincuenta convive con nosotros y, de hecho, el apelativo "Inteligencia" está progresivamente desapareciendo como rasgo diferenciador de nuevas aplicaciones, soluciones, sistemas, productos y servicios basados o dotados de la misma, concibiéndose como algo ya presupuesto, propio, inherente y cada vez más común. Muchos de ellos no se perciben ya como dotados realmente de una inteligencia a significar o se les presupone la misma.

[56] GARDNER, H. *Frames of Mind: The Theory of Multiple Intelligences*. Nueva York: Basic Books.2004.

Además, el concepto inteligencia artificial califica una realidad y efectúa un atributo a la misma que está en constante cambio ante la incesante innovación y desarrollo de la tecnología y su aplicación, por lo que lo que hoy consideramos "inteligente", cuanto menos socialmente, es posible que mañana no lo sea o, cuando menos, no tengamos esa percepción.

Todo ello se haya en directa relación con lo que se denomina el "efecto IA" y que el precitado McCarthy destacaba en el sentido de que "tan pronto como funciona, ya nadie lo llama IA". Barrio Andrés[57] lo resume en que cada vez que un proyecto de investigación de inteligencia artificial hizo un nuevo descubrimiento útil, ese producto generalmente cobró autonomía para formar una nueva especialidad científica o comercial con su propio nombre distintivo, como sistemas expertos, *Machine Learning*, etc.

Lo cierto es que los sistemas más avanzados de reconocimiento facial o reconocimiento óptico de caracteres no suelen reconocerse o percibirse en general como sistemas inteligentes a pesar de que lo son y del elevado nivel de conocimiento e innovación asociado a su diseño, desarrollo, despliegue, interacción, funcionamiento y aplicación.

Como he referido anteriormente, la inteligencia artificial forma parte de nuestra vida diaria, convivimos, usamos e interaccionamos con ella física y virtualmente a través de nuestros *smartphones, smartwatches, apps*, ordenadores, *tablets*, redes sociales, *smartTVs,* sistemas de domótica, coches, controles de tráfico, controles de acceso biométrico, aspiradoras, neveras, etc.

Los sistemas de inteligencia artificial están presentes de manera cada vez más habitual y rutinaria en todos los sectores como la medicina, el transporte, las comunicaciones, el bancario y financiero, la educación o los videojuegos y nuestra vida diaria a través de cualquier asistente virtual personal o doméstico o un traductor inteligente.

Del mismo modo, empezarán a estar más presentes en otros ámbitos, como en el periodístico -por ejemplo, el sistema "Heliograf" del Washington Post-, o en el jurídico, donde ya tenemos experiencias piloto de "jueces virtuales" en China y Estonia, si bien,

[57] BARRIO ANDRÉS, M. (2020). *Manual de Derecho Digital*. Tirant lo Blanch. Valencia 2020. P. 56.

en sus primeras fases, constituyen más bien una herramienta de apoyo para el juez tradicional que es quien tiene la decisión, y sujeto a su supervisión. Y también abogados-robots como "Ross"[58], asesores digitales para el análisis contractual básico como "Lynn"[59] o "LawGeex"[60], o las herramientas de mediación "Family Winner" y "Asset Divider"[61].

En cualquier caso, considero que actualmente, a nivel social y ético es menos relevante determinar si el sistema es inteligente o no, que determinar si el mismo tiene capacidad de decidir por sí mismo, es decir, si tiene "supuesta" autonomía, así como su capacidad de autoaprendizaje y de decisión, lo que hace necesario el análisis de su diseño, funcionamiento, aplicabilidad y usos desde un punto de vista ético y jurídico. Estos aspectos los abordaré con detalle en el capítulo III.

Sorprendentemente, el concepto de autonomía que constituía un elemento esencial para la definición de los sistemas de inteligencia artificial en las propuestas de Reglamento de octubre de 2020 del Parlamento Europeo, ha quedado a un lado en la nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, que analizaré con profundidad en el capítulo IV.

A modo de conclusión, en mi opinión, efectivamente, cuando hablamos de inteligencia artificial a nivel técnico deberíamos referimos a la misma como la "capacidad" de un sistema o programa informático instalado en un *hardware*, máquina o dispositivo, operativo aisladamente o integrado con otros sistemas y tecnologías, que permite la identificación, calificación, captación, almacenamiento, análisis e interpretación de los

[58] SOURDIN, T. (2018). "Judge v Robot? Artificial Intelligence and judicial decision-making". *UNSW Law Journal*. Vol. 41 (4), 2018.
[59] https://juriblox.nl/toepassingen/contract-review/
[60] El promedio de precisión en la revisión contractual fue del 94% del sistema y de 85% en el caso de los abogados participantes que competían contra la máquina. El especial valor del sistema se identificó en la eficiencia y velocidad, dado que el sistema precisó 26 segundos mientras que el promedio de los abogados participantes osciló entre los 156 y 51 minutos. Recuperado de Vidal, Marc. "*La era de la humanidad*". Posición 3466 Versión Kindle. Ediciones Deusto. Barcelona. 2019.
[61] ZELEZNIKOW, J. (2017) "Can Artificial Intelligence and Online Dispute Resolution enhance efficiency and effectiveness in Courts". *International Journal for Court Administration*. Vol. 8, N. 2. May 2017. Disponible en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2999339.

datos y conocimientos de entrada, su tratamiento y gestión conforme a unos parámetros predefinidos, que pueden incluir su capacidad de autoaprendizaje automatizado y profundo, y su utilización para proporcionar datos o conocimientos de salida, que pueden ir asociados a una propuesta de resolución, decisión o ejecución de tareas o acciones igualmente predefinidas por el ser humano. Y si a dicho concepto añadimos el componente ético, adicionaría "y llevadas a cabo bajo la supervisión y control humano".

En definitiva, si al concepto técnico adicionamos los principios, valores y normas éticas esenciales que pretenden ser consensuadas a nivel europeo e internacional, en especial, la seguridad y el control y la supervisión humana, quizás en la actualidad, no deberíamos hablar tanto de sistemas realmente autónomos sino más bien automáticos o automatizados o, en todo caso, de autonomía restringida, limitada o reducida. Otra cosa distinta sería la creación de sistemas de inteligencia artificial fuerte dotados de una mayor autonomía, nunca plena por aplicación de las normas éticas precitadas.

Sin embargo, este concepto técnico no coincide con el concepto político, económico, social y jurídico sobre el que se están identificando los principales riesgos y temores que se asocian a la inteligencia artificial y sobre el que se están construyendo marcos éticos y futuros marcos normativos.

La Comunicación de la Comisión "Inteligencia *artificial para Europa*"[62] definió en 2018 la inteligencia artificial como "los sistemas que manifiestan un comportamiento inteligente, pues son capaces de analizar su entorno y pasar a la acción -con cierto grado de autonomía- con el fin de alcanzar objetivos específicos".

Estos sistemas pueden consistir simplemente en un programa informático (por ejemplo, asistentes de voz, programas de análisis de imágenes, motores de búsqueda, sistemas de reconocimiento facial y de voz) o hallarse integrados en dispositivos de hardware (por ejemplo, robots avanzados, automóviles autónomos, drones u otros dispositivos del Internet de las Cosas -IOT por sus siglas en inglés -*Internet of Things*-). De esta

[62] Comunicación de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones "Inteligencia artificial para Europa". 25 de abril de 2018. COM/2018/237 final

definición, destacar la omisión del rol del ser humano en su generación y la inclusión del concepto autonomía por lo que luego se dirá.

El *Grupo de Expertos de alto nivel en inteligencia artificial* de la Comisión Europea, del que posteriormente hablaré, concretó más esta definición: "los sistemas de inteligencia artificial (IA) son programas informáticos (y posiblemente también equipos informáticos) diseñados por seres humanos que, dado un objetivo complejo, actúan en la dimensión física o digital mediante la percepción de su entorno mediante la adquisición de datos, la interpretación de los datos estructurados o no estructurados, el razonamiento sobre el conocimiento o el tratamiento de la información, fruto de estos datos y la decisión de las mejores acciones que se llevarán a cabo para alcanzar el objetivo fijado".

De esta definición, destacar que aparece el rol desarrollado por el ser humano en su creación: "Diseñados por seres humanos". Y lo destaco por su especial relevancia a efectos de la interpretación y aplicación de las normas sobre responsabilidad que integren los futuros marcos normativos que partan de estas definiciones.

El Grupo de Expertos precitado distingue dos dimensiones distintas de la inteligencia artificial, como sistemas y como disciplina.

De un lado, los sistemas de inteligencia artificial son sistemas esencialmente de *software* (y en su caso de *hardware*) que, ante un objetivo complejo, actúan en la dimensión física o virtual mediante la percepción del entorno a partir de la adquisición de datos -estructurados o no- que se hayan recabado, el procesamiento de la información derivadas de dichos datos y la adopción de la mejor decisión para alcanzar el objetivo precitado. "Los sistemas de inteligencia artificial pueden usar reglas simbólicas o aprender un modelo numérico, y pueden también adaptar su comportamiento analizando cómo sus acciones previas afectan al ambiente".

De otro, la inteligencia artificial constituye una disciplina científica que incluye diversos enfoques y técnicas, como el *Machine Learning* – y dentro del mismo el *Deep Learning* o el *Reinforcement Learning*-, el *Machine Reasoning* y la *robótica*. Posteriormente abordaré estos conceptos.

De la definición precitada, destacar la doble dimensión de la inteligencia artificial, de un lado, como disciplina científica y, de otro, como "sistema de información" -en mi opinión más adecuado que "sistema basado en *software* o en *software* y *hardware"*-, que integra igualmente un elemento esencial para su funcionamiento: Los datos.

Por último, el *Libro blanco sobre la inteligencia artificial* precitado de la Comisión Europea, que analizaré con mayor detalle posteriormente, significa la necesidad de que su definición sea lo suficientemente flexible, para adaptarse al progreso técnico, a la vez que precisa para garantizar la seguridad jurídica necesaria. La define como una combinación de tecnologías que agrupa datos, algoritmos y capacidad informática.

Siguiendo mis reflexiones sobre este concepto, considero necesario abordar el concepto evolucionado al que se hace referencia generalmente cuando utilizamos la expresión "Inteligencia artificial" y se asocia a un conjunto de retos y riesgos desde un punto de vista jurídico.

Sin duda, el tratamiento de todas estas cuestiones abriría un tema de interesantísima discusión desde distintas ópticas, pero que excedería del objeto y alcance de la presente investigación, sin perjuicio de que me permita abordar algunos aspectos a continuación, por su indudable transcendencia, especialmente para su conceptualización jurídica y su futura regulación.

A mi juicio, buena parte de las preocupaciones por los retos y riesgos que puede comportar van más allá del concepto tradicional de inteligencia artificial y se orientan cada vez más a un concepto evolucionado y a una inteligencia más avanzada o superior, centrado en la capacidad de tomar decisiones o ejecutar acciones sin supervisión humana sobre la base de su supuesta "autonomía", capacidad de autoaprendizaje e impredecibilidad, que en cualquier caso deberían estar predefinidas y restringidas en su diseño y desarrollo por el ser humano, especialmente en el marco de los denominados "sistemas inteligentes considerados de alto riesgo" para las personas, instalaciones y cosas.

Además, la inteligencia artificial se construye mediante algoritmos que, de modo muy simple, me permito definir como el conjunto metodológico de pasos para hacer cálculos,

previsiones, resolver problemas y/o tomar decisiones. Esos pasos junto con el método diseñado, instrucciones definidas, capacidades otorgadas y calidad de los datos de entrada, determinan los principales retos y riesgos de la inteligencia artificial.

Por consiguiente, me permito abordar en el siguiente apartado el concepto jurídico de inteligencia artificial, para lo que debo acudir a los documentos e instrumentos normativos propuestos que abordan su definición.

3.2. Inteligencia artificial y autonomía

La UE está liderando las propuestas de regulación de la inteligencia artificial a nivel internacional, sin embargo, no ha definido de manera uniforme este concepto en sus distintas propuestas, informes, comunicaciones y documentos de trabajo hasta la fecha.

En unos casos la define como sistemas[63]: "Los sistemas que muestran un comportamiento inteligente al analizar su entorno y tomando medidas, con cierto grado de autonomía, para lograr objetivos específicos".

En otros como habilidad o capacidad[64]: "La habilidad de una máquina de presentar las mismas capacidades que los seres humanos, como el razonamiento, el aprendizaje, la creatividad y la capacidad de planear". De este modo, un sistema dotado de inteligencia artificial percibe su entorno, se relaciona con él, resuelve problemas y actúa con fines específicos.

La Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas[65], integra una Propuesta de Reglamento del

[63] Informe del Parlamento Europeo: "Artificial intelligence: How does it work, why does it matter, and what can we do about it?". UE Junio 2020. Accesible en: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641547/EPRS_STU(2020)641547_EN.pdf

[64] https://www.europarl.europa.eu/news/es/headlines/society/20200827STO85804/que-es-la-inteligencia-artificial-y-como-se-usa

[65] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))

Parlamento Europeo y del Consejo sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas.

Esta Propuesta de Reglamento define en su artículo 4.I.a) la "Inteligencia artificial" como un sistema basado en programas informáticos o incorporado en dispositivos físicos que manifiesta un comportamiento inteligente al ser capaz, entre otras cosas, de recopilar y tratar datos, analizar e interpretar su entorno y pasar a la acción, con cierto grado de autonomía, con el fin de alcanzar objetivos específicos.

Se trata de una definición adaptada de la definición efectuada en la Comunicación de la Comisión Europea de 25 de abril de 2018[66].

Por su parte, la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial[67], integra la Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial.

Esta Propuesta de Reglamento recoge en su artículo 3.I.a) una definición de "sistema de inteligencia artificial", considerando como tal, todo sistema basado en programas informáticos o incorporado en dispositivos físicos que muestra un comportamiento que simula la inteligencia, entre otras cosas, mediante la recopilación y el tratamiento de datos, el análisis y la interpretación de su entorno y la actuación, con cierto grado de autonomía, para lograr objetivos específicos.

Se trata de una definición no idéntica pero similar a la incluida en la precitada Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, coetánea en el tiempo con aquélla.

[66] Comunicación de la Comisión Europea COM (2018) 237 final, de 25.4.2018. P.1.
[67] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))

De ambas definiciones, destacar que plantea una disyuntiva, a mi modo de ver, algo imprecisa, al indicar que se trata de un "sistema" basado en programas informáticos o un "sistema" incorporado en un dispositivo físico, cuando en ambos casos entiendo que está haciendo referencia a una sola realidad, esto es, un sistema, no de cualquier tipo, sino a sistemas "de información" o "informáticos" que operan independientemente o incorporados a dispositivos físicos, cuyo principal elemento que lo conforma es el *software*, conjunta y necesariamente con otros, como *hardware*, procesadores, algoritmos y datos.

Asimismo, ambas definiciones asocian a los sistemas de inteligencia un comportamiento inteligente (Resolución del Parlamento Europeo sobre aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas) o un comportamiento que simula la inteligencia (Resolución del Parlamento Europeo sobre un régimen de responsabilidad civil en materia de inteligencia artificial) como rasgo definitorio y diferenciador, y ambas definiciones especifican en que consiste dicha inteligencia -como atributo de un comportamiento en el caso de la primera y como objeto de emulación en el caso de la segunda-, asociándola en ambos casos a la capacidad, entre otras cosas, de recopilar y tratar datos, analizar e interpretar su entorno y pasar a la acción, con cierto grado de autonomía, con el fin de alcanzar objetivos específicos.

Es decir, la inteligencia simulada se asocia de manera enunciativa que no limitativa, a recopilar y tratar datos, analizar e interpretar su entorno y realizar acciones para alcanzar objetivos concretos, con otra característica necesaria asociada, esto es, un cierto grado de autonomía.

A mi juicio, a pesar de similitud contienen matices con cierta relevancia técnica en virtud de su interpretación, y consecuente relevancia jurídica, en particular la referencia a "comportamiento inteligente" y "comportamiento que simula la inteligencia".

En la primera definición, considero que se está asociando una capacidad o atributo inherente al sistema no definido en la Propuesta de Reglamento, aunque incorporando algunos de los elementos que lo conforman en la propia definición de manera enunciativa, no limitativa.

En la segunda, se está hablando de simulación de dicha inteligencia, que considero posiblemente se ajusta más a la realidad que pretende definir, que obviamente no se presenta como una capacidad o atributo propio e inherente, sino simulado. Ningún sistema o *software* actual puede pensar como un humano ni tiene consciencia.

Del mismo modo, como abordaré posteriormente, ambas definiciones -al igual que las anteriores-, hacen referencia al concepto de "autonomía", como uno de los rasgos que define esa inteligencia simulada y el Parlamento Europeo también define en sus dos Propuestas de Reglamento el concepto de "autonomía" asociado a los sistemas de inteligencia artificial.

En particular, en el artículo 3.I.b) de la Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial, define "autonomía" como todo sistema de inteligencia artificial que funciona interpretando determinados datos de entrada y utilizando un conjunto de instrucciones predeterminadas, sin limitarse a ellas, a pesar de que el comportamiento del sistema esté limitado y orientado a cumplir el objetivo que se le haya asignado y otras decisiones pertinentes de diseño tomadas por su desarrollador.

Sin embargo, de nuevo, la definición en la Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas es similar, pero no idéntica, en la medida que mejora su redacción en algún matiz ortográfico al indicar que, el comportamiento del sistema deberá estar orientado a cumplir el objetivo que se le hubiere asignado y estará limitado por este objetivo y por otras decisiones de diseño pertinentes tomadas por su desarrollador, cuando en la primera, parece desprenderse de su tenor literal que el comportamiento del sistema, más que limitado por otras decisiones pertinentes de diseño, debería hallarse orientado por las mismas.

En ambas definiciones, la autonomía se circunscribe a la no limitación por parte del sistema a las instrucciones predeterminadas del mismo para cumplir el objetivo atribuido o adoptar las decisiones adecuadas conforme a su diseño confiriéndole, en consecuencia, una cierta libertad para seguir instrucciones no predeterminadas en dicho diseño, aunque limitadas por el objetivo definido y decisiones permitidas en su diseño.

En cualquier caso, se concibe como un sistema dotado de la capacidad para para interpretar datos y actuar en base instrucciones predeterminadas e integradas en el mismo u otras distintas "adquiridas".

Este concepto de autonomía asociado a un sistema de inteligencia artificial supone superar el automatismo inicial asociado a estos sistemas que podría incluir desde el desarrollo de un proceso o funcionamiento de un mecanismo o aparato por sí solo, o la ejecución mecánica de actos sin ser consciente de ellos, hasta la sustitución del operador humano en un proceso por dispositivos mecánicos o electrónicos.

La autonomía, tal y como ha sido concebida en las definiciones jurídicas precitadas, va más allá de las meras instrucciones predefinidas a aplicar para que se produzca un resultado de manera inmediata y consecuente al acaecimiento de un hecho, en la medida que el sistema no estaría sujeto y limitado por las instrucciones preconcebidas, aunque si por los objetivos definidos y permisos atribuidos.

La inclusión en la definición de que el sistema opere utilizando un conjunto de instrucciones predeterminadas, pero sin limitarse a ellas, comporta a mi juicio la definición de una inteligencia artificial más avanzada que no es la que pretenden regular ambas propuestas, las cuales se orientan a una inteligencia artificial más básica o "débil" como luego definiré. En consecuencia, estas definiciones deberán ser revisadas en adelante. En el capítulo V abordaré con mayor detenimiento esta cuestión al analizar con detalle la Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial.

En cualquier caso, ambas definiciones -"sistema de inteligencia artificial" y "autonomía"-, ponen de manifiesto el control sobre el riesgo del diseñador o desarrollador -algo de indudable trascendencia a los efectos de responsabilidad-, al limitar el comportamiento del sistema aunque pueda operar sin limitarse estrictamente a todas las instrucciones predeterminadas para la interpretación de los datos de entrada, es decir, el necesario control y limitaciones sobre los datos de salida y las decisiones o acciones consecuentes

para la que pueda estar programado el sistema conforme a sus objetivos y decisiones de diseño.

Y me parece un matiz técnico relevante desde el punto de visto jurídico, especialmente por lo que se refiere a la posible imputación y determinación del grado de responsabilidad de los distintos agentes intervinientes, en relación con los daños causados por la inteligencia artificial y en función de su grado de control sobre sus riesgos, que abordaré posteriormente.

Sin duda, la definición integrada en las dos propuestas indicadas evidencia un concepto de inteligencia artificial más evolucionado que los anteriores, identificándola con un sistema informático que muestra o emula un comportamiento "inteligente", en la medida que no es un atributo del que disponga verdaderamente, y considera que lo es al tener, entre otras capacidades, las de recopilar, tratar datos, analizar e interpretar su entorno y actuar con cierto grado de autonomía, obviamente predefinida y restringida, para alcanzar objetivos concretos asignados, que nada tiene que ver con la inteligencia humana y los rasgos y capacidades inherentes a la misma que la diferencia.

Aun así, desde mi punto de vista, constituye una definición todavía más próxima a su funcionamiento automático o automatizado que autónomo en sí mismo o "fuerte" en los términos que definiré más adelante, y ello en base a su actuación conforme instrucciones predefinidas y nuevas instrucciones lógicas -generadas lógicamente en base a su programación- conforme a los nuevos datos de entrada analizados y en base igualmente, de un lado, a sus capacidades conferidas en su diseño de captación, análisis, "interpretación", autoaprendizaje y entrenamiento y, de otro, de los nuevos datos de entrada generados por cada contexto particular, que serán, en cualquier caso, tratados e interpretados conforme a instrucciones informáticas predefinidas por el diseñador y/o programador o no, pero restringido a los objetivos asignados y con las limitaciones, autorizaciones y permisos pertinentes predefinidas en su diseño y concepción.

De este modo, las decisiones y acciones de un sistema de inteligencia artificial deberían ser en todo caso previsibles o previstas si partimos de este concepto de inteligencia artificial y se garantizan los principios y atributos éticos esenciales ya en su concepción y diseño, esto es, la *Ethics by design*, en especial, la seguridad, la responsabilidad y la

supervisión y control humano, sin perjuicio de su supervisión y control coetánea o incluso posterior durante todo su ciclo de vida que permita incluso, entre otras cosas, la reversibilidad.

Dicho control, constituye un componente esencial para el marco de seguridad y confiabilidad pretendido por la UE para todas las partes implicadas en la inteligencia artificial, que permitirá un mejor desarrollo y aplicación de la misma.

Las definiciones analizadas no contemplan una inteligencia artificial con consciencia o pensamiento crítico sino la simulación parcial de la inteligencia "humana", nunca en toda su amplitud, circunscrita a un conjunto acciones susceptibles de ser llevadas a cabo con cierto grado de autonomía, nunca plena, e incluso, conforme he argumentado anteriormente, restringida conforme a instrucciones predefinidas, aunque con cierta libertad para no limitarse a las mismas.

El concepto que recogen ambas propuestas, en mi opinión y conforme a su tenor literal, se aparta de la realidad que pretenden regular, conforme he referido, en la medida que supone sistemas que no solo pueden tomar decisiones o actuar en base a conocimientos adquiridos con posterioridad a su diseño mediante el autoaprendizaje y entrenamiento, sino que podrían actuar a margen de las instrucciones embebidas en su diseño.

La inteligencia artificial que se está desarrollando y aplicando extendidamente en la actualidad no incluye sistemas más avanzados y evolucionados que no son todavía una realidad y que son los que mayor debate están suscitando -como la inteligencia artificial "fuerte" o general a las que luego haré referencia-, especialmente en relación con sus riesgos, retos y regulación.

En definitiva, conforme a las definiciones precitadas objeto de análisis y de los futuros marcos regulativos, los sistemas de inteligencia artificial estarían dotados de ambos rasgos, cierta autonomía y automatismo, sin embargo, algunas voces del ámbito científico

prefieren hablar de sistemas automáticos o automatizados más que autónomos en algunos sistemas como, por ejemplo, los coches sin conductor[68].

Conforme he comentado anteriormente, este concepto evolucionado de inteligencia artificial sustentado en la autonomía es el que está despertando grandes inquietudes en la actualidad ante sus riesgos, conforme analizaré más adelante, y más mayores cuando se abordan posibles sistemas de inteligencia artificial más avanzados categorizados como "fuertes" o "generales", conforme abordaré en el próximo apartado.

En la actualidad estamos todavía en el despliegue de una inteligencia artificial "estrecha" o "débil", de modo que la misma pueda resolver tareas concretas en situaciones predecibles, donde hay muchos datos disponibles para entrenar al algoritmo.

Los mayores retos se vislumbran en relación con una supuesta autonomía, relativa libertad y relativa impredecibilidad asociada a sistemas inteligentes más avanzados, aunque previsiblemente muy alejados de toda expectativa de conciencia y voluntad.

Y, es más, considero que no debemos obviar algunas de las conclusiones del estudio llevado a cabo por el *Grupo Europeo de Ética en Ciencia y Nuevas Tecnologías para la Comisión Europea*[69], en el que destaca que el concepto "autonomía" tiene una base filosófica y viene referido a la capacidad del ser humano de legislarse a sí mismo, de formular, pensar y elegir normas y leyes que seguir por sí mismos, lo que incluye el derecho a ser libre para elegir sus propias normas, metas y propuestas.

Según el mismo, esta capacidad y derecho se halla sustentada en un proceso cognitivo que integra funciones de autoconciencia, así como aspectos como el sentimiento de culpabilidad y la propia autoría en relación con razones y valores. De este modo, el estudio significa que "autonomía, en el sentido ético relevante del término, puede por ello atribuirse solamente al ser humano", por lo que considera un error atribuírselo a meros artefactos, pese a su muy avanzada adaptabilidad o, incluso, inteligencia de sus sistemas.

---

[68] MARÍN, S. (2019). "Ética e inteligencia artificial*"*. *Cuadernos de la Cátedra Caixabank de Responsabilidad Social Corporativa*. IESE 2019.
[69] SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret.* Barcelona. Enero 2018. P. 13.

El estudio concluye que el término "sistemas autónomos" se ha ido extendiendo entre la literatura científica y el público para referirse al más alto nivel de autonomía e independencia de los seres humanos en términos de autonomía operativa y en la toma de decisiones, si bien, "autonomía", en su sentido original, es uno de los aspectos más importantes de la dignidad humana que no debe relativizarse.

En consecuencia, desde una perspectiva ética, partiendo de este concepto y conforme a la tecnología actual, nunca podría atribuirse una autonomía plena a un sistema inteligente en el sentido indicado, en la medida que carece de autoconciencia, así como aspectos como el sentimiento de culpabilidad y la propia autoría en relación con razones y valores, sin perjuicio de que en el futuro se le pudiese atribuir una "autonomía artificial", restringida y sujeta a control y supervisión humana, nunca plena, en virtud de los principios y atributos éticos esenciales pretendidos para la inteligencia artificial por los futuros marcos reguladores objeto de análisis. Todo ello enlaza directamente con los aspectos de responsabilidad objeto de esta investigación, inimputabilidad civil y penal, personalidad jurídica y capacidad de obrar.

Hechas estas consideraciones sobre las definiciones contempladas en las Resolución del Parlamento Europeo precitadas, la nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[70], parte de una definición de inteligencia artificial muy distinta a las propuestas, informes y documentos referenciados anteriormente.

Conforme expondré en su análisis en el capítulo IV de esta investigación, esta última propuesta parte de que los sistemas de inteligencia artificial pueden diseñarse para funcionar con distintos niveles de autonomía y utilizarse de forma independiente o como componente de un producto, independientemente de si el sistema está integrado físicamente en el producto (integrado) o sirve a la funcionalidad del producto, sin estar integrado en el mismo.

[70] COM (2021) 206 final 2021/0106 (COD)

De este modo, la Propuesta de Reglamento define de manera amplia en su artículo 3 el concepto de "Sistema de inteligencia artificial", en particular, como el software desarrollado con una o más de las técnicas y enfoques enumerados en el Anexo I del Reglamento propuesto y que puede, para un conjunto dado de objetivos definidos por el ser humano, generar resultados como contenido, predicciones, recomendaciones o decisiones que influyen en los entornos con los que interactúan.

El concepto de autonomía desaparece de su definición sin perjuicio de que consideremos incluido implícitamente como atributo asociado a este tipo de sistemas. De inicio, la definición se aleja de la deseable claridad, concreción y sencillez a la hora de definir un concepto jurídico esencial y tecnológico, sin duda complejo, pero que constituye el objeto de regulación de esta propuesta y de todo un conjunto de normas derivadas tanto de *hard law* como *soft law*, como referiré más adelante.

Adicionalmente, la definición precitada se debe integrar y completar necesariamente con las complejas técnicas y enfoques listados en el precitado anexo que, además, como en las propuestas predecesoras en materia ética, constituyen un *numerus clausus*, si bien, la Comisión Europea quedaría facultada por el futuro Reglamento para adoptar actos delegados para actualizar dichos listados conforme a la evolución de la tecnología y del mercado.

Las técnicas y enfoques de inteligencia artificial que integra la definición y se relacionan en el precitado Anexo I del Reglamento propuesto, son los siguientes:

a) Enfoques de aprendizaje automático, incluido el aprendizaje supervisado, no supervisado y por refuerzo, utilizando una amplia variedad de métodos, incluido el aprendizaje profundo.

b) Enfoques basados en la lógica y el conocimiento, incluida la representación del conocimiento, la programación inductiva (lógica), las bases de conocimiento, los motores de inferencia y deductivos, el razonamiento (simbólico) y los sistemas expertos.

c) Enfoques estadísticos, estimación bayesiana, métodos de búsqueda y optimización.

En definitiva, una definición que se aparta nuevamente del concepto sistema de información o informático y se focaliza exclusivamente en el *software* que sea desarrollado con cualquiera de estas técnicas o enfoques, o con una combinación de todas o algunas de ellas.

En la medida que este *software* tenga la capacidad de generar resultados como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúan, conforme a los objetivos definidos, se considerará un sistema de inteligencia artificial.

Sin duda, una definición compleja a nivel técnico y, a mi juicio, excesivamente genérica y amplia respecto de los sistemas que pretende contemplar, pero no exhaustiva, donde parece que el grado de autonomía del sistema o su mera concurrencia es irrelevante para su calificación como tal, por lo que cualquier *software* común con cualquiera de esas capacidades y características sería considerado inteligencia artificial.

En definitiva, aún consciente de la dificultad, especialmente por la propia falta de consenso a nivel científico en su definición, precisamos una definición jurídica lo más clara, precisa y consensuada posible de una realidad compleja como es la inteligencia artificial, sin perjuicio de considerar sus tipologías y especificidades, para construir los marcos que pretenden regularla.

A nivel internacional, la primera definición jurídica de inteligencia artificial en EE.UU. se efectuó en la sección 238 (g) de la Ley de Autorización de la Defensa Nacional John McCain para el año fiscal 2019[71], la cual la define como cualquier sistema artificial que realiza tareas bajo circunstancias variables e impredecibles, sin una supervisión humana significativa o que puede aprender de la experiencia y mejorar el rendimiento cuando se expone a conjuntos de datos.

De inicio, se presenta como una definición sencilla y concreta, aunque "sesgada", ante una realidad tan compleja y diversa como lo es la inteligencia artificial, la cual difiere de

[71] Recuperado de: https://www.congress.gov/115/bills/hr5515/BILLS-115hr5515enr.pdf. Consultado el 10.02.2021.

la definición utilizada en el Plan Estratégico Nacional de Investigación y Desarrollo de la IA de EE.UU. de 2019.

No obstante, incluye también dentro del concepto precitado otras tipologías de sistemas, en particular, los siguientes: a) Sistemas artificiales desarrollados en software informático, hardware físico u otro contexto que resuelven tareas que requieren percepción, cognición, planificación, aprendizaje, comunicación o acción física similares a las de los humanos; b) Los sistemas artificiales diseñados para pensar o actuar como un humano, incluidas las arquitecturas cognitivas y las redes neuronales; c) Los conjuntos de técnicas, incluido el aprendizaje automático, que estén diseñados para aproximarse a una tarea cognitiva; d) Los sistemas artificiales diseñados para actuar racionalmente, incluido un agente de software inteligente o un robot incorporado que logra objetivos utilizando la percepción, la planificación, el razonamiento, el aprendizaje, la comunicación, la toma de decisiones y la actuación.

Del mismo modo, nos encontramos otras definiciones a nivel internacional en las que figura o no el componente humano, con su repercusión correlativa a efectos de responsabilidad conforme al concepto que finalmente acojan los marcos que la regulen.

Australia la definió en su Plan de Acción de IA[72] como un conjunto de tecnologías interrelacionadas que se utilizan para resolver problemas de forma autónoma y realizar tareas para alcanzar objetivos definidos, en algunos casos sin la orientación explícita de un ser humano.

Singapur la definió en su Estrategia Nacional de IA[73] de manera muy sencilla como la capacidad de simular un comportamiento inteligente, similar al humano, en los ordenadores.

Por su parte, Canadá la definió en su Estrategia sobre IA[74] como la tecnología de la información que realiza tareas que normalmente requerirían la fuerza cerebral biológica

[72] Recuperado de: https://consult.industry.gov.au/digital-economy/ai-action-plan/. Consultado el 15.11.2020.
[73] Recuperado de: https://www.smartnation.gov.sg/docs/default-source/default-document-library/national-ai-strategy.pdf?sfvrsn=2c3bd8e9_4. Consultado el 15.11.2020.
[74] Recuperado de: https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592#appA. Consultado el 15.11.2020.

para llevarlas a cabo, como dar sentido al lenguaje hablado, aprender comportamientos o resolver problemas.

En definitiva, una realidad compleja con definiciones distintas y que conforman el objeto de regulación de nuevos marcos jurídicos.

De las múltiples definiciones, podemos extraer algunas notas comunes, como la autonomía, el tratamiento de datos, la complejidad o la opacidad que, a su vez, constituyen los grandes retos jurídicos de la misma. Sin embargo, las definiciones incorporadas en las distintas propuestas reguladoras de la UE, analizadas en esta investigación, ofrecen conceptos parciales al objeto de acotar el tipo de inteligencia artificial que pretende ser regulada por las mismas, sirviendo la última Propuesta de Reglamento de Parlamento Europeo y del Consejo de 21 de abril de 2021 como claro ejemplo de ello, en la que incluso no se hace una referencia expresa a la autonomía.

El concepto jurídico de la realidad que se pretende regular es determinante para la construcción de los nuevos marcos jurídicos, especialmente los de responsabilidad.

3.3. Disciplinas y ramas de la inteligencia artificial

La inteligencia artificial no constituye una única tecnología, sino que engloba un conjunto de métodos, algoritmos y tecnologías.

La inteligencia artificial es un concepto genérico o "supraconcepto" que puede integrar distintas tecnologías inteligentes, con aspectos comunes y estrechamente interrelacionados entre sí, pero que han ido adquiriendo sustantividad propia como disciplinas científicas.

Las principales ramas de la misma que destaca Barrio Andrés[75], siguiendo a Mills, serían aprendizaje automático (*Machine Learning o ML*), procesamiento del lenguaje natural

[75] BARRIO ANDRÉS, M. (2020). *Manual de Derecho Digital*. Tirant lo Blanch. Valencia 2020. Pp. 58-59.

(*Natural Language Processing o NLP*), diseño de sistemas expertos, visión artificial (*Computer visión*), reconocimiento del habla y planificación automática.

A las mismas debemos de añadir el denominado *Deep Learning o ML*, redes neuronales (*Artificial neural networks o ANNs*), generación de lenguaje natural (*Natural Language Generation o NLG*), asistentes virtuales digitales (*Virtual Digital Assistants*) o análisis predictivo (*Predictive analysis*).

En el apartado 3.5 abordaré algunos de estos conceptos.

3.4. Otros conceptos relacionados

La Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas precitada, abordó otros conceptos asociados inexorablemente a la inteligencia artificial, en especial, autonomía, robótica y tecnologías conexas que, por las razones expuestas, considero oportuno abordarlos, cuanto menos, sucintamente.

La propuesta precitada definió el concepto de "robótica", concebida como las tecnologías que permiten que las máquinas controladas automáticamente, reprogramables y multifuncionales realicen en el mundo físico acciones tradicionalmente realizadas o iniciadas por los seres humanos, en particular mediante la inteligencia artificial o las tecnologías conexas, es decir, se concibe como el *software*, los algoritmos y demás tecnología que se incorpora a un elemento, máquina o *hardware* para realizar acciones en el mundo físico mediante la inteligencia artificial o las denominadas "tecnologías conexas".

La definición incorporada a la propuesta, en mi opinión, plantea algunas preguntas y resulta algo incompleta a la vista de la realidad actual. Por ejemplo, ¿sólo incluye las tecnologías que permiten a las máquinas realizar acciones en el mundo físico?, ¿y las acciones que realizan las máquinas en el mundo virtual o en el ciberespacio?, ¿y las acciones de manera programada y reactiva o proactiva y autónoma por el sistema?

Considero que esta definición deberá revisarse en el futuro para contemplar una concepción integrada más amplia, máxime si nos planteamos la existencia de robots basados exclusivamente en software o en incluso *Xenobots*[76], presentados científicamente en la fecha muy próximas al cierre de esta investigación como máquinas vivientes sintéticas.

Sin duda, la realidad tecnológica y avance continúo nos obligará a revisar y evolucionar continuamente éstas y otras definiciones para abarcar toda la realidad que pretende definir y regular. A modo de ejemplo y, conforme he indicado, a mi juicio no tendría sentido dejar fuera del concepto "robótica" las acciones lógicas realizadas en el mundo virtual o ciberespacio.

La Propuesta de Reglamento precitada también define el concepto de "tecnologías conexas" como aquellas tecnologías que permiten que los programas informáticos controlen, con un grado de autonomía parcial o total, un proceso físico o virtual, aquellas que son capaces de detectar los datos biométricos, genéticos o de otro tipo, así como aquellas que copian o utilizan de otro modo características humanas. Es decir, *software* para el control autónomo total o parcial de procesos físicos o virtuales, tecnología de detección de datos y tecnologías de emulación o utilización de características humanas.

Y, por último, de manera asociada a todo ello, la precitada Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas define las denominadas "tecnologías de alto riesgo".

Según el texto normativo propuesto, estas tecnologías son aquellas cuyo desarrollo, despliegue y uso entrañen un riesgo significativo de causar lesiones o daños a particulares o a la sociedad, vulnerando los derechos fundamentales y las normas de seguridad establecidas en el Derecho de la Unión, a cuyos efectos deben tenerse en cuenta el sector en el que se desarrollan, despliegan o utilizan, su uso o finalidad específica y la gravedad

[76] BLACKISTON, D. ET ALT. "A cellular platform for the development of synthetic living machines". Publicado en *Science Robotics*. 31 de marzo de 2021. Vol. 6 Edición 52, eabf1571. DOI: 10.1126 / scirobotics.abf1571. Disponible en: https://robotics.sciencemag.org/content/6/52/eabf1571.Consultado el 31.03.2021.

de la lesión o daño que cabe esperar que se produzca. De este modo, para la categorización del riesgo como significativo y alto, deberá considerarse el sector, uso, finalidad y gravedad del potencial impacto del mismo, pero no se incluye expresamente su probabilidad, criterio necesario a considerar en cualquier evaluación de riesgos.

De esta definición me permito destacar ya, por lo que luego analizaré a lo largo del capítulo IV, que esas tecnologías de alto riesgo sí incluyen una mención expresa a, de un lado, la paralela vulneración junto a los daños y perjuicios causados, de derechos fundamentales de las personas y, de otro, la vulneración de las normas de seguridad vigentes en la UE. De este modo, son estas tecnologías de alto riesgo las que mayor preocupación provocan desde un punto de vista del derecho de daños.

La definición precitada es similar a la definición sobre sistemas de inteligencia artificial de alto riesgo que incorpora la Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial. Esta considera en su artículo 3.I.c) que un sistema de inteligencia artificial que funciona de forma autónoma es de alto riesgo en atención a su potencial significativo para causar daños o perjuicios a una o más personas de manera aleatoria y que excede lo que cabe esperar razonablemente.

Es decir, un sistema de inteligencia artificial se considera de alto riesgo, entre otros factores, siempre que comporte autonomía (sin especificar el grado), lo que difiere radicalmente del concepto de sistemas de alto riesgo que regula la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, que analizaré más adelante.

Prosiguiendo con la definición incorporada al precepto indicado, la calificación de ese "potencial significativo" se hace depender de la gravedad del posible daño o perjuicio, el grado de autonomía de la toma de decisiones, la probabilidad de que el riesgo se materialice y el modo y el contexto en que se utiliza el sistema, cuando en la Propuesta antedicha de Reglamento del Parlamento Europeo y del Consejo sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, para determinar su entidad no se hacía referencia ni al grado de

autonomía del sistema ni a la probabilidad de materialización de riesgo y, además, adiciona la necesidad de tener en cuenta el sector en el que se desarrolla, despliega o utiliza. De nuevo, a pesar de tratarse de propuestas coetáneas podemos observar la falta de armonización en los conceptos esenciales.

No obstante, ambas propuestas requieren considerar el contexto en el que se utilice el sistema, en particular su uso, finalidad y sector específicos -en el caso de la Propuesta antedicha de Reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas-, y el modo y el contexto en los que se utiliza el sistema -en el caso de la Propuesta de Reglamento sobre responsabilidad civil derivada del funcionamiento de los sistemas de inteligencia artificial-, que posteriormente vincula los mismos a su categorización objetiva.

La primera propuesta habla de tecnologías del alto riesgo en general mientras que la segunda habla de tecnologías específicas que conforman sistemas de alto riesgo.

Si partimos de la máxima precisión en la definición de los conceptos jurídicos, ambas deberían hallarse alineadas y completas, al objeto de no dejar fuera de su definición, según el instrumento normativo utilizado, aspectos como el grado de autonomía del sistema, la probabilidad de materialización de riesgo o la necesidad de tener en cuenta el sector en el que se desarrolla, despliega o utiliza.

3.5. Del aprendizaje automático (*Machine Learning*) al profundo (*Deep Learning*)

En este apartado pretendo definir de la manera más sencilla posible algunos de los conceptos básicos relacionados con la inteligencia artificial que aparecerán en distintos apartados de esta investigación, especialmente en la medida que cualifican la tipología y características de los sistemas y determinan los posibles riesgos y retos a abordar.

El desarrollo de la inteligencia artificial se está sustentando en este momento en múltiples técnicas y enfoques, en especial, en el denominado *Machine Learning* -ML por sus siglas en inglés-, consistente en el aprendizaje que llevan a cabo estos sistemas mediante el

análisis del mayor conjunto de datos posible. El aprendizaje permite que los sistemas puedan realizar predicciones y tareas específicas de forma autónoma.

De inicio, desde un punto de vista técnico, podemos definir el aprendizaje automático como el método de aprendizaje de la inteligencia artificial que toma como *input* datos y genera como *output* un modelo que puede ser usado para resolver un problema o para la toma de decisiones. El proceso de generación del modelo entrenado con datos sería el aprendizaje.

Del mismo modo y desde esta misma óptica podemos igualmente definirlo, en lugar de como método, como conjunto de algoritmos que proporcionan a los sistemas la capacidad de aprender y mejorar automáticamente a partir de la experiencia y acumulación de datos sin necesidad de estas programados previamente o de tener reglas previamente codificadas. Un ejemplo de ello sería el *software* de reconocimiento de imágenes.

Este método de aprendizaje identifica patrones en conjuntos de datos objeto de análisis, construye modelos y hace predicciones sin disponer de reglas o modelos predefinidos.

La falta de datos fue precisamente uno de los grandes problemas que tuvo la inteligencia artificial para desarrollarse durante sus primeros años, como he expuesto al analizar su origen y evolución.

La disrupción del Big data que ha facilitado Internet y el aumento de la capacidad de procesamiento y almacenamiento de los sistemas, han permitido la eclosión de la inteligencia artificial a nivel global.

Un ejemplo claro de ello es China, que se ha posicionado como una potencia mundial en inteligencia artificial no sólo por su capacidad de desarrollo sino de aplicación, especialmente, por su mayor disponibilidad de datos masivos frente a otros países.

Sin embargo, los datos por sí mismos no generan conocimiento, sino que este se genera gracias a su análisis, procesamiento e interpretación, lo que comporta también retos y riesgos, en la medida que los conjuntos de datos pueden contener datos erróneos o falsos que pueden pasar desapercibidos, así como datos asociados a contextos culturales, temporales o personales concretos que, fuera del mismo, no pueden ser valorados de la

misma manera y que, en consecuencia, pueden "viciar" el entrenamiento del sistema y, consecuentemente determinadas las decisiones y conducta del mismo.

Los sistemas deben ser entrenados con datos adecuados para permitirles adoptar la decisión más adecuada posible en cada contexto. Cuantos más datos se analicen por el sistema, mayor grado de precisión y adecuación tendrá su decisión.

Del mismo modo, dicho análisis debe permitir la comparación, análisis y extracción de patrones, lo que a su vez permite al sistema clasificar, discriminar y agrupar en categorías para su posterior utilización y decisión en cualquier contexto al que se someta el sistema. Todo ello lo cualifica para realizar predicciones a través de cálculos de probabilidad con un alto índice de acierto, acorde a la calidad y fiabilidad de los datos con los que se ha sido entrenado.

La utilidad de estas capacidades para el ser humano es infinita, en la medida en que aprenden y resuelven problemas gracias a la experiencia. Entre otros proyectos de inteligencia artificial en los que he tenido la suerte de colaborar desde mi óptica jurídica, significo uno de los más recientes cuyo objeto es la creación de un sistema de inteligencia artificial orientado al diagnóstico precoz de cáncer de piel:

> Un ejemplo:
>
> Actualmente en España, es muy frecuente que una persona que acude a su médico de familia con una peca, mancha o lunar para su examen, ni tan siquiera sea derivado inicialmente al dermatólogo para su examen, sino que incluso es el propio médico de familia el que obtiene una fotografía del mismo y envía electrónicamente una interconsulta al especialista, adjuntando la misma.
>
> Días, semanas o incluso meses después, el paciente recibe una respuesta de su médico de familia indicándole que no se aprecia sospecha de riesgo en la peca, mancha o lunar o, en caso contrario, se le remite al especialista para su examen, lo que sucederá semanas o meses después.

No obstante, el paciente derivado al especialista, cuando acuda al mismo, será valorado en base a su conocimiento y experiencia respecto de la literatura consultada y casos examinados durante su vida profesional.

Si un sistema fuese capaz de almacenar de manera anónima todas las imágenes de todos los casos vistos durante toda su vida por un millón de especialistas junto con su diagnóstico inicial y posterior confirmación, con el objetivo de entrenar al sistema y que pueda analizar toda esa información para su comparación con la muestra inicial y emitir un primer diagnóstico en segundos, para su posterior examen y validación por un médico especialista, pensemos en la agilidad, eficacia y optimización de los recursos limitados de los servicios de salud que el mismo puede comportar.

El sistema trabajaría sobre un volumen enorme de casos tratados por miles de profesionales en todo el mundo y esto permitiría compartir el conocimiento y el *expertise* profesional para activar o no, en función del pre-diagnóstico, todos los protocolos necesarios para confirmar el posible riesgo y, en su caso, inicio de las actuaciones médicas sin demora que, en su caso, procedan.

Y no sólo esto, estos sistemas convertirían al profesional médico en un especialista cualificado para el diagnóstico asistido con un espectro de experiencia en millones de casos, es decir, constituiría una herramienta de asistencia y mejora de la cualificación profesional humana, aumentando sus capacidades.

El aprendizaje automático puede ser supervisado, semisupervisado o no supervisado, y también aprendizaje por refuerzo según la naturaleza de los datos que recibe o *Reinforcement Learning*, en el que el sistema aprende por prueba y error.

En el aprendizaje automático supervisado -*Supervised Machine Learning*- se lleva a cabo un aprendizaje previo en base a un conjunto amplio y representativo de datos previamente etiquetado y clasificado a modo de muestra para entrenar al modelo y que permiten hacer predicciones y tomar decisiones. De este modo, el sistema aprenda a clasificar las muestras de entrada, generando así un modelo predictivo y realizando las adecuaciones necesarias en el modelo conforme a cada error detectado en la estimación del resultado.

Es decir, se trabaja con datos "etiquetados" y se genera un modelo predictivo basado en datos de entrada y salida.

Actualmente, se utiliza para el pre-diagnóstico médico o la detección de correos comerciales o publicitarios no deseado -*spam*-.

El aprendizaje automático no supervisado -*Unsupervised Machine Learning*- funciona de manera similar al supervisado, si bien, el sistema no dispone de un conocimiento previo, no dispone de datos etiquetados para el entrenamiento y solo tiene acceso los datos de entrada, pero ni etiquetados ni clasificados, por lo que ajusta su modelo predictivo tomando como única referencia los datos de entrada sin etiquetas extrayendo patrones. Operan "sobre la marcha" con el objetivo de encontrar patrones que permitan organizarlos de alguna manera.

Actualmente se utiliza, por ejemplo, en marketing para extraer patrones de datos masivos provenientes de redes sociales y crear campañas publicitarias totalmente segmentadas.

El aprendizaje por refuerzo -*Reinforcement Learning*- tiene como objetivo que el sistema aprenda a partir de la propia experiencia, de modo que se mejora la respuesta del modelo mediante su retroalimentación. Reagrupa un conjunto de técnicas para que los modelos aprendan secuencias de acciones en un entorno posiblemente incontrolado y/o desconocido. El sistema aprende observando y experimentando con el contexto le rodea, siendo capaz de tomar la mejor decisión ante diferentes situaciones, de acuerdo a un proceso de prueba y error en el que se recompensan las decisiones correctas, es decir, aprende en base a prueba-error.

Actualmente se utiliza para posibilitar el reconocimiento facial o el diagnóstico médico.

La evolución tecnológica ha permitido llegar a un aprendizaje más avanzado, el denominado *Deep Learning* o aprendizaje profundo, de modo que el sistema integra un algoritmo estructurado que emula el aprendizaje humano con el fin de obtener ciertos conocimientos, no requiriendo reglas programadas previamente, sino que el propio sistema es capaz de aprender por sí mismo para efectuar una tarea, a través de una fase de entrenamiento previo. Es decir, el propio sistema tiene la capacidad de aprender de los

datos y de la experiencia y, a partir de los mismos, puede definir nuevos criterios y tomar decisiones autónomas.

La principal distinción respecto de los anteriores radica en la estructura y procesamiento de la información que imita las redes neuronales del cerebro humano, donde una señal de entrada es procesada y tramitada por cientos de neuronas entrelazadas para extraer una conclusión.

Actualmente se utiliza para el diagnóstico médico, análisis financiero o detección de fraudes, entre otros usos, y en cosas tan cotidianas como traductores inteligentes, lenguaje natural hablado y escrito por parte de asistentes virtuales en *smartphones* o equipos, reconocimiento de voz, interpretación semántica o el reconocimiento facial.

3.6. Sistemas basados en conocimiento o sistemas expertos.

En congruencia de las razones por las que he abordado la definición y características básicas de algunas técnicas y enfoques sobre los que se sustentan los sistemas inteligentes, considero igualmente necesario hacer una breve referencia a los caracteres esenciales de los sistemas expertos.

Algunos sistemas de inteligencia artificial no se basan en el aprendizaje según el volumen masivo de datos que asimilan y analizan, es decir, "sistemas basados en datos", sino en el conocimiento y en un motor de inferencia deductiva que extrae un juicio o conclusión a partir de unos hechos. Se trata de los denominados "sistemas basados en conocimiento", "sistemas basados en reglas" o "sistemas expertos".

La diferencia entre los sistemas de inteligencia artificial basados en datos y los basados en reglas o expertos, radica en que, mientras que en unos entran los datos, el algoritmo procesa esos datos según sus instrucciones y sale el resultado, en los otros se generan algoritmos en base a datos para conseguir el resultado deseado.

3.7. La inteligencia artificial como ciencia, ingeniería del conocimiento y tecnología habilitadora.

Por último y como he apuntado al abordar su definición, la inteligencia artificial está siendo abordada en la actualidad desde dos enfoques principales, como ciencia y como ingeniería del conocimiento[77].

La primera, como la ciencia encargada de buscar una teoría computable del conocimiento humano, es decir, de comprensión de los procesos cognitivos y de desarrollo de modelos conceptuales que constituye uno de los primeros objetivos de la inteligencia artificial.

La segunda, pretende reproducir formalmente las inferencias obtenidas del análisis de los procesos cognitivos y programar los sistemas.

No obstante, en la actualidad se está también abordando no sólo como ciencia y campo de investigación, sino como un sistema tecnológico habilitador[78] que permite la innovación en distintos sectores y permite generar modelos económicos disruptivos, incluyendo el desarrollo de procesos de transformación digital.

4. Clases de inteligencia artificial

La inteligencia artificial puede ser clasificada desde distintos enfoques y criterios, si bien, me focalizaré en su clasificación básica desde la perspectiva científico-técnica, ética y jurídica.

---

[77] PALMA, J. Y MARÍN, R. (2008). *Inteligencia artificial. Técnicas, métodos y aplicaciones*. McGraw-Hill - Interamericana de España, 2008.
[78] BONFANTI, M. E. (2020). *Artificial Intelligence and Cybersecurity: A Promising but Uncertain Future*. ARI 139/2020. Real Instituto Elcano. 09.12.2020. P. 1; COOMBS, T. (2018), "Artificial Intelligence & Cybersecurity for Dummies". IBM 2018. https://hosteddocs.ittoolbox.com/ai_cybersecurity_dummies.pdf. Consultado el 29.01.2021.

La inexistencia de una definición consensuada de inteligencia artificial en el ámbito científico y la existencia de un cierto consenso respecto de sus objetivos han llevado a las voces más autorizadas a nivel científico, ético y jurídico a diferenciar y clasificar distintos tipos de inteligencia artificial.

Desde un punto de vista científico-técnico, Rusell y Norvig[79] diferencian en sus estudios distintos sistemas de inteligencia artificial en atención a sus objetivos, citando a otros expertos posicionados del mismo modo en la conceptualización de algunos de ellos:

a) Los sistemas que piensan como humanos, es decir, sistemas que tratan de emular el pensamiento humano, por ejemplo, las redes neuronales artificiales. Se trata de sistemas que automatizan actividades que se vinculan con procesos de pensamiento humano u operaciones mentales como la toma de decisiones, la resolución de problemas y el aprendizaje. Alineados con esta categorización y definición destacar a científicos como Bellman[80] y Haugeland[81], conforme a sus argumentos y posicionamientos recogidos en algunas de sus publicaciones.

b) Los sistemas que actúan como humanos o, al menos, tratan de actuar como tales, es decir, imitan el comportamiento humano, por ejemplo, la robótica. Se trata de sistemas con capacidad de procesar el lenguaje natural, representar conocimiento, razonar automáticamente y aprender para adaptarse a nuevas circunstancias. Alineados con esta categorización y definición, destacar a investigadores de referencia como Kurzweil[82], Rich y Knight[83]. conforme recogen en algunas de sus publicaciones.

c) Los sistemas que piensan racionalmente, es decir, que tratan de imitar el pensamiento lógico racional del ser humano aplicando la lógica para alcanzar

[79] RUSSELL, STUART J. Y NORVIG, P. (2009). "*Artificial intelligence: a modern approach*". 3.ª edición. Upper Saddle River, N.J. Prentice Hall 2009. Pp 4-5.
[80] BELLMAN, R. (1978). *An Introduction to Artificial Intelligence: Can Computers Think?*. Boyd & Fraser Publishing Company. 1978.
[81] HAUGELAND, J. (1985) *Artificial Intelligence: The Very Idea*. MIT Press, Cambridge. 1985.
[82] KURZWEIL, R. (1990). *The Age of Intelligent Machines*. MIT Press, Cambridge.1990.
[83] RICH, E. Y KNIGHT, K. (1991). *Artificial Intelligence*. McGraw-Hill, 1991.

conclusiones, por ejemplo, los sistemas expertos. Alineado con esta categorización y definición, destacar, entre otros investigadores de referencia, a Winston[84].

d) Los sistemas que actúan racionalmente, es decir, sistemas que tratan de emular el comportamiento humano de forma racional, ampliando la racionalidad más allá de la lógica e incluir elementos como la incertidumbre, la autonomía, el cambio, por ejemplo, los agentes o máquinas inteligentes. Alineados con esta categorización y definición, destacar a investigadores de referencia como Poole, Goebel, Mackworth[85] y Nilsson[86], conforme evidencian en algunas de sus publicaciones.

Desde un punto de vista ético y filosófico, siguiendo a Dreyfus[87], podemos distinguir entre inteligencia artificial "simbólica" -cognitiva psicológica- o "subsimbólica" -contextual-. Este autor considera que la verdadera inteligencia debe integrar el contexto con el texto -cognitivismo por medio de la manipulación de los símbolos y el lenguaje computacional-.

Por su parte, el filósofo John Searle[88] distingue entre inteligencia artificial "fuerte" -creativa e improvisadora, con autonomía y capacidad de improvisación- y "débil" -automatizada y mecánica, que carece de autonomía y capacidad de improvisación-. Aspecto fundamental para la determinación de la responsabilidad ante daños causados por sistemas inteligentes.

Otras definiciones que complementan esta diferenciación, identifican la inteligencia artificial "fuerte" o "general" como aquella que tiene las mismas características que la inteligencia humana y que tendría la capacidad de aprender, razonar, entender, aplicar sentido común y comprender la relación de causalidad, mientras que la "débil" o "estrecha" la identifican con aquella con capacidad de resolver una tarea muy concreta y limitada pero que no podría resolver otras relacionadas con la misma desde una

[84] WINSTON, P. H. (1992). *Artificial intelligence*. Addison-Wesley, Reading, MA , 3ª Edición. 1992.
[85] POOLE, D. I.; GOEBEL, R. G. Y MACKWORTH, A. K. (1998). *Computational Intelligence: A Logical Approach*. Op. cit.
[86] NILSSON, N. (1998). *Artificial Intelligence: A New Synthesis*. Elsevier 1998.
[87] DREYFUS, H. (1972). *What computers can't do: The limits of artificial intelligence*. Editorial HarperCollins, Londres, 1972.
[88] SEARLE J. (1980). *Minds, brains, and programs», Behavioral and Brain Science*. 3 (3). Cambridge (UK), septiembre 1980. Pp.417-457. Y también se recoge en los argumentos de este autor recogidos en SEARLE, J. (2000). *El misterio de la conciencia*. Editorial Paidos Ibérica, Madrid, 2000.

perspectiva humana[89]. Según esta definición, la inteligencia artificial "fuerte" o "general" sería lo que todavía no existe, en la medida que los sistemas actuales no pueden "entender".

Otros expertos[90] en el ámbito jurídico como Barrio Andrés, hablan de inteligencia artificial "débil" o "estrecha" para referirse a aquella que se centra en una tarea concreta y limitada en sectores diversos, esto es, una aplicación específica, y sin autoconsciencia, así como de inteligencia artificial "fuerte" -asimilada igualmente a los conceptos "Inteligencia artificial general" (IAG) o "Inteligencia profunda"-, para referirse a aquella que igualaría o incluso excedería la inteligencia humana, siendo definida como "la capacidad de razonar, representar el conocimiento, planificar, aprender, comunicarse en lenguaje natural e integrar todas estas habilidades hacia un objetivo común". En relación con ambas tipologías los expertos adicionan el concepto de "súper inteligencia" para referirse a la capacidad de razonamiento y resolución de problemas notablemente superior a la de cualquier ser humano y que se asocia a la denominada "singularidad tecnológica" a la que hice referencia anteriormente al analizar el origen y evolución histórica del concepto de inteligencia artificial.

Gallego Sánchez[91], siguiendo a Dillon y Shemtov, diferencia todavía más las mismas, considerando que la inteligencia artificial débil -o *narrow AI*- sería capaz de llevar a cabo tareas predefinidas, la general o *general AI* tareas intelectuales similares a las humanas y la súper inteligencia o *super AI* sería la que superaría a la inteligencia humana en todos los aspectos.

A la vista de las distintas concepciones, considero que "Inteligencia artificial fuerte" e "Inteligencia artificial general" no son términos exactamente iguales, en la medida que toda inteligencia artificial "fuerte" será "general", pero no a la inversa, en la medida que la primera pretende simular más el comportamiento humano o la inteligencia humana, pero no necesariamente igualarla o incluso superarla.

[89] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Op.cit. Pos. 3844.
[90] BARRIO ANDRÉS, M. (2020). *Manual de Derecho Digital*. Tirant lo Blanch. Valencia 2020. P. 57.
[91] GALLEGO SÁNCHEZ, E. (2019). "La patentabilidad de la inteligencia artificial. La compatibilidad con otros sistemas de protección". *La Ley Mercantil*. Nº 59, de 1 de junio 2019. Wolters Kluwer 2019. P. 4.

Un ejemplo de inteligencia artificial débil sería "Alexa", "Cortana", "Siri" o "Google Translate". La inteligencia artificial fuerte sería la que se espera que llegue en el futuro conforme evolucione el desarrollo de ésta, y la "súper inteligencia" es la que muchos auguran en unas décadas y que otros consideran inviable.

Conforme a esta distinción científica, ética y filosófica, la inteligencia simbólica y débil sería la que estaría presente en un ordenador, *smartphone* o electrodoméstico, y es la subsimbólica y fuerte la que se aproximaría al concepto más avanzado de inteligencia artificial y que comportará los mayores desafíos éticos y jurídicos en el futuro.

Desde un punto de vista jurídico, tomando como referencia los últimos documentos de trabajo, informes, comunicaciones y propuestas de la UE relativas a la definición de un marco ético y jurídico europeo de la inteligencia artificial, la inteligencia artificial también puede clasificarse en dos tipos:

a) La inteligencia artificial no integrada "basada en *software"*, como asistentes virtuales, software de análisis de imágenes, motores de búsqueda o sistemas de reconocimiento de voz y rostro.

b) La inteligencia artificial integrada en máquinas u otros productos, como robots, drones, vehículos autónomos o Internet de las Cosas -IoT por sus siglas en inglés-.

Una vez expuestas algunas de las principales clasificaciones de la inteligencia artificial desde un punto de vista científico, filosófico-ético y jurídico, desde un enfoque de riesgos, podríamos igualmente clasificar los sistemas de inteligencia artificial, en primer lugar, como sistemas clasificados o no clasificados. Y, en caso de clasificación y en atención al su nivel de riesgo, estos podrían clasificarse como de riesgo alto, medio o moderado o bajo.

La Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial[92] pretende regular la responsabilidad derivada del funcionamiento de cualquier sistema de

[92] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))

inteligencia artificial, sea se "débil" o "fuerte", sin embargo, como se ha expuesto, la definición de "sistema de inteligencia artificial" que contiene en su artículo 3 se sitúa en la órbita de sistemas más avanzados, con autonomía y posibilidad de actuación no limitada a las instrucciones dadas en su concepción y, en consecuencia, plantea un concepto de inteligencia artificial superior a la básica o "débil" que es la que se haya en aplicación en la actualidad y a la que pretende ser orientado el marco regulatorio que propone, como he expuesto anteriormente.

Dicha Propuesta define los sistemas de alto riesgo en su artículo 3.I.c) en los términos anteriormente indicados, esto es, como aquellos que funcionan de forma autónoma con potencial significativo para causar daños o perjuicios a una o más personas de manera aleatoria y que excede lo que cabe esperar razonablemente, partiendo que la entidad del potencial dependerá de la relación entre la gravedad del posible daño o perjuicio, el grado de autonomía de la toma de decisiones, la probabilidad de que el riesgo se materialice y el modo y el contexto en que se utiliza el sistema de inteligencia artificial.

Sin embargo, la reciente Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial *-Artificial Intelligence Act-*, de 21 de abril de 2021[93], introduce una nueva clasificación de los sistemas inteligentes bajo un enfoque de riesgos, como lo expondré en su análisis en el capítulo IV, en particular, los sistemas de inteligencia artificial de riesgo inadmisible (prohibidos), salvo bajo determinados condicionantes, y los de riesgo admisible (permitidos), entre los que se limitada a diferenciar los considerados de alto riesgo, regulados minuciosamente en el mismo y el resto, que incluirían los de nivel medio o limitado, los de nivel bajo o mínimo, y los no clasificados. Del mismo modo, incorpora definiciones específicas de concretos sistemas, como expondré al analizar esta propuesta en el capítulo IV.

[93] COM (2021) 206 final 2021/0106 (COD)

5. Consideraciones finales

La inteligencia artificial forma parte de nuestra vida cotidiana y su despliegue y aplicación en todos los ámbitos de nuestra vida crece de manera incesante y de manera exponencial, especialmente ante la mayor capacidad de computación, su interacción con otras tecnologías, aumento de sus capacidades y la disposición de datos masivos.

El potencial que supone para satisfacer necesidades, resolver problemas, mejorar nuestras vidas y aumentar la capacidad innovadora y competitiva de las empresas, ya sea sola o integrada con otras tecnologías, la están convirtiendo en un medio o recurso esencial para ciudadanos, empresas y gobiernos, pero no debe ser concebida como un fin en sí misma sino como un medio.

Su cualificación constante, el aumento de sus capacidades, su desarrollo exponencial, su creciente accesibilidad y su aplicabilidad casi infinita en todo tipo de actividades y procesos la han convertido en una utilidad cada vez más imprescindible para el ser humano, como lo fue la rueda o el fuego.

Sin embargo, en paralelo, todo ello comporta riesgos crecientes que deben ser adecuadamente identificados y gestionados, para lo que el desarrollo, consenso y adhesión a marcos éticos aportan un enorme valor, si bien, a mi juicio, insuficiente para poder garantizar seguridad, confiabilidad y un adecuado equilibrio entre todos los intereses en juego de los distintos agentes involucrados en el desarrollo, despliegue, uso, aplicación y explotación comercial de la misma.

Las principales potencias mundiales han iniciado una carrera desenfrenada por la supremacía mundial de la inteligencia artificial a nivel mundial, con China, EEUU y UE a la cabeza, si bien, ante la ausencia de marcos jurídicos que regulen la misma, considero necesario acompañar su desarrollo, despliegue, uso, aplicación y explotación con marcos reguladores que conviertan en vinculantes muchos de los principios, valores y normas éticas sobre la misma, que contribuyan a minorar los retos y riesgos que supone y garanticen la seguridad para todos en su desarrollo, uso aplicación, incluyendo gobiernos, empresas y ciudadanos.

En este sentido, la UE está teniendo un protagonismo internacional por sus esfuerzos en promover el consenso de estos principios y valores éticos.

Las propuestas regulatorias deberán partir de una definición previa clara y concisa de la realidad compleja que pretenden regular, esto es, la inteligencia artificial y distintos conceptos relacionados con la misma, para concretar su objeto y alcance con precisión, las cuales deberán ser flexibles, evolutivas e ir adecuándose a la evolución que experimente la inteligencia artificial en lo sucesivo.

Además, en mi opinión, deberá empezar a valorarse no sólo la regulación de sistemas dotados de inteligencia artificial débil, sino posibles sistemas con capacidades más avanzadas, así como contemplar un conjunto de normas básicas esenciales para cualquier sistema inteligente, cualquiera que sea su riesgo inicial, por sus propias características, posibles capacidades y riesgos asociados.

Uno de los aspectos fundamentales que, a mi juicio, estos nuevos marcos reguladores deberán contemplar y abordar adecuadamente con las finalidades indicadas, además del carácter vinculante de determinados principios y normas éticas, es la seguridad y la responsabilidad por los daños derivados del uso de la inteligencia artificial, adaptando, complementando o modificando las disposiciones europeas o locales reguladoras de la misma.

El Parlamento Europeo aprobó en octubre de 2020 tres resoluciones sobre inteligencia artificial, dos de las cuales incorporaban sendas Propuestas de Reglamentos en materia ética y de responsabilidad civil, que serán igualmente abordadas y objeto de análisis en los próximos capítulos. Recientemente, se ha publicado una nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, pretendiendo establecer un conjunto de normas armonizadas sobre inteligencia artificial.

Todas estas novedosas iniciativas parten de una definición, objeto y alcance distintos, incluso de una definición de inteligencia artificial diferente, por lo que considero esencial armonizar y definir jurídicamente la realidad que se pretende regular y los objetivos

pretendidos con ello, dado que la opción elegida en la última propuesta referenciada se centra en una inteligencia artificial débil y, dentro de esta, en los sistemas de riesgo inadmisible y en los de alto riesgo tipificados conforme a la misma, no regulando el resto de sistemas que podrían integrar capacidades susceptibles de generar riesgos de distinto y cambiante alcance durante todo su ciclo de vida, desde el diseño hasta su aplicación, y que podrían integrar de origen cierto grado de autonomía que, conforme a las propuestas previas los podría clasificar de inicio como de alto riesgo y, consiguientemente, sujetos a las normas éticas y/o obligaciones jurídicas previstas en aquéllas.

Capítulo II

Riesgos, retos y regulación: Seguridad física, lógica, moral y jurídica

1. Introducción

La tecnología aporta indudables ventajas al ser humano y a la sociedad en general, solucionando problemas, satisfaciendo necesidades y mejorando la vida y el mundo en el que vivimos.

La inteligencia artificial en mayor medida ante su potencial, especialmente ante sus capacidades asociadas de automatización, "autonomía" y autoaprendizaje y de relación, interacción, gobierno y utilización de distintas tecnologías, y en todo tipo de ámbitos de nuestra vida.

La mejora de la atención médica, la disposición de medios de transporte más seguros, la accesibilidad a una mejor información, educación y formación, mejora en la atención de personas mayores y con discapacidad, la obtención de productos y servicios personalizados más ágiles y a un menor coste, la mejora en la gestión de seguridad de la información, de la seguridad laboral, la gestión responsable y eficiente de infraestructuras, redes y recursos, la mejora de los procesos industriales, el aumento de la eficiencia de la agricultura, la mejora de la sostenibilidad medioambiental o la mitigación del cambio climático, la mejora de la lucha contra la delincuencia, la mejora en la administración de justicia o en la prestación de servicios públicos, son simplemente algunas de las ventajas que supone la inteligencia artificial.

La inteligencia artificial supone una ayuda para la mejora de la vida de las personas, pero también grandes oportunidades para empresas y Administraciones públicas, tanto para la mejora de su organización, gestión y procesos, como para el desarrollo de nuevos modelos de negocio, productos y servicios basados en la misma.

También para gobiernos, mediante el refuerzo de la democracia, la diversidad y la transparencia, el acceso a la información de calidad, el gobierno de los datos, la prevención de la desinformación y de los ciberataques.

También para fuerzas y cuerpos de seguridad del estado para la prevención de delito. Para la administración de justicia, esto es, para juzgados y tribunales como medio para agilizar y mejorar la justicia. Y también para la defensa del Estado y los servicios de inteligencia.

Sin embargo, su evolución, despliegue y aplicación comporta riesgos y retos de distinta naturaleza para todos los sujetos individuales o colectivos relacionados con la misma, lo que exige, de un lado, su adecuada identificación y gestión y, de otro la construcción de un marco regulador adecuado que, entre otros objetivos, contribuya a ello, al objeto de alcanzar una inteligencia artificial segura y confiable en beneficio de ciudadanos, empresas, mercados, Administraciones públicas, gobiernos o estados.

En mi opinión, el desarrollo y aplicación de la inteligencia artificial requiere una adecuada reflexión, valoración y gestión de lo que denomino "las 4R", esto es, riesgos, retos, responsabilidad y regulación, con un objetivo fundamental, que es conseguir una inteligencia artificial confiable y segura desde el punto de vista físico, lógico, moral y jurídico.

En este sentido, destacar que el riesgo no es la tecnología en sí misma, sino el uso que se hace de ella, quién la utiliza y para qué.

La inteligencia artificial puede usarse en el campo militar y de la inteligencia para anticiparse a la comisión de delitos como medio de prevención, detección, contención, defensa, respuesta y recuperación, también para minimizar la probabilidad e impacto de las amenazas cibernéticas, pero también incluso para investigar y actuar en materia de piratería o de *phishing*.

La inteligencia artificial está siendo también utilizada por el lado más oscuro del ciberespacio para realizar ciberataques más precisos y con mayor impacto a gobiernos, infraestructuras, Administraciones públicas, empresas y ciudadanos, si bien, en relación con la ciberseguridad, la inteligencia artificial constituye, a su vez, una herramienta

imprescindible y esencial para precisamente prevenir y detectar precozmente este tipo de ataques, adelantándose a los mismos, así como para minorar su probabilidad, impacto y efectos, en lo que ya es una ciberguerra en un nuevo entorno, esto es, el ciberespacio, y con nuevas armas, entre otras, la inteligencia artificial.

El uso y aplicación creciente de sistemas de inteligencia artificial cada vez más complejos y con mayores capacidades en todas las esferas de nuestra vida, comporta paralelamente riesgos potenciales que deberán ser adecuadamente identificados, valorados y gestionados, y desde el diseño y por defecto, y asegurando en todo momento la supervisión y control humano, con el objetivo de garantizar el desarrollo, despliegue y uso seguro y confiable de aplicaciones, sistemas, productos y servicios dotados de inteligencia artificial.

Los riesgos que plantea no son siempre visibles sino, en ocasiones, son invisibles y pueden revestir mayor gravedad para la sociedad en general como, por ejemplo, el sesgo en los algoritmos.

De hecho, cuando hablamos del temor que en ocasiones han suscitado los robots para parte de la sociedad, se debe significar que lo más peligroso no son los robots en sí mismos, sino determinados sistemas inteligentes que puedan estar detrás de los mismos o basados exclusivamente en *software*, invisibles, intangibles y, en ocasiones, opacos.

Los bienes jurídicos a proteger frente a estos riesgos o amenazas son la propia persona, su vida, su salud y su integridad física y moral, así como sus derechos, personalísimos o no, incluidos los fundamentales, pero también los bienes e infraestructuras (críticas o no) así como, a mi juicio, la libre competencia, la innovación y la competitividad empresarial, la democracia y la capacidad de liderazgo de los estados.

De este modo estos riesgos nos afectan a todos y todas, personas, empresas, Administraciones Públicas, gobiernos, etc.

La UE ha fijado como objetivo estratégico su consolidación como líder mundial en la economía de los datos y sus aplicaciones, así como en inteligencia artificial. En este contexto, el Parlamento Europeo está liderando distintas propuestas normativas en

materia de inteligencia artificial a las que ya he hecho una referencia general en el capítulo anterior.

La Propuesta de Reglamento europeo sobre el régimen de responsabilidad civil en materia de inteligencia artificial incorporado en la Resolución del Parlamento Europeo, de 20 de octubre de 2020[94], recoge en su ámbito de aplicación algunos de estos bienes a proteger, en particular, la vida, la salud y la integridad física de una persona física, así como los bienes de una persona física o jurídica.

Por su parte, la Propuesta de Reglamento europeo sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas incorporado en la Resolución del Parlamento Europeo, de 20 de octubre de 2020[95], entre otros aspectos, establece principios y normas éticas para garantizar la seguridad de los sistemas e inteligencia artificial.

En los próximos apartados analizaré los principales riesgos y retos asociados al despliegue, funcionamiento, aplicación y uso de la inteligencia artificial, con especial énfasis en su seguridad, física, lógica, moral y jurídica, como base para la creación de un ecosistema fiable y confiable para todas las partes interesadas.

2. Evolución de la inteligencia artificial: Necesidad de marcos éticos, jurídicos y de seguridad.

Cuando surge una nueva tecnología en constante desarrollo y evolución es difícil saber qué tipo de impacto tendrá en el mundo a medio y largo plazo, especialmente ante su interacción con otras.

[94] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))
[95] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))

La inteligencia artificial existe desde hace décadas si bien, no ha sido hasta fechas recientes cuando se está produciendo su eclosión en su despliegue y aplicación con todo su potencial, poder y bondades, pero también con sus retos y sus riesgos.

Compartiendo una vez más las palabras del astrofísico Stephen Hawking "La inteligencia artificial puede ser lo mejor o lo peor que nos ha sucedido a la humanidad, todavía no lo sabemos"[96].

La mejora de sus capacidades y de las tecnologías asociadas junto a la mayor disponibilidad de datos permite el aumento exponencial de su velocidad de desarrollo y, sobre todo, de aplicación en todo tipo de sectores y ámbitos de la sociedad.

La inteligencia artificial avanza de manera incesante, también su supuesta "autonomía", su capacidad de aprendizaje -*Machine Learning* y *Deep Learning*-, y en consecuencia, su eficiencia, agilidad y, en ocasiones, su impredecibilidad consecuente.

Del mismo modo, crece de forma exponencial su despliegue y aplicación en todo tipo de ámbitos -domésticos, administrativos, empresariales y gubernamentales- y sectores, estrechando su relación con el ser humano de manera cada vez más cotidiana.

La capacidad de autoaprendizaje está conllevando la relativa imprevisibilidad de la inteligencia artificial en su relación e interacción con su entorno, contexto y personas, que debería ser escasa o nula si realmente sometemos la inteligencia artificial a la calidad, supervisión y control humano y establecemos las adecuadas limitaciones tanto en su diseño como durante todo su ciclo de vida, y en cualquier caso susceptible de anulación y reversibilidad conforme a los marcos éticos que se están debatiendo a nivel internacional y europeo.

Constituye un paradigma tecnológico complejo con un potencial de impacto elevadísimo en el ser humano, en especial, en los principios y valores europeos, en los derechos

[96] HAWKING, S. (2016). Discurso a cargo del astrofísico inglés durante la inauguración del *Centro Leverhulme para el futuro de la inteligencia* -CFI por sus siglas en inglés-, en la Universidad de Cambridge, el 20.10.2016. Recuperado de: https://www.rtve.es/noticias/20161020/stephen-hawking-inteligencia-artificial-sera-mejor-peor-pase-humanidad/1428740.shtml. Consultado el 10.12.2020.

fundamentales, así como en bienes e intereses de todos los sujetos que puedan verse afectados por la misma.

El objetivo de alcanzar una inteligencia artificial plenamente autónoma atentaría frontalmente contra los principios y normas éticas que hoy se están consensuando a nivel internacional, especialmente en materia de seguridad, responsabilidad y la precitada supervisión y control humano.

El objetivo científico de crear sistemas de inteligencia artificial dotados de consciencia, incluso de sentimientos y emociones, y gobernados por algoritmos que operen como la mente humana es actualmente irrealizable dado el estado de desarrollo de la tecnología, conforme se expone en esta investigación, lo que no significa que sea posible en el futuro, por lo que todavía hoy forma parte de la ciencia-ficción más que de la ciencia.

No obstante, hemos visto como la historia de la humanidad refleja la inquietud del ser humano de superarse y convertir lo imposible en posible. Los expertos no son unánimes a la hora de desechar plenamente esta posibilidad.

A mi juicio, la primera de las preguntas a contestar sobre esta posible evolución y realidad futura sería sobre la necesidad de crear estos sistemas con estas capacidades.

Otro de los principales aspectos a valorar sería la posibilidad de dotar en el futuro de sentimientos o emociones a algunos sistemas inteligentes.

La gran mayoría de sistemas inteligentes actuales no necesitan sentimientos o emociones para desarrollar sus funciones y tareas. No obstante, en algunos sistemas con mayor interacción humana, la detección de sentimientos y estado emocional de las personas con las que se relacione para adoptar una decisión o conducta adaptada al contexto, puede ser de gran valor.

A modo de ejemplo, pensemos en asistentes virtuales para personas de edad avanzada o aplicaciones para la prevención precoz del acoso escolar.

En la actualidad existen sistemas inteligentes capaces de detectar emociones en el rostro de una persona a la que se está haciendo una entrevista de trabajo a través de videoconferencia[97], lo que posibilita identificar el entusiasmo u honestidad del candidato.

Del mismo modo, se están trabajando en proyectos piloto[98] y desarrollando sistemas de inteligencia artificial que podrían llegar a detectar el estado mental y emocional de los conductores para prevenir accidentes.

Las emociones y los sentimientos son propios del ser humano.

Conforme significa García Cuesta[99], el sentimiento es el estado del ánimo y las emociones reflejarían el estado de este. Según el mismo, un sistema de inteligencia artificial podría ser capaz de mimetizar las conductas asociadas a una emoción e incluso desarrollar su propio modo artificial de "sentirlas" para poder integrarse con el ser humano. Conforme destaca este autor, algunos sistemas actuales tienen la capacidad de detectar y generar diferentes emociones con la voz, generar un lenguaje con expresiones emocionales o definir estados afectivos, y que podrían llegar a considerarse en el futuro "sentimientos artificiales" o, como suelo utilizar "sentimientos sintéticos", diferentes a los que caracterizan a los seres humanos.

Desde mi modesto punto de vista, concibiendo una emoción como un proceso fisiológico que dispara una serie de respuestas en el organismo con el objetivo de cumplir una función protectora o adaptativa, resulta difícil, por no decir que impensable con el estado de la tecnología actual, que pueda llegar a crearse un sistema inteligente con sentimientos o emociones, sin perjuicio de que las detecte y las simule.

La mayoría de los expertos[100] consideran que, si en el futuro se consigue replicar en un algoritmo el modo en que el cerebro genera las emociones, es muy complicado que pueda llegar a sentirlas o a desarrollar una "Inteligencia emocional", dado que las emociones

[97] NILSSON, P. (2018). "How AI helps recruiters track jobseekers' emotions". Publicado en *Financial Times*. Febrero 2018. Disponible en: https://www.ft.com/content/e2e85644-05be-11e8-9650-9c0ad2d7c5b5
[98] Recuperado de www.affectiva.com
[99] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Op.cit. Pos. 30.
[100] BENJAMINS, R. Y SALAZAR, I. (2020). "*El mito del algoritmo*". Op.cit. Pos. 3069.

humanas dependen tanto de la percepción que tenemos del exterior, como de nuestro interior y estos sistemas tienen una realidad muy diferente.

Sus estímulos llegan a través de sus sensores externos -audio, video o infrarrojos- y no tienen sensores internos, por lo que no podrían generar emocionas propias de la introspección.

Es más, para algunos investigadores como Raúl Arrabales[101], es poco útil intentar dotar de consciencia a la inteligencia artificial para las tareas que se quieran automatizar, salvo en supuestos muy concretos.

Los sistemas de inteligencia artificial actuales carecen de emociones o sentimientos, y no parece que las necesiten, aunque estos sistemas pueden ser un complemento del ser humano en labores asistenciales y terapéuticas que pueden exigir que lo parezca y que puedan mostrarse como empáticos.

Otro de los aspectos a considerar y reflexionar sería la consciencia, concebida como el estado de la mente que nos permite apercibirnos de nuestra propia existencia, del resto del mundo y de las cosas que pasan o, como indica el precitado Arrabales, "algo así como una pantalla mental donde el cerebro presenta continuamente la información que necesitamos conocer en cada momento para guiar el comportamiento". Sin embargo, como destaca este experto, no todo lo que procesa el cerebro humano acaba en un resultado consciente.

La cuestión es que todavía hoy no sabemos cómo se crea la consciencia en el ser humano por lo que resulta imposible replicarla. Además, el estado de la técnica actual no permitiría dotar de consciencia real a un sistema inteligente sin perjuicio de que pudiera llegar a parecer que la tiene, si bien, en cualquier caso, considero que, de nuevo, la primera pregunta o reflexión a efectuar sería sobre su necesidad y pertinencia de manera previa a continuar dedicando esfuerzos perseverantes en dicha dirección, especialmente en

[101] BENJAMINS, R. Y SALAZAR, I. (2020). "*El mito del algoritmo*". Op. cit. Pos. 3069.

contextos específicos. Todo ello no obsta a que en el futuro sea posible lo que hoy es imposible.

Según el investigador citado, si en el futuro se descubre el modo de procesar la información sobre el que se base la consciencia, podría ser replicado en sistemas inteligentes, con las reflexiones que ello abriría en relación con la responsabilidad jurídica y la posible atribución a sistemas inteligentes, así como en relación a la posible creación futura de una personalidad electrónica conforme analizaré en capítulos posteriores.

Para distintos científicos nunca será posible crear sistemas inteligentes con consciencia, en la medida que la misma se genera en virtud de elementos propios de los sistemas biológicos. Para otros, como Christoph Koch, jefe científico del *Allen Institute de Seatle*, sí podrían llegar a tenerla en el futuro.

En cualquier caso, tampoco existe un consenso a nivel científico sobre si la consciencia es un requisito necesario para llegar a una "Inteligencia artificial general", la cual no sería necesaria para exhibir un comportamiento lógico, pero sí para distinguir entre lo bueno y lo malo, lo que de nuevo permitiría abrir un debate alrededor de una posible imputabilidad de responsabilidad civil y/o penal, en el caso de una futura personalidad jurídica.

Por último, otro aspecto importante a valorar, relacionado con los anteriores, sería la posibilidad de dotar de empatía artificial a los sistemas inteligentes, aspecto que, para algunos expertos, entre otros, el precitado García-Cuesta[102], constituye un gran reto para que los sistemas inteligentes puedan coexistir con el ser humano, en la medida que mejoraría la relación con los mismos y tendría mayor grado de aceptación. No obstante, también comportaría aspectos negativos como los cambios sociológicos y su influencia en las decisiones individuales humanas.

En relación con todos estos aspectos analizados surge la posibilidad de que la inteligencia artificial evolucionara hasta el punto de llegar a la denominada "Inteligencia artificial

[102] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Op. cit. Pos. 3151.

general" o "súper inteligencia". Tampoco existe consenso en la comunidad científica sobre si ello será factible y, en caso afirmativo, sobre cuándo.

Los expertos que si consideran factible una inteligencia artificial general que iguale o supere la inteligencia humana tampoco coinciden en cuándo. Martin Ford[103] recoge algunas de las predicciones recogidas por distintos expertos como Barrat que, entre 200 expertos encuestados, significaba que sólo un 2% considera que nunca ocurrirá, mientras que un 42% la sitúa en 2030, el 25% en 2050 y el resto en 2100.

Según Ford, citando a Hawking, Tegmark y Wilczek[104], significa que un ordenador que superara el nivel de inteligencia humana podría ser capaz de "ser más astuto que los mercados financieros, de inventar más que los investigadores humanos, de manipular más que los líderes humanos y de desarrollar armas que ni siquiera podemos entender". Descartar todo esto como ciencia ficción podría resultar ser "potencialmente nuestro peor error en la historia".

Por su parte, según Nick Bostrom[105] la predicción media de los investigadores participantes en su estudio sitúa la inteligencia artificial general (AGI) en 2040 y la superinteligencia para su irrupción en las tres décadas siguientes.

Ray Kurzweil[106], un experto de referencia en inteligencia artificial que he citado previamente, concibe un futuro de interrelación más que de sustitución, en el que los seres humanos y las máquinas se habrán fusionado por completo, es decir, la síntesis hombre-máquina a la que he hecho referencia en el capítulo anterior. Entre sus predicciones, considera que en 2029 dispondremos de ordenadores con una inteligencia comparable a la de los seres humanos -inteligencia artificial general- y que se llegará a alcanzar la singularidad hacia 2045, cuando el ser humano multiplicará su inteligencia efectiva por mil millones al fusionarse con la inteligencia que ha creado.

---

[103] FORD, M. (2015). *Rise of the Robots*. Basic Books. Nueva York, 2015. P. 231.
[104] FORD, M. (2015). *Rise of the Robots*. Op. cit. P. 229.
[105] BOSTROM, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press. Oxford. 2014. P. 19.
[106] REEDY, C. (2017). "Kurzweil Claims That the Singularity Will Happen by 2045". Publicado en *Futurism* el 10 de mayo de 2017.

Bernhardt L. Trout[107], un experto en inteligencia artificial de referencia mundial, considera que sí es posible la denominada "singularidad" y que, de hecho, ya hemos empezado ese camino ignorando cuál es su objetivo, por lo que debemos reflexionar ya sobre la autonomía y libertad que queremos en el futuro, en lugar de dejar nuestro destino en manos de algoritmos. Estoy muy alienado en este sentido con algunas de las reflexiones de este experto, no en todas.

En mi opinión, es el ser humano a quién corresponde crear su futuro con la inteligencia artificial, en base a las decisiones y medidas que adopte hoy.

En este sentido me permito citar una de las frases más célebres atribuidas a Alan Kay, uno de los pioneros de la computación que conocemos hoy: "La mejor forma de predecir el futuro es inventarlo".

Por su parte, Elon Musk[108] ha manifestado públicamente en diversas ocasiones que la denominada "superinteligencia" es el mayor riesgo al que nos enfrentamos como civilización, y que debe ser regulada.

Para otros expertos como López de Mántaras, al que aludiré con mayor amplitud posteriormente, la complejidad del cerebro humano sería prácticamente imposible de replicar en el futuro, considerando que la singularidad está muy lejos y quizás sea imposible de alcanzar.

Todas estas cuestiones constituyen retos transformados en objetivo para parte de la comunidad científica, cuestionada en este sentido ante la necesaria reflexión previa sobre la conveniencia y necesidad de algunos de estos desafíos para su posterior identificación como objetivo a conseguir, ante la imparable evolución de la inteligencia artificial.

---

[107] Experto en IA en el Massachusetts Institute of Technology (MIT). Director del Novartis-MIT Center for Continuous Manufacturing y copresidente del Singapore-MIT Alliance Program on Chemical and Pharmaceutical Engineering. Recuperado de BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*". Op.cit. Pos. 3487.

[108] PALAZUELOS, F. (2017). "Elon Musk: `La inteligencia artificial amenaza la existencia de nuestra civilización'". Publicado en *El País* el 18.07.2017. Disponible en: https://elpais.com/tecnologia/2017/07/17/actualidad/1500289809_008679.html. Consultado el 25.02.2021.

Tanto el desarrollo y estado actual de la inteligencia artificial, principalmente categorizada como "débil", como el potencial desarrollo y evolución de la misma en los próximos años para alcanzar una inteligencia artificial más avanzada, "fuerte", "general" o una "súper inteligencia", comporta enormes riesgos y retos de variada naturaleza y entidad, requieren importantes reflexiones globales hoy y exigen articular mecanismos para su adecuada identificación, tratamiento y gestión, en especial para garantizar principalmente la seguridad física, lógica, moral y jurídica e incluso la existencia del propio ser humano en el futuro, por lo que la regulación jurídica constituye un instrumento esencial para determinar y delimitar qué inteligencia artificial queremos y permitiremos hoy y mañana, y bajo qué requisitos y condicionantes éticos, jurídicos y de seguridad.

Si la finalidad principal de un ordenamiento jurídico es establecer un conjunto de normas por las que se rija una sociedad, debemos ser conscientes de que ésta ha cambiado.

Vivimos en una sociedad cada vez más digital. En una sociedad en la que convergen y se fusionan su dimensión física y su dimensión virtual, y que convive en un espacio virtual -el ciberespacio- donde se relacionan e interaccionan miles de millones de máquinas, personas, empresas, asociaciones, fundaciones, entes con/sin personalidad jurídica y gobiernos.

Lo que ocurre en el mundo virtual puede afectar al mundo físico y lo que ocurre en el mundo físico puede afectar al mundo virtual, es decir, se trata de una afectación global y/o bidireccional en la medida que forman parte de una misma cosa.

En consecuencia, la evolución de la inteligencia artificial exige la creación de marcos éticos, legales y de seguridad que, de un lado, regulen de inicio la posibilidad de que existan de sistemas inteligentes autónomos (o con cierto grado de autonomía), sus requisitos y condiciones, y, de otro, que regulen las normas de convivencia y simbiosis entre seres humanos y sistemas inteligentes que deben regir esa nueva sociedad digital de las que todos formamos parte, cualquiera que sea el grado de autonomía o interacción de los sistemas inteligentes.

La persona y sus atributos no puede acabar siendo un número asociado a otros expresados en dólares, euros o bitcoins.

El Derecho tiene la finalidad de garantizar la seguridad humana, previendo y previniendo, al máximo posible, los posibles conflictos, si bien, no puede prever todo, como destacan autores como Martínez Rey y Pazos Sierra[109].

La confiabilidad y seguridad en la inteligencia artificial permitirá su desarrollo, despliegue y uso y, de manera paralela la inversión, investigación y la innovación en el misma, por lo que la misma debe sustentarse en la ética, la seguridad y las normas jurídicas.

Durante esta investigación abordaré los marcos jurídicos vigentes y las propuestas regulatorias europeas en relación con todos estos aspectos, en especial, en materia ética, de responsabilidad y de seguridad.

3. Percepción de la inteligencia artificial: Una perspectiva social y económica.

La percepción sobre la inteligencia artificial ha ido evolucionando conforme lo ha hecho su desarrollo y aplicación.

Las personas más alejadas de la tecnología identifican frecuentemente inteligencia artificial con robots y dispositivos físicos electrónicos que pueden hacer todo de tipo de cosas automatizadamente que podría llevar a cabo una persona.

El concepto de robot no conlleva necesariamente asociado el de inteligencia artificial.

La percepción social de ambas realidades y la perspectiva alrededor de las mismas será objeto también de análisis en este apartado.

[109] MARTÍNEZ REY, M.A. Y PAZOS SIERRA, J. (2019). "La inteligencia artificial y el derecho: Pasado, presente y futuro", en Monterrosso Casado, E. (Dir.), *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch 2019. P. 559

La aceptación de la inteligencia artificial y los robots inteligentes es mucho mayor en Oriente que en Occidente, especialmente ante sus peculiaridades culturales, históricas, sociales y económicas.

Japón fue uno de los primeros países que identificó el reto de enfrentarse una un envejecimiento insostenible de la población y la absoluta necesidad de la inteligencia artificial y la robótica para mantener la economía y cuidar a sus ancianos. Fruto de ello, fue la aprobación en 2014 de su *Estrategia de Revitalización de Japón* con este propósito. A nivel social y cultural, la sociedad japonesa tiene muy interiorizada y relativamente normalizada su relación con sistemas y robots inteligentes[110].

En Europa la percepción y aceptación no son uniformes, pero, en cualquier caso, radicalmente distinta a países del entorno citado.

Según el informe de la Comisión Europea denominado *Actitudes sobre el impacto de la digitalización y la automatización en la vida cotidiana*[111] de mayo de 2017, casi la mitad de la población europea encuestada había oído, leído o visto algo sobre la inteligencia artificial en los últimos 12 meses, el resto no. Estoy convencido que la realización de la misma encuesta en la actualidad arrojaría resultados radicalmente distintos.

De las conclusiones del precitado informe, destacar que la mayoría de la población encuestada estaba de acuerdo en que los robots y las tecnologías de inteligencia artificial requieren una cuidadosa gestión. El 61% de los europeos estaba a favor de la inteligencia artificial y de los robots, pero casi nueve de cada diez encuestados (88%) evidenciaba un temor y consideraba que necesitan un cuidado particular en su gestión.

[110] La sociedad japonesa tiene la consciencia de que algunas cosas tienen alma, como pone de relieve la antropóloga, Jennifer Robertson, autora del libro "*Robo Sapiens Japanicus: Robots, Gender, Familiy, and the Japanese Nation*". Un ejemplo del concepto de la Sociedad japonesa sobre la inteligencia artificial y la robótica es el templo de Kodaiji en Kioto, donde el sacerdote budista que recita las oraciones es un robot humanoide llamado Mindar.

[111] Informe *Attitudes towards the impact of digitisation and automation on daily life*. Encuesta solicitada por la Comisión Europea, Dirección General de Redes de Comunicación, Contenido y Tecnología y coordinado por la Dirección General de Comunicación. Mayo 2017. Accesible en: https://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/SPECIAL/surveyKy/2160. Consultado el 04.12.2020.

Por lo que se refiere a la percepción del impacto de los robots y la inteligencia artificial en el mercado laboral, la mayoría evidenciaba estar de acuerdo en que los robots y la inteligencia artificial son buenos para la sociedad, pero también estaban de acuerdo en que eliminarán puestos de trabajo.

De las percepciones generales de la población encuestada, destacar que más de dos tercios de los encuestados (68%) estaban de acuerdo en que los robots y la inteligencia artificial son algo bueno para la sociedad porque ayudan a la gente a hacer su trabajo o a llevar a cabo las tareas diarias en casa, si bien más de un cuarto evidenciaba estar en desacuerdo (26%).

Aunque más de seis de cada diez encuestados evidenciaba una opinión positiva de los robots y la inteligencia artificial, una proporción aún mayor (72%) estaba de acuerdo en que los robots y la inteligencia artificial quitarían trabajo a las personas.

La mayoría evidenciaba la creencia que el uso de robots e inteligencia artificial tendrá un impacto negativo en el número de empleos y sólo una minoría pensaba que su trabajo podría ser hecho por un robot o inteligencia artificial.

Si analizamos algunas de las últimas encuestas y estudios internacionales, destacar la última encuesta llevada a cabo a nivel mundial por el *Pew Research Center*[112] de New York, a finales de 2019 y principios de 2020.

De la encuesta[113], destacar que, en general, aproximadamente el 53% de los encuestados considera que el desarrollo de la inteligencia artificial y el desarrollo de sistemas informáticos para imitar el comportamiento humano ha sido algo positivo para la sociedad, mientras que el 33% lo considera algo negativo para la misma. Respecto del uso de robots para automatizar trabajos, un 48% lo considera algo positivo, frente al 42% que considere que ha tenido un impacto negativo.

[112] JOHNSON, C Y TYSON, A. "People globally offer mixed views of the impact of artificial intelligence, job automation on society". Pew Research Center. 15.12.2020. Recuperado de: https://www.pewresearch.org/fact-tank/2020/12/15/people-globally-offer-mixed-views-of-the-impact-of-artificial-intelligence-job-automation-on-society/. Consultado el 22.02.2021.
[113] *International Science Survey 2019-2020. Q11b*. Pew Research Center.2020

Las opiniones son generalmente positivas en el área asiática, como Singapur (72%), Corea del Sur (69%), India (67%), Taiwán (66%) y Japón (65%). En países como EE.UU. y Reino Unido los porcentajes en ambos sentidos son prácticamente similares. En España y Suecia un 60% de los encuestados se inclina hacia esa positividad, si bien Francia, las opiniones negativas superan en más de un 10% las positivas.

Del informe, destacar que las personas con mayor conocimiento y educación, así como las de menor edad suelen tener una visión más positiva de estas tecnologías.

En definitiva, las conclusiones son muy variadas, especialmente en función del conocimiento sobre estas tecnologías y la edad de las personas.

Sin embargo, la perspectiva de los trabajadores que comparten entornos de trabajo con el uso de inteligencia artificial evidencia algunos aspectos muy interesantes.

Un estudio llevado a cabo por Oracle[114] en 2019 concluyó que los trabajadores confían más en los robots que en sus gerentes. El estudio se realizó entre 8.370 empleados, gerentes y líderes de RR.HH. de 10 países. El 82% consideraba que los robots pueden hacer las cosas mejor que sus gerentes y en países como China, Indica, Singapur o Japón, entre el 75% y el 89% de los encuestados confían más en los robots.

Sin embargo, todas estas perspectivas confluyen con las de los empresarios, que tradicionalmente identifican la inteligencia artificial como una de las tecnologías de mayor impacto, conforme ha sido expuesto anteriormente, y base de su transformación digital.

Algunas de las perspectivas empresariales son absolutamente apocalípticas, otras simplemente objetivas y proactivas con el fin de reflexionar ya sobre cómo será el mañana y como deberemos prepararnos para llegar al mismo.

[114] *AI Is Winning Hearts & Minds in the Workplace: Global research highlights how AI is changing the relationship between people and technology at work*. Oracle, 2019. https://go.oracle.com/LP=86149?elqCampaignId=230263.

A modo de ejemplo, el investigador y empresario Kai-Fu Lee[115] augura que, dentro de quince años, la inteligencia artificial será técnicamente capaz de reemplazar entre el 40% y el 50% de los puestos de trabajo en EE.UU., e incluso llega más allá, vislumbrando un tremendo desorden social y colapso político derivados del desempleo generalizado y de la enorme desigualdad. Según el mismo la inteligencia artificial exacerbará la desigualdad económica mundial, concentrado la riqueza en unos pocos.

La realidad es que la gran mayoría de las empresas que integran la lista Fortune 500 utilizan ya sistemas inteligentes para entrevistar a futuros candidatos a incorporar en su plantilla. Otros sistemas permiten analizar los perfiles en redes sociales de los candidatos para identificar los rasgos de su personalidad. Otros sistemas integrados en robots humanoides ya atienden y asisten personalmente en hoteles, sustituyendo al personal humano que habitualmente venía desarrollando esas funciones.

La inteligencia artificial y la robótica afectará al empleo inevitablemente y la forma de trabajar, pero no todo será destrucción, sino reconversión y también creación.

Hoy, países como Alemania, tradicionalmente con niveles próximos al pleno empleo, son los que tienen más robots per cápita. Sin embargo, países como España o Grecia, con mayores tasas de paro son los que menos robots tienen. Ello puede constituir una evidencia de que una automatización planificada, gestionada y organizada aporta eficiencia y puede ser compatible con el empleo y la cualificación.

El mundo va a seguir transformándose gracias a la tecnología, como ha venido haciéndolo hasta ahora, y no va a parar.

Las claves para la adecuada preparación y gestión del impacto consecuente a estos cambios se hallan en las buenas estrategias y en el liderazgo político.

[115] LEE, KAI-FU (2018). *Superpotencias de la inteligencia artificial*. Op.cit. Pos. 31.

La diferencia es que la velocidad a la que se transforma es cada vez mayor, con lo que disponemos de menos tiempo para la planificación y nos exigirá disponer de capacidades de adaptación más rápidas a los mismos, si queremos mitigar su impacto.

Se están automatizado ya tareas que requieren esfuerzo físico, pero también las que requieren esfuerzo mental e incluso creatividad, desaparecerán puestos de trabajo para siempre, aparecerán otros nuevos y se reducirán las jornadas laborales de los trabajadores.

Esto no sólo sucederá en mayor medida, sino que ya está sucediendo. Estamos rodeados con asistentes digitales en nuestros hogares, nuestros vehículos y nuestros dispositivos. Las industrias llevan años integrando la automatización robótica de procesos -RPA por sus siglas en inglés-, que están automatizando la actividad humana, y cada vez más estos sistemas están usando la inteligencia artificial para aumentar sus capacidades.

La inteligencia artificial, el aprendizaje profundo o los sistemas expertos van a modificar el modo de trabajar, pero no porque puedan hablar o pensar, sino porque son más rápidos, eficientes y económicos que el ser humano.

La discusión sobre el impacto de las revoluciones tecnológicas y la automatización no es nueva, y la mayoría de las cuestiones no son tanto relativas a si ocurrirán o no, sino cuándo.

Según Jack Ma, fundador y presidente del gigante asiático Alibaba con sede en Hangzhou (China), y defensor abierto de la cultura del trabajo extremo, afirmó sorprendentemente en la *Conferencia Mundial de inteligencia artificial (IA) de 2019*[116] que en el futuro sólo trabajaremos 12 horas a la semana.

De hecho, esta es una de las posibles soluciones que se barajan en el futuro para la previsible pérdida de empleo y redistribución, en la medida que incluso reducir la jornada laboral a menos jornadas laborales y con menos horas diarias ha sido ya testeada en

[116] "Jack Ma, presidente de Alibaba, augura que en el futuro se trabajará solo 12 horas a la semana". Publicado en *El País* el 29.08.2019. Disponible en: https://elpais.com/economia/2019/08/29/actualidad/1567076867_110475.html. Consultado el 19.02.2021.

proyectos piloto por empresas como Microsoft[117], mostrando que incluso pueden incrementar de manera exponencial la productividad, y no es la única, dado que desde hace años se han evidenciado estos resultados en distintos estudios[118].

Según algunos autores[119], muchos de los trabajos que actualmente desempeña la clase media se automatizarán. Sin embargo, algunos expertos y economistas consideran que la pérdida de puestos de trabajo por la tecnología puede carecer de fundamento y sostienen que la inteligencia artificial aumentará enormemente la productividad y promoverá un crecimiento saludable del empleo y del bienestar de las personas[120]. Los argumentos de estos últimos han sido acogidos también por algunos de los gigantes tecnológicos.

El debate sobre el impacto de la automatización en el empleo tampoco es nuevo. Al respecto, Frey y Osborne, investigadores de la Universidad de Oxford, publicaron ya en 2013 un artículo científico denominado *The future of employment: How susceptible are jobs to automation*[121], en el que se auguraba sin demasiada precisión temporal, que en una o dos décadas se podrá automatizar el 47% de los empleos en EE.UU.

Un estudio posterior de 2016[122] llevado a cabo por investigadores de la *Organización para la Cooperación y el Desarrollo Económicos* (OCDE), concluyó que sólo el 9% de los empleos en EE.UU. corría un riesgo elevado de automatización, en la medida que no serán los empleos en su totalidad los que serán automatizados, sino tareas específicas dentro de los mismos. El porcentaje de alto riesgo oscilaba entre el 6% en Corea y el 12%

---

[117] PAUL, K. (2019). "Microsoft Japan tested a four-day work week and productivity jumped by 40%". Publicado en *The Guardian* el 04.11.2019. Disponible en: https://www.theguardian.com/technology/2019/nov/04/microsoft-japan-four-day-work-week-productivity. Consultado el 21.02.2021.

[118] GLAVESKI, S. (2018). "El caso de la jornada laboral de 6 horas". Publicado en *Harvard Business Review* el 11.12.2018. Disponible en: https://hbr.org/2018/12/the-case-for-the-6-hour-workday?language=es. Consultado el 21.02.2021.

[119] GRATTON, L. Y SCOTT, A. J. (2016). *The 100-Year Life: Living and Working in an Age of Longevity*. Bloomsbury Publishing. London. 2016.

[120] LEE, KAI-FU (2018). *Superpotencias de la inteligencia artificial*. Op.cit. Pos. 174.

[121] FREY, C. B. Y OSBORNE, M.A. (2013). "The future of employment: How susceptible are jobs to automation". *Oxford Martin Programme on Technology and Employment*. 17 de septiembre de 2013. P. 47. Disponible en https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf?link=mktw. Consultado el 25.02.2021.

[122] ARNTZ, M.; GREGORY, T. Y ZIERAHN, U. (2016). "The risk of automation for jobs in OECD countries: A comparative analysis". OECD Social. Publicado en *Employment and Migration Working Papers*, Nº 189. 14 de mayo de 2016. Pp. 14 y 25. Disponible en: https://www.oecd-ilibrary.org/social-issues-migration-health/the-risk-of-automation-for-jobs-in-oecd-countries_5jlz9h56dvq7-en. Consultado el 25.02.2021.

en Austria. La OCDE llevó a cabo su análisis, pero no desde un enfoque de puestos de trabajo sino de tareas automatizables y no automatizables.

De entre los estudios "apocalípticos" más recientes, me permito también significar el informe *Artificial Intelligence, Automation, and the Economy*[123] publicado el 20 de diciembre de 2016, promovido por Barack Obama y realizado por su Oficina Ejecutiva. Este informe prosigue su informe previo elaborado en octubre del mismo año bajo el título *Preparing for the Future of Artificial Intelligence*, con la finalidad de continuar la investigación de los efectos de la automatización impulsada por la inteligencia artificial en el mercado de trabajo y la economía de EE.UU.

Dicho informe destacaba la baja probabilidad de que los sistemas inteligentes presenten una inteligencia de aplicación general comparable o superior a la de los humanos en los próximos 20 años, sin perjuicio de que sigan alcanzando y superando el rendimiento humano en determinados aspectos.

El informe destaca los grandes cambios económicos y sociales que se avecinan -especialmente el desempleo y la desigualdad-, y la necesidad de medidas políticas agresivas para ayudar a los estadounidenses que se vean perjudicados por estos cambios, y para garantizar que los enormes beneficios de la inteligencia artificial y la automatización sean desarrollados y disfrutados para todos. Conforme recoge, la inteligencia artificial ya ha empezado a transformar el lugar de trabajo estadounidense, a cambiar los tipos de puestos de trabajo disponibles y a remodelar las habilidades que los trabajadores necesitan para prosperar.

El informe propone distintas estrategias contra los efectos de la inteligencia artificial y la automatización en el desempleo y la desigualdad, especialmente para educar y capacitar a los nuevos trabajadores para los puestos del futuro en el proceso de transición, ayudar a los trabajadores que pierdan su empleo, mantenerlos vinculados a la fuerza laboral y combatir la desigualdad.

[123] *Artificial Intelligence, Automation, and the Economy*. 20.12.2016. Executive Office of the President. U.S. Government. Disponible en: https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.PDF. Consultado el 28.02.2021

Asimismo, el informe destaca especialmente el valor que supone la inteligencia artificial para la ciberdefensa y la mejora de los sistemas de detección de transacciones y mensajes fraudulentos.

Según el estudio independiente llevado a cabo posteriormente por la consultora *PriceWaterhouseCoopers* en 2017[124], el 38% de los empleos en EE.UU. corría un alto riesgo de ser automatizados a principios de la década de 2030, lo que difiere notablemente del estudio y previsión realizado por la OCDE.

Un informe posterior de *McKinsey*[125] diferencia de nuevo entre puestos de trabajo y tareas que se automatizan. Según la misma, son pocos los trabajos que son completamente automatizables, en particular, menos del 5%, pero que el 60% de los puestos de trabajo tienen al menos un 30% de tareas que sí lo son. Según esta consultora los trabajos cuya automatización será más probable serán los que implican actividades físicas repetibles -un 81%-, procesamiento de datos -69%- y colección de datos -64%-.

Otro informe ulterior de la *American Economic Association*[126], parte de las mismas conclusiones que *McKinsey* respecto de que no se automatizarán puestos de trabajo sino tareas, de modo que no se producirá una situación de sustitución masiva de puestos por sistemas inteligentes. Asimismo, abordó las condiciones para poder automatizar una tarea.

El investigador y empresario Kai-fu Lee[127], anteriormente citado, estima que en diez o veinte años seremos técnicamente capaces de automatizar entre el 40% y el 50% de los puestos de trabajo de EE.UU. Para los empleados que no sean sustituidos por completo, la creciente automatización de su volumen de trabajo continuará reduciendo su valor

[124] BERRIMAN, R. Y HAWKSWORTH, J. (2017). *Will robots steal our jobs? The Potential Impact of Automation on the UK and other major economies*. PWC. Marzo 2017. Pp. 1, 7, 39 y 45. Disponible en: https://www.pwc.co.uk/economic-services/ukeo/pwcukeo-section-4-automation-march-2017-v2.pdf. Consultado el 25.02.2021.
[125] MANYIKA, J.; CHUI, M. Y OTROS (2017). *A future that works: Automation, employment, and productivity*. McKinsey Global Institute. Chicago. 2017.
[126] BRYNJOLFSSON, E.; MITCHELL, T. Y ROCK, D. (2018). "What can machines learn, and what does it mean for occupations and the economy?" *AEA Papers and Proceedings*. Vol. 108. May 2018. Pp. 43-47.
[127] LEE, KAI-FU (2018). *Superpotencias de la inteligencia artificial.* Op.cit. Pos. 174.

añadido para las empresas, reduciendo su poder de negociación sobre los salarios y es posible que conduciendo a despidos a largo plazo.

Un nuevo informe posterior de febrero de 2018 de la consultora *Bain and Company* concluyó que para 2030, se necesitarán entre un 20% y un 25% menos de trabajadores.

En una reciente entrevista[128], Amy Webb, nombrada por la revista Forbes como una de las cinco mujeres que están cambiado el mundo mediante la tecnología, asesora de alguna de las empresas más influyentes del mundo y autora del libro *Los nueve gigantes: Cómo las grandes tecnológicas amenaza el futuro de la humanidad*, afirmaba como nos podemos estar acercando a un escenario catastrófico con la inteligencia artificial.

En definitiva, la percepción sobre la inteligencia artificial ha ido progresivamente cambiando y han irrumpido connotaciones muy negativas hasta el punto de considerarla una amenaza, por lo que los gobiernos tienen mucho por hacer en relación con todo ello al objeto de ir cambiando esta percepción y convertir una amenaza, en algunos aspectos inevitable, en una oportunidad para su verdadero despliegue, aplicación y uso.

Las tres revoluciones industriales que han precedido a la revolución constante en la que nos hallamos cambiaron el mundo por completo, creando nuevos modelos y extinguiendo otros.

La primera vino de la mano de la mecanización de la producción gracias a la máquina de vapor, la segunda de la división del trabajo, la producción en masa, la electricidad, los hidrocarburos, nuevos materiales y nuevos sistemas de transporte y, la tercera, de la automatización vinculada a la informática y la electrónica.

No podemos frenar el avance tecnológico y el progreso. En este contexto me permito citar una de las frases más conocidas atribuidas a Henry Ford: “El verdadero progreso es el que pone la tecnología al alcance de todos”. Este considero que debe ser el objetivo final, una tecnología de todos y para todos.

[128] VILLAMOR, N. (2021). Amy Webb: “Nos acercamos a un escenario catastrófico con la inteligencia artificial”. Publicado en *The Objective* el 09.07.2021.

Todas estas revoluciones conllevaron un proceso de transición complejo y duro pero el resultado final fue una mejora de la sociedad y de nuestro mundo.

La revolución actual, más compleja, sustentada en la transformación digital propiciada por los sistemas inteligentes interconectados y en la velocidad con las que intercambiamos el conocimiento, así como las siguientes que la sucedan, deberían tener esa finalidad última.

No obstante, en mi opinión, el mayor problema es la velocidad de los cambios constantes de estas nuevas revoluciones sustentadas en la tecnología, en la información y en el conocimiento que, de un lado, pueden acelerar los procesos, pero sin las adecuadas previsiones, planes y estrategias, su impacto puede ser mucho más profundo y, de manera consecuente, conllevar procesos de transición más largos y absolutamente asíncronos hacia esa supuesta mejora global pretendida.

La tecnología inevitablemente elimina y eliminará puestos de trabajo y no es algo nuevo como he referido. Igualmente transformará los modelos económicos y la propia sociedad.

Desde la antigüedad, el mero descubrimiento de la rueda eliminó mano de obra para mover o transportar cargas pesadas, desde bloques de piedra a mercaderías, optimizando recursos.

La automatización agraria, la revolución industrial y la cadena de producción supuso la eliminación de mano de obra en las fábricas.

Posteriormente la informática e Internet supuso una nueva oleada de cambios.

Hoy la tecnología sigue y seguirá eliminando mano de obra humana y, desde luego, la inteligencia artificial conllevará inevitablemente y de manera inherente esta eliminación, pero exigirá nuevos perfiles profesionales y creará también nuevos puestos.

La Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica[129], efectuó una

[129] https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_ES.html

lectura positiva sobre este asunto al considerar que, si bien es posible que el uso generalizado de robots no acarree automáticamente la sustitución de puestos de trabajo, sí que es probable que los empleos menos cualificados en sectores intensivos en mano de obra sean más vulnerables a la automatización.

En este sentido, añadió que "la investigación ha demostrado que el crecimiento del empleo es considerablemente más rápido en los puestos de trabajo que hacen un mayor uso de la informática". Asimismo, consideró que la automatización de los puestos de trabajo puede liberar a las personas de tareas manuales monótonas y permitirles que se dediquen a otras más creativas y significativas, si bien, esta automatización obligará a los gobiernos a invertir en educación y a acometer otras reformas con el fin de mejorar la redistribución en los tipos de capacidades que necesitarán los trabajadores en el futuro.

El propio Parlamento Europeo destacaba en dicha Resolución algunas de las previsiones de la Comisión, según las cuales, hasta finales del año 2020 Europa se enfrentaría a una escasez cercana a un millón de profesionales en el sector de las TIC, a la vez que el 90 % de los puestos de trabajo requerirá al menos unas capacidades digitales básicas. Y fueron ciertas sus previsiones, la demanda de perfiles de profesionales en estos campos crece de manera incesante y las empresas se encuentran ante la escasez de recursos cualificados y la dificultad de cubrirlos con empleados provenientes de otros países del mundo, dada las restricciones para la libre circulación de trabajadores a nivel internacional.

No obstante, el Parlamento Europeo también advertía en las consideraciones previas de la precitada Resolución que el progreso de la robótica podría traducirse en una elevada concentración de la riqueza y el poder en manos de una minoría, instando a la Comisión a iniciar ya entonces el análisis y supervisión de la evolución a medio y largo plazo del empleo, con especial atención a la creación, la deslocalización y la pérdida de puestos de trabajo en los diferentes campos, con el fin de determinar en qué ámbito se estará creando empleo y en cuáles se estará perdiendo como consecuencia de la mayor utilización de los robots.

Del mismo modo, el Parlamento Europeo significó en dicha Resolución que la inteligencia artificial y la robótica también provocaría cambios sociales y afectaría a la viabilidad de los actuales sistemas de seguridad sociales en los Estados miembros, advirtiendo a la Comisión sobre la necesidad de analizar los posibles escenarios y sus consecuencias.

De hecho, uno de los principales retos y riesgos que plantea la inteligencia artificial no es sólo la pérdida de empleo y sus efectos consecuentes, sino la creciente desigualdad tanto interna como entre países, de la que se viene alertando por los expertos, como he destacado anteriormente.

Las previsiones hablan de una concentración sin precedentes de la riqueza en manos de unos pocos, principalmente empresas chinas y estadounidenses. Pero, además, en mi opinión, es previsible también un fuerte impacto psicológico del desempleo ante la posible perspectiva para las personas de no sólo de quedarse temporalmente sin empleo sino de poder ser excluidas permanentemente del mundo laboral.

En conclusión, la inteligencia artificial contribuirá a aumentar los niveles de eficiencia, ahorro y calidad de los servicios, pero también transformará el mercado laboral, los perfiles, las modalidades, las formas y las condiciones de trabajo. De este modo, la inteligencia artificial impactará en la naturaleza y características de muchos puestos de trabajo, en la oferta y demanda de empleo y en los sistemas de previsión social.

Además, su impacto involucrará tanto a puestos más físicos como intelectuales, especialmente aquellos que requieran unas habilidades cognitivas sustituibles o mejorables por los sistemas, por lo que se verán afectados puestos cualificados, como no cualificados.

Se perderán empleos sin perspectivas de reemplazo y se crearán nuevos perfiles acordes a las nuevas demandas, pero también se generarán nuevas formas de trabajar y nuevos espacios de relación entre hombre-máquina, enriqueciéndose mutuamente y permitiendo alcanzar logros inalcanzables para el ser humano.

A modo de ejemplo, pensemos en la aplicación del sistema inteligencia "Watson" de IBM aplicado a la ciberseguridad, el cual constituye un complemento para los profesionales de la seguridad para la detección precoz de ataques para su bloqueo y la activación de los mecanismos de respuesta, o el uso de sistemas inteligentes para el diseño de niveles de videojuegos o de diseños gráficos como herramienta para los creadores del sector en el ámbito de diseño, el arte y la programación.

Las claves para el futuro, como comentaba en anteriores apartados, considero que deben situarse en el necesario matrimonio indisoluble entre hombre-máquina, en su necesaria simbiosis, integración y colaboración para satisfacer necesidades, solucionar problemas, mejorar la vida y nuestro planeta, y ello de manera eficaz y eficiente, complementándose, no sustituyéndose, y explotando lo mejor de ambas.

Según el *Foro Económico Mundial* o "Foro de Davos", la cuarta revolución o revolución robótica podría tener un impacto en el empleo de más de 5,1 millones de puestos de trabajo destruidos en las 15 economías más grandes del mundo durante el período 2015-2020 (EE.UU., China, Alemania, Australia, Brasil, Francia, India, Italia, Japón, México, Sudáfrica, Turquía, Reino Unido y los grupos conformados por la Asociación de Naciones del Sudeste Asiático -ASEAN- y el Consejo de Cooperación para los Estados Árabes del Golfo -GCC-).

Según el mismo, los sistemas robóticos podrían ocupar los puestos de trabajo en los que en la actualidad (en la fecha de su informe) trabaja el 47% de la población activa, lo cual significaría la destrucción de 1.600 millones de empleos, y afectaría a países desarrollados como en vías de desarrollo.

La consultora *McKinsey Global* en su informe de 2019 concluye que el 20% de los trabajadores del mundo serán reemplazados por robots en solo 12 años y, en un estudio de la *Universidad de Oxford* sostiene que, en un intervalo de 10 a 20 años, el 47% de los trabajos serán realizados por máquinas.

Algunos autores como Gómez Salado[130] hablan del reto que definirá el siglo XXI como el "siglo de la destrucción masiva de empleo".

La automatización y sustitución además se está acelerando sustancialmente.

Según el informe *The Future of Jobs 2018* del Foro Económico Mundial, en 2025 habrá más máquinas inteligentes trabajando que personas. Según el mismo, los robots destruirán 75 millones de empleos en el mundo en los próximos cuatro años, pero que surgirán nuevas funciones que permitirán crear 133 millones, lo que supone un neto de creación de empleo de 58 millones.

Del mismo modo, según la UE, gracias a la inteligencia artificial y la robótica podrían crearse para 2025 más de 60 millones de nuevos puestos de trabajo en el mundo.

De manera consecuente a todo lo expuesto, los sistemas de previsión y protección social se verán igualmente afectados, en la medida que previsiblemente se mermarán los recursos para financiar los mismos provenientes de las cuotas de cotización, ante una población cada vez más envejecida y longeva, y se incrementarán las prestaciones por desempleo. De este modo, se verá comprometida la viabilidad de estos sistemas y su sostenimiento, especialmente para poder abonar las prestaciones futuras derivadas de jubilaciones.

Frente a los posibles escenarios futuros, se están planteando y debatiendo posibles opciones como la reconversión o el reciclaje profesional, la reducción de la jornada de trabajo, los impuestos y contribuciones especiales vinculadas a la sustitución de personas o uso de sistemas inteligentes fuertes o la redistribución de los ingresos, especialmente a través de renta básica garantizada.

Por lo que se refiere a impuestos y contribuciones especiales, se ha propuesto la cotización a la Seguridad Social por parte de las empresas por los sistemas de inteligencia artificial que utilicen, especialmente ante el ahorro en cotizaciones y costes de empleo para las

[130] GÓMEZ SALADO, M.A. (2018). "Robótica, empleo y seguridad social. La cotización de los robots para salvar el actual estado de bienestar", en *Revista Internacional y Comparada de Relaciones Laborales y Derecho del Empleo*. ISSN-e 2282-2313. Vol. 6, Nº. 3, 2018. Pp. 139-170.

mismas, lo que justificaría contribuir al gasto social y fortalecimiento de la protección social y su responsabilidad social corporativa.

Algunos empresarios como Bill Gates -fundador de Microsoft-, proponen la introducción de impuestos a programas de automatización para compensar la pérdida de impuestos que generará la pérdida de empleos, o tasas al capital concentrado -gigantes tecnológicos, fabricantes e inventores de sistemas de inteligencia artificial-[131]. El Parlamento Europeo ha rechazado esta posibilidad inicialmente, aunque cada vez hay más partidarios de ello.

Por lo que se refiere a una renta básica garantizada como posible solución, en la actualidad se está planteando el establecimiento de dicha renta o subsidio universal -UBI por sus siglas en inglés- que garantice a todas las personas de manera automática e incondicionada un ingreso periódico de subsistencia, evitando las desigualdades sociales y asegurando a todos los ciudadanos la satisfacción de las necesidades esenciales en aras de la dignidad de la persona[132].

Son muchos los expertos que recomiendan el establecimiento de una renta básica universal y ya se están llevando a cabo algunos proyectos piloto en algunos países[133], por ejemplo, Canadá (Ontario), Escocia o Finlandia.

Otra modalidad alternativa sería el denominado ingreso mínimo garantizado -IMG- de la que únicamente resultarían beneficiados las personas en una situación económica más débil.

En relación con todo ello, se ha planteado en diversas ocasiones reflexionar sobre la posible creación de una personalidad jurídica a los robots o sistemas dotados de inteligencia artificial para la asociación de impuestos o contribuciones, si bien, no concibo esta opción como solución, como expondré más adelante al abordar los aspectos de responsabilidad, especialmente ante la imposibilidad de hacerlo, conforme al marco

[131] FORD, M. (2015). *Rise of the Robots*". Basic Books, Nueva York, 2015. Op. cit.
[132] MERCADER, J. R. (2017). *El futuro del trabajo en la era de la digitalización y la robótica*. Editorial Tirant lo Blanch. Valencia 2017. Pp. 39 y 40.
[133] LANT, K. (2019). "Universal Basic Income: UBI Pilot Programs Around the World". *Futurism*. 2019. Disponible en: https://futurism.com/images/universal-basic-income-ubi-pilot-programs-around-the-world. Consultado el 20.02.2021.

jurídico vigente, particularmente ante la ausencia de una autonomía, consciencia y libertad plena.

En mi opinión, de un lado, los sistemas inteligentes serán más eficientes, eficaces, incansables y rentables por lo que posiblemente serán quienes nos paguen nuestras pensiones en el futuro. De otro, las personas no dejarán de trabajar, sino que trabajarán menos horas y en puestos más cualificados o especializados.

Algunos expertos como Jacques Bughin[134], significa que cuando "el mayor ruido que se ha hecho alrededor de los riesgos de la inteligencia artificial ha sido el relacionado con el empleo", si bien, el resto de riesgos como el sesgo, uso no ético, consumo de energía o la desigualdad, deben ser igualmente considerados para la definición de una estrategia correcta para el desarrollo de la inteligencia artificial, en la medida que estos riesgos, en opinión de este experto, "también tendrán un peso indirecto en la perspectivas de empleo", el cual propone un escenario de modelo de cooperación y de complementariedad entre el trabajo humano y los sistemas inteligentes.

La preocupación por el impacto de la inteligencia artificial en la economía, trabajo y la sociedad forma parte de la agenda digital de la mayoría de países desarrollados y durante los últimos cuatro años se ha abierto el debate a expertos en el ámbito científico, jurídico, ético, económico o social, pero también a empresas, consumidores u ONGs.

Entre otras y junto a las europeas, destacar las consultas públicas llevadas a cabo por la *Oficina de Políticas sobre Ciencia y Tecnología* (OSTP) en EE.UU. así como la iniciativa *National Robotics Initiative 2.0*[135] financiada por la *National Science Foundation*, las consultas públicas llevadas a cabo por el Parlamento británico y su informe *Robotics and artificial intelligence*[136] o las iniciativas del *Comité de la Agenda Digital* del Parlamento alemán sobre los efectos de la robótica en la economía, el trabajo y la sociedad.

El mundo está cambiando y cada vez más rápido, las claves es haberlo previsto, implementar planes y estrategias en el ámbito político, social, económico, empresarial y

[134] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Op. cit. Pos. 1164.
[135] Recuperado de https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503641
[136] Recuperado de https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/14502.htm

personal, y prepararnos para ello, para vivir en un entorno cada vez más cambiante y a mayor velocidad donde tendremos que tener las habilidades necesarias para adaptarnos a los cambios constantes lo más temprana y rápidamente posible y ser más "camaleónicos" que nunca. Y es el ser humano quién deberá elegir su futuro.

En definitiva, si la inteligencia es la capacidad de adaptarse a los cambios como defendía Stephen Hawking, es el momento en el que el hombre va a tener que ser más inteligente que nunca, sin perjuicio de que esta adaptación pueda llevarla a cabo con apoyo en la inteligencia artificial.

Por todo ello y respecto de las distintas percepciones sobre la inteligencia artificial desde un punto de vista social y económico, obviamente la inteligencia artificial plantea desafíos y riesgos, pero no viene a sustituir al ser humano sino a mejorarlo, a mejorar nuestro mundo, nos complementaremos, nos necesitaremos y nos ayudaremos en ambas direcciones, considerando a la tecnología un medio y no un fin.

4. Principales retos y riesgos asociados a la inteligencia artificial

4.1. Consideraciones generales

El *Libro blanco sobre la inteligencia artificial*[137] de la Comisión Europea identifica y agrupa los principales riesgos de la misma en: a) Riesgos para los derechos fundamentales; b) Riesgos para la seguridad y; c) Riesgos para el funcionamiento eficaz del régimen de responsabilidad civil. Es decir, sitúa lo riesgos en el ámbito ético y de derechos, de la seguridad y de la responsabilidad consecuente, lo que constituye precisamente el enfoque principal de esta investigación.

a) Riesgos para los derechos fundamentales.

[137] "*Libro Blanco sobre la inteligencia artificial - Un enfoque europeo orientado a la excelencia y la Confianza*". Comisión Europea. Bruselas, 19.2.2020. COM(2020) 65 final. Pp. 13-15.

Los riesgos para los derechos fundamentales pueden tener su origen en el diseño de los sistemas, el uso de datos y en su aplicación. En este sentido, me permito citar el trabajo de investigación llevado a cabo por del Consejo de Europa[138].

Estos riesgos pueden afectar a valores y derechos fundamentales como la dignidad humana, la protección de datos personales y la privacidad, la libertad de expresión, la libertad de reunión, la no discriminación por razón de sexo, raza u origen étnico, religió o creo, discapacidad edad u orientación sexual, el derecho a una tutela judicial efectiva y a un juicio justo o la protección de los consumidores.

Su origen puede estar en su diseño defectuoso, en el uso, disponibilidad o calidad de los datos, en el aprendizaje de los sistemas o en su aplicación.

En relación con los mismos debe significarse la falta de explicabilidad y opacidad asociada a la inteligencia artificial y las capacidades y utilidades potenciales de seguimiento y análisis de conductas y costumbres individuales, vigilancia de empleados, vigilancia masiva o rastreo y desanonimización de datos, entre otras.

Lo cierto es que la inteligencia artificial potencia estos riesgos, pero no lo es menos que están también en nuestro entorno físico y/o virtual ajeno a estos sistemas inteligentes y en nuestras relaciones con otras personas. Aspectos como el sesgo, prejuicio o discriminación son riesgos inherentes a toda actividad social o económica y la toma de decisiones, la cual no es ajena al error o la subjetividad en su adopción, si bien, la inteligencia artificial puede tener efectos muchos más amplios a nivel cualitativo y cuantitativo y, en función del sistema, con posibles menos controles que los que puedan utilizarse en las relaciones humanas, personales, profesionales, económicas, administrativas o empresariales no automatizadas o utilizando como medio sistemas inteligentes. No obstante, como

[138] Consejo de Europa. https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5

abordo en esta investigación, la inteligencia artificial puede ser empleada precisamente con el objetivo de su prevención, detección, gestión y mitigación.

No obstante, lo cierto es que las características asociadas a los sistemas inteligentes, en función de su tipología y capacidades, aumentan estos riesgos, de nuevo, tanto cualitativa como cuantitativamente, especialmente la opacidad y falta de explicabilidad, y ello ante aspectos como su complejidad, comportamiento parcialmente autónomo o imprevisibilidad, lo que puede complicar la comprobación del cumplimiento de los marcos reguladores vigentes sobre protección de derechos fundamentales e impedir su cumplimiento efectivo.

No obstante, marcos más evolucionados como el precitado Reglamento General de Protección de Datos (RGPD) en relación con el derecho a la privacidad exigen no sólo protegerlo de manera efectiva y cumplir el marco regulador, sino estar en condiciones de demostrarlo, es decir, un cumplimiento y seguridad proactiva, gestionada y eficazmente implementada, en el diseño y por defecto, esto es cumplimiento, privacidad y seguridad *by design & by default*.

Además, todos estos riesgos generan que las personas afectadas pueden tener dificultades para el acceso a la justicia en caso de verse afectadas por los mismos, incluyendo la reclamación de los daños y perjuicios que haya podido sufrir como consecuencia de los incumplimientos y vulneración de sus derechos y libertades.

b) Riesgos para la seguridad (incluyendo la física y la lógica).

La integración de la inteligencia artificial en determinados productos y servicios puede comportar importantes riesgos para las personas físicas o jurídicas que pueden desembocar en daños personales y materiales, pero también para gobiernos, estados y entidades supranacionales.

Igualmente, los sistemas de inteligencia artificial pueden aumentar cualitativa y cuantitativamente los riesgos del mundo físico, aumentando o agravando los riesgos existentes, especialmente ante su uso malicioso y delictivo contra la seguridad de personas, instalaciones y cosas.

De igual modo que los anteriores, su origen puede estar en su diseño defectuoso, en el uso, disponibilidad o calidad de los datos, en al aprendizaje de los sistemas, su entrenamiento o en su uso/aplicación.

A todo ello, debemos adicionar la ausencia actual de un marco regulador claro y específico en materia de seguridad para abordar estos riesgos, como expresamente recoge el precitado *Libro blanco sobre la inteligencia artificial*, lo que genera desconfianza e inseguridad jurídica a las empresas que comercializan productos que integran inteligencia artificial.

De manera consecuente, las autoridades competentes pueden encontrarse ante situaciones en las que carecen de facultades para intervenir o carezcan de la capacidad para ello.

No obstante, respecto de los riesgos relacionados con la ciberseguridad, la Comisión Europea se ha remitido para su evaluación por la Agencia de Seguridad de las Redes y de la Información de la UE (ENISA), conforme analizaré con posterioridad, habiendo ya emitido los primeros informes.

Esta cuestión enlaza directamente con el tercer grupo de riesgos.

c) Riesgos para la eficacia de la responsabilidad civil.

La falta de marcos de referencia y requisitos claros y las propias características de los sistemas inteligentes precitados, esto es, su opacidad y falta de explicabilidad, complejidad, autoaprendizaje, entrenamiento, comportamiento autónomo y/o imprevisibilidad pueden complicar la trazabilidad de las decisiones tomadas mediante sistemas inteligentes, la identificación de los sujetos intervinientes con mayor control sobre el riesgo dentro en la cadena y ciclo de vida de estos sistemas y la imputabilidad de la responsabilidad derivada de los daños causados por los mismos, dificultando así el derecho a un resarcimiento efectivo en los marcos de responsabilidad civil vigentes en el seno de la UE y de los distintos países que la integran. En cualquier caso, esta cuestión será abordada con detalle en el capítulo V de esta investigación.

La interrelación entre todas estas categorías de riesgos, es decir, entre ética, seguridad y responsabilidad es evidente, en la medida que de materializarse estos riesgos y causar un impacto, la ausencia o incumplimiento de los requisitos éticos, jurídicos y de seguridad, previa y claramente definidos, podría determinar, sin perjuicio del régimen sancionador que resulte de aplicación a nivel administrativo o incluso penal, el régimen de responsabilidad a aplicar, sus consecuencias, así como dificultades asociadas para su exigencia en función del contexto y marco aplicable.

La ausencia de estos marcos éticos, jurídicos y de seguridad debilita la posición de la persona afectada por los daños y perjuicios causados por o mediante sistemas inteligentes, genera desconfianza e inseguridad jurídica a la sociedad en general y al sector empresarial en particular y, sin duda, puede afectar negativamente a la inversión e innovación en inteligencia artificial en Europa por parte del sector privado y a su posterior despliegue y aplicación.

La falta de transparencia, trazabilidad y rendición de cuentas (*accountability*) de la inteligencia artificial y la dificultad de seguimiento retrospectivo de las decisiones problemáticas es aplicable a todas las categorías de riesgos mencionados e impactan directamente en los regímenes de responsabilidad vigentes, por lo que su exigencia es una necesidad.

Estos aspectos mencionados constituyen principios y normas éticas esenciales, objeto de un creciente tratamiento internacional para su consenso, pero que no son exigibles jurídicamente y, consecuentemente, por el momento, integran marcos éticos más o menos consensuados a nivel internacional o europeo, no vinculantes.

Si partimos de la eficacia de los marcos vigentes, sin perjuicio de su análisis profundo en el capítulo V, veamos un mero ejemplo.

Conforme al marco vigente de responsabilidad por daños por productos defectuosos tanto en la UE[139] como en España[140], puede resultar muy complicado demostrar la existencia de un defecto en un producto, el daño causado y el nexo causal entre ambos, sin perjuicio de las dudas previas sobre la aplicación de estos marcos en el caso de algunos defectos, entre otros y como recoge expresamente el *Libro blanco sobre inteligencia artificial*, ante un fallo de ciberseguridad del producto.

En este sentido, la exigencia de ciberseguridad de estos sistemas sí se contempla expresamente en la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, si bien, exclusivamente respecto de los sistemas inteligentes considerados de alto riesgo, no respecto del resto.

Las propuestas europeas en materia ética, jurídica y de responsabilidad civil llegan en el momento oportuno y considero que, con independencia de las modificaciones que se produzcan en su próxima tramitación, se han convertido en una referencia internacional para la regulación de los aspectos esenciales de la inteligencia artificial.

4.2. Consideraciones particulares

La inteligencia artificial junto con sus bondades y fortalezas, presenta debilidades, desafíos y riesgos como los indicados anteriormente que deberán gestionarse adecuadamente.

Su naturaleza y características comporta debilidades inherentes o relacionadas con las mismas, que también pueden erigirse en fortalezas, como su complejidad, invisibilidad,

[139] Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos. OJ L 210, 7.8.1985. Pp. 29–33.

[140] Real Decreto Legislativo 1/2007, de 16 de noviembre (Texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias).

transversalidad, escalabilidad, capacidad de predicción, así como la necesidad de actualización constante.

De los riesgos generales expuestos en el apartado anterior, me permito abordar a continuación algunos de ellos junto con otros más específicos o relacionados con aquéllos que trataré con mayor profundidad y detalle por su actualidad y/o impacto potencial, sin perjuicio de su posterior agrupación bajo una o varias de las categorías indicadas por sus implicaciones.

Los principales riesgos asociados a la inteligencia artificial que pretendo significar por los motivos expuestos serían, entre otros, los asociados a su asimetría, infrautilización, su uso excesivo, los errores de diseño, defectos de programación o errores funcionales en su desarrollo, despliegue, integración e interoperabilidad con otras aplicaciones, sistemas, tecnologías, productos o servicios, utilización inadecuada o mal uso, datos o conocimiento de entrada sin calidad o viciados, falta de transparencia en su operatividad y funcionamiento, opacidad consecuente en la toma de decisiones, discriminación, intromisión en la vida privada, uso con fines ilegítimos o ilícitos, otorgamiento de capacidades de decisión o ejecución de acciones físicas o digitales autónomas o semiautónomas sin controles adecuados o ausencia de aprobación o control humano.

4.2.1. Asimetría

Los sistemas de inteligencia artificial deben ser diseñados y entrenados por expertos que necesitan acceder a grandes cantidades de datos y de computación, sin embargo, solo una minoría puede beneficiarse de la inteligencia artificial, especialmente grandes tecnológicas y algunas de las grandes potencias, en la medida que tengan acceso a datos, dispongan de la capacidad de computación, el conocimiento y la experiencia. Todo ello enlaza con la desigualdad en el mercado y afecta igualmente a la accesibilidad a la misma.

La asimetría enlaza con el problema de la asincronía en la disponibilidad de estas capacidades, de su ofrecimiento al mercado y consiguiente desigualdad competitiva.

De un lado, minimizar el desafío que supone la asimetría beneficiará a la sociedad en su conjunto y no a una parte. Y, de otro, minimizar esa asincronía, permitirá la posibilidad de competir nacional e internacionalmente al sector de la inteligencia artificial.

4.2.2. Infrautilización de la inteligencia artificial

La infrautilización de la inteligencia artificial se considera una amenaza en la medida que supondría no aprovechar las oportunidades que ofrece la misma para cualquier sector y esfera de nuestra vida, y la imposibilidad de alcanzar objetivos en campos como la sostenibilidad.

A modo de ejemplo, la infrautilización de la inteligencia artificial podría suponer la imposibilidad de alcanzar objetivos estratégicos e implementar de forma deficiente programas gubernamentales, por ejemplo y en el marco de la UE, el Pacto Verde Europeo. El impacto de este riesgo sería especialmente significativo para garantizar la ciberseguridad y la competitividad, por ejemplo, en la medida que estaríamos en desigualdad de armas.

La infrautilización puede suponer la pérdida de ventajas competitivas de las propias regiones, estados y empresas, el estancamiento económico y una pérdida de calidad de vida para la ciudadanía.

Esta infrautilización puede venir provocada por múltiples razones, por ejemplo, por la desconfianza de los ciudadanos frente a su uso -especialmente ante la falta de transparencia y seguridad-, o por el propio sector privado, que puede ver desincentivada su inversión en innovación e inteligencia artificial ante la irrupción de nuevos marcos normativos que impongan nuevos requerimientos, con costes asociados, que hagan inviable la misma o constituyan un obstáculo frente a su desarrollo en otras jurisdicciones con ausencia de regulación o marcos más laxos.

Esta infrautilización puede igualmente tener su origen en la falta de infraestructuras digitales y de telecomunicaciones que la soporte, la falta de iniciativa empresarial, las bajas inversiones, la falta de apoyo a la innovación y aplicación, así como la

fragmentación de los mercados digitales que repercute en el acceso a datos de los que depende el aprendizaje automático de la inteligencia artificial.

4.2.3. Uso o aplicación excesiva

El uso o aplicación excesiva es también un riesgo, especialmente ante inversiones en tecnologías de moda o identificadas como de mayor impacto para los próximos años, en base a criterios de marketing o comerciales más que respondiendo a objetivos de negocio o del bien común, lo que podrían dar como resultado generar aplicaciones de la inteligencia artificial que no aportan valor, o usar la inteligencia artificial en tareas que no la requieren.

La historia de la tecnología y de sus creadores nos ha dejado algunos ejemplos de lo que constituye innovar y desarrollar para mejorar la vida humana. George Westinghouse inventó todo tipo de cosas innovadoras, pero lo hizo pensando en mejorar la vida de las personas, es decir, como un medio para mejorar nueva vida y nuestro mundo, no como un mero fin.

Asimismo, otro de los riesgos que puede comportar es el creciente uso y dependencia de la misma y la posible amenaza de los procesos públicos de toma de decisiones en relación con su opacidad, que podría afectar a los pilares de su gobernanza y a los valores democráticos, de manera asociada a otros de los riesgos expuestos abordados más adelante, como la manipulación informativa.

4.2.4. Daños y perjuicios derivados de uso y funcionamiento.

Otro de los riesgos asociados son los daños o perjuicios derivados de su uso y aplicación y, en especial, la responsabilidad derivada de los mismos y su resarcimiento efectivo, que integra el objeto de esta investigación.

El problema principal en esta materia se suscita a la hora de determinar quién es el responsable de un daño causado por un elemento, dispositivo, aplicación, sistema, producto o servicio dentro de la cadena: El diseñador, el desarrollador, el fabricante del sistema, el fabricante del *hardware* o del elemento que la integra, el entrenador, el operador, el propietario, el usuario, el comercializador, el distribuidor u otros operadores dentro de la cadena de suministro o de su ciclo de vida.

A modo de ejemplo, un *robo-advisor* de segunda generación que nos aconseja una inversión de riesgo mínimo, un dron que dirige el vuelo autónomamente de origen a destino que atraviesa una cristalera de una casa o un vehículo autónomo que causa un accidente.

Los futuros marcos reguladores en materia de responsabilidad deberán asegurar el resarcimiento efectivo y definir las responsabilidades dentro de la cadena, al objeto de garantizar la seguridad, transparencia, trazabilidad y la rendición de cuentas desde el diseño y por defecto, armonizando intereses de todas las partes implicadas y buscando el equilibrio entre la confianza y la seguridad, con los incentivos adecuados para su uso, la inversión y la innovación.

La sociedad precisa una inteligencia artificial segura y confiable. Los diseñadores y productores deberán tener los suficientes incentivos para ofrecer productos seguros y de calidad, libre de errores y defectos funcionales.

4.2.5. Incapacidad de la IA para representar una realidad social compleja

Otro posible riesgo asociado a la inteligencia artificial sería la incapacidad de la misma para representar una realidad social compleja representada por meros valores numéricos, que puede conllevar a una apreciación y valoración errónea, alejada de toda objetividad, es decir, lo que conocemos como *mathwashing*, especialmente en la dimensión digital, donde, como he referido, las personas pasan a convertirse en números que suelen llevar asociado otros valores numéricos en dólares, euros u otras monedas.

4.2.6. Inadecuada o negligente definición, configuración, funcionamiento o uso.

Los riesgos asociados a una inadecuada o negligente definición, configuración, captación y tratamiento de datos, generación de resultados, toma de decisiones, ejecución de acciones o uso -"consciente" o "inconsciente"-, pueden comportar consecuencias para su propietario o usuario y en terceros, que incluyen gobiernos y sistemas democráticos, así como impactar en los derechos fundamentales de las personas afectadas, especialmente ante la existencia de sesgo y discriminación en el diseño o en los datos,.

El Convenio Europeo de Derechos Humanos (CEDH) no recoge una definición de lo que constituye discriminación directa o indirecta, si bien, la jurisprudencia del Tribunal Europeo de Derechos Humanos (TEDH)[141] la define como una diferencia de trato entre personas que se encuentra en situaciones análogas o similares, basada en una característica o condición identificable.

Estos riesgos pueden afectar también a la igualdad y la equidad.

En relación con el sesgo y la discriminación, significar que el razonamiento humano está impregnado de prejuicios y múltiples sesgos cognitivos y psicológicos, de modo que los mismos influyen en la inteligencia artificial, la cual puede ser igualmente utilizada para corregirlos o, cuanto menos, mejorarlos. Los sesgos son también el reflejo de la sociedad.

De hecho, el sesgo puede combatirse detectando y mitigando el mismo en los datos de entrada, haciendo al sistema consciente del sesgo y que lo aprenda a resolver, o eliminarlo en los datos de salida y resultado.

Según López de Mántaras[142], es absolutamente imposible que la inteligencia artificial esté libre de sesgo, "porque nosotros los tenemos. De hecho, el problema somos nosotros, no los algoritmos". En su opinión "los sesgos no tienen solución".

---

[141] *TEDH, Guía sobre el artículo 14 del Convenio (prohibición de la discriminación) y el artículo 1 del Protocolo N. o 12 (prohibición general de la discriminación*. Consejo de Europa/Tribunal Europeo de Derechos Humanos, 2020. P. 11.

[142] PASCUAL, M.G. (2021). "Cuando el algoritmo se equivoca". Publicado en *El País* el 27.06.2021.

En fechas próximas al cierre de esta investigación el National Institute of Standards and Technology (NIST), del Departamento de Comercio de los EE.UU., ha publicado un borrador de informe sujeto a consulta pública para la reducción del riesgo de sesgo en sistemas inteligentes bajo el título *A Proposal for Identifying and Managing Bias in Artificial Intelligence*[143], en el que destaca la necesidad de identificar y gestionar el sesgo en sus diferentes momentos del ciclo de vida de un sistema, y de aplicar control en el mismo.

El artículo 15.3 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión, de 21 de abril de 2021, hace referencia explícita al sesgo y la necesidad de mitigarlo durante su ciclo de vida, estableciendo que los sistemas de inteligencia artificial de alto riesgo que continúan aprendiendo tras su introducción en el mercado o puesta en servicio se desarrollarán de tal modo que los posibles sesgos en la información de salida debidos al uso de esta como datos de entrada en futuras operaciones -bucle de retroalimentación-, se subsanen debidamente con las medidas de mitigación oportunas.

La presencia de sesgos en los sistemas inteligentes puede producirse en la propia recogida de datos y muestras -cuando se recogen datos que pretende sean la representación más fiel posible de la realidad-, así como en su posterior tratamiento mediante la exclusión de algunos datos del conjunto por lo que no serán considerados en su procesamiento -sesgo de exclusión-.

Adicionalmente, otros sesgos comunes de la inteligencia artificial son los prejuicios que se producen cuando los datos de entrenamiento incurren en estereotipos y prejuicios comunes existentes en la sociedad, y el denominado sesgo algorítmico, que puede incluir distintos sesgos, que pueden provenir de errores en el diseño del algoritmo, errores en la recogida de datos con la forma en la que éstos se seleccionan y procesan.

[143] Recuperado de: https://www.nist.gov/artificial-intelligence/proposal-identifying-and-managing-bias-artificial-intelligence-sp-1270. Consultado el 29.06.2021.

Como expresamente recoge el precitado *Libro blanco sobre la inteligencia artificial*[144], pueden producirse situaciones en las que el uso de determinados algoritmos para predecir, por ejemplo, la reincidencia delictiva, de lugar a prejuicios raciales o de género y puede prever una probabilidad distinta de reincidencia para hombres y mujeres o para nacionales o extranjeros[145].

Del mismo modo, algunos programas de inteligencia artificial utilizados para el análisis facial pueden integrar con facilidad prejuicios raciales o de género[146].

El riesgo de sesgo y de opacidad ya se ha materializado en múltiples ocasiones en casos muy notorios de los que se ha hecho eco la prensa internacional, siendo especialmente significativos en casos como el caso B. Borden[147], el de Amazon[148] o Apple[149].

El primero de los casos se trataba de una joven de raza negra de 18 años que fue denunciada y posteriormente encarcelada por coger una bicicleta y un patinete de un niño de seis años con un amigo para circular por la calle y que abandonaron instantes después tras los gritos de una persona que les recriminó la acción.

B. Borden fue llevada a prisión junto con otras personas condenadas y fue sometida, como el resto, a calificación por un algoritmo, el cual debía determinar el nivel de riesgo de futura comisión de nuevos delitos. El algoritmo determinó un nivel de riesgo 8 para B. Borden mientras que, simultáneamente, determinó el nivel de riesgo de V. Prater, un

---

[144] *Libro Blanco sobre la inteligencia artificial: Un enfoque europeo orientado a la excelencia y la confianza*. Comisión Europea. 2020. COM (2020) 65 final.

[145] TOLAN S., MIRON M., GOMEZ E. Y CASTILLO C. (2019)."Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia*"*. Best Paper Award, *International Conference on AI and Law*, 2019.

[146] BUOLAMWINI, J. Y GEBRU, T. (2018). "Proceedings of the 1st Conference on Fairness, Accountability and Transparency*"*. *PMLR* 81. Pp 77-91. 2018.

[147] LARSON, A.J.; MATTU, S. Y KIRCHNER, L. (2016). "Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks". *ProPublica*. 2016. Recuperado de: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[148] Publicada por agencias y medios de comunicación como *Reuters*: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G?feedType=RSS&feedName=topNews&utm_source=twitter&utm_medium=Social] o el *El País* [https://elpais.com/tecnologia/2018/10/11/actualidad/1539278884_487716.html]. Consultado el 12.2.2021.

[149] Publicada, entre otros medios, por la BBC: https://www.bbc.com/mundo/noticias-50375172

delincuente de raza blanca con antecedentes y condenas por delitos a mano armada, en un nivel 3.

El segundo de los casos indicados, reconocido por la compañía estadounidense, se planteó en relación con el algoritmo diseñado y desarrollado por Amazon para la contratación de empleados que, tras su entrenamiento, empezó a discriminar los currículos de mujeres frente a los de hombres, que motivó la retirada del mismo.

El tercero, negado por la compañía e imputado por la misma al "inexplicable algoritmo", hace referencia a las noticias que se publicaron en medios de comunicación denunciando que la tarjeta Apple Card estaba dando ventajas en el crédito a los hombres frente a sus esposas, a pesar de hallarse ambos en las mismas condiciones financieras.

La situación se hizo a pública a raíz de un tuit publicado por un desarrollador danés con el siguiente tenor literal: "*Mi esposa y yo hacemos declaraciones de impuestos conjuntas, vivimos en una propiedad compartida y llevamos mucho tiempo casados. Sin embargo, los algoritmos de Apple creen que yo merezco un límite de crédito 20 veces mayor que ella*". Ello motivó una investigación por el Departamento de Servicios Financieros del Estado de Nueva York -DFS por sus siglas en inglés-.

Asimismo, en relación con estos aspectos, debo significar el uso del algoritmo COMPAS -*Correctional Offender Management Profiling for Alternative Sanctions*- utilizado por el sistema judicial de Wisconsin (EE.UU.), como ayuda para los jueces para dictar sentencia y valorar aspectos como el riesgo de reincidencia, como en otros casos precitados. La investigación llevada a cabo por un grupo de periodistas evidenció que atribuía a las personas de raza negra mayor probabilidad de ser juzgados incorrectamente, atribuyéndoles un mayor riesgo de reincidencia que a las personas de raza blanca[150].

Otro ejemplo evidente se produjo en relación con Word2Vec[151], la red neuronal de Google Brain, que aprendió que "hombre es a rey lo que mujer a reina", pero también que

[150] LARSON, J.; MATTU, S.; KIRCHNER, L. Y ANGWIN, J. (2016). "How we analyzed the COMPAS recidivism algorithm". *ProPublica*. 23 de mayo de 2016. Disponible en: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

[151] PASCUAL, M.G. (2021). "Cuando el algoritmo se equivoca". Publicado en *El País* el 27.06.2021.

"padre es a médico lo que madre a enfermera" o que "hombre es a programador lo que mujer a ama de casa".

El sesgo ha provocado incluso la caída de gobiernos, como ha ocurrido recientemente en Holanda por su programa de gestión de ayudas y subsidios, que obligó a que su ejecutivo presentará su dimisión en bloque[152].

El Convenio Europeo de Derechos Humanos y la propia Constitución Española, al igual que otras constituciones, prohíbe la discriminación y son múltiples los marcos normativos que igualmente prohíben la misma.

En relación con todo ello, el propio Reglamento Europeo de Protección de Datos (RGPD)[153], en su artículo 22, reconoce el derecho de las personas a no ser objeto de decisiones basadas únicamente en el tratamiento automatizado de la información, con efectos jurídicos o afectación relevante, exigiéndose en determinados contextos las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado -como mínimo el derecho a obtener intervención humana por parte del responsable de su tratamiento, a expresar su punto de vista y a impugnar la decisión-, con la consiguiente reclamación de responsabilidad por los daños y perjuicios sufridos derivados de su infracción.

Las nuevas propuestas reguladoras europeas en el ámbito de la inteligencia artificial, objeto de esta investigación incluyen, regulan expresamente la no discriminación. Del mismo modo, son múltiples las iniciativas privadas[154] que la contemplan entre sus principios éticos para su autorregulación, como abordaré posteriormente, si bien, en ocasiones, surgen dudas de algunas iniciativas y su intención cuando, de un lado se anuncia por una empresa su compromiso con la privacidad y la no discriminación, pero

[152] FERRER, I. (2021). "El Gobierno holandés dimite en bloque por el escándalo en las ayudas al cuidado de niños". Publicado en *El País*, el 16.01.2021.
[153] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos).
[154] BENJAMINS, R. (2020). "Los líderes mundiales de la inteligencia artificial ética". *Thing Big / Empresas*. 08.01.2020. Disponible en: https://empresas.blogthinkbig.com/los-lideres-mundiales-de-la-inteligencia-artificial-etica/. Consultado el 24.02.2021

de otro, en su práctica y funcionamiento vulnera estos principios, es decir, lo que se denomina "*Ethics washing*"[155].

No obstante, merece destacarse que la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre inteligencia artificial - Ley sobre inteligencia artificial- de 21 de abril de 2021, no contiene una definición de "sesgo" entre sus múltiples definiciones, sin perjuicio de que sea contemplado a lo largo del mismo, tanto en sus Considerandos 40 y 44 como en sus artículos 10, 14 y 15.

En una reciente comunicación bajo el título *Mitigating Bias in Artificial Intelligence* de IBM Policy Lab, éste propone algunas medidas al legislador para minimizar los casos de sesgo.

4.2.7. Falta de calidad de calidad de los datos y conocimientos de entrada

La falta de calidad de los datos o conocimientos de entrada o que los mismos se hallen viciados, es otro de los riesgos que puede convertir la tecnología en ineficaz, el cual se halla directamente relacionado con los anteriores.

Un ejemplo de ello, es la introducción de sesgos por razón de la raza, sexo o edad en los datos de entrada para la toma de decisiones, como conceder un préstamo, autorizar un contrato, proponer un despido o establecer indicios de relación, culpabilidad o imputación de una conducta ilícita civil o penal.

[155] WAGNER, B. (2018). *Ethics as an Escape from Regulation: From ethics-washing to ethics-shopping?*. In M.Hildebrandt (Ed.), Being Profiling. Cogitas ergo sum. Amsterdam University Press. 2018.

4.2.8. Riesgos para la intimidad, la privacidad y la protección de datos personales.

Del mismo modo, la aplicación de la inteligencia artificial puede comportar importantes riesgos para la intimidad, la privacidad y la protección de datos personales, especialmente antes sistemas de reconocimiento facial, seguimiento en línea o "*profiling*".

Además, la interacción de la inteligencia artificial con Big data y otras tecnologías para la relación de datos proporcionados por la persona con datos captados, sin su conocimiento o asociados con datos anónimos masivos, puede generar perfiles personales de los que ni tan siquiera la persona afectada es consciente o pudiera ni tan siquiera imaginar.

Las brechas de privacidad e incumplimientos flagrantes de los marcos protectores de la misma han comportado la consiguiente inquietud y preocupación social por la privacidad que, a su vez, conlleva una correlativa falta de confianza.

La inteligencia artificial se nutre de datos. Son su sustento. La mayor disponibilidad de datos permite un mayor desarrollo y aplicación de la inteligencia artificial, como está ocurriendo en China.

Los datos se consideran "el petróleo del siglo XXI", estructurados y desestructurados, sean personales o no personales, pero especialmente los primeros, en un mundo donde las personas nos hemos convertido para las grandes tecnológicas en un número en dólares o en euros, incluso no importa tanto nuestro nombre y apellidos en la medida que el valor lo constituyen el conjunto de datos que generamos, en el que somos el producto, en el que "donamos", en ocasiones "permutamos" y hasta "vendemos" nuestra privacidad de forma continua y habitualmente de manera imperceptible e inconsciente, y donde se lleva a cabo una vigilancia excesiva y manipulación de la información que, de nuevo, fue reflejada por la ciencia-ficción, en particular en la novela de George Orwell bajo el título "1984", publicada el 8 de junio de 1949, en la que ya hablaba del "Gran Hermano" que vigila.

Las gigantes tecnológicas que conforme las *GAFAM* (Google, Amazon, Facebook, Apple y Microsoft), con miles de millones de usuarios en el mundo, son las que concentran la mayoría de los datos que se generan, lo que comporta varios riesgos, en especial el enorme

impacto que puede generar cualquier brecha de seguridad o privacidad, así como el poder de llegar a dominar los servicios de inteligencia artificial en gran parte del mundo.

Por el momento, los marcos reguladores lo dificultan, pero el acceso y disponibilidad a todos esos datos masivos de ciudadanos se está mostrando como el esencial para la eclosión de la aplicación de la inteligencia artificial, como está ocurriendo en China.

El Reglamento General de Protección de Datos (RGPD), el Reglamento europeo sobre datos no personales y la Directiva europea de ciberseguridad plantean en determinados contextos el problema de la disponibilidad de los datos en la UE.

Actualmente, algunas empresas utilizan sistemas de inteligencia artificial para monitorizar a sus empleados, sus comunicaciones, sus perfiles en redes sociales o evaluar su personalidad. Y en todos estos contextos es donde debemos reflexionar ahora si debemos ir más allá y autorizar el uso de nuestros datos por parte de empresas y administraciones públicas con la finalidad de mejorar nuestras vidas, incluso la posibilidad de que compartan nuestros datos con terceros con esta misma finalidad, es decir, el uso y tratamiento de nuestros datos en beneficio personal e incluso de la sociedad en general o bien común.

Esta reflexión ha cobrado todavía mayor relevancia en los tiempos de pandemia mundial que estamos viviendo y en cómo hemos visto que se ha gestionado la pandemia en determinados países como China y Corea de Sur, de un lado, y en países como EE.UU. o Europa, de otro.

Para ello es necesaria la confianza absoluta en ese uso y tratamiento gubernamental y empresarial con dicha finalidad y es la base de la denominada economía de los datos, sobre lo que hay distintas iniciativas, como la del *Massachusetts Institute of Technology* (MIT)[156], liderada por Sandy Pentland, sobre la confianza como base para sacar el máximo provecho de los datos para las personas individualmente consideradas como para la sociedad y el planeta.

[156] *MIT Trust::data consortium.* 2019. Recuperado de: https://trust.mit.edu/. Consultado el 15.02.2021.

De hecho, deberíamos realizar la reflexión de si sería "no ético" el hecho de no usar tecnologías disponibles para resolver los grandes problemas del ser humano y del planeta, como las pandemias o el cambio climático.

Además de todo ello, el potencial de los sistemas de inteligencia artificial para el tratamiento de datos de manera masiva no solamente está siendo utilizada con fines legítimos sino también con finalidades delictivas, especialmente para llevar a cabo acciones como el denominado "*spear phishing*" selectivo[157] basado en inteligencia artificial, en las que los delincuentes se benefician de estos sistemas para rastrear las redes y para encontrar información útil sobre la persona específica a suplantar, analizarla e imitar su lenguaje, estilo de comunicación o incluso el tono y timbre de voz.

Los marcos de protección de los datos no son uniformes a nivel mundial, si bien, la UE ha sentado un marco de referencia respecto de la protección de datos personales a través, principalmente, del Reglamento General de Protección de Datos (RGPD)[158], que se ha erigido en una referencia y modelo legislativo internacional en esta materia, así como sobre la protección de datos no personales y protección del secreto empresarial, a través principalmente, de su Reglamento relativo a un marco para la libre circulación de datos no personales en la Unión Europea[159] y de la Directiva relativa a la protección de los conocimientos técnicos y la información empresarial no divulgados (secretos comerciales) contra su obtención, utilización y revelación[160], transpuesta al ordenamiento jurídico español a través de la Ley 1/2019, de 20 de febrero, de Secretos Empresariales[161].

[157] MUÑOZ VELA, J.M. (2021). "Las estrategias delictivas en el ciberespacio se perfeccionan*"* Publicado el 17.02.2021 en *Valencia Plaza*. Recuperado de: https://valenciaplaza.com/estrategias-delictivas-ciberespacio-ciberataques.

[158] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respeta al tratamiento de datos personales ya la libre circulación de estos datos y por el que se deroga la Directiva 95/46 / CE (Reglamento general de protección de datos) (Texto pertinente a efectos del EEE). DO L 119 de 4.5.2016. Pp. 1-88

[159] Reglamento (UE) 2018/1807 del Parlamento Europeo y del Consejo de 14 de noviembre de 2018 relativo a un marco para la libre circulación de datos no personales en la Unión Europea. DO L 303 de 28.11.2018. Pp. 59-68.

[160] Directiva (UE) 2016/943 del Parlamento Europeo y del Consejo, de 8 de junio de 2016, relativa a la protección de los conocimientos técnicos y la información empresarial no divulgados (secretos comerciales) contra su obtención, utilización y revelación ilícitas. OJ L 157, 15.6.2016. Pp. 1-18

[161] BOE Núm. 45, de 21 de febrero de 2019. Pp. 16713 a 16727

Los marcos reguladores de la privacidad exigen transparencia y seguridad proactiva en el diseño y por defecto, así como medidas de seguridad adecuadas a los riesgos y contexto que pueden incluir la anonimización (disociación irreversible), la disociación de los datos y medidas como el cifrado en su almacenamiento y comunicación de los datos siempre que sea posible conforme al contexto, en especial, en relación con el tratamiento por sistemas inteligentes. Aun así, medidas como la anonimización completa ante sistemas de inteligencia artificial no es una cuestión pacífica, dadas sus capacidades de tratamiento de datos no personales o anónimos que podrían, en su análisis conjunto, permitir asociarlos a una persona en determinados supuestos.

Estos marcos legislativos son considerados por algunos como un obstáculo para el desarrollo de la inteligencia artificial, de los que se están beneficiando otros países sujetos a marcos más laxos y permisivos.

En mi opinión, debemos encontrar el punto de equilibrio entre regulación, transparencia, seguridad y protección de derechos fundamentales, sin olvidar el fomento de la innovación y la competitividad, que permita vislumbrar un futuro donde las personas tengamos la confianza necesaria en el uso, tratamiento y compartición de datos a nivel empresarial o gubernamental para el bien individual y común.

De este modo, los datos podrían permitir el despliegue y aplicación de sistemas inteligentes que se alimentarían de los mismos para luchar contra el hambre, la pobreza, la salud, la predicción y control de pandemias, el cambio climático, la sostenibilidad de las ciudades, etc.

El movimiento que está promoviendo este tipo de uso de los datos se conoce como "*Data for good*", "*AI for good*", "*Business to government data sharing*" o "*Big data por social good*".

Existen distintas iniciativas de entidades de investigación y empresas, como el GovLab[162] de la Universidad de Nueva York, que gestiona un inventario con casi un centenar de

[162] *Govlab*. Recuperado de: https://datacollaboratives.org/. https://datacollaboratives.org/explorer.html. Consultado el 20.03.2021

iniciativas[163] donde se usan datos de empresas privadas para el bien social, la iniciativa BIDA para el cambio climático, los proyectos en España de Telefónica[164] y Vodafone, el proyecto de vocación internacional *BBVA Data & Analytics para el Bien Social* del BBVA con proyectos de salud pública desplegados en Bangladesh y Pakistán, *Orange Data for Development* con despliegue en Senegal y Costa de Marfil o *Turkcell Data for Refugees Challenge* con el objetivo de compartir datos agregados y anonimizados de la operadora Turkcell para ayudar a resolver la crisis de los refugiados.

Junto con algunos de los grandes operadores de telecomunicaciones, las grandes tecnológicas también se han sumado con iniciativas sobre el uso de datos e inteligencia artificial para el bien social, como Facebook[165], Google[166] y Microsoft[167].

Asimismo, destacar otras iniciativas sobre el *Big data* y la inteligencia artificial para el bien común, como el *New Deal on Data* desde el Foro Económico Mundial, *Flowminder* en Suecia, el *World Data Forum* y el *Global Partnership for Suitainable Development Data* liderado por Naciones Unidas, *OPAL*, *Partnership on AI*, *GSMA Big Mobile Data for Social Good*, *AI for Good Global Summit of the ITU* o el *Center for Humane Technology* en California.

La Comisión Europea ha creado un *Grupo de Expertos para definir la estrategia del "Business to Government -B2G- data sharing*"[168] en Europa, que ha emitido un primer documento de recomendaciones[169] que incluye la adopción de medidas legislativas sobre la gobernanza de los datos, el acceso y su reutilización para el interés público. Todo ello

---

[163] *Data Collaboratives Explorer*. Govlab, NYU, 2019. Op. cit. Disponible en: https://datacollaboratives.org/. Consultado el 20.03.2021

[164] *Data for Good - Giving back the value of data to society in line with the 2030 Sustainable Development Goals*. LUCA, Telefónica, 2016. Recuperado de: https://luca-d3.com/data-for-good. Consultado el 20.03.2021

[165] *Facebook Data for Good. 'We use data to address some of the world's greatest humanitarian issues'*. Recuperado de: https://dataforgood. Consultado el 20.03.2021.
fb.com/

[166] *Google AI*. Recuperado de: https://ai.google/social-good/impact-challenge/. Consultado el 20.03.2021.

[167] *Microsoft AI*. Recuperado de: https://www.microsoft.com/en-us/ai/ai-for-good. Consultado el 20.03.2021.

[168] *Towards a European strategy on business-to-government data sharing for the public interest*. Expert Group on B2G Data Sharing, 2020. Recuperado de: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=64954. Consultado el 15.02.2021.

[169] Recuperado de: https://ec.europa.eu/digital-single-market/en/news/meetings-expertgroup-business-government-data-sharing. Consultado el 14.12.2020.

alineado con la *Estrategia de Datos*[170] europea de la UE, orientada a crear un único mercado europeo para datos que asegure la competitividad de Europa y la soberanía de sus datos.

Existe consenso en el grupo de expertos respecto de la necesidad de incrementar y mejorar el uso de los datos de las empresas privadas por el interés común, destacando tres bloques de medidas que consideran imprescindibles para ello:

a) Medidas de gobernanza y explorar una posible regulación;

b) Medidas de transparencia, implicación de la ciudadanía y ética, con garantías como privacidad, explicabilidad y ausencia de sesgos y;

c) Definición de modelos operacionales, infraestructura y herramientas con incentivos a las empresas para compartir datos y la asignación de fondos europeos para la creación de herramientas e infraestructura.

En mi opinión, a todo ello debería adicionarse necesariamente información, concienciación y formación, para generar confianza y seguridad en todos los agentes involucrados, desde la persona interesada hasta la empresa responsable, cedente o cesionaria de los datos.

La mayoría de estos proyectos y aplicaciones que he comentado ni tan siguiera necesitan datos personales, sino datos anonimizados y agregados, de modo que no resultarían de aplicación inicial algunas posibles restricciones por parte de los marcos regulativos en materia de privacidad.

Por ello, considero que la situación actual de pandemia que estamos viviendo a nivel mundial evidencia más que nunca la necesidad de adoptar un posicionamiento al respecto.

[170] Recuperado de: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digitalage/european-data-strategy_en. Consultado el 14.12.2020.

Es innegable el valor que podría estar aportando el uso global de la inteligencia artificial y el Big data para ayudar a la gestión y minoración de la situación actual, en términos de eficacia y eficiencia, sirviendo como indicador las experiencias en este sentido de países como China o Corea del Sur y las del resto del mundo.

De hecho, a finales de diciembre una plataforma de inteligencia artificial[171] canadiense -"Bluedot"-, ya había detectado basándose en *Big data* un número importante de neumonías raras alrededor de un mercado de Wuhan y alertó a las autoridades.

Asimismo, durante la expansión de la pandemia, algunas *start-ups* de inteligencia artificial evidenciaron su valor y potencial para el bien social, entre otras, la valenciana "Quibim"[172], que desarrolló una red neuronal con aprendizaje profundo supervisado para aconsejar si un contagiado necesita ser hospitalizado o no.

En el ámbito europeo, los primeros pasos se focalizaron en las operadoras de telecomunicaciones para usar su *Big data* para la lucha contra la pandemia, mediante el uso de datos anonimizados y agregados.

Pero el valor de la inteligencia artificial no sólo se halla relacionado con la gestión y control eficaz de una pandemia, sino incluso con el propio desarrollo acelerado de una posible vacuna y su eficacia.

Y el uso de datos con estas finalidades de salud pública y emergencia sanitaria no es nuevo. Se han utilizado datos de empresas de telecomunicaciones para combatir el ébola en África o la gripe porcina en México. También se han utilizado datos de redes sociales para comprender las percepciones sobre el zika en Brasil o datos obtenidos vía satélite para rastrear los brotes estacionales de sarampión en Níger.

---

[171] STIEG, C. (2020). "How this Canadian start-up spotted coronavirus before everyone else knew about it". Publicado en *CNBC* el 03.03.2020. Disponible en: https://www.cnbc.com/2020/03/03/bluedot-used-artificial-intelligence-to-predict-coronavirus-spread.html. Consultado el 19.02.2021.

[172] RUBIO, I. (2020). "Inteligencia artificial valenciana para aconsejar si hospitalizar a un contagiado". Publicado el 18.03.2020 en *El País*. Disponible en: https://retina.elpais.com/retina/2020/03/17/innovacion/1584450877_681658.html. Consultado el 19.02.2021.

La inteligencia artificial se alimenta de datos y el crecimiento de su aplicación y usos para el bien común depende de los mismos.

Actualmente y hasta que dispongamos de un nuevo contexto regulador, en la mayoría de los supuestos un tratamiento lícito de datos personales requiere el consentimiento expreso de su titular para las finalidades concretas para las que vayan a ser tratados o cedidos a terceros, sin perjuicio de que puedan concurrir otras más excepcionales como la concurrencia de obligaciones legales, la necesidad para proteger intereses vitales del interesado o implementación de misiones en interés públicos o en el ejercicio de poderes públicos, y en cualquier caso sujeto a límites temporales. Sin embargo, el tratamiento de datos verdaderamente anónimos no estaría sujeto a todo ello.

Este contexto genera temor a las entidades que traten datos personales y que puedan cederlos a terceros para un bien social, debiendo cumplir los marcos precitados, anonimizando los datos (disociación irreversible) de manera efectiva, aplicando seguridad y considerando los riesgos asociados a cualquier incumplimiento o brecha que puede suponer. Y dentro de los riesgos no sólo deben considerar las importantísimas sanciones económicas que podrían derivarse sino, sobre todo, los daños en su reputación e imagen corporativa y su impacto en su relación con socios, inversores y cuenta de resultados.

A mi juicio, ahora y más que nunca es el momento de poner el *Big data* y la inteligencia artificial al servicio de la humanidad y cambiar ciertos paradigmas con ciertas garantías que nos permitan concebir los datos como un servicio para el bien común o *Data as a Service* -DaaS por sus siglas en inglés-, restringido a usuarios y contextos legítimos.

Su valor social está, en gran medida, desaprovechado, lo que supone que la sociedad podría estar abordando y, en su caso, resolviendo grandes problemas más rápido y mejor. Las estrategias europeas de datos deberán acometer estos aspectos.

La UE es consciente de todas estas cuestiones y, en paralelo, las iniciativas para vertebrar una estrategia de sobre gobierno de los datos han sido especialmente intensas en los últimos dos años.

En febrero de 2020 la UE publicó sus estrategias para impulsar la transformación digital y garantizar el desarrollo de la inteligencia artificial y, en particular, su *Estrategia Europea de Datos* [173] que busca convertir a la UE en líder de una sociedad impulsada por los mismos.

La *Estrategia Europea de Datos* se presentó al mismo tiempo que la Comunicación de la Comisión bajo el título *Modelar el futuro digital de Europa* y el *Libro Blanco sobre la inteligencia artificial* que refleja la manera en que la Comisión apoyará y promoverá el desarrollo y el uso generalizado de la inteligencia artificial en toda la UE.

Los datos son considerados por la UE como un elemento vital del desarrollo económico, en la medida que constituyen la base de muchos nuevos productos y servicios, son un recurso esencial para las empresas emergentes y para las pequeñas y medianas empresas (PYMES) a la hora de desarrollar productos y servicios, y "su disponibilidad es fundamental para entrenar a los sistemas de inteligencia artificial, dado que los productos y servicios están evolucionando rápidamente desde el reconocimiento de patrones y la generación de información hasta técnicas de predicción más sofisticadas y, por tanto, mejores decisiones".

Los datos también nutren la amplia aplicación de prácticas transformadoras como el uso de gemelos digitales[174] en la fabricación. Y, por último, para la UE, es esencial disponer de más datos y mejorar la manera en que se utilizan para hacer frente a los retos sociales, climáticos y relacionados con el medio ambiente, contribuyendo a unas sociedades más sanas, más prósperas y más sostenibles.

La Comisión Europea ha reconocido que el valor de la economía de los datos va a alcanzar los 829.000 millones de euros en 2025 (5,8 % del PIB de la UE).[175]

[173] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones Una Estrategia Europea de Datos. 19.2.2020. COM/2020/66 final

[174] Los gemelos digitales crean una réplica virtual de un producto, proceso o sistema. La réplica puede prever, por ejemplo, cuándo fallará una máquina a la luz del análisis de los datos, lo que permite aumentar la productividad mediante el mantenimiento predictivo.

[175] *Estrategia Europa de Datos*. Comisión Europea 2020. Disponible en: https://ec.europa.eu/commission/presscorner/api/files/attachment/862111/European_data_strategy_es.pdf.pdf. Consultado el 01.03.2021.

Conforme destaca la Comisión Europea en su *Comunicación sobre Una Estrategia Europea de Datos de 19.02.2020*[176], el valor de los datos reside en su uso y reutilización y, en la actualidad, no hay suficientes datos disponibles para una reutilización innovadora, en particular, en el caso del desarrollo de la inteligencia artificial. Además, aborda la necesidad de la disponibilidad y uso de los datos en beneficio del bien público.

Según la estrategia de la EU, durante el período 2021-2027, la Comisión invertirá en un proyecto de gran impacto sobre los espacios de datos europeos y las infraestructuras federadas de computación en la nube, financiando infraestructuras, herramientas de intercambio de datos, arquitecturas y mecanismos de gobernanza con vistas a unos ecosistemas *florecientes* para la puesta en común de datos y la inteligencia artificial.

La UE pretende que los datos sean disponibles, de calidad e interoperables y que se dispongan de las infraestructuras para almacenarlos y tratarlos, lo que requiera no sólo decisiones políticas y medidas consecuentes, sino también un acompañamiento legislativo, que se ha instrumentado al través del futuro Reglamento de gobernanza de datos -Ley de Gobernanza de Datos- de la que hablaré posteriormente.

La gobernanza de los datos pretender conformar un conjunto de normas y medios para su utilización, por ejemplo, a través de mecanismos de intercambio, acuerdos y normas técnicas. Implica estructuras y procesos para compartir datos de manera segura, incluso a través de terceros de confianza.

En este sentido, la Comisión Europea ha propuesto una nueva forma europea de gobernanza de los datos para facilitar el intercambio de los mismos entre sectores y Estados miembros con la finalidad de generar riqueza para la sociedad, proporcionar control a los ciudadanos y confianza a las empresas.

El objetivo es crear un mercado único de datos en el que los datos pueden circular por toda la UE y entre sectores en beneficio de todos, así como que se respeten los marcos

[176] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones Una Estrategia Europea de Datos. 19.2.2020. COM/2020/66 final.

regulatorios vigentes, especialmente en materia de privacidad y competencia, y que las normas para el acceso a datos y su utilización sean justas, prácticas y claras.

Como consecuencia de todo ello y de la necesidad de una intervención legislativa armonizada en la UE, se ha formalizado la precitada Propuesta de Reglamento para la gobernanza de los datos en la UE[177], en tramitación en la fecha de cierre de esta investigación y que complementará la Directiva sobre datos abiertos, de junio de 2019[178], a la que ha proseguido la tramitación de la iniciativa conocida como *Data Act* para facilitar la accesibilidad y utilización de los datos, en particular entre empresas y entre éstas y las Administraciones Públicas, sobre la que se ha abierto una consulta pública y que pretende revisar la Directiva 96/9/EC sobre la protección legal de bases de datos.

El futuro Reglamento sobre la gobernanza de los datos pretende facilitar y garantizar el acceso y utilización de más datos para la economía y la sociedad de la UE, y ofrecer un mayor control a los ciudadanos y las empresas sobre los datos que generan. Esto reforzará la soberanía digital de Europa en el ámbito de los datos.

De este modo, será más fácil para los europeos permitir, en beneficio de la sociedad, el uso de los datos que les conciernen, respetando plenamente las normas de protección de los datos personales. Y también las pequeñas como las grandes empresas se beneficiarán de nuevas oportunidades de negocio, así como de una reducción de los costes de adquisición, integración y tratamiento de datos, de la reducción de los obstáculos para acceder a los mercados, y de un acortamiento del plazo de comercialización de nuevos productos y servicios.

En este sentido, coincido con algunos de los expertos más autorizados a nivel europeo sobre la necesidad de federar los recursos europeos sobre datos, entre otros, Luis Filipe

[177] Documento de Trabajo de los Servicios de la Comisión Resumen del Informe de la Evaluación de Impacto que Acompaña al Documento Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la Gobernanza Europea de Datos (Ley de Gobernanza de Datos). 25.11.2020. SWD/2020/296 final.

[178] Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativa a los datos abiertos y la reutilización de la información del sector público. PE/28/2019/REV/1. OJ L 172, 26.6.2019. Pp. 56-83

Coelho[179], catedrático del Departamento de Ciencias de la Computación de la *Facultad de Ciencias de la Universidad de Porto* y Director del *Competence Centre for Cybersecurity and Privacy* de la misma universidad.

Según el mismo, la UE debe pensar y actuar en sus estrategias internacionales y competencia como grupo, en especial, articular una estrategia europea para el *cloud computing*, dado que, de otro modo, los enfoques estatales pueden lastrar el avance. Según este experto, considera una buena estrategia federar los recursos europeos en torno al dato y el aprovechamiento de datos entre compañías.

La estrategia de la Comisión Europea se acompañará por dos marcos regulativos más, la denominada *Digital Services Act* y la *Digital Markets Act*, con el objetivo de limitar el poder de las grandes plataformas de servicios a través de internet y reforzar los derechos de los consumidores limitando las condiciones injustas que aquellas imponen.

Del mismo modo, me permito destacar la técnica legislativa aplicada y la necesidad de mantener el equilibrio para evitar una excesiva regulación, pero hacerlo de manera "inteligente", garantizando la seguridad y lo derechos fundamentales, e impulsando la innovación, la productividad y el crecimiento económico.

4.2.9. Afectación de otros derechos fundamentales

La inteligencia artificial puede poner en riesgo la dignidad personal y dentro de ésta la propia identidad personal y, en particular, la identidad en línea, conformada por los propios algoritmos en función del flujo de información de la persona en la red fuera de su control.

La aplicación de la inteligencia artificial también plantea importantes retos y riesgos para otros derechos fundamentales, como la libertad de expresión, especialmente de crítica, y

[179] Intervención en el "*XXI Seminario Permanente de la Cátedra Google de Privacidad, Sociedad e Innovación*" sobre los "*Desafíos y oportunidades que plantea la estrategia Europea de Datos y transformación digital*" celebrada el 22.10.2020.

la libertad de reunión, especialmente ante el rastreo y control de personas, en particular, asociadas a determinadas convicciones o acciones.

La inteligencia artificial puede ser utilizada para modelar la opinión pública generando y distribuyendo información falsa -*fake news*- de forma gestionada para su manipulación e influencia en procesos democráticos.

La inteligencia artificial permite la acotación de información y contenidos en línea en base al comportamiento de las personas afectadas, lo que permite mostrar únicamente contenidos específicos y afectar a los valores democráticos. Esto está sucediendo diariamente en la actualidad y ha motivado incluso la restricción del uso o no distribución de algunos sistemas por su potencial y riesgos de uso malicioso[180].

Del mismo modo, las capacidades más avanzadas de la inteligencia artificial están siendo utilizadas contra estos derechos precitados para crear imágenes, declaraciones, noticias, comunicaciones, audios y videos falsos con absoluta apariencia de realidad y verosimilitud, es decir, los llamados *deep fakes*.

En este caso se trata de contenidos audiovisuales generados gracias al *Deep Learning*, mediante técnicas propias del cine que permiten sintetizar la imagen humana, combinando imágenes creadas digitalmente con ya existentes y que pueden utilizarse para resucitar a alguien ya fallecido o generar un video pornográfico de un personaje público[181].

[180] A modo de ejemplo, GPT-2, un sistema de aprendizaje profundo creado por los investigadores de Open AI -entidad de investigación sin ánimo de lucro-, y capaz de generar texto coherente a partir de frases escritas por personas, debido a sus capacidades para automatizar la creación de noticias falsas, llevó a sus creadores a no publicar su investigación y restringir el mismo por los riesgos de uso malicioso. Hern, Alex. HERN, A. (2019). "New AI fake text generator may be too dangerous to release, say creators". Publicado el 14.02.2019 en *The Guardian*. Disponible en: https://www.theguardian.com/technology/2019/feb/14/elon-musk-backed-ai-writes-convincing-news-fiction. Consultado el 23.02.2021.

[181] Se han visto afectadas *celebrities* como la actriz Scarlett Johansson o la cantante Taylor Swift, entre otras muchas. ZAMORANO E. (2020). "Los 'deepfakes' del porno: así han reaccionado las actrices". Publicado en el *El Confidencial* el 23.01.2020. Disponible en: https://www.elconfidencial.com/alma-corazon-vida/2020-01-23/deepfake-porno-sexualidad-internet-internet_2420819/. Consultado el 17.02.2021.

Se trata de una imitación prácticamente perfecta del aspecto físico de una persona en toda su dimensión, incluyendo gestos, volumen, timbre y tono de voz, ritmo, pausas, etc.

Estos contenidos pueden impactar en la opinión pública, desestabilizar gobiernos y sociedades democráticas, influenciar procesos electorales, determinar la toma de decisiones, afectar a la cotización de valores, comportar riesgos financieros, dañar la imagen corporativa y la reputación de personas y organizaciones.

La Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*- , de 21 de abril de 2021, regula específicamente una previsión respecto de los *deep fakes* dentro de las obligaciones generales de transparencia de cualquier sistema, en particular en su artículo 52.3, exigiendo a los usuarios de los sistemas inteligentes que generen o permitan los mismos revelar que el contenido ha sido generado o manipulado artificialmente.

De nuevo el problema no es la tecnología, sino el uso que se hace de la misma. De hecho, la inteligencia artificial se está utilizando en la actualidad en sentido inverso, para combatir de manera muy efectiva las noticias y contenidos falsos.

4.2.10. Eliminación de puestos de trabajo

El despliegue y uso de la inteligencia artificial a nivel empresarial conlleva también el riesgo de eliminación de un gran número de puestos de trabajo[182], si bien, crea otros, conforme he analizado con profundidad en apartado 3 de este capítulo.

En este sentido, los gobiernos deberán promover planes de estabilización del empleo y de educación y formación sobre nuevas competencias y perfiles que demandará el mercado

[182] El 14 % de los empleos en los países de la OCDE son altamente automatizables y un 32 % podría enfrentar cambios sustanciales (estimación). Think Tank del PE 2020.

conforme aumente dicho uso, lo que ayudará a prevenir el desempleo a largo plazo y garantizará disponer de los perfiles y competencias cualificadas que requerirá el mercado.

La aceptación social de los sistemas dotados de inteligencia artificial constituye un gran reto por sus implicaciones en materia de empleo y en la sostenibilidad de los sistemas de previsión social.

4.2.11. Riesgos para el mercado y los consumidores

En relación con lo anterior, el uso de la inteligencia artificial también puede comportar riesgos para el mercado y la competencia, especialmente ante la acumulación de la información y su tratamiento interesado por parte de unos pocos actores que dispondrían de más información y les generaría una ventaja competitiva en perjuicio del resto.

Asimismo, los gobiernos deben considerar el riesgo de excesivo proteccionismo de determinados derechos, de regulación excesiva de la tecnología y de establecimiento de marcos obligaciones y de responsabilidad restrictivos aplicables únicamente a los sujetos incluidos en su ámbito de aplicación, que podrían colocarlos en una situación de desventaja competitiva inicial a nivel internacional.

Esta situación podría suponer una disyuntiva para empresas fabricantes o operadoras en materia de competencia ética y leal, así como una disyuntiva para gobiernos y ciudadanos de optar por el avance tecnológico o el cumplimiento regulatorio, con el riesgo consecuente de que entidades ubicadas en países terceros puedan continuar avanzando en el desarrollo y despliegue de la tecnología -frente a las empresas locales-, ante la sujeción de aquéllas a ordenamientos jurídicos más laxos y favorecedores de la innovación, lo que a su vez comporta el dilema para gobiernos, empresas y ciudadanos de renunciar a la protección y cumplimiento del marco local en aras de importar, explotar y/o usar una tecnología que, en determinados contextos podría incluso estar salvando vidas diariamente.

En este sentido, recordemos que la tecnología siempre va por delante del derecho y no es nada fácil encontrar y situarse en un contexto de equilibrio en el que se garantice la seguridad física, lógica y jurídica, la innovación y el desarrollo tecnológico.

A modo de ejemplo, pensemos en modelos de negocio como el de Google, Facebook, Whatsapp o Uber, en mi opinión, posiblemente, nunca podrían haberse desarrollado y consolidado a la velocidad en la que lo han hecho si se hubieran aplicado con rigor los marcos jurídicos vigentes a nivel local o regulado con mayor profundidad este tipo de servicios en la mayor parte de países europeos.

Por otra parte, también se plantean importantes retos y riesgos para las personas en su calidad de consumidores, algunos generales como los que están siendo objeto de análisis, otros más específicos, como las asimetrías de poder y plano de desigualdad entre la inteligencia artificial y el consumidor, especialmente ante colectivos más vulnerables y falta de conocimientos técnicos, así como prácticas de esclavización del consumidor por determinadas plataformas, ajustes algorítmicos de precios y discriminación de precios.

4.2.12. Riesgos para la seguridad física y moral de las personas

Los riesgos para la seguridad de las personas, para su integridad física y moral son especialmente evidentes en aquellos dispositivos o aplicaciones vinculadas al control de su salud, asistencia vital o que se hallen en contacto físico o integradas en el cuerpo humano, con especial significación de los interfaces neuronales controlados por sistemas dotados de inteligencia artificial, prótesis robóticas, neuroestimuladores, etc.

Un mal diseño, integración, interacción, uso, *hackeo o crackeo* del sistema puede comportar directamente daños personales e incluso la muerte.

Pueden no tener un impacto tan grave y directo en aquellos sistemas concebidos para la asistencia, cuidado o compañía de personas especialmente vulnerables, si bien, sus reacciones o conductas inadecuadas pueden provocar desequilibrios y daños emocionales en las personas.

A modo de ejemplo, el robot japonés "Paro"[183]. Una foca de peluche que ofrece asistencia a personas con necesidades especiales, como personas con autismo, síndrome de Down o edad avanzada, y que se está usando para terapias afectivas y en casos de depresión, ansiedad o demencia, ayudando a reducir la ingestión de fármacos. Una simple inacción o conducta no adecuada puede comportar graves desequilibrios y conductas reactivas graves por parte de cualquier persona de especial vulnerabilidad y sensibilidad adscrita a estos colectivos.

4.2.13. Riesgos para la información, bienes e instalaciones

La aplicación de la inteligencia artificial también implica riesgos para la información, bienes e instalaciones, especialmente infraestructuras críticas, derivados de errores en su diseño, integración, uso, *hackeo o crackeo*.

La inteligencia artificial puede ser utilizada para detectar vulnerabilidades y romper arquitecturas de seguridad en redes y sistemas, para introducirse en los mismos, para secuestrar, cifrar o sustraer información, para alterar sistemas o incluso para manipular algoritmos de inteligencia artificial ya en uso y confiables. De estas cuestiones hablaré con más detalle cuando aborde los aspectos de ciberseguridad.

Sin embargo, el uso de la inteligencia artificial no debe concebirse como una mera amenaza o debilidad, sino que puede constituir a su vez una fortaleza u oportunidad, en la medida que constituye una herramienta muy eficaz para la seguridad de todo ello, tanto con fines preventivos y detectivos, como posteriores de bloqueo, contención, gestión, minoración y recuperación ante un incidente e impacto.

[183] MONJE, C. (2018). "Salud El poder terapéutico de un robot-peluche". Publicado en *El País* el 11.11.2018. Disponible en: https://retina.elpais.com/retina/2018/11/09/tendencias/1541790426_183947.html. Consultado el 21.02.2021.

La inteligencia artificial es y será utilizada cada vez más para luchar contra el *software* malicioso o *malware* y para la protección de los sistemas, especialmente su activo más importante, esto es, la información.

Del mismo, pueden verse afectados bienes e instalaciones, especialmente las críticas, por el uso malicioso de la inteligencia artificial. Estas cuestiones las trataré con mayor detalle cuando abordé las cuestiones de ciberseguridad.

4.2.14. Riesgos asociados a su uso con fines de defensa

Los riesgos del uso de la inteligencia artificial con finalidades de defensa para personas, bienes e instalaciones son obvios, especialmente ante errores en su diseño, desarrollo, funcionamiento y/o control, o su manipulación o uso posterior indebido.

El uso militar entraña los mayores riesgos ante el bien jurídico más preciado: La vida. En este sentido, hay abierto un debate internacional sobre si se debe permitir que la decisión de quitar una vida humana pueda ser delegada en un sistema armamentístico autónomo.

La inteligencia artificial tiene numerosas aplicaciones en el ámbito militar[184], especialmente para la coordinación y facilitación de maniobras y para salvar vidas en caso de accidentes.

La inteligencia artificial supone ventajas en un conflicto armado[185], entre otras, la reducción de bajas militares, multiplicación de la fuerza empleada, pero con posibilidad de reducir la letalidad mediante la orientación a la inmovilización o desarme, ampliación del campo de batalla, menor tiempo de reacción, un aumento de la precisión, reducción de costes, acciones ajenas al miedo, a las emociones, a los odios o sesgos.

[184] GÓMEZ ÁGREDA, A.; MOLINO, J. Y OTROS (2019). *Usos militares de la inteligencia artificial, la automatización y la robótica*. Instituto Español de Estudios Estratégicos (ieee.es), 2019.
[185] KAHN, PAUL W., "Imagining Warfare". *The European Journal of International Law*. Vol. 24 no. 1, Pp. 199-226, Published by Oxford University Press (on behalf of EJIL Ltd) 2013.

También puede suponer una mejora el rendimiento de los ejércitos proporcionando sistemas aliados e incluso podría ayudar a generar robots-soldados o "supersoldados" -como ya están haciendo distintos países-, sustituyendo al ser humano en determinadas misiones para preservar la vida humana, como incursiones en terrenos minados.

Sobre la inteligencia artificial y su aplicación en la defensa en España, me permito citar el *Documento de Seguridad y Defensa del Instituto Español de Estudios Estratégicos y el Ministerio de Defensa*, publicado en 2019 bajo el título *inteligencia artificial aplicada a la defensa*[186].

Sobre su aplicación a nivel internacional, me permito igualmente citar el reciente informe elaborado por el *Congressional Research Service* de los EE.UU. bajo el título *Artificial Intelligence and Nacional Security*[187], en el que se refleja la estrategia de EE.UU., China y Rusia en este campo, y se significan sus oportunidades y ventajas como, por ejemplo, facilitar las operaciones autónomas, conducir a una toma de decisiones militares más informada y aumentar la velocidad y la escala de la acción militar, pero también sus retos, como su imprevisibilidad o vulnerabilidad ante su posible manipulación. El documento plantea algunos de los problemas potenciales que la inteligencia artificial supone y que deberá abordar el Congreso estadounidense, especialmente en materia ética y regulatoria.

Para los expertos en el ámbito militar, los denominados sistemas armamentísticos autónomos letales -SALA o LAWS por las siglas en inglés de *Lethal Autonomous Weapon Systems*-, están creando la denominada *tercera revolución bélica*, tras la pólvora y las armas nucleares. De nuevo, el cine ya nos ha planteado algunos de los posibles conflictos que plantean, desde la absoluta ficción como "*Soldado Universal*" (1992) a la realidad actual como "*Espías desde el cielo*" (2015).

En este último caso, la decisión final sobre la actuación de un dron armamentístico (lanzamiento de un misil para la eliminación de un objetivo) recaía en el ser humano, que

---

[186] "Inteligencia artificial aplicada a la defensa*".* En *Documento de Seguridad y Defensa* 79. Instituto Español de Estudios Estratégicos. Departamento de Seguridad Nacional (DSN). Ministerio de Defensa. Marzo 2019.

[187] *Artificial Intelligence and Nacional Security*. Congressional Research Service (CRS). EE.UU. 10.11.2020. Disponible en: https://crsreports.congress.gov/

es quién está en mejor posición para valorar el contexto y el riesgo real desde un punto de vista global, ético, jurídico y de seguridad.

Sin embargo, si el sistema hubiera estado dotado de autonomía y hubiera tomado la decisión de actuar en el contexto planteado, lo hubiera hecho en base a criterios objetivos, estadísticos y matemáticos, nunca humanos ni morales, al objeto de obtener el resultado predefinido bajo la mejor opción en términos de probabilidad de consecución del objetivo y, supuestamente, de bajas y otros daños colaterales.

¿Quién decide quién vive y quién muere? ¿Quién decide si los daños colaterales están justificados? ¿Quién define los parámetros éticos en estos contextos? ¿Quién asume la responsabilidad por los daños colaterales? Si estamos ante sistemas inteligentes, ¿Podría aplicarse el aprendizaje automático supervisado para enseñarle a la máquina cómo valorar los daños colaterales? La respuesta de los expertos a esta pregunta, en la actualidad es negativa[188], especialmente ante el numero casi infinito de situaciones posibles que puede motivar la existencia de situaciones no previstas y la inexistencia de suficientes datos de todas las posibles situaciones.

La repuesta a estas preguntas las podemos encontrar también a nivel europeo en el informe del Grupo de Expertos publicado el 8 de abril de 2019 bajo el título *Directrices éticas para una IA fiable*[189], que abordaré más adelante.

Conforme a dichas directrices compartidas por la Comisión Europea, la necesaria supervisión de los sistemas debe lograrse mediante la gobernanza de los sistemas inteligentes, en particular, sustentando la misma en un enfoque de supervisión humana o también denominada "*Human on the loop*" (HOTL), de participación humana o "*Human in the loop*" (HITL) y/o de control humano o "*Human in command*" (HIC).

Los sistemas informáticos pueden contener errores y defectos en su diseño, funcionamiento y gobernanza[190], y aquellos que integran inteligencia artificial, además,

[188] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Op. cit. Pos. 1071.
[189] Recuperado de: https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1. Consultado el 08.02.2021
[190] Como ejemplo de ausencia de seguridad y gobernanza adecuada, significar algunos de los escándalos que han afectado al nuevo paradigma que supone el *Blockchain.* El ejemplo sobre el sustento mi reflexión

pueden errar y equivocarse tanto en la captación de información de entrada, en su procesamiento, así como en la definición y ejecución de la conducta consecuente conforme al objetivo y aplicación para los que fueron concebidos. Podemos encontrarnos con falsos positivos y con falsos negativos. Cuando hablamos de que los riesgos puedan afectar a las personas, su integridad física y moral e incluso a su vida, no considero viable ni compatible una autonomía real alejada de todo control y supervisión humana, especialmente porque en estos casos no hay posibilidad de revertir la acción y su resultado, ante daños físicos que pueden incluir la muerte.

El error debe ser esperado y previsto, en la medida que se haya asociado al ser humano e inevitablemente a los entes artificiales creados por el mismo, por lo que el uso de la inteligencia artificial, a mi juicio, requiere la intervención humana para determinadas decisiones y acciones. El comportamiento no esperado debe ser igualmente previsto. La historia de la humanidad nos ha dejado múltiples ejemplos para la reflexión.

---

tiene su origen en la creación de una DAO u *Organización Autónoma Descentralizada* por un grupo de desarrolladores mediante Ethereum. Estos "entes" permiten organizar y hacer funcionar organizaciones de manera revolucionaria mediante el uso de los denominados "*smart contracts*" y la tecnología *Blockchain* para ofrecer transparencia, inmutabilidad, autonomía y seguridad a las mismas, además del anonimato. La organización creada pasa a ser controlada en su totalidad por algoritmos computacionales, pero no por ninguna inteligencia artificial. Todo lo que ocurre o puede ocurrir está definido en el código o programa informático, es decir, en el *smart contract*, que determina las reglas e instrucciones informáticas de cómo deben relacionarse las partes implicadas en la DAO y de todas las cosas que van a ocurrir automática y automatizadamente, es decir, un programa informático de acceso público que tiene instruccionado todo lo que se puede hacer o dejar de hacer. Los *smart contracts* definen todo lo que puede pasar en ese ecosistema, son la ley, transparentes, públicos e inmutables desde el momento que sean incorporados en la red informática descentralizada (*Blockchain*), por lo que todo el mundo puede revisar su funcionamiento y las reglas que se han programado dentro, teniendo la seguridad de que no podrán ser modificadas en el futuro. De este modo, las DAO permiten a una persona intercambiar sus fondos con cualquier otra en cualquier lugar del mundo, bajo la forma de inversión, donación, contribución, patrocinio, préstamo u otros, sin ningún intermediario. Una omisión, error o agujero de seguridad en su código inicial puede permitir su explotación por un ciberdelincuente y, posiblemente, no puede ser subsanado hasta que exista un consenso de la mayoría que integra esa red, lo que ofrece un abanico temporal amplio para la explotación de la vulnerabilidad detectada. Pero a veces no es necesaria la explotación de una vulnerabilidad en el código por un tercero, sino que es suficiente que un miembro simplemente haga algo que el propio código informático o algoritmo permite. Esto es lo que sucedió hace tres años, en una DAO creada por un grupo de desarrolladores y sustentada en un *smart contract* desplegado en la red, permitiendo que cualquiera pudiera vincular su moneda virtual "*Ether*" al mismo, lo que hicieron 11.000 personas anónimas de todo el mundo que metieron allí su "dinero" con la intención de utilizarlo como un ahorro o una inversión a largo plazo. El conjunto de instrucciones informáticas que conformaban el *smart contract* era la norma a cumplir pero nadie se apercibió de que tenía un error en su concepción que permitía la extracción de *Ethers* sin el permiso de los demás conforme a "ley" (*smart contract*). Sin embargo, alguien anónimo sí se dio cuenta y fue retirando cantidades crecientes de criptomonedas hasta un importe, todavía no aclarado, de unos 50 millones de dólares. Posteriormente, esta persona publicó que todo lo que había hecho estaba en el código informático. Recuperado de: https://www.incibe-cert.es/alerta-temprana/bitacora-ciberseguridad/50-millones-dolares-ether-robados.

Durante la crisis de los misiles cubanos en 1962 el mundo estuvo a punto de sufrir la Tercera Guerra Mundial. Conforme recoge Tegmark[191] en su libro *Vida 3.0: Qué significa ser humano en la era de la inteligencia artificial*, el 27 de octubre de 1962, once destructores de la Marina estadounidense y un portaaviones habían arrinconado al submarino soviético B-59 muy cerca de Cuba en aguas internacionales. Sin embargo, desconocían que estaba sucediendo en el interior del submarino, dado que la temperatura sobrepasaba los 45 grados al haber dejado de funcionar el aire acondicionado y estar acabándose las baterías. Los tripulantes se estaban quedando sin aire y algunos empezaron a perder el conocimiento. No tenían contacto con Moscú y no sabían que estaba pasando fuera. El ejército estadounidense empezó a lanzar cargas de profundidad para que el submarino saliese a la superficie y abandonase la zona, según había informado a Moscú, sin que la tripulación supiese nada de ello. El problema es que los estadounidenses desconocían que el submarino tenía un misil nuclear y la autorización para lanzarlo sin permiso de Moscú y así lo decidió hacer el capitán del mismo. Afortunadamente, la decisión de lanzamiento requería la autorización de tres oficiales a bordo y uno de ellos se negó a hacerlo, lo que quizás evitó el estallido de una Tercera Guerra Mundial. ¿Qué hubiera sucedido si se hubiese tratado de un submarino autónomo controlado por inteligencia artificial y sin ninguna intervención humana?

Y no es el único ejemplo que el experto citado refleja en su libro. El 9 de septiembre de 1983, un sistema automatizado de alerta temprana soviético avisó que EE.UU. había lanzado cinco misiles nucleares hacia la Unión Soviética (URSS) y el oficial responsable, Stanislav Pedrov, disponía de unos minutos para decidir si se trataba de una falsa alarma. Se comprobó que el satélite funcionaba correctamente por lo que, según el protocolo, éste debería haber informado que se trataba de un ataque nuclear y actuado en consecuencia. Sin embargo, siguió su instinto considerando que era poco probable que se tratase de un ataque al ser sólo cinco misiles. Informó a sus superiores de que era una falsa alarma sin saber que era así. Más tarde se consideró que fue un efecto óptico de la luz solar. De nuevo, ¿qué hubiera sucedido si se hubiese tratado de un sistema de defensa autónomo

[191] TEGMARK, M. (2018). *Vida 3.0: Qué significa ser humano en la era de la inteligencia artificial*. Penguin Random House Grupo Editorial. 2018. Pp. 142-144.

controlado por inteligencia artificial y sin ninguna intervención humana, el cual hubiera seguido el protocolo informatizado en base a los inputs?

Posteriormente, el 3 de julio de 1988 durante la guerra de Irán e Irak, el USS Vincennes -un crucero dotado de misiles guiados y apodado "Robocruiser" por su sistema "Aegis" que detecta automáticamente, rastrea y neutraliza amenazas como misiles antibuque y aviones-, informó que se aproximaba un avión y su capitán consideró que se trataba de un caza iraní, por lo que dio permiso al sistema "Aegis" para disparar, derribando un avión civil de pasajeros con 290 personas a bordo, que fallecieron. Posteriormente se comprobaron errores en el funcionamiento del sistema, entre otros, la asignación de maniobras de vuelo indicadoras de la supuesta amenaza, pero correspondientes a otro avión distinto y además estadounidense, que motivaron información de salida incorrecta, lo que motivó una decisión final humana errónea.

Y me permito adicionar algún ejemplo más a los citados por este experto.

Durante los Juegos Olímpicos de Múnich de 1972 se produjo un atentado terrorista reivindicado por el grupo denominado *Septiembre Negro* vinculado a la Organización para la Liberación de Palestina -OLP- que se materializó en el secuestro de varios deportistas israelíes en la propia Villa Olímpica, el intento de fuga de los terroristas desde el aeropuerto con un resultado desastroso de 11 rehenes, 1 policía y 5 terroristas muertos y 3 terroristas detenidos. Sorprendentemente, se tomó la decisión de continuar con los juegos y tras aplazar la ceremonia de clausura un día llevándola al 11 de septiembre, ese mismo día, el ministro de defensa recibió una llamada informándole del secuestro de un avión comercial finlandés con 150 pasajeros a bordo en dirección Múnich bajo la sospecha de atentar contra la ceremonia de clausura con 70.000 personas dentro del estadio donde se estaba celebrando la misma.

Se trasladó la información al locutor del estadio que optó por guardar silencio para no provocar una situación de pánico descontrolado entre la multitud. En ese instante, el ministro de defensa ordenó que dos cazas asegurarán el espacio aéreo y con autorización para ordenar derribar el avión secuestrado. Ante la proximidad del avión, el ministro tuvo claro que tenía que tomar la decisión de derribar el avión sin más de demora. Afortunadamente, el avión consiguió en ese instante comunicarse con el aeropuerto para

informar de un problema electrónico y en las comunicaciones, no existiendo ningún secuestro. Si finalmente el problema no hubiera sido resuelto, los cazas hubieran derribado el avión con 150 pasajeros a bordo. Es un ejemplo que evidencia la relevancia de la comunicación y conectividad en determinados contextos. Pensemos que los sistemas inteligentes dependen de las mismas. ¿Cuáles podrían ser la consecuencia de simples fallos en las mismas en relación con sistemas de inteligencia artificial de cualquier tipo y especialmente de aquéllos de los que dependa en momentos determinados salvar o quitar vidas? Ni los servicios de conexión ni las comunicaciones son infalibles, ni los sistemas informáticos ni de inteligencia artificial que dependen de aquéllos son infalibles y están libres de errores.

Actualmente hay más de 130 sistemas de armas autónomas letales -que sepamos- y más de 80 países están desarrollando soluciones militares basadas en la inteligencia artificial[192]. La mayor parte, no obstante, están orientadas a la toma de decisiones, con participación humana y manteniendo el control.

En este sentido, son muchas las voces que se han alzado a nivel internacional para impedir la creación de sistemas de inteligencia artificial con finalidades bélicas y especialmente si no son dependientes de un ser humano.

El *Future of Life Institute*[193] publicó en 2015 una carta abierta suscrita por más de 16.000 personas, entre otras, Elon Musk -empresario y fundador de SpaceX y Tesla-, Stephen Hawking o el filósofo estadounidense Noam Chomsky. En la misma se alerta sobre las amenazas que los sistemas de combate basados en inteligencia artificial suponen para la población civil, sobre el riesgo de una carrera armamentística y sobre el peligro de sus consecuencias no deseadas para la humanidad.

Dos años más tarde, en agosto de 2017, el precitado Elon Musk y un centenar de diseñadores de sistemas de robótica y de inteligencia artificial enviaron a la ONU una

[192] BENJAMINS, R. Y SALAZAR, I. (2020). *El mito del algoritmo*. Op.cit. Pos. 1164.
[193] *The Future of Life Institute* es un instituto de investigación sin fines de lucro y con objetivos divulgativos que trabaja para mitigar los riesgos existenciales a los que se enfrenta la humanidad, especialmente el riesgo existencial de la inteligencia artificial avanzada.

petición[194] solicitando que se prohíba totalmente el desarrollo y las pruebas de armas ofensivas autónomas[195].

Según estos expertos, la creación de ejércitos de robots capaces de realizar hostilidades de forma autónoma conllevará el surgimiento de sentimientos de poder absoluto e impunidad para quienes dispongan de los mismos. Asimismo, afirman la ausencia de actitudes morales, sentimientos y emociones en sus acciones propias del ser humano y que surgen de la observación directa del sufrimiento humano y que, en el caso de estos sistemas no existiría. En su conferencia de presentación advirtieron que "Cuando se abra la caja de Pandora, será difícil cerrarla".

Y lo cierto es que este es el escenario, debate y conflicto ético consecuente en el que ya nos encontramos, en el que la decisión de matar a un humano puede recaer directamente en el sistema o, a través del mismo, en la persona que, sentada en su silla y a 12.000 kilómetros de la situación de conflicto, visualiza y recibe a través de una pantalla la imagen en tiempo real que un sistema le suministra, con la supuesta identificación por el mismo de cinco posibles objetivos en el interior de una vivienda, mostrados como un conjunto volumétrico de pixeles rojos que se mueven mientras el operador debe decidir si eliminar o abortar. Desgraciadamente no es algo que el cine nos haya ya mostrado, sino que es real. La guerra no es ética por naturaleza.

Los sistemas no piensan ni sienten, sólo procesan, carecen de las capacidades y valores propios del ser humano, por lo que no pueden hacer juicios de valor humano en sus actuaciones, y sus acciones, en caso de ser permitidas en su construcción, no pueden considerarse ni legítimas ni moralmente justificables.

Los sistemas no tienen la capacidad de entender la naturaleza de sus acciones desde un punto de vista humano, pueden llegar a imitar emociones pero no comprenden, no sienten, no "viven" las situaciones, sólo razonan lógicamente en base a los datos e instrucciones incorporadas en su diseño y desarrollo, y conforme a los posteriores datos extraídos del

[194] Recuperado de https://futureoflife.org/autonomous-weapons-open-letter-2017/. Consultado el 29.01.2021.
[195] Recuperado de https://es.unesco.org/courier/2018-3/amenaza-robots-asesinos. Consultado el 29.01.2021.

contexto en el que operen y generados por los mismos, en función de la propia autonomía, libertad e independencia que el propio ser humano le haya otorgado.

Todo ello provoca retos adicionales, como la posibilidad de desviación de los objetivos, patrones y criterios éticos predefinidos por el ser humano y de servir a los intereses generales y al bien común, incluso de dañar y destruir a al propio ser humano.

La debatida y pretendida autonomía nunca debería ser real.

A modo de ejemplo, en 2016 se utilizó un robot no autónomo -MARCbot-[196] diseñado para las operaciones militares de EE.UU. en Irak y Afganistán, para neutralizar a un francotirador en Dallas (EE.UU.), acabando con su vida con la detonación de un explosivo que integraba.

Los sistemas de inteligencia artificial deberían estar sujetos en todo momento al control y supervisión humana, tanto en su concepción como en su posterior despliegue y aplicación, es decir, hallarse sometida a controles previos, coetáneos y posteriores durante todo su ciclo de vida, y tanto reactivos como proactivos, y de carácter preventivo, detectivo, reactivo, correctivo y evolutivo.

Y dentro de estos controles, las acciones de estos sistemas deberían ser revocables y los mismos deberían susceptibles de reprogramación, suspensión, apagado o destrucción, si bien, en función de su diseño, programación, capacidades y libertades conferidas, podría plantearse el riesgo adicional de su autoprotección o autodeterminación, de modo que fuese imposible llevar a cabo las acciones precitadas sobre los mismos, impedir por sí mismo su autodestrucción o su reprogramación, suspensión, apagado o destrucción por un tercero.

Todo ello comporta retos asociados, especialmente en el caso de que el sistema pueda alcanzar un nivel de autoprotección o autodeterminación irreversible.

[196] "Cómo funciona el MARCbot, el robot con el que la policía mató al francotirador de Dallas". Redacción. Publicado el 9 de junio de 2016 por la *BBC*. Disponible en: https://www.bbc.com/mundo/noticias-36751451. Consultado el 01.03.2021.

Algunos autores como Martínez Quirante[197] y Núñez Zorrilla[198], en el contexto armamentístico, muestran como la "metacognición" del sistema podría llevar a poner en peligro a toda la sociedad y conllevar el fin de la humanidad. De nuevo, estas reflexiones nos evocan futuros apocalípticos que ya el cine nos ha mostrado en numerosas ocasiones en películas como *Terminator* o *2001: Una odisea del espacio*.

Por otra parte, los trabajadores de algunas de las grandes tecnológicas se han alzado en contra del desarrollo de armas autónomas letales. Un ejemplo de ello fueron los miles de trabajadores de Google que se posicionaron en este sentido y obligó a la misma a retirarse del proyecto "Maven"[199] del Departamento de Defensa de los EE.UU.

Del mismo modo, distintas iniciativas internacionales promovidas por organizaciones no gubernamentales están promoviendo campañas, entre otras, la *Campaña contra los Robots de Combate*[200] promovida por una coalición formada por 172 organizaciones no gubernamentales en 65 países que está trabajando para prohibir las armas totalmente autónomas.

Por lo que se refiere al posicionamiento de algunos países sobre la dotación de autonomía a estos sistemas, los posicionamientos son dispares incluso en el marco europeo.

Inicialmente, la UE se posicionó al respecto mediante una Resolución del Parlamento Europeo de 12 de septiembre de 2018, para limitar el uso, la producción y el desarrollo de cualquier arma capaz de matar sin necesidad de control humano[201]. Los europarlamentarios concluyeron que "Las armas sin un control humano significativo sobre la selección y el ataque de objetivos deberían prohibirse antes de que sea demasiado tarde". La resolución fue aprobada por 566 votos contra 47 y 73 abstenciones.

---

[197] MARTÍNEZ QUIRANTE, R. (2018). *Inteligencia artificial y armas letales autónomas. Un nuevo reto para Naciones Unidas*. Tres Ensayos. Asturias 2018. Pp. 119, 120, 129 y 130.
[198] NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Reus Editorial, Madrid 2019, P. 18.
[199] HARWELL, D. (2018). "Google to drop Pentagon AI contract after employee objections to the 'business of war'". Publicado en *The Washington Post*. 01.06.2018. Recuperado de: https://www.washingtonpost.com/news/the-switch/wp/2018/06/01/google-to-drop-pentagon-ai-contract-after-employees-called-it-the-business-of-war/. Consultado el 18.02.2021.
[200] Recuperado de: https://www.stopkillerrobots.org/. Consultado el 12.02.2021
[201] Recuperado de: https://www.europarl.europa.eu/news/en/press-room/20180906IPR12123/european-parliament-speaks-out-against-killer-robots. Consultado el 12.02.2021

Posteriormente, en la *Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre inteligencia artificial: Cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal*[202], considera que la inteligencia artificial utilizada en un contexto militar y civil debe estar sujeta a un control humano apropiado, de modo que el ser humano tenga en todo momento los medios para corregir su curso, detenerla o desactivarla en caso de comportamiento imprevisto, intervención accidental, ciberataque o interferencia de terceros con tecnología basada en la inteligencia artificial, o cuando terceros adquieran dicha tecnología.

Asimismo, introduce un aspecto de nuevo, muy relevante a los efectos de imputación de la responsabilidad, considerando que "la toma de decisiones autónoma no debe eximir a los seres humanos de su responsabilidad y que, siempre, las personas deben ser responsables en última instancia de los procesos de toma de decisiones de modo que pueda identificarse al ser humano responsable de una decisión".

La precitada Resolución finaliza, de un lado, recordando que el Parlamento ha pedido la elaboración y adopción urgentes de una posición común sobre los sistemas armamentísticos autónomos letales -SAAL- para prevenir el desarrollo, la producción y la utilización de este tipo de sistemas capaces de realizar ataques sin un control humano significativo, así como el inicio de negociaciones eficaces para su prohibición y, de otro, recordando la necesidad de adoptar una estrategia a escala de la Unión contra los SAAL y de prohibir los denominados "*killer robots"* (robots asesinos) pidiendo un posicionamiento común y propone evitar cualquier confusión entre una persona y un robot, de modo que se prohíba la antropomorfización de los SAAL, entre otras cuestiones.

La mayoría de los países que participan en la *Convención sobre Ciertas Armas Convencionales* -CCW por sus siglas en inglés-[203] coinciden en que es necesaria una regulación que asegure la presencia del ser humano al mando, entre los que se encuentra

[202] Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre inteligencia artificial: cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal (2020/2013(INI)).

[203] Recuperado de: https://www.unog.ch/80256EE600585943/(httpPages)/4F0DEF093B4860B4C1257180004B1B30?OpenDocument. Consultado el 12.02.2021.

España, si bien, no se encuentra en el grupo de los países que piden directamente la prohibición directa de armas letales autónomas como Pakistán, Ecuador, Cuba, Bolivia, Brasil, Irak, Palestina e incluso China, uno de los principales fabricantes de armas letales autónomas y que plantea prohibir el uso de armas totalmente autónomas, pero no su desarrollo o producción.

Entre los países que se han opuesto a la prohibición de las armas letales autónomas figuran países como España, Israel, Francia, Turquía, Reino Unido, Rusia o Estados Unidos.

La realidad actual es que desde hace años se están fabricando robots y vehículos dotados de inteligencia artificial con fines de "defensa" pero también con capacidad para matar, por ejemplo, el robot-centinela "SGR-A1"[204] fabricado por Samsung Techwin para reemplazar a las tropas que vigilan las zonas desmilitarizadas en la frontera entre las dos Coreas, el cual tiene una capacidad de disparo a más de tres kilómetros de distancia, pero sujeto a la última orden de un supervisor humano.

Del mismo modo, citar los vehículos no tripulados o *Unmanned Surface Vehicle* -USV por sus siglas en inglés- dotados de armas operadas desde centros de mando remotos, como el buque Sea Hunter estadounidense o el USV fabricado por la compañía francesa Thales para su empleo contra las minas en conflictos armados.

Por el camino, el Departamento de Defensa estadounidense presentaba públicamente, el pasado 26 de mayo de 2021, su memorándum y compromiso para una inteligencia artificial responsable.

Las grandes potencias están mostrando públicamente la carrera armamentística "inteligente" en las que se hayan inmersas. EE.UU. se halla inmersa en una carrera armamentística sustentada en la inteligencia artificial, en la que participan países como China, Rusia o Corea del Norte.

[204] PRIGG. M. (2014). *Who goes there? Samsung unveils robot sentry that can kill from two miles away*. Publicado en el *Daily Mail*, 16.09.2014. Recuperado de: https://www.dailymail.co.uk/sciencetech/article-2756847/Who-goes-Samsung-reveals-robot-sentry-set-eye-North-Korea.html. Consultado el 18.02.2021

EE.UU. presentó recientemente el sistema de inteligencia artificial "Skyborg" para pilotar de manera autónoma aviones de combate, habiendo realizado ya los primeros vuelos de prueba en cazas de combate totalmente autónomos. Simultáneamente, Rusia difundía el inicio de la producción de robots de combate autónomos basados en inteligencia artificial.

Asimismo, existen otras aplicaciones inicialmente concebidas para defensa y seguridad, pero que están empezando a ser desplegadas en el ámbito civil y comercial, con sus ventajas y riesgos, por ejemplo la tecnología "VibraImage" para el escaneado de expresiones faciales utilizada durante los XXII Juegos Olímpicos de Invierno en Sochi (Rusia) en 2014, y que fue empleada para detectar personas asistentes que pudieren mostrar facialmente un estado mental agitado y que pudiera ser una amenaza inminente.

Dicha tecnología está siendo ya utilizada en el sector del comercio minorista para identificar el estado emocional de sus potenciales clientes. Es más, distintas redes sociales permiten predecir el comportamiento de sus usuarios y definir su perfil y personalidad mediante el estudio con algoritmos inteligentes del habla, acciones o expresiones.

La *Comisión de Seguridad Nacional sobre inteligencia artificial* constituida en EE.UU. con la finalidad de valorar y asesorar sobre estos aspectos e integrada por un grupo de expertos en la materia, concluyó su borrador de informe en enero de 2021 para su presentación al Congreso[205], en el que determinó que EE.UU. no debería aceptar prohibir el uso o desarrollo de armas autónomas desarrolladas por *software* de inteligencia artificial y lo considera además una "cuestión moral", en la medida que cometen menos errores que los humanos, lo que reduciría las bajas causadas por la identificación errónea de objetivos.

Por el camino, en fechas próximas al cierre de esta investigación se iniciaron los trámites para la prórroga por parte de Estados Unidos y Rusia del Tratado de Reducción de Armas Estratégicas (Start III) que expiraba el 5 de febrero de 2021, lo que sin duda evidencia la posible postura de la nueva administración estadounidense tras las recientes elecciones.

[205] Reuters. Publicado en *The Guardian*. 26.01.2021. Disponible en https://www.theguardian.com/science/2021/jan/26/us-has-moral-imperative-to-develop-ai-weapons-says-panel

La necesidad de un marco regulador y de compromisos internacionales resulta evidente, así como de definir los marcos de responsabilidad en contextos tan complejos a nivel técnico y ético.

4.2.15. Uso malintencionado y delictivo

El uso malintencionado e incluso delictivo de la inteligencia artificial constituye otro de los riesgos a considerar.

Anteriormente he reflejado un ejemplo en relación con el *spear phishing* basado en inteligencia artificial, sin perjuicio de los supuestos que igualmente abordaré el analizar la responsabilidad penal relacionada con la inteligencia artificial.

Los sistemas de inteligencia artificial pueden ser el medio o instrumento para la comisión de un delito u objeto de los mismos como analizaré en el capítulo precitado, especialmente ante delitos de *hacking, cracking, ransomware* y otros relacionados.

La inteligencia artificial puede ser utilizada para afectar e incluso paralizar sistemas críticos de empresas, ciudades, sociedades y gobiernos. El cine, de nuevo, nos ha mostrado algunos ejemplos, entre otros, "*La Jungla 4.0*" (2007) que, desgraciadamente, no se alejan demasiado de la realidad, como ocurrió en New Orleans[206], en EE.UU.

Estas conductas, calificables o no como "ciberdelitos" en función del ordenamiento jurídico aplicable, no dejan de crecer, especialmente el *ransomware*, que permite el secuestro o sustracción de datos, y no respeta ni a nada ni a nadie, ni tan siquiera a los hospitales de campaña en plena pandemia mundial[207].

[206] WINDER D. (2019). "New Orleans Declares State Of Emergency Following Cyber Attack". Publicado en *Forbes* el 14.12.2019. Recuperado de: https://www.forbes.com/sites/daveywinder/2019/12/14/new-orleans-declares-state-of-emergency-following-cyber-attack/?sh=65b824246a05. Consultado el 17.02.2021.

[207] FERNÁNDEZ, J. (2020). "Los ciberataques amenazan con colapsar los hospitales: 'Sería terrorífico'", Publicado en *Expansión* el 24.10.2020. Recuperado de: https://www.expansion.com/economia-digital/companias/2020/10/24/5f915917e5fdea64298b45e1.html. Consultado el 17.02.2021.

A modo de ejemplo, un ataque de esta naturaleza a nivel internacional en 2017, paralizó al menos a 16 hospitales y centros de salud de Reino Unido[208] de modo que los profesionales sanitarios no podían acceder a los sistemas que almacenaban la información de sus pacientes. Recientes ataques en 2020[209] afectaron a distintas infraestructuras hospitalarias y han causado incluso muertes. En la fecha de cierre de esta investigación, el sistema de salud irlandés ha sido objeto de otro ataque[210].

Las organizaciones criminales no sólo cifran los datos de las personas y organizaciones afectadas, sino que amenazan con publicar información confidencial si no se avienen a pagar un rescate.

Estos ataques han aumentado exponencialmente durante el último año y que algunas firmas especializadas en ciberseguridad[211] cifran en un 160% y la tendencia es orientarse hacia el *Ransomware como Servicio* o RaaS por sus siglas en inglés -Ransomware as a Service-, cada vez más perfeccionado y profesional, y el *Ransomware of Things*, lo que significaría no sólo bloquear el dispositivo conectado, desde un *smartwach* hasta un vehículo autónomo, sino incluso tomar el control del mismo mediante la inserción de un *malware* con este propósito.

De las iniciativas para luchar contra esta lacra, destacar el proyecto *No More Ransom*[212] avalado por la Europol, en el que participan también la Policía Nacional Holandesa y las compañías de ciberseguridad McAfee y Kaspersky, que integra soluciones de descifrado para más de 140 familias de *ransomware*.

---

[208] GUIMÓN, P. (2017). "Un ciberataque paraliza 16 hospitales de Reino Unido y les exige dinero". Publicado en *El País* el 12.05.2017. Disponible en: https://elpais.com/tecnologia/2017/05/12/actualidad/1494602389_458942.html. Consultado el 16.02.2021.

[209] CARBAJOSA, A. Y PÉREZ, J. (2020). "Ciberataque a un hospital alemán en tiempos de pandemia". Publicado en *El País* el 04.10.2020. Recuperado de: https://elpais.com/internacional/2020-10-03/ciberataque-a-un-hospital-aleman-en-tiempos-de-pandemia.html. Consultado el 16.02.2021.

[210] DE MIGUEL, R. (2021). "Un ciberataque obliga a Irlanda a cerrar el sistema informático de la sanidad pública". Publicado en El País el 14.05.2021. Recuperado de: https://elpais.com/internacional/2021-05-14/un-ataque-cibernetico-en-irlanda-obliga-a-cerrar-el-sistema-informatico-de-la-sanidad-publica.html. Consultado el 30.05.2021.

[211] Check Point. Recuperado de: https://research.checkpoint.com/. Consultado el 17.02.2021.

[212] Recuperado de: https://www.nomoreransom.org/. Consultado el 16.02.2021

4.2.16. Uso ilegítimo o inadecuado

No toda conducta lesiva, inadecuada o ilegítima es ilícita, ni toda conducta ilícita es delictiva. En relación con las conductas citadas en el apartado precedente y en otros anteriores también relacionados con estos aspectos, significar que el uso ilegítimo, inadecuado o "mal uso" de la inteligencia artificial constituye otro de sus principales retos, y no sólo por parte de personas o empresas, sino especialmente por parte de gobiernos, supuestamente los menos democráticos y autoridades.

A modo de ejemplo, significar el debate que ha generado la proporcionalidad de la aplicación de sistemas inteligentes en el mundo oriental, por ejemplo, el reconocimiento facial masivo en China para controlar y asignar puntuaciones a sus ciudadanos por su buen comportamiento, lo que chocaría frontalmente a nivel cultural y de protección de los derechos fundamentales en el mundo occidental. Incluso la utilización de reconocimiento facial masivo exigible como exigencia obligatoria a todo adquirente de un terminal móvil.

En la ciudad de Shenzen se están tecnologías de reconocimiento facial utilizando de reconocimiento facial para identificar a los peatones y, en caso de infracción, imponer una sanción y notificarla instantáneamente.

Estos sistemas han sido considerados en la nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, conforme analizaré en el capítulo IV, como de riesgo inadmisible y, en consecuencia, prohibidos, tal y como expresamente regula en su artículo 5.

La cuestión es que determinados usos gubernamentales pueden considerarse ilegítimos o inadecuados en función de quién los valore y según en qué marco jurídico, ético y cultural se produzca.

En relación con todo ello nos encontramos con el perfilado masivo o *profiling* a través de sistemas inteligentes y el Big data, que puede romper el anonimato en sistemas basados en el tratamiento de datos no personales, así como invadir las esferas más profundas de

la persona a partir de sus datos personales, agregados o no, con datos no personales. Por ejemplo, saber si una persona está embarazada en función de sus compras.

También nos encontramos estos riesgos en relación con las noticias y contenidos falsos -*fake news* y *deep fakes*-, a los que he hecho referencia anteriormente.

Los *riesgos de desinformación*, entendida como la información falsa, inexacta o engañosa diseñada, presentada y promocionada para causar "intencionadamente" un daño público o con fines lucrativos, no dejan de aumentar, con la correlativa afectación, como he expuesto anteriormente, de la libertad de expresión, el pluralismo de los medios de comunicación y la democracia.

Algunos ejemplos de ello fueron las elecciones presidenciales de 2016 en EE.UU. o el referéndum sobre la salida de la UE del Reino Unido -*Brexit*-. Todo ello ha motivado distintas iniciativas a nivel internacional, entre otras, el informe del Parlamento Europeo *Regulating disinformation with artificial Intelligence*[213] de marzo de 2019.

Este informe analiza la posición de distintos expertos, informes y distintas iniciativas políticas y tecnológicas precedentes y muestra su cautela a la hora de utilizar la inteligencia artificial para abordar estas amenazas, especialmente ante la precisión limitada de estas herramientas y la necesidad de probar la ilegalidad de la desinformación antes de considerar adecuado el filtrado o bloqueo que podría acabar una censura por parte de la inteligencia artificial, y que no debería quedar en manos de prestadores de servicios de la Sociedad de la Información o prestadores en línea.

La inteligencia artificial puede ser un mecanismo eficaz para luchar contra determinados tipos de contenidos en línea, pero no todos, y es necesaria mucha cautela para evitar una afectación de la libertad de expresión y el pluralismo de los medios.

En este sentido, el informe precitado plantea distintas opciones orientadas a la futura regulación de la inteligencia artificial para combatir la desinformación y, de nuevo, considero que la solución debe ser híbrida, esto es, mediante la promoción de la

[213] MARSDEN, C. ET ALT. *Regulating disinformation with artificial Intelligence*. EPRS European Parliamentary Research Service. Marzo 2019.

autorregulación, la regulación y las normas técnicas de apoyo, con la colaboración de los intermediarios en la línea del Reglamento (UE) 2021/784 del Parlamento Europeo y del Consejo, de 29 de abril de 2021, sobre la lucha contra la difusión de contenidos terroristas en línea[214] en el ámbito de la UE, actualmente en vigor pero no aplicable hasta el 7 de junio de 2022.

Según el mismo, las autoridades competentes de los Estados miembros estarán facultadas para emitir órdenes de retirada a los prestadores de servicios, retirar contenidos terroristas o bloquear el acceso a ellos en todos los Estados miembros. Los prestadores de servicios tendrían que retirar o bloquear el acceso al contenido en el plazo de una hora. Las autoridades competentes de los Estados miembros en los que estuviera establecido el prestador de servicios tendrían derecho a examinar las órdenes de retirada emitidas por otros Estados miembros.

Los prestadores de servicios de alojamiento de datos expuestos a contenidos terroristas también deberían adoptar medidas específicas para hacer frente al uso indebido de sus servicios y protegerlos contra la difusión de contenidos terroristas. El Reglamento deja muy claro que la decisión sobre la elección de las medidas corresponde al prestador de servicios de alojamiento de datos.

El marco aprobado pretende garantizar la libertad de expresión y de información y la libertad de empresa, para lo que incluye la opción de recurrir por parte de los usuarios cuyos contenidos hayan sido retirados y también para reclamar por parte de los prestadores de servicios.

El informe precitado anteriormente también hace referencia a la necesidad de revisar el artículo 13 de la entonces Propuesta de Directiva sobre derechos de autor en el mercado único digital[215] en relación con la protección y responsabilidad de los intermediarios en base a la utilización de tecnologías de filtrado. La redacción de la actual Directiva sobre

[214] Reglamento (UE) 2021/784 del Parlamento Europeo y del Consejo, de 29 de abril de 2021, sobre la lucha contra la difusión de contenidos terroristas en línea. DOUE N. 172, de 17 de mayo de 2021. Pp. 79 a 109

[215] Propuesta de Directiva del Parlamento Europeo y del Consejo sobre los derechos de autor en el mercado único digital. COM/2016/0593 final - 2016/0280 (COD). Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52016PC0593. Consultada el 03.01.2021.

los derechos de autor y derechos afines en el mercado único digital[216] difiere notablemente de lo previsto en la propuesta precitada, en particular, su artículo 17.

El informe considera que "la legislación para proteger la libertad de expresión puede ser prematura y potencialmente peligrosa con respecto a los derechos fundamentales" siendo preferible la colaboración entre diferentes grupos de interesados con el examen público.

En cuanto a las opciones políticas para afrontar estos retos, me permito destacar por su relevancia algunas consideraciones y propuestas de los autores de este informe, entre otras:

- La mejor forma de abordar la desinformación es a través de iniciativas de pluralismo y alfabetización de los medios de comunicación, ya que éstas permiten la diversidad de expresión y elección.

- Desaconsejan la adopción de medidas reguladoras que fomenten el uso de la inteligencia artificial para la moderación de contenidos, sin que haya procesos de revisión y apelación humanos sólidos.

- Apuestan por la revisión y la auditoría independientes de las condiciones de las plataformas sobre sus usuarios. Cuando los intermediarios técnicos tengan que moderar los contenidos y las cuentas, son esenciales las políticas detalladas y transparentes, los procedimientos de notificación y apelación y los informes periódicos. Y considera que esto también es útil para las eliminaciones automáticas.

- Estandarización de los aspectos básicos de los procedimientos de notificación y revisión y la presentación de informes, y de crear un organismo multisectorial de autorregulación o corregulación.

[216] Directiva (UE) 2019/790 del Parlamento Europeo y del Consejo, de 17 de abril de 2019, sobre los derechos de autor y derechos afines en el mercado único digital y por la que se modifican las Directivas 96/9/CE y 2001/29/CE. PE/51/2019/REV/1. OJ L 130, 17.5.2019. Pp. 92–125

- Introducir una mayor transparencia en la variedad de técnicas de inteligencia artificial y de reducción de la desinformación utilizadas por las plataformas en línea y los proveedores de contenidos.

Estos aspectos en relación con otros comentados anteriormente se ponen en evidencia ante los riesgos de manipulación de la opinión pública, vigilancia y control generalizado de la ciudadanía, especialmente significativa en países "no democráticos.

A modo de ejemplo, un reciente informe del *Center for Security and Emerging Technology* de Georgetown, publicado en mayo bajo el título *Truth, Lies and Automation* pone en evidencia el impacto de la inteligencia artificial y las redes neuronales actuales si se orientaran para generar campañas de desinformación.

En este sentido, distintos gobiernos han concebido los *deep fakes*, incluyendo bajo este concepto fotografías, audios, vídeos como otras falsificaciones realistas generadas con tecnologías de inteligencia artificial, como fuente de problemas de seguridad nacional en los próximos años.

Algunos de estos contenidos falsos pueden detectarse frecuentemente sin herramientas de detección especializadas, si bien, la sofisticación de la tecnología está progresando rápidamente hasta un punto en el que la detección humana sin ayuda será muy difícil o imposible como destacan algunos informes[217].

La industria está invirtiendo ya en herramientas automatizadas más complejas de detección de falsificaciones profundas para responder a estas amenazas y también los gobiernos.

EE.UU., a través de su *Agencia de Proyectos de Investigación Avanzada de Defensa* -DARPA-, tiene actualmente dos programas dedicados a la detección de falsificaciones profundas: *Media Forensics -MediFor-* y *Semantic Forensics -SemaFor-*.

[217] *Deep Fakes and National Security*. Congressional Research Service (CRS). EE.UU. 26.08.2020. Recuperado de: https://crsreports.congress.gov/. Consultado el 23.03.2021.

MediFor está desarrollando algoritmos para evaluar automáticamente la integridad de las fotos y vídeos y proporcionar a los analistas información sobre cómo se generaron los contenidos falsificados.

SemaFor pretende desarrollar algoritmos que detecten, atribuyan y caractericen automáticamente (es decir, que identifiquen como benignos o maliciosos) varios tipos de falsificaciones profundas.

4.2.17. Riesgo de confusión

Otro aspecto relacionado con todo lo anterior es el riesgo de confusión, especialmente ante la inclusión del lenguaje natural humano, de modo que pueda llevar a la persona a verse afectada en sus emociones o incluso ser objeto de engaño al estar interactuando con un sistema inteligente sin saberlo. Recordar que las emociones son "*hackeables"* y "*crackeables*".

El artículo 52 del Reglamento propuesto sobre normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, regula específicamente entre las obligaciones de transparencia para determinados sistemas de inteligencia artificial que, en el caso de sistemas inteligentes destinados a interactuar con personas físicas, deberán estar diseñados y desarrollados de tal manera que las personas físicas estén informadas de que están interactuando con un sistema de inteligencia artificial, a menos que esto sea obvio por las circunstancias y el contexto de uso, siendo obligación de los proveedores de estos sistemas asegurar el cumplimiento de esta obligación.

No obstante, según la propuesta precitada, esta obligación no se aplicaría a los sistemas de inteligencia artificial autorizados por la ley para detectar, prevenir, investigar y perseguir delitos, a menos que esos sistemas estén disponibles para que el público pueda denunciar un delito.

Además del riesgo de confusión respecto de la naturaleza humana o artificial de con quién se esté interactuando, también se produce el riesgo de confusión respecto de creaciones o innovaciones generadas mediante sistemas inteligentes, como fotografías o retratos, tal y

como abordaré en el último capítulo relativo a las implicaciones de la inteligencia artificial en materia de Propiedad Intelectual e Industrial.

4.2.18. Falta de transparencia e información

Otro de los principales riesgos de la utilización de la inteligencia artificial relacionados con los precedentes es la falta de transparencia y falta de información en un ecosistema de desequilibrio informativo, y no sólo respecto de su utilización, explicabilidad y responsabilidad, conforme he expuesto, sino de su funcionamiento y aplicación, lo que puede provocar que una persona en determinados momentos no sepa si está interactuando con una persona o con una máquina dotada de inteligencia artificial, como he expuesto en el anterior apartado, o que se esté analizando en línea su poder adquisitivo/tesorería o sus preferencias para la recepción de determinados mensajes de contenido político o electoral.

Las obligaciones de transparencia e información han sido reguladas en las distintas propuestas europeas elaboradas en el seno de la UE hasta la fecha, tanto en las de 20 de octubre de 2020, como en la reciente de 21 de abril de 2021, conforme abordaré en los próximos capítulos durante su análisis, si bien, con distinto alcance y ámbito de aplicación como expondré.

La opacidad en el funcionamiento de un sistema inteligente lo puede convertir en una *caja negra -black box-*, de modo que se dificulte o imposibilite la explicabilidad de sus decisiones y acciones, así como la prueba y exigencia de responsabilidad, sin perjuicio de las dificultades que presentan por naturaleza los sistemas de aprendizaje profundos.

Sin duda, las propias características de la inteligencia artificial y de las posibles capacidades de las que pueda estar dotada, especialmente el autoaprendizaje y más el profundo, contribuyen, de inicio, a dicha falta de transparencia y opacidad, junto con otros factores de carácter tecnológico -como la propia complejidad del código algorítmico o los datos utilizados durante su entrenamiento o aprendizaje que pueden no estar disponibles-, económico -como el impacto que puede tener a nivel de costes la transparencia

incluyendo el acceso a secretos comerciales-, e incluso social -dado que podrían vulnerar el derecho a la privacidad de algunas personas-.

La adecuada gestión de estos riesgos de los sistemas inteligentes, ya sea en el sector público o privado, se haya ya contemplada por otros marcos regulatorios, especialmente los de privacidad.

En este sentido, significar la Sentencia del Tribunal de Distrito de la Haya de 5 de febrero de 2020 sobre un sistema de inteligencia artificial holandés llamado "Systeem Risico Indicatie" -SyRI-[218], utilizada por el gobierno de los Países Bajo para detectar diversas formas de fraude a las Administraciones públicas.

La sentencia resolvió la conflictividad planteada sustentándose, entre otros aspectos, en la transparencia, información y posible discriminación de un sistema utilizado para una entidad pública para determinar el riesgo de los ciudadanos, con infracción del artículo 8 del Convenio Europeo de Derechos Humanos (CEDH) sobre el derecho a la intimidad y la discriminación. Considera que el tratamiento de datos que efectuaba el sistema llevaba a la exclusión injustificada, estigmatización y la discriminación de determinados barrios.

A nivel internacional, destacar la Propuesta de Ley de Responsabilidad Algorítmica de 2019 presentada en EE.UU. en abril de 2019, pendiente de tramitación y aprobación.

4.2.19. Riesgos medioambientales

Por último, también deben ser considerados los riesgos medioambientales como significa el *Libro blanco sobre la inteligencia artificial* de la Comisión Europea, referenciado anteriormente, ante las repercusiones medioambientales de los sistemas de inteligencia

[218] COTINO, L. (2020). "SyRI, ¿a quién sanciono? 'Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público y la sentencia holandesa de febrero de 2020'". *La Ley Privacidad*. Wolters Kluwer. Nº 4. Mayo 2020.

artificial a lo largo de su ciclo de vida y durante toda la cadena de suministro, por ejemplo, ante los recursos que precisa para el entrenamiento y el almacenamiento de datos.

La inteligencia artificial constituye una herramienta de ayuda para gestionar el clima, especialmente para construir modelos predictivos de cambio climático y de consumo energético, pero también comporta en enorme consumo de recursos energéticos.

Algunos expertos como Anders Andrae[219] estima en un reciente estudio que, si no mejoramos la eficiencia energética de los procesadores, en 2025, la industria TIC podría consumir el 20% de toda la electricidad del planeta y emitir hasta un 5,5% de las emisiones de CO_2. El despliegue y aplicación masiva de la inteligencia artificial comportará una demanda energética de alto impacto.

4.2.20. Ausencia de normas éticas vinculantes

Como he anticipado en la introducción y capítulos precedentes, la construcción de la inteligencia artificial para el futuro no puede llevarse a cabo al margen de la ética sino considerando la misma desde su diseño y por defecto, junto a la seguridad, el cumplimiento regulatorio y el respeto de los valores y derechos fundamentales.

Los valores y normas éticas esenciales, como el control y la supervisión humana, la seguridad, la transparencia, la trazabilidad, la explicabilidad, la no discriminación o el respeto de los derechos fundamentales, constituyen un marco indiscutible para el futuro de la humanidad. Su ausencia sólo conllevará la materialización de las amenazas y riesgos precitados que pueden afectar hasta incluso a nuestra propia supervivencia.

Sin perjuicio de que estos aspectos los abordaré con mayor profundidad en el capítulo III, destacar ya que necesitamos ahora y en este instante unas normas éticas esenciales y vinculantes en el diseño, funcionamiento y uso de los sistemas inteligentes, cualquiera que sea su nivel de riesgo ante sus capacidades potenciales, sin perjuicio de la promoción

[219] ANDRAE, A. (2017). "Total Consumer Power Consumption Forecast". *Nordic Digital Business Summit*, 2017.

simultánea de la autorregulación y de los códigos de buenas prácticas que contemplen otras normas éticas o herramientas para su cumplimiento. En este sentido, los nuevos marcos reguladores deben integrar estas normas éticas esenciales como exigencia para alcanzar su carácter vinculante.

Esa base ética deberá incluir, junto a la transparencia y la explicabilidad, la rendición de cuentas y la responsabilidad, para garantizar en la mayor medida posible un resarcimiento efectivo por parte de las personas afectadas por los daños causados por sistemas inteligentes, sin perjuicio de que los futuros marcos pueden contemplar mecanismos, como la responsabilidad objetiva absoluta para determinados sistemas u otros, para la consecución de estos objetivos que analizaré en el capítulo V.

La propuesta regulatoria incorporada en la Resolución del Parlamento Europeo de 20 octubre de 2020 en materia ética, incluye un conjunto de principios y normas básicas vinculantes para cualquier sistema inteligente, si bien, la última Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, únicamente exige algunas de ellas pero exclusivamente para sistemas calificados conforme al mismo de alto riesgo, no para el resto, a excepción de algunas obligaciones de transparencia e información para determinados sistemas. Todas estas cuestiones serán objeto de análisis detallado en los capítulos posteriores.

4.2.21. Ausencia de marcos reguladores específicos y falta de adecuación de los existentes

La inteligencia artificial plantea una realidad muy compleja no contemplada en la mayoría de los ordenamientos jurídicos de manera global y específica, de modo que la UE se ha convertido en la actualidad en referencia mundial en sus propuestas regulatorias.

No obstante, conforme será objeto de análisis en los próximos apartados, las mismas deben ir acompañadas de la revisión de distintos marcos reguladores existentes, ante su insuficiencia para resolver los problemas que los sistemas inteligentes pueden plantear ya

en la actualidad y en un fututo inminente, especialmente en materia de responsabilidad por daños derivados del funcionamiento y uso de estos sistemas.

4.2.22. Apreciaciones finales

En conclusión, todos estos riesgos suponen importantes retos para los gobiernos y, especialmente, para el legislador, que deberá acometer una necesaria revisión de los marcos vigentes para identificar su adecuación para abordar, evitar o mitigar, o en su caso trasladar estos riesgos -por ejemplo, mediante seguros y constitución de fondos de compensación-, completándolos en todo aquello que sea necesario, especialmente mediante el establecimiento de mecanismos de verificación y certificación.

La adecuada gestión de estos riesgos, a mi juicio, exigirá nuevos marcos regulativos con visión y eficacia transnacional y global, que impidan el lanzamiento de sistemas al mercado que no sea seguros y que permitan garantizar el cumplimiento de unos principios y normas éticas esenciales que, desde luego, garanticen los valores y los derechos fundamentales como la vida, la intimidad, la privacidad, la libertad, la dignidad, la libertad de expresión, el derecho a la información, la protección de la salud o del medio ambiente, cuestiones sobre las que la UE ha dado distintos pasos, como se analiza en esta investigación.

Pero, estos derechos y bienes y los sujetos titulares de los mismos no son los únicos a proteger frente a los retos y riesgos precitados, sino que pueden verse afectadas empresas, infraestructuras críticas, servicios esenciales, Administraciones públicas, gobiernos o estados.

Y, sin perjuicio de ello, los marcos precitados deberán acompañarse por información, concienciación y formación orientada no solo a operadores y usuarios, sino a la sociedad en general, que no sólo deberá implicar a gobiernos y autoridades competentes en materia digital, tecnología, privacidad y seguridad para su difusión, sino también a las autoridades educativas, que deberían incorporarla en los sistemas educativos en las edades más

tempranas, a lo que también puede contribuir y mucho la propia tecnología implicada: La inteligencia artificial.

No obstante, la solución a estos riesgos y retos no puede ser meramente jurídica, sino un conjunto de actuaciones y marcos desde un enfoque global a nivel ético, jurídico -*hard law* y *soft law*- y de seguridad que mantengan el justo equilibrio para la salvaguarda de todos los intereses en juego, y en el que la promoción de códigos éticos y estándares de buenas prácticas en la industria contribuirá enormemente a su involucración desde el diseño y concepción de los sistemas para asegurar la denominada *Etichs, privacy, security and compliance by design*.

Ese justo equilibrio debería prevenir y evitar una serie de riesgos adicionales.

En primer lugar, el de una hiperregulación de la inteligencia artificial a nivel local, que podría favorecer a economías extranjeras sujetas a regímenes más favorecedores frente al sector empresarial y economía local.

En segundo lugar, todo ello podría hacer surgir otro riesgo adicional, el de superación e inaplicación del derecho vigente para primar el avance tecnológico y la beneficencia de la tecnología disponible para el ser humano y el mundo.

Y, por último, un exceso de regulación podría conllevar control y restricción, por lo que las preguntas obligadas de cara al futuro son: ¿Quién debe controlar una inteligencia que podría ser igual o superior a la humana en su emulación? ¿Quiénes deben establecer los límites a la misma? ¿Los políticos, los científicos, los filósofos?

5. Seguridad de personas, cosas e infraestructuras

Si durante los apartados precedentes he reflexionado sobre los principales retos y riesgos de la inteligencia artificial, la seguridad física, lógica, moral y jurídica constituye posiblemente el principal riesgo y reto asociado a la misma.

Abordarlos requiere su comprensión previa y su adecuada gestión para garantizar estos objetivos, a fin de generar una inteligencia fiable y segura que permita su desarrollo y aplicación.

5.1. Cuestiones generales

La seguridad de las personas, cosas e infraestructuras conforma la base de cualquier ordenamiento jurídico y es el requisito imprescindible para el desarrollo de la sociedad, en especial, de la denominada Sociedad Digital.

La seguridad incluye diversos bienes y derechos, es un concepto amplio, de modo que no sólo incluye la vida, la salud o los derechos fundamentales.

La Sociedad Digital se despliega en un nuevo entorno en el que conviven miles de millones de personas, empresas, asociaciones, fundaciones, Administraciones públicas, gobiernos y otros entes con/sin personalidad jurídica, incluso agentes con personalidad electrónica, sistemas y máquinas, y que no conoce de fronteras ni límites territoriales, lo que impacta frontalmente con el concepto de ordenamiento jurídico tradicional, concebido como el conjunto de normas que rigen en un territorio determinado en un tiempo concreto.

No obstante, la seguridad digital o lógica no puede desligarse de la seguridad física, dada la convergencia cada vez mayor entre la dimensión física y virtual de nuestro mundo.

El estudio *Ciberseguridad, un desafío mundial*[220] de la Fundación Innovación Bankinter, que recogió trabajos e intervenciones de expertos en seguridad, nacionales e internacionales en el foro *Future Trends Forum* celebrado en diciembre de 2015, concluyó con una hoja de ruta para la ciberseguridad a modo "Decálogo", cuyos primeros

[220] El estudio completo fue presentado el 5 de mayo de 2016 [en línea]: https://www.fundacionbankinter.org/documents/20183/42758/Ciberseguridad/f22bf829-7bb5-447c-9c88-9f7bac51f482 El estudio es también una publicación del Future Trends Forum donde se analizan los desafíos de la ciberseguridad en el entorno global y se plantea una hoja de ruta de diez propuestas para regular y mejorar nuestra ciberseguridad.

dos puntos eran, de un lado, la necesidad de reducir los costes globales de los ciberataques y cibercrimen, instando a los países a nivel internacional a trabajar conjuntamente bajo el liderazgo de ENISA, y de otro, garantizar la integridad de las infraestructuras y soluciones tecnológicas.

Por su parte, el Instituto Nacional de Ciberseguridad de España -INCIBE- también ha publicado un "Decálogo de Ciberseguridad"[221], basado en el análisis de riesgos, protección de la información y seguridad gestionada.

La seguridad asociada a los sistemas de inteligencia artificial es un requerimiento ético que integra la mayoría de marcos propuestos a nivel mundial, especialmente el europeo, y pretende incorporarse específicamente como requerimiento regulativo en las propuestas de la UE para regular la inteligencia artificial, en particular, en la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas, que incorpora una Propuesta de Reglamento del Parlamento Europeo y del Consejo, sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, así como en la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial, que incorpora una Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial.

Sin embargo, la Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[222], regula la ciberseguridad y la robustez de los sistemas, pero exclusivamente como un requerimiento de los sistemas calificados de alto riesgo conforme al mismo, lo que de inicio, evidencia lo que considero una omisión de un requerimiento ético y jurídico que

[221] INCIBE. *Decálogo de ciberseguridad. El camino hacia la ciberseguridad de su empresa*. Recuperado de: https://www.incibe.es/extfrontinteco/img/File/empresas/blog/2014Octubre/decalogo_ciberseguridad_empresas.pdf. 2015

[222] COM (2021) 206 final 2021/0106 (COD)

debería ser esencial para cualquier sistema inteligente (por ejemplo, un *chatbot*), sea considerado de nivel medio o bajo o simplemente todavía no clasificado. Abordaré esta cuestión con mayor detalle más adelante.

No obstante, la exigencia de la seguridad ya fue abordada previamente a las propuestas precitadas en la Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica[223], en la que el Parlamento Europeo se focalizó en la seguridad como uno de los aspectos fundamentales a considerar y contemplar por los marcos normativos específicos y no sólo del propio sistema, sino de los agentes con los que se relaciona.

De hecho, en su apartado 21º, destacó que la utilización adecuada de la robótica y la inteligencia artificial requería un alto grado de seguridad de estos sistemas, significando la necesidad de garantizar la seguridad de redes de robots y sistemas de inteligencia artificial interconectados para evitar posibles quiebras de seguridad.

En este sentido, el Parlamento Europeo ya destacó en dicha Resolución de 2017 la responsabilidad de los diseñadores de robots y sistemas de inteligencia artificial de desarrollar productos que sean seguros, fiables y que cumplan su función, e instó a la Comisión Europea y a los Estados miembros a que apoyen e incentiven el desarrollo de la tecnología necesaria con este propósito, en especial, la seguridad desde el diseño -*Security by design*-.

Con este propósito, la precitada Resolución incorporó en su anexo una *Carta sobre Robótica* como código de conducta ética en el campo de la robótica que integraba un modelo de licencia para los diseñadores de robots o sistemas dotados de inteligencia artificial avanzada, consistente en un conjunto de principios y normas que los diseñadores debían considerar y que ya contemplaba la denominada "*Security by design*" y la "*Privacy by design*", exigida por el precedente Reglamento General de Protección de Datos europeo (RGPD).

[223] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))

Además, el Parlamento Europeo destacó en la misma el enorme potencial de la robótica a la hora de mejorar la seguridad en el entorno laboral mediante la transferencia a los robots de tareas peligrosas y perjudiciales que, en la actualidad, estaban desempeñando personas, sin perjuicio de advertir en paralelo del peligro que podría entrañar la robotización al crear una serie de nuevos riesgos como consecuencia del creciente número de interacciones entre los seres humanos y los robots en el lugar de trabajo, lo que exigía un marco normativo que regule las interacciones entre los seres humanos y los robots con la finalidad, entre otras, de garantizar la salud, la seguridad y el respeto de los derechos fundamentales de las personas y, en especial, de quienes se hallen desempeñando su actividad laboral en su entorno de trabajo.

En cualquier caso, en ausencia actual de marcos específicos reguladores de la seguridad requerida por los sistemas inteligentes, la seguridad y, en especial, la seguridad en el diseño -*Security by design*-, como analizaré posteriormente, es ya exigida por el ordenamiento jurídico vigente para determinados productos, servicios y sistemas, en particular por los marcos reguladores vigentes en materia de seguridad de los productos, ciberseguridad o privacidad en el sector público y privado.

5.2. La seguridad: Un requerimiento ético y jurídico

La seguridad constituye uno de los principios y valores éticos esenciales que integran los distintos marcos éticos que se han ido conformando y consensuando a nivel internacional en materia de inteligencia artificial.

Los denominados *Principios de Asilomar* aprobados por la Conferencia organizada por el *Future of Life Institute* en dicha localidad californiana en enero de 2017, recogían dentro de los mismos la seguridad, considerando que los sistemas de inteligencia artificial siempre deben ser seguros y fiables a lo largo de todo su ciclo de vida útil, así como verificables. Además, su aplicación a un ámbito de la realidad debe ser factible.

Del mismo modo, Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre

robótica[224], incluyó en la seguridad como uno de los aspectos éticos esenciales a considerar y contemplar por los futuros marcos normativos.

Un año más tarde, el *Grupo Europeo sobre Ética de la Ciencia y las Nuevas Tecnologías* publicó sus reflexiones y conclusiones en materia ética el 9 de marzo 2018, en la denominada "*Declaración sobre inteligencia artificial, robótica y sistemas autónomos*", en la que se recogían los principios éticos que deben inspirar la futura regulación de los mismos en la Unión.

Como se abordará más adelante, el documento pretendía establecer las bases para la construcción de un futuro marco ético y legal común e internacionalmente reconocido para el diseño, producción, uso y gobernanza de la inteligencia artificial, la robótica y los sistemas "autónomos", proponiendo un conjunto de principios éticos fundamentales que pueden servir de guía para el desarrollo de este marco ético y legal, basados en los valores establecidos en los Tratados de la UE y en la Carta de Derechos Fundamentales de la UE.

Uno de estos principios era y es el de seguridad, protección, e integridad física y mental que concreta en tres aspectos:

a) La seguridad externa, que se ofrece al entorno y a los usuarios;

b) La confiabilidad y la robustez interna y;

c) La seguridad emocional en la interacción humano-máquina.

Conforme recoge la declaración precitada, estas tres dimensiones de la seguridad y la protección deben ser consideradas por diseñadores y desarrolladores de inteligencia artificial y deben ser estrictamente evaluadas antes del lanzamiento de cualquier sistema de inteligencia artificial considerado "fuerte", con la finalidad de garantizar que estos sistemas no infrinjan el derecho de los seres humanos a la integridad física y mental, y a un entorno seguro.

[224] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))

En este sentido, el documento destaca la necesidad de prestar especial atención a aquellas personas más vulnerables, así como al posible doble uso de los sistemas y al uso militar de la inteligencia artificial, especialmente en el ámbito de la ciberseguridad -como luego abordaré específicamente-, las finanzas, las infraestructuras y los conflictos armados.

Además de recoger específicamente la seguridad como un principio esencial sobre el que se deben construir los futuros marcos éticos y jurídicos, también incorpora otros principios relacionados, como el de "Estado de derecho y rendición de cuentas" que deben proteger y garantizar la seguridad o la privacidad, así como el principio de "Protección de datos y privacidad" de modo que se garantice el respecto de los marcos reguladores de protección de datos por parte de los sistemas de inteligencia artificial, ya sean integrados en robots físicos o *softbots* basados en *software*.

En definitiva, la confianza construida sobre la seguridad jurídica, técnica y organizativa junto con la transparencia, es la base para el desarrollo y crecimiento de la sociedad y la denominada economía digital.

De hecho, la seguridad y la confiabilidad en la inteligencia artificial han pasado a considerarse un objetivo también estratégico para la UE, con el objetivo de garantizar el desarrollo y aplicación de la misma en beneficio de la sociedad y la economía.

Pero no debemos olvidar un aspecto fundamental del que he puesto ejemplos muy gráficos en esta investigación, los sistemas inteligentes precisan de sistemas de comunicaciones y de conexión a redes.

Un error, suspensión o ataque a los mismos (ni tan siquiera al propio sistema inteligente), comportaría la caída de todos los sistemas inteligentes que gestionen o gobiernen en remoto servicios o dispositivos críticos o no críticos, lo que puede afectar a *Smartcities* enteras, vehículos "autónomos", suministro de agua, electricidad, sistemas de monitorización vital, intervenciones quirúrgicas, etc. Esta dependencia constituye una de las vulnerabilidades y riesgos más importantes de la inteligencia artificial que debe se adecuadamente considerada y gestionada, especialmente mediante sistemas de redundancia, que no he visto recogida específicamente en los múltiples documentos de trabajo y propuestas analizadas en esta investigación.

De este modo, las recientes propuestas regulatorias del Parlamento Europeo precitadas de octubre de 2020, en materia ética y de responsabilidad civil de la inteligencia artificial, integran la seguridad como exigencia y desde el diseño. En particular, la relativa a la ética, la contempla como una exigencia para cualquier sistema inteligente, sea o no de alto riesgo.

Por su parte, la nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021 precitada, regula de manera general en su artículo 15 los requisitos de precisión, robustez y ciberseguridad de los sistemas de inteligencia artificial considerados de alto riesgo, pero no respecto del resto. En este sentido, conforme dispone su artículo 6.1, incluso cuando un sistema inteligente vaya a ser un componente de seguridad o sea el propio producto en el ámbito del *Nuevo Marco Legislativo* previsto en el mismo, se considerará de alto riesgo.

La inteligencia artificial puede afectar y dañar a la persona a nivel físico o psicológico, así como a sus derechos y bienes.

Los principales derechos de las personas físicas que pueden verse afectados de forma significativa en estos contextos serían el derecho a la vida y la integridad física y moral, la libertad de pensamiento, la movilidad y el desplazamiento, el derecho al honor, la intimidad, a la dignidad, a la privacidad y a la propia imagen, la libertad de expresión e información, todos ellos derechos fundamentales reconocidos por la Constitución Española.

Los derechos de las personas jurídicas que también pueden verse afectados de forma más significativa podrían ser los derechos sobre los secretos empresariales e innovaciones, propiedad intelectual e industrial o libertad de empresa.

La salud, la integridad física, la protección de la información, de los bienes y de los derechos de las personas físicas o jurídicas o la protección del medio ambiente se hayan ya regulados en distintas normas nacionales, autonómicas e incluso locales, pero también en convenios y tratados internacionales y en el Derecho de la Unión.

El objeto y alcance de esta investigación me impide poder abordar con la profundidad necesaria todos los marcos jurídicos vigentes reguladores de la seguridad de los bienes y derechos que pueden verse afectados por los sistemas de inteligencia artificial, si bien, ante la propia naturaleza de los sistemas inteligentes, me permito significar cuanto menos algunos de los marcos generales de referencia a nivel nacional y europeo en materia de seguridad relativos a información personal, información empresarial y seguridad en infraestructuras críticas y servicios esenciales, seguridad de las máquinas y de los productos y seguridad de sistemas inteligentes, que deberán ser considerados en el diseño y desarrollo de sistemas de inteligencia artificial orientados a su gestión.

Adicionalmente, disponemos de marcos sectoriales y específicos, por ejemplo, en el ámbito médico-sanitario, vehículos, etc.

En el capítulo IV de esta investigación haré también referencia a otros marcos de seguridad específicos a nivel internacional.

5.3. Seguridad de la información

Sin perjuicio de la amplísima normativa reguladora de la seguridad de la información en el ámbito público y privado, dado el objeto y alcance limitados de esta investigación, me permito referenciar cuanto menos los generales relativos a información personal e información corporativa empresarial.

5.3.1. Información personal

La seguridad y protección de los datos relativos a personas físicas, identificadas o identificadas, se regula en la UE en el Reglamento General de Protección de Datos (RGPD)[225].

El RGPD tiene un enorme impacto en las organizaciones y, sin duda, sus principios y obligaciones son para cualquier empresa una necesidad pero, como me he permitido comentar en alguna publicación previa[226], también un reto y una oportunidad para la mejora de la seguridad de su información corporativa, de su organización, para mejorar la gestión de sus activos de información y del *Big data*, para la reingeniería y organización de sus procesos, para mejorar el propio valor y calidad de la información, para gestionar datos de valor, para aprovechar de forma eficaz y legal el valor de los datos, para permitir una mayor orientación al cliente, para asegurar que la organización está preparada para la sociedad y economía digital, para mejorar y generar un ecosistema de confianza a sus clientes y usuarios (internos o externos) pero también a otras partes interesadas, como socios, trabajadores, inversores, proveedores, Administraciones Públicas y otros terceros.

La norma exige la responsabilidad proactiva y gestionada. No sólo aplicar seguridad de manera efectiva sino estar en todo momento en condiciones de demostrarlo, monitorizada, revisada y controlada e integrada en el diseño y por defecto, esto es, la denominada *Privacy by design and by default*.

Asimismo, pretende tener una eficacia transnacional, de modo que es exigible a entidades ubicadas dentro y fuera de la UE, cuando las actividades de tratamiento estén relacionadas con la oferta de bienes o servicios a interesados en la UE o con el control de su comportamiento, si éste se produce en la UE.

[225] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento General de Protección de Datos.

[226] MUÑOZ VELA, J.M. (2019). "Los retos del Derecho en una Sociedad Digital", en GIMÉNEZ, I. (Coord.). *Retos de la sociedad digital y medios de pago*. Colección Tratados y Manuales de Economía. Editorial Civitas - Thomson Reuters Aranzadi. Navarra 2019. P. 65.

El objeto de protección de este marco es la información relativa a cualquier persona física identificada o identificable, considerando persona física identificable "toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un identificador, como por ejemplo un nombre, un número de identificación, datos de localización, un identificador en línea o uno o varios elementos propios de la identidad física, fisiológica, genética, psíquica, económica, cultural o social de dicha persona".

Los principios y obligaciones esenciales sobre los que sustenta la norma son la licitud, lealtad y transparencia, limitación de la finalidad, minimización de datos (adecuados, pertinentes y no excesivos), exactitud y actualidad, limitación del plazo de conservación, integridad y confidencialidad y responsabilidad proactiva.

Entre las principales obligaciones y medidas, la norma exige la identificación de tratamientos que lleve a cabo, la elaboración de un registro de actividades de tratamiento en determinados casos, la identificación de las finalidades y las bases jurídicas de los tratamientos, la realización de análisis de riesgos y evaluaciones de impacto en la privacidad -PIA- en determinados contextos, la implantación efectiva de medidas técnicas, jurídicas y organizativas adecuadas a los riesgos, la verificación de las medidas y controles implantados, protocolos de actuación y comunicación ante quiebras de seguridad, monitorización, revisión y control de cumplimiento continuo.

El RGPD incorpora en sus artículos 77 a 84 un régimen sancionador aplicable en caso de incumplimiento de sus prescripciones.

El régimen sancionador administrativo previsto en el mismo difiere en Dinamarca y Estonia, en la medida que sus respectivos ordenamientos jurídicos no permiten las multas administrativas, por lo que, en Dinamarca, conforme recoge el propio RGPD, las multas podrían ser impuestas por los tribunales nacionales competentes como sanción penal y, en Estonia, por la autoridad de control, en el marco de un juicio de faltas.

En adición al RGPD, España aprobó su propia Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD), para regular aquellos aspectos específicos que aquél dejó a los Estados miembros para su regulación, por ejemplo, la mayoría de edad en protección de datos.

La LOPDGDD adaptó el ordenamiento jurídico español al RGPD, derogando distintas normas, entre otras, la anterior LOPD y el Real Decreto Ley 5/2018, de 27 de julio (RCL 2018\1123), de medidas urgentes para la adaptación del Derecho español a la normativa de la Unión Europea en materia de protección de datos.

En adición a las normas precitadas, destacar la Directiva 2016/680 UE, del Parlamento Europeo y del Consejo, de 27 de abril, que regula la protección de las personas físicas respecto al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención investigación, detección y enjuiciamiento de infracciones penales.

El plazo de trasposición de la Directiva al ordenamiento jurídico español de dos años finalizó en 2018 sin que se hubiera aprobado una iniciativa legislativa con tal finalidad. De hecho, la LOPDGDD hace referencia a la misma y a la futura norma de transposición, estableciendo en su Disposición Transitoria Cuarta que, los tratamientos sometidos a la precitada Directiva continuarán rigiéndose por la Ley Orgánica 15/1999, de 13 de diciembre, y en particular el artículo 22, y sus disposiciones de desarrollo, en tanto no entre en vigor la norma que trasponga al Derecho español lo dispuesto en la citada directiva.

Desgraciadamente, a pesar de la aprobación extemporánea del proyecto de Ley Orgánica aprobado por el Consejo de Ministros, mediante el que se pretende transponer esta Directiva al ordenamiento jurídico español, en fechas casi coincidentes con la misma se notificó la Sentencia de 25.02.2021 del Tribunal de Justicia de la Unión Europea (TJUE) en el asunto C-658/19 Comisión/España, en la que se condenó a España por no haber transpuesto ni comunicado las medidas de transposición de dicha Directiva, en particular al pago de una suma de 15,5 millones de euros a tanto alzado y una multa coercitiva diaria de 89.548,20 euros por cada día de retraso en la transposición.

Destacar que es la primera vez que el Tribunal de Justicia de la Unión Europea (TJUE) impone estos dos tipos de multas de manera simultánea, conforme al artículo 260.3 del Tratado de Funcionamiento de la Unión Europea (TFUE). España es uno de los

Estados miembros con más procedimientos de infracción abiertos según el *Informe de aplicación del derecho europeo de 2019*[227].

La Ley Orgánica 7/2021, de 26 de mayo, de protección de datos personales tratados para fines de prevención, detección, investigación y enjuiciamiento de infracciones penales y de ejecución de sanciones penales[228] fue finalmente aprobada en la fecha de cierre de esta investigación.

Por último, en relación con la *Iniciativa para Construir una Economía de los Datos Europea -Building a European Data Economy*, 2017- para la regulación de la libre circulación de datos no personales y las restricciones a la localización de los datos, se aprobó el Reglamento (UE) 2018/1807 del Parlamento Europeo y del Consejo, de 14 de noviembre de 2018, relativo a un marco para la libre circulación de datos no personales en la Unión Europea.

Este Reglamento contempla un enfoque de autorregulación a través de códigos de conducta para facilitar la portabilidad de los datos, así como para cambiar entre proveedores de servicios en la nube, consideradas estas condiciones necesarias para garantizar una Economía de los Datos competitiva en el marco de Mercado Único Digital (DSM).

También aborda otros retos legales inherentes a las nuevas tecnologías sustentadas en datos, tales como el acceso y la transferencia de datos no personales generados de forma automática, la responsabilidad de los datos y la portabilidad, interoperabilidad y estándares.

El Reglamento establece un principio equivalente al del RGPD para los datos no personales, prohibiendo la imposición de restricciones a su circulación, salvo aquellas que resulten justificadas por motivos de seguridad.

---

[227] Informe de la Comisión: *Control de la aplicación del Derecho de la Unión Europea Informe anual de 2019*. COM/2020/350 final. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52020DC0350.

[228] BOE. N. 126, de 27.05.2021.

5.3.2. Información empresarial

La seguridad y protección de la información empresarial debería ser un objetivo estratégico de cualquier organización. Para la protección de esta información disponemos de distintos marcos reguladores, incluyendo la propia legislación en materia de competencia desleal, propiedad intelectual e industrial, así como incluso instrumentos reactivos como el Código Penal.

Los riesgos de seguridad relacionados con la información empresarial pueden afectar a la competitividad, la creatividad, la iniciativa empresarial, la inversión e, incluso, la continuidad, por lo que la adecuada protección de la información empresarial no divulgada de manera preventiva es una necesidad.

Las divergencias nacionales existentes en materia de protección de información empresarial no divulgada, llevó a la aprobación de la Directiva UE 2016/943, del Parlamento Europeo y del Consejo, de 8 de junio de 2016, con el fin de armonizar la legislación en Europa, que fue transpuesta al ordenamiento jurídico español mediante la Ley 1/2019, de 20 de febrero, de Secretos Empresariales, completando así la regulación de la Ley 3/1991, de 10 de enero, de Competencia Desleal, tanto a nivel sustantivo como especialmente procesal.

Ante el objeto y alcance limitados de esta investigación, no puedo detenerme en el análisis de estos marcos, si bien, cuanto menos, considero necesario matizar su objeto, esto es, el concepto de "secretos empresariales".

La norma española opta por traducir el término "*trade secrets*" por "*secretos empresariales*", en la medida que su traducción por "*secretos comerciales*" no parecía la más adecuada, en la medida que el calificativo "*comercial*" podría dejar fuera el concepto "*secreto industrial*" regulado en la directiva.

La Ley tiene como objeto de protección el *know-how* y la información corporativa no divulgada considerada "*secreto de empresa*", definida a los efectos de aplicación de la norma española como: "Cualquier información o conocimiento, incluido el tecnológico, científico, industrial, comercial, organizativo o financiero, que reúna las siguientes

condiciones: a) Ser secreto, en el sentido de que, en su conjunto o en la configuración y reunión precisas de sus componentes, no es generalmente conocido por las personas pertenecientes a los círculos en que normalmente se utilice el tipo de información o conocimiento en cuestión, ni fácilmente accesible para ellas; b) Tener un valor empresarial, ya sea real o potencial, precisamente por ser secreta, y c) Haber sido objeto de medidas razonables por parte de su titular para mantenerlo en secreto".

La norma amplía su definición de "secreto de empresa" respecto del concepto que incorporaba su Anteproyecto, incorporando no sólo "información" sino también "conocimiento", adicionando expresamente los conocimientos "científicos" y matizando que el valor empresarial podrá ser real o potencial.

En consecuencia, la información o conocimiento secreto, a los efectos de la aplicación de esta norma, "será aquella información o conocimiento empresarial que no haya sido divulgado, tenga valor empresarial y se hayan establecido medidas razonables para preservar su secreto por parte de la empresa[229]".

5.4. Infraestructuras críticas y servicios esenciales

La seguridad y protección de las infraestructuras críticas y los servicios esenciales se regula principalmente en la UE, en la Directiva (UE) 2016/1148 del Parlamento Europeo y del Consejo, de 6 de julio de 2016, relativa a las medidas destinadas a garantizar un elevado nivel común de seguridad de las redes y sistemas de información en la Unión ("Directiva NIS") y en su Reglamento de Ejecución (UE) 2018/151 de la Comisión, de 30 de enero de 2018, por el que se establecen normas para la aplicación de la Directiva (UE) 2016/1148 del Parlamento Europeo y del Consejo en lo que respecta a la especificación de los elementos que han de tener en cuenta los proveedores de servicios digitales para gestionar los riesgos existentes para la seguridad de las redes y sistemas de

[229] MUÑOZ VELA, J.M. (2019). "Los retos del Derecho en una Sociedad Digital", en GIMÉNEZ, I. (Coord.). "*Retos de la sociedad digital y medios de pago*". Colección Tratados y Manuales de Economía. Editorial Civitas - Thomson Reuters Aranzadi. Navarra 2019. Pp. 55 y 56.

información, así como de los parámetros para determinar si un incidente tiene un impacto significativo.

En España se regula principalmente en el Real Decreto-ley 12/2018, de 7 de septiembre, de seguridad de las redes y sistemas de información, que transpuso la Directiva NIS precitada.

La Directiva NIS fue concebida para dar una respuesta efectiva a los problemas de seguridad de las redes y sistemas de información en la UE, mediante el establecimiento de unos requisitos comunes de seguridad para operadores de servicios esenciales y proveedores de servicios digitales, así como unas bases comunes para el desarrollo de capacidades y planificación, intercambio de información y cooperación. Su transposición se llevó a cabo por el precitado Real Decreto-Ley 12/2018, de 7 de septiembre, de seguridad de las redes y sistemas de información, ampliando los servicios incluidos en la Directiva, regulando nuevas obligaciones, así como estableciendo un régimen sancionador en caso de incumplimiento de sus disposiciones.

La Memoria de Impacto Normativo que acompañó al Anteproyecto de Ley ya recogía algunos de los datos publicados el año anterior a su tramitación en materia de ciberseguridad, con referencia al informe anual de ciberseguridad de Cisco de 2017, en el que se destacaba que un 29% de las organizaciones sufrieron pérdidas de ingresos causadas por ataques contra la seguridad de la información, siendo estas pérdidas superiores al 20% de los ingresos en un 38% de los casos. Asimismo, hacía referencia al informe de IBM de 2016 sobre impacto en el negocio de incidentes de seguridad que estimaba en 4 millones de dólares el coste medio de cada incidente, con un coste anual per cápita promedio cercano a 200 dólares. Y finalizaba reflejando los datos a nivel nacional de la gestión de incidentes llevada a cabo por el CERT de Seguridad e Industria, del Instituto Nacional de Ciberseguridad (INCIBE) español, que pasó de unos 18.000 en 2014 a 50.000 en 2015 y más de 106.000 en 2016. En 2017, el número de incidentes fue de 123.064. Sin comparamos aquellos datos con los más actuales correspondientes, el diferencial es abrumador.

La norma española establece el marco legal de seguridad exigida que deben observar tanto los operadores de servicios esenciales (energía, transporte, salud, servicios públicos

o transporte) como los proveedores de servicios digitales, así como las facultades de supervisión y auditoría de las autoridades y un régimen sancionador en caso de incumplimiento de las mismas.

Los paralelismos entre este marco normativo en materia de gobierno, organización y gestión de la seguridad y cultura del riesgo con el previsto en el Reglamento General de Protección de Datos (RGPD) en materia de privacidad son evidentes.

La norma obliga a que los operadores de servicios esenciales y los proveedores de servicios digitales adopten medidas técnicas y organizativas adecuadas y proporcionadas para gestionar los riesgos para la seguridad de las redes y sistemas de información utilizados para la prestación de los servicios regulados en la misma, aunque su gestión sea externalizada, con el objetivo de prevenir y reducir al mínimo el impacto de los incidentes. Asimismo, al igual que el RGPD, las medidas deben estar basadas en una evaluación previa de riesgos, si bien no las define, por lo que se remite a su desarrollo reglamentario donde se regularán las mismas, así como la elaboración de instrucciones y guías técnicas.

La norma también prevé la supervisión de los proveedores de servicios esenciales y digitales por parte de las autoridades competentes, aunque no del mismo modo en ambos casos, que podría incluir la exigencia de acreditación de medidas, evaluaciones y auditoría.

Por su parte, el Reglamento de Ejecución de la Directiva NIS[230] citado establece las normas para la aplicación de dicha Directiva por parte de los proveedores de servicios digitales, especificando las medidas exigidas a los mismos para garantizar la seguridad de las redes y sistemas de información que se utilicen en el marco de la oferta de los servicios regulados en el anexo III de la Directiva (UE) 2016/1148, esto es, energía,

[230] Reglamento de Ejecución (UE) 2018/151 de la Comisión, de 30 de enero de 2018, por el que se establecen normas para la aplicación de la Directiva (UE) 2016/1148 del Parlamento Europeo y del Consejo en lo que respecta a la especificación de los elementos que han de tener en cuenta los proveedores de servicios digitales para gestionar los riesgos existentes para la seguridad de las redes y sistemas de información, así como de los parámetros para determinar si un incidente tiene un impacto significativo

transporte, banca, Infraestructuras de los mercados financieros, sector sanitario, suministro y distribución de agua potable, infraestructura digital.

El Reglamento regula las medidas y elementos de seguridad generales de sistemas, redes y entornos físicos a adoptar por los proveedores de servicios digitales, en particular, las siguientes:

a) La gestión sistemática de redes y sistemas de información;

b) La creación de políticas adecuadas de gestión de la seguridad de la información, incluyendo análisis de riesgos, recursos humanos, seguridad de las operaciones, arquitectura de la seguridad, gestión segura del ciclo de vida de datos y sistemas y, cuando proceda, el cifrado y su gestión;

c) La seguridad física y del entorno;

d) La seguridad del abastecimiento de modo que se dispongan y mantengan políticas adecuadas para garantizar la accesibilidad y, en su caso, trazabilidad de los suministros críticos utilizados en la prestación de los servicios;

e) El control de accesos a las redes y sistemas de información, disponiendo de un conjunto de medidas que garanticen el acceso físico y lógico a las redes y los sistemas de información;

f) Gestión de incidentes;

h) Continuidad de actividades;

i) Supervisión, auditorías y pruebas.

El Reglamento regula también los parámetros que deben ser tenidos en cuenta para determinar si el impacto de un incidente es “significativo” a los efectos de evaluación del riesgo, y lo hace en función de número de usuarios afectados, duración del incidente, zona afectada, grado de perturbación del funcionamiento del servicio y alcance del impacto sobre las actividades económicas y sociales.

No obstante, como abordaré en el siguiente apartado, la UE pretende revisar los marcos precitados de seguridad para adaptarse al contexto actual mediante sendas propuestas de Directiva en fase de tramitación en la fecha de cierre de esta investigación, en particular, la propuesta de Directiva sobre medidas para un alto nivel común de ciberseguridad en toda la Unión[231], también conocida como "Directiva NIS revisada" o "NIS 2", y una nueva Directiva sobre resiliencia de entidades críticas[232].

5.5. Seguridad de las máquinas y los productos

Como he expuesto, ante la ausencia actual de marcos específicos reguladores de la seguridad requerida por los sistemas inteligentes, la seguridad es también exigida por el ordenamiento jurídico vigente para determinadas máquinas y productos, que no contemplan la posibilidad de que puedan ser gobernados o gestionados por sistemas inteligentes.

La seguridad sobre los productos se regula principalmente en la UE la Directiva 2001/95/CE del Parlamento Europeo y del Consejo, de 3 de diciembre de 2001 relativa a la seguridad general de los productos[233], en relación con la Directiva (UE) 2019/1937 del Parlamento Europeo y del Consejo, de 23 de octubre de 2019, relativa a la protección de las personas que informen sobre infracciones del Derecho de la Unión[234] y el Reglamento (UE) 2019/1020 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativo a la vigilancia del mercado y la conformidad de los productos y por el que se modifican la Directiva 2004/42/CE y los Reglamentos (CE) n.º 765/2008 y (UE) n.º 305/2011[235].

El objetivo de la Directiva 2001/95/CE es garantizar que los productos que se pongan en el mercado sean seguros, si bien, se circunscribe a los productos destinados al

[231] Proposal for a Directive of the European Parliament and of the Council on measures for a high common level of cybersecurity across the Union, repealing Directive (EU) 2016/1148. Bruselas. 16.12.2020. COM (2020) 823 final.
[232] Proposal for a Directive of the European Parliament and of the Council on the resilience of critical entities. Bruselas. 16.12.2020. COM (2020) 829 final.
[233] DOUEL 15.01.2002
[234] DOUEL 26.11.2019
[235] DOUEL 25.06.2019

consumidor. La misma define conforme a su objeto los conceptos de "producto", "producto seguro", "productor" o "riesgo grave".

El Reglamento (UE) 2019/1020, tiene como objetivo mejorar el funcionamiento del mercado interior mediante el fortalecimiento de la vigilancia del mercado de productos a los que se aplica la legislación de armonización de la UE enumerada en su Anexo I -que incluye máquinas, vehículos, productos sanitarios, eléctricos o electrónicos-, al objeto de garantizar que solamente se comercialicen en la UE productos conformes que cumplan los requisitos que proporcionan un nivel elevado de protección de intereses públicos, como la salud y la seguridad en general, la salud y la seguridad en el trabajo, la protección de los consumidores, del medio ambiente y la seguridad pública y cualquier otro interés público protegido por dicha legislación.

El Reglamento citado es íntegramente aplicable desde el 1 de enero de 2021 y se aplica a la vigilancia del mercado y a los productos sujetos a dicha legislación de armonización, en la medida en que no existan en la misma, disposiciones concretas con el mismo objetivo que regulen de manera más específica determinados aspectos de la vigilancia del mercado y la ejecución de las normas.

Por otra parte, sus artículos 25 a 28 relativos a los controles de los productos que entran en el mercado de la UE, se aplican por defecto a los productos regulados por el Derecho de la UE, en la medida en que no existan disposiciones específicas en éste relativas a la organización de controles sobre los productos que entren en el mercado de la Unión.

De este Reglamento me permito significar la definición que incorpora en su artículo 3 de "producto que presenta un riesgo" y "producto que presenta un riesgo grave", donde se parte de los criterios de probabilidad e impacto para su calificación y no meramente de este último.

En primer lugar, considera "producto que presenta un riesgo" a aquel producto que puede afectar negativamente a la salud y la seguridad de las personas en general, a la salud y la seguridad en el trabajo, a la protección de los consumidores, al medio ambiente, a la seguridad pública o a otros intereses públicos protegidos por la legislación de armonización de la UE aplicable, en un grado que vaya más allá de lo que se considere

razonable y aceptable en relación con su finalidad prevista o en las condiciones de uso normales o razonablemente previsibles del producto en cuestión, incluida la duración de su utilización y, en su caso, los requisitos de su puesta en servicio, instalación y mantenimiento

En segundo lugar, considera "producto que presenta un riesgo grave" aquél que presenta un riesgo para el que, sobre la base de una evaluación del riesgo y teniendo en cuenta el uso normal y previsible del producto, se considere que la combinación de la probabilidad de que se produzca un peligro que cause un daño o perjuicio y su gravedad requiera una rápida intervención de las autoridades de vigilancia del mercado, incluidos los casos en que el riesgo no tenga efectos inmediatos.

De ambas definiciones, destacar la amplitud de los derechos y bienes objeto de protección susceptibles de ser afectados por los riesgos a gestionar conforme al marco regulatorio, no circunscritos a la salud, la seguridad o los derechos fundamentales como algunas de las propuestas en materia de inteligencia artificial, así como la consideración de los criterios de probabilidad e impacto en el marco del contexto concreto para determinar su nivel de riesgo.

En España, la seguridad de los productos se regula en el Real Decreto 1801/2003, de 26 de diciembre[236], sobre seguridad general de los productos, y en otras normas relacionadas, especialmente sectoriales como, por ejemplo, en el ámbito sanitario.

Conforme regula el mismo en su artículo 1, su objetivo es garantizar que los productos que se pongan en el mercado sean seguros, si bien, sus disposiciones únicamente se aplicarán a los productos destinados al consumidor, incluidos los ofrecidos o puestos a disposición del mismo en el marco de una prestación de servicios para que este los consuma, maneje o utilice directamente o que, en condiciones razonablemente previsibles, "pueda ser utilizado por el consumidor aunque no le esté destinado, que se le suministre o se ponga a su disposición, a título oneroso o gratuito, en el marco de una actividad comercial, ya sea nuevo, usado o reacondicionado".

[236] BOE 10.01.2004

En relación con la seguridad sobre máquinas, la Comisión Europea anunció la nueva regulación sobre máquinas que sustituirá a la antigua Directiva 2006/42/CE de 17 de mayo de 2006, en la que se incluyen tanto productos de consumo como profesionales, a diferencia de las Directiva de Productos Defectuosos que se circunscribe a los primeros, ampliando su ámbito, e incluyendo robots, cortadoras de césped, impresoras 3D, máquinas de construcción y líneas de producción industrial.

El futuro Reglamento sobre Máquinas pretende contemplar la nueva generación de maquinaria que garantice la seguridad de usuarios y de consumidores, promoviendo en paralelo la innovación.

El Reglamento de Máquinas complementaría el Reglamento de inteligencia artificial propuesto de 21 de abril de 2021, en la medida que este aborda los riesgos para la seguridad de los sistemas de inteligencia artificial que asuman las funciones de seguridad en las máquinas, conforme analizaré en el capítulo IV, mientras que aquél se focalizará en la integración de sistemas inteligentes, al objeto de garantizar, en su caso, la integración segura del sistema de inteligencia artificial en la maquinaria en general, al objeto de no poner en peligro la seguridad de la máquina en su conjunto.

Conforme se ha anunciado, el futuro Reglamento de Máquinas contemplará algunos requisitos adicionales en materia de seguridad al objeto de garantizar la integración segura de la inteligencia artificial en la maquinaria a nivel general. En paralelo, pretende simplificar la documentación necesaria para obtener los permisos para vender en la UE.

El Reglamento de Máquinas contemplará igualmente normas de clasificación aplicables a las máquinas de alto riesgo y una evaluación de conformidad de la maquinaria que se haya modificado de forma sustancial.

La coordinación y complementariedad de ambos Reglamentos facilitará los requerimientos para la industria, en la medida que únicamente se efectuará una evaluación de conformidad para ambos, conforme expondré con mayor detalle al analizar la Propuesta de Reglamento de inteligencia artificial en el capítulo IV, reduciendo a su vez la carga administrativa y financiera a los fabricantes y permitiendo la utilización de un soporte o formato digital para las instrucciones o declaración de conformidad. Asimismo,

en caso de PYMES, también podrán solicitar una adaptación de las tasas cuando se necesite un tercero para la evaluación de conformidad de la máquina.

El marco actual de seguridad de los productos y maquinaria no contempla disposiciones específicas en materia de responsabilidad, sin perjuicio de la aplicación de los marcos reguladores de la responsabilidad del fabricante. Esto no debería suponer un problema si el producto es defectuoso, como se desprende de las directivas europeas vigentes, por ejemplo, la Directiva de Máquinas o la Directiva de Seguridad General de los Productos. El problema, entre otros, es que en estas normas no hay criterios de seguridad específicos para los productos que integre *software* o sistemas inteligentes, conforme abordaré en el capítulo V.

Por otra parte, a nivel internacional son múltiples las acciones orientadas a fijar bases para la seguridad de los productos y, en particular, relacionadas con el funcionamiento y uso de la inteligencia artificial, especialmente de consumo. A modo de ejemplo, U.S. Consumer Products Safety Commission ("CPSC") publicó el 19 de mayo de 2021 un informe sobre cómo asegurar la seguridad en productos al consumo basados en soluciones de inteligencia artificial y *Machine Learning*, bajo el título *Artificial Intelligence and Machine Learning in Consumer Products*[237].

5.6. Seguridad de los sistemas inteligentes

La seguridad exigible a los sistemas de inteligencia artificial no se haya expresamente regulada jurídicamente, sin perjuicio de las exigencias generales de seguridad exigidas por los marcos generales precitados respecto de productos y máquinas, así como, especialmente, en los marcos reguladores europeos en materia de privacidad y ciberseguridad.

[237] *Artificial Intelligence and Machine Learning in Consumer Products.* U.S. Consumer Products Safety Commission. 19.05.2021. Recuperado de: https://cpsc.gov/s3fs-public/Artificial-Intelligence-and-Machine-Learning-In-Consumer-Products.pdf?rKrx2abZ.EbrnFBga0j583KS0KX81Uxh. Consultado el 30.05.2021.

No obstante, la precitada Propuesta de Reglamento del Parlamento Europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión, de 21 de abril de 2021, regula con carácter general los requisitos de ciberseguridad de los sistemas considerados conforme al mismo de alto riesgo, en particular en sus Considerandos 43, 49 y 51, y en sus artículos 13 y 42, en relación con el Reglamento (UE) 2019/881 del Parlamento Europeo y del Consejo, de 17 de abril de 2019, relativo a ENISA (Agencia de la Unión Europea para la Ciberseguridad) y a la certificación de la ciberseguridad de las tecnologías de la información y la comunicación y por el que se deroga el Reglamento (UE) n.º 526/2013 (Reglamento sobre la Ciberseguridad)[238].

El artículo 13 del precitado Reglamento propuesto establece que las instrucciones preceptivas que deben acompañar a los sistemas inteligentes de alto riesgo, deben informar del nivel de precisión, solidez y ciberseguridad regulado en el artículo 15 del mismo, con respecto al cual debe haberse probado y validado el sistema inteligente, así como que puede esperarse de este y las circunstancias conocidas o previsibles que podrían afectar al nivel de precisión, solidez y ciberseguridad esperado.

Por su parte, su artículo 15 del Reglamento propuesto se limita a regular de manera general los requerimientos de ciberseguridad de los sistemas de inteligencia artificial de alto riesgo, de modo que sean diseñados y desarrollados de manera que, en atención a su finalidad, alcancen un nivel "adecuado" de precisión, solidez y ciberseguridad, y funcionen de manera consistente con este objetivo durante todo su ciclo de vida.

Ese nivel adecuado, al igual que el RGPD, deberá definirse en las evaluaciones de riesgos previas, que deberá determinar el nivel de seguridad a aplicar en función de los riesgos identificados y contexto.

Estas disposiciones las complementa en su inciso final, estableciendo que las soluciones técnicas encaminadas a garantizar la ciberseguridad de los sistemas inteligentes de alto riesgo serán de nuevo las "adecuadas" a las circunstancias y los riesgos pertinentes, es decir, al contexto. Y, entre las soluciones técnicas orientadas a subsanar vulnerabilidades,

[238] DO L 151 de 7.6.2019

en función el contexto, recoge las medidas para prevenir y controlar los ataques que traten de manipular el conjunto de datos de entrenamiento -contaminación de datos-, los datos de entrada diseñados para hacer que el modelo cometa un error -ejemplos adversarios- o los defectos en el modelo.

El precepto citado adiciona que los sistemas inteligentes de alto riesgo deberán ser resistentes a los errores, fallos e incoherencias que pueden surgir en los propios sistemas o en el entorno donde operan, en particular, a causa de su interacción con personas físicas u otros sistemas. Y, del mismo modo, exige que sean resistentes a los intentos de terceros no autorizados de alterar su uso o funcionamiento aprovechando las vulnerabilidades del sistema.

Asimismo, el artículo 42.2 del Reglamento propuesto regula una presunción relativa sobre los sistemas de inteligencia de alto riesgo que hayan sido certificados o para los que se haya expedido una declaración de conformidad, con arreglo a un esquema de ciberseguridad en virtud del precitado Reglamento (UE) 2019/881 del Parlamento Europeo y del Consejo y cuyas referencias se hayan publicado en el Diario Oficial de la Unión Europea, los cuales se considerará que cumplen los requisitos de ciberseguridad establecidos en el artículo 15 precitado, en la medida en que el certificado de ciberseguridad o la declaración de conformidad, o partes de estos, prevean estos requisitos

Por último, significar que estos requerimientos únicamente son exigibles a los sistemas de inteligencia artificial calificados conforme al mismo de alto riesgo, así como a los sistemas prohibidos excepcionados en su caso, pero no respecto del resto.

No obstante, como luego abordaré, la Agencia de la Unión Europea para la Ciberseguridad -ENISA- ha publicado distintos informes, a los que aludiré posteriormente, destacando el relativo a los retos de la inteligencia artificial publicado bajo el título *AI Cybersecurity Challenges.*

5.7. Otros marcos

Y todo ello, sin perjuicio de la seguridad y protección que pueden exigir y garantizar los propios fabricantes y proveedores de sistemas de inteligencia artificial en virtud de estándares, códigos tipo y códigos de buenas prácticas, corporativos, sectoriales o generales a los que pueda estar adherido, o que pueda exigirse a los fabricantes y proveedores de sistemas de inteligencia artificial a nivel contractual, especialmente a través de requerimientos contractuales específicos y acuerdos de nivel de servicio o ANS -*Service Level Agreement* o SLA por sus siglas en inglés-.

En definitiva, deberán considerarse las tres dimensiones de requerimientos en materia de seguridad y protección para la evaluación de los riesgos de un sistema de inteligencia artificial, el corporativo, el legal y el contractual, para determinar su clasificación y nivel de riesgo resultante, para decidir de manera previa y consciente su uso y aplicación, sin perjuicio de que la futura regulación de los marcos de seguridad específicos de los mismos faciliten la calificación del riesgo y decisión, especialmente ante la exigencia de mecanismos de evaluación, certificación y de seguridad.

5.8. Retos y riesgos de seguridad asociados a sistemas de inteligencia artificial

La inteligencia artificial y la seguridad -y estrictamente la denominada "ciberseguridad"- se hayan estrechamente relacionadas, en la medida que la inteligencia artificial puede servir para garantizar o mejorar la segunda, especialmente mediante medidas y controles preventivos, detectivos y reactivos de gestión, contención y mitigación, así como para atentar con la misma con absoluta precisión, personalización y efectividad.

La ciberseguridad puede reforzar el despliegue y aplicación de la inteligencia artificial pero también impactar negativamente en la misma.

La información se ha convertido en uno de los principales activos de cualquier persona u organización y objeto de deseo para el lado más oscuro del llamado "ciberespacio". Personas, organizaciones, gobiernos y máquinas se relacionan e interaccionan en un

espacio virtual que no conoce de fronteras y donde no existe un marco legal único que defina sus reglas[239].

Conforme anticipé anteriormente, la seguridad digital no puede desligarse de la seguridad física, dada la convergencia cada vez mayor entre el mundo físico y el virtual y su fusión, a lo que sin duda contribuirá en mayor medida la inteligencia artificial. La seguridad lógica o informática tiene impacto en el mundo físico para personas, instalaciones -especialmente alto en infraestructuras críticas y servicio esenciales-, empresas y gobiernos.

En 2010 fuimos testigos como un virus informático (gusano) denominado "Stuxnet"[240] fue introducido en los sistemas de una central nuclear en Natanz (Irán)[241] para paralizar la misma, consiguiéndolo, y ello utilizando el eslabón más débil en la seguridad, el humano.

El medio utilizado fue digital si bien, el impacto fue físico y el medio de entrada del virus en las instalaciones también -mediante un USB "perdido" en el exterior de la planta, encontrado y conectado indebidamente a sus sistemas informáticos por su propio personal-. Se vio comprometida la seguridad física de personas, instalaciones y cosas, así como la lógica de sus sistemas. Este ataque se produjo en el contexto de tensiones diplomáticas internacionales por el desarrollo del programa nuclear iraní que se vio retrasado casi dos años. Posteriormente se extendió a otros países como Indonesia, India, Paquistán, Rusia, Cuba e incluso EE.UU.

---

[239] MUÑOZ VELA, J.M. (2019). "Los retos del Derecho en una Sociedad Digital", en GIMÉNEZ, I. (Coord.). *Retos de la sociedad digital y medios de pago*. Colección Tratados y Manuales de Economía. Editorial Civitas - Thomson Reuters Aranzadi. Navarra 2019. Pp. 23-24.

[240] Symantec, «W32.Stuxnet Dossier», febrero de 2011. Recuperado de: http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/w32_stuxnet_dossier.pdf. Consultado el 15.03.2021.

[241] Publicado en *BBC News* el 11.10.2015. Recuperado de: https://www.bbc.com/mundo/noticias/2015/10/151007_iwonder_finde_tecnologia_virus_stuxnet#:~:text=Seg%C3%BAn%20la%20firma%20de%20seguridad,sistema%20inform%C3%A1tico%20de%20la%20planta. Consultado el 17.02.2021.

Idéntico ataque con impacto en el mundo físico lo sufrió una fábrica de acero alemana[242], pero los ataques han sido incesantes, antes y después.

En 2003 el ataque denominado "Titan Rain"[243] duró 3 años y tuvo como objetivo múltiples objetivos gubernamentales y empresariales de EE.UU. y Reino Unido, incluyendo también a la NASA, el FBI o el Departamento de Defensa en EE.UU. El gobierno estadounidense acusó al gobierno chino de estar detrás del mismo.

En 2007 Estonia fue objeto de un ataque de Denegación de Servicio Distribuida (DDoS) dirigido a la caída de sistemas y servicios clave, incluyendo servicios públicos, teléfonos móviles y tarjetas bancarias.

El ciberataque en Ucrania en 2014[244] llevado a cabo durante las protestas prorrusas y la crisis de Crimea anuló los sistemas de comunicaciones, dejando al país aislado del mundo exterior. Un año después, en 2015, otro ataque dejó sin servicio diversas centrales eléctricas del país. Estos ataques no generaron exclusivamente daños informáticos sino físicos, psicológicos y económicos.

En 2017 un ataque de *ransomware* etiquetado como "Wannacry"[245] mediante un gusano "bautizado" con el mismo nombre infectó, más de 230.000 computadoras en más de 150 países.

---

[242] *Die Lage der IT-Sicherheit in Deutschland 2014*. (El estado de la seguridad informática en Alemania 2014, título del informe traducido por el autor). Bundesamt für Sicherheit in der Informationstechnik. Publicado en noviembre de 2014. Recuperado de: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2014.pdf?__blob=publicationFile. Consultado el 17.02.2021.

[243] SÁNCHEZ, F. Y LÓPEZ, J. (2017). "Cooperación público-privada en la protección de infraestructuras críticas". Cuadernos de Estrategia 185. *Ciberseguridad: La cooperación público-privada*. Instituto Español de Estudios Estratégicos. Departamento de Seguridad Nacional (DSN). Ministerio de Defensa. Marzo 2017. Pp. 179-180

[244] CASTELLÓN, J. Y LÓPEZ, M.M. (2017). "Crisis y ciberespacio: Hacia un modelo integral de respuesta en el Sistemas de Seguridad Nacional". Cuadernos de Estrategia 185. "*Ciberseguridad: La cooperación público-privada*". Instituto Español de Estudios Estratégicos. Departamento de Seguridad Nacional (DSN). Ministerio de Defensa. Marzo 2017. P. 76.

[245] *WannaCry: Ransomware Attacks Show Strong Links to Lazarus Group*. Symantec. 22.05.2017. Recuperado de: https://www.symantec.com/connect/blogs/wannacry-ransomware-attacksshow-strong-links-lazarus-group. Consultado el 21.01.2021.

También en 2017 se produjo un ciberataque masivo a Equifax[246], una de las tres mayores compañías estadounidenses de información sobre solvencia crediticia, mediante la explotación de una vulnerabilidad de una aplicación para acceder a datos de los clientes, produciendo una brecha de seguridad que afectó a más de 143 millones de estadounidenses (el 44% de la población del país) así como a información sensible de un número no determinado de consumidores canadienses y británicos. La brecha fue reconocida públicamente por la empresa. La extrema gravedad de este ataque no fue sólo por el número de usuarios afectados, inferior al ataque masivo que sufrió previamente Yahoo!, sino por los datos a los que los ciberdelincuentes tuvieron acceso que incluían número de la seguridad social, fecha de nacimiento y dirección -datos que facilitan la suplantación de identidad- pero también números de tarjetas de crédito y otros datos personales.

Y por el camino los ciberataques contra la reputación, la democracia, la influencia política, el ciberterrorismo, ataques distribuidos de denegación de servicio (DDos), mediante redes *botnet* (como "Mirai atack"), ataques de ransomware, ataques silenciosos e incesantes de ciberespionaje mediante virus, especialmente para acceder y sustraer información confidencial, como "Moonlight Maza", "Titan Rain", "Duqu", "Flame", "Red October" o "Gauss"[247], entre muchos otros.

Según Robert Mueller, Director del FBI, solamente hay dos tipos de empresas: Aquellas que han sido hackeadas y aquellas que lo serán.

El aumento cuantitativo y cualitativo de los ataques es exponencial. A modo de ejemplo, CISCO informó en 2018[248] que se bloquearon siete billones de amenazas en nombre de sus clientes.

[246] PRIETO, M. (2017). "Equifax reconoce un ciberataque masivo que afecta a 143 millones de clientes". *Expansión* 08.09.2017. Recuperado de: https://www.expansion.com/economia-digital/companias/2017/09/08/59b23dd822601dc97c8b4655.html. Consultado el 01.02.2021.
[247] CASTELLÓN, J. Y LÓPEZ, M.M. (2017). "Crisis y ciberespacio: Hacia un modelo integral de respuesta en el Sistemas de Seguridad Nacional". Cuadernos de Estrategia 185. *Ciberseguridad: La cooperación público-privada*. Instituto Español de Estudios Estratégicos. Departamento de Seguridad Nacional (DSN). Ministerio de Defensa. Marzo 2017. P. 77.
[248] *Visual Networking Index(VNI) Forecast Highlights Tool.* Cisco. Diciembre 2018

Cybersecurity Ventures[249] ha calculado que los daños por delitos cibernéticos costarán al mundo 6 billones de dólares anuales para 2021, exponencialmente más que el daño infligido por desastres naturales en un año, y más rentable que el comercio mundial de todas las principales drogas ilegales combinadas.

Según la misma, los ataques de *ransomware* experimentaron un aumento del 350 por ciento en 2018 y predice que los costos globales de daños por *ransomware* alcanzarán los 20 mil millones de dólares en 2021. Asimismo, las 10 mayores violaciones de datos de todos los tiempos, en función del número de cuentas pirateadas y el año en que se produjeron, fueron: Yahoo, 3 mil millones (2013); Marriott, 500 millones (2014-2018); Adult FriendFinder, 412 millones (2016); MySpace, 360 millones (2016); Under Armour, 150 millones (2018); Equifax, 145,5 millones (2017); eBay, 145 millones (2014); Objetivo, 110 millones (2013); Heartland Payment Systems, más de 100 millones (2018); LinkedIn, 100 millones (2012).

Del mismo modo, durante los últimos años se está produciendo un incesante aumento del *phishing*[250] y lo seguirá haciendo en lo sucesivo, especialmente el denominado *spear phishing* o *phishing* selectivo, orientado a destinatarios específicos. Y en particular, dentro de éste, está siendo especialmente significativo el aumento de dos modalidades:

De un lado, el denominado "*whaling*" o "*fraude del CEO*", orientado a altos cargos y dirigido a obtener información confidencial de una organización o dinero y, de otro, el denominado "*spear phishing basado inteligencia artificial*"[251], en la que los delincuentes se benefician de estos sistemas para rastrear las redes y para encontrar información útil sobre la persona a suplantar, analizarla e imitar su lenguaje, estilo de comunicación o el timbre de voz.

[249] *2019 Cybersecurity Almanac: 100 Facts, Figures, Predictions and Statistics*. Cybersecurity Ventures, Febrero 2019.
[250] El *phishing* es una técnica de ingeniería social dirigida a obtener de forma ilegítima y mediante engaño datos confidenciales de una persona.
[251] MUÑOZ VELA, J.M. (2021) "Las estrategias delictivas en el ciberespacio se perfeccionan". Artículo de opinión publicado en *Valencia Plaza*. Recuperado de: https://valenciaplaza.com/estrategias-delictivas-ciberespacio-ciberataques. Consultado el 17.02.2021.

El objetivo final es obtener una cantidad de dinero o información confidencial de una entidad, dirigiendo una orden o mandato con esta finalidad a un subordinado, suplantando a la persona y falsificando las comunicaciones y sus contenidos.

De hecho, la inteligencia artificial se está utilizando desde hace años para incluso suplantar la voz de una persona y cometer fraudes mediante la misma.

En marzo de 2019, unos delincuentes utilizaron software de inteligencia artificial para exigir telefónicamente una transferencia fraudulenta por importe de 220.000€ a un director ejecutivo (CEO) de una empresa energética en Reino Unido, y ello imitando la voz del director ejecutivo de su matriz alemana, quien le pidió que la efectuara con carácter urgente y dirigida a un proveedor húngaro[252].

De este modo, los ciberdelincuentes han visto en la inteligencia artificial la herramienta perfecta para llevar a cabo sus actos a nivel mundial, especialmente ante sus distintas capacidades que les permiten, entre otras cosas, realizar una labor previa de recopilación de toda aquella información que pueda ser relevante para los mismos, como dirección de correo electrónico en directorios corporativos, redes sociales, antivirus de la empresa, estilos, perfiles de escritura, etc, así como para, posteriormente, llevar a cabo la suplantación mediante una comunicación y contenido falso.

Del mismo modo, para este mismo fraude empiezan a utilizar sus capacidades más evolucionadas, especialmente los *deep fakes* comentados anteriormente, simulando la imagen y/o la voz de un alto cargo de una entidad para ordenar o convencer a un subordinado para que facilite determinada información confidencial o haga una transferencia dineraria a una cuenta bancaria cuyo supuesto titular suele ser un proveedor real de la entidad, para vestir de la mayor credibilidad la acción.

Y a todo ello contribuye la enorme cantidad de datos que las personas comparten en Internet, incluyendo direcciones electrónicas, imagen o incluso voz, lo que facilita la labor de estas organizaciones criminales, dado que ya no se trata de meros individuos que

[252] STUPP, C. (2019). "Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case". *The Wall Street Journal*. 30.08.2019. Recuperado de https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402. Consultado el 11.03.2021.

actúan independientemente, sino redes perfectamente organizadas a nivel internacional, algunas de las cuales operan como verdaderas multinacionales bajo un modelo de CaaS o "*Crime as a Service*".

Por todo ello, considero que el uso de la inteligencia artificial como instrumento o medio para la comisión de actos ilícitos o delictivos no va a dejar de crecer en los próximos años, especialmente no sólo mediante el análisis de comportamiento de los sistemas, sino, sobre todo, es previsible que se focalice en "*la parte más débil de la muralla"*, esto es, el ser humano, analizando sus patrones de comportamiento para llevar a cabo las acciones ilegítimas, ilícitas o delictivas que permitan franquearla.

Pero de nuevo, recordemos que no todo lo ilegítimo es ilícito, ni todo lo ilícito es delito, en función del ordenamiento jurídico o jurisdicción que resulte de aplicación en actos cometidos en el ciberespacio.

La inteligencia artificial y la ciberseguridad se hayan pues estrechamente relacionadas en distintas dimensiones y distintas interdependencias y, de manera consecuente, la responsabilidad derivada de los daños causados por errores, defectos, vulnerabilidades o ataques relacionados con sistemas inteligentes.

La inteligencia artificial tiene un impacto cuantitativo y cualitativo en las amenazas y riesgos de ciberseguridad.

De un lado puede ampliar el número de amenazas existentes, de otro, modificar sus características y potencial lesivo y, por último, introducir nuevas amenazas y desconocidas en número y en características[253].

La inteligencia artificial también puede ampliar el tipo y número de actores que pueden llevar a cabo usos maliciosos, la velocidad con la que pueden llevar a cabo sus

[253] PATEL, A. ET AL. (2019). "Security Issues, Dangers and Implications of Smart Information Systems*"*. *Sherpa Project* D1.3. 2019. Recuperado de: https://dmu.figshare.com/articles/D1_3_Cyberthreats_and_countermeasures/7951292. Consultado el 02.02.2021.

actividades, su opacidad y el número de personas afectadas. Y también el grado de probabilidad y la gravedad del impacto de una amenaza.

Por todo ello, la ciberseguridad debe responder adecuadamente a los nuevos retos y amenazas sobre la base de una estrategia basada en los pilares tradicionales, pero apoyada por la inteligencia artificial. Estos pilares son cuatro:

a) La prevención, que debe incluir políticas, normas y procedimientos, concienciación, formación, control de acceso y de autorización, medidas técnicas, organizativas y legales;

b) La detección, que debe incluir monitorización continua, identificación y gestión de vulnerabilidades;

c) La respuesta, que incluye sistemas de recuperación y aplicación de contramedidas, esto es, nuevas medidas para evitar la reproducción de la brecha;

d) La inteligencia y compartición, es decir compartir la información con empresas e instituciones conocer los ataques y mejorar las respuestas.

Actualmente, existen diversos estudios publicados sobre identificación de amenazas en el ámbito de la ciberseguridad y la inteligencia artificial, entre otros, el llevado a cabo por ENISA al que luego haré referencia, si bien, los expertos destacan que existe una amplia gama de posibles explotaciones maliciosas que aún no se han explorado por completo[254].

La eficacia, escalabilidad y adaptabilidad de la inteligencia artificial puede ser usada con estas finalidades aumentando enormemente su potencial lesivo.

La inteligencia artificial se ha convertido en esencial para la ciberseguridad, si bien, no es la única tecnología que promete cambiar la ciberseguridad, dado que otras tecnologías y técnicas evidencian una capacidad potencial de transformación y afectación similar, como la computación cuántica o el cifrado *homomórfico*, y más todavía cuando la propia

[254] BONFANTI, M. E. (2020). *Artificial Intelligence and Cybersecurity: A Promising but Uncertain Future*. ARI 139/2020. Real Instituto Elcano. 09.12.2020, P. 5.

inteligencia artificial interaccione con estas tecnologías de una manera que hoy es imposible predecir con exactitud.

Por otra parte, la inteligencia artificial puede utilizarse tanto para usos defensivos como ofensivos y no necesariamente ilegítimos, generando lo que se denomina *ciberinteligencia*, es decir, conocimientos procesables para apoyar la toma de decisiones en cuestiones relacionadas con el ciberespacio y la seguridad.

Como significa Bonfanti[255], la inteligencia artificial es capaz de integrar varias funciones del proceso de *ciberinteligencia*, en particular, la recopilación, el procesamiento y el análisis de la información, puede potenciar la recogida de información y ampliar su alcance a múltiples fuentes y varios puntos finales, puede mejorar la selección de la información y corroborarla con datos adicionales proporcionados por otras fuentes, y también puede apoyar el análisis encontrando patrones y correlaciones ocultas en los datos procesados.

Según este experto, desde un punto de vista operativo, la inteligencia artificial podría utilizarse para recuperar y procesar los datos recogidos por los programas de análisis de seguridad de la red y correlacionarlos con otra información disponible.

Desde el punto de vista táctico, en mi opinión, al igual que la de expertos como el precitado, la inteligencia artificial apoyará cada vez más la detección, el análisis y, posiblemente, la prevención de *ciberamenazas*, mejorará los sistemas de detección de intrusiones y los sistemas de detección de *spam* y *phishing*, así como con las herramientas de detección y análisis de *malware*.

Los sistemas inteligentes también integrarán sistemas de autenticación o verificación multifactoriales que ayudarán a detectar un patrón de comportamiento de un usuario concreto para identificar cambios en esos patrones. Y también podrán ser aplicados en

[255] GALYARDT A. ET AL. (2019). *Artificial Intelligence and Cyber Intelligence: An Implementation Guide*. 2019. Recuperado de: https://resources.sei.cmu.edu/asset_files/EducationalMaterial/2019_011_001_548767.pdf. Consultado el 04.02.2021.

defensiva mediante pruebas de vulnerabilidad automatizadas -*fuzzing*-. El reto que plantean estas aplicaciones son los falsos positivos y negativos.[256]

En mi opinión, la ciberseguridad no solamente permite a las organizaciones proteger sus activos críticos, su actividad y negocio, sino incluso, en ocasiones, su propia existencia. Además, asegura su reputación e imagen en el ámbito interno y externo, como socios, inversores, empleados, proveedores, clientes y frente al público en general. También le ayuda a cumplir los requerimientos reguladores y le evita incurrir en responsabilidades. Pero también puede ser un factor diferencial para proporcionarle una ventaja competitiva en el mercado y en el ciberespacio, proveyendo soluciones seguras y confiables.

En consecuencia, es necesaria una "Ciberseguridad para la inteligencia artificial" o *Cybersecurity for AI*, mediante la adopción de buenas prácticas de ciberseguridad, así como la promoción de amplios programas de *ciberhigiene* con requisitos específicos para la investigación, el desarrollo y la aplicación de la inteligencia artificial[257].

Por el momento, la reciente Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, que analizaré en el capítulo IV, regula la exigencia de ciberseguridad, así como la existencia de sistemas de gestión de riesgos, pero exclusivamente para los sistemas de inteligencia artificial calificados conforme al mismo de alto riesgo, pero no respecto del resto.

Su artículo 9 regula con detalle cómo deben ser los sistemas de gestión de riesgos y su artículo 15 regula los requisitos de precisión, solidez y ciberseguridad que deben integrar estos sistemas en su diseño y desarrollo como durante su desempeño. El precepto exige su resiliencia y su seguridad, especialmente ante intentos de alteración por parte de terceros no autorizados.

---

[256] XIN Y. ET AL. (2018). *Machine Learning and Deep Learning Methods for Cybersecurity*. IEEE. 15.05.2018. Disponible en: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8359287. Consultado el 03.03.2021.

[257] SPRING, J.M. ET AL. (2019). "Machine Learning In Cybersecurity: A Guide". *SEI-CMU Technical Report*. Nº 5, Software Engineering Institute-Carnegie Mellon University, September 2019. P. 11. Recuperado de: https://resources.sei.cmu.edu/asset_files/TechnicalReport/2019_005_001_633597.pdf. Consultado el 02.02.2021.

La Propuesta de Reglamento no define las medidas de seguridad concretas que deben cumplir, sino que, partiendo de un enfoque de riesgos y siguiendo la técnica utilizada en instrumentos jurídicos precedentes como el Reglamento General de Protección de Datos (RGPD), establece que las soluciones y medidas deberán ser desde el diseño y las adecuadas a las circunstancias y riesgos, las cuales deberán incluir cuando apliquen, conforme al precepto citado, medidas de prevención y control de ataques que intenten manipular el conjunto de datos de entrenamiento, las entradas diseñadas para hacer que el modelo cometa un error o los defectos del modelo.

En cualquier caso, hecho este inciso y prosiguiendo con mi análisis, de lo expuesto hasta este momento, considero conveniente aglutinar y diferenciar las cuatro dimensiones de interrelación entre ciberseguridad e inteligencia artificial: a) La ciberseguridad para la inteligencia artificial; b) La inteligencia artificial para apoyar la ciberseguridad; c) Uso malintencionado de la inteligencia artificial; d) La inteligencia artificial como objeto de ciberataques.

a) La ciberseguridad para la inteligencia artificial

La ciberseguridad tendría como objeto proporcionar fiabilidad y robustez a los sistemas inteligentes, especialmente mediante una adecuada protección de datos, validación de procesos, cadena de suministro de software o del rendimiento, y evitar vulnerabilidades de los modelos y algoritmos de inteligencia artificial, por ejemplo mediante la inferencia o manipulación de modelos por terceros, manipulación y alteración de los datos utilizados por los sistemas o ataques contra sistemas *ciberfísicos* impulsados por inteligencia artificial.

b) La inteligencia artificial para apoyar la ciberseguridad

La inteligencia artificial sería un medio o instrumento para crear una ciberseguridad avanzada mediante medidas y controles de seguridad más eficaces, por ejemplo cortafuegos activos y dinámicos, antivirus inteligentes, operaciones automatizadas de inteligencia sobre amenazas cibernéticas -*Cyber Threat Intelligence* o CTI por sus

siglas en inglés-, *fuzzing*[258] de inteligencia artificial, análisis forense inteligente, escaneo de correo electrónico, análisis automatizado de malware, ciberdefensa automatizada o *sandboxing* adaptativo.

Y todo ello facilitando los esfuerzos de las fuerzas y cuerpos de seguridad del Estado y otras autoridades en sus actuaciones preventivas, detectivas, de investigación y de persecución, especialmente mediante el uso del *Big data* generado para estas tareas y con respeto de los nuevos marcos regulatorios precitados en el ámbito español y de la UE para el tratamiento de datos con esta finalidades, en particular, la Directiva 2016/680 UE, del Parlamento Europeo y del Consejo, de 27 de abril, que regula la protección de las personas físicas respecto al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención investigación, detección y enjuiciamiento de infracciones penales, y la norma de tardía transposición de ésta al ordenamiento jurídico español, en particular, la Ley Orgánica 7/2021, de 26 de mayo, de protección de datos personales tratados para fines de prevención, detección, investigación y enjuiciamiento de infracciones penales y de ejecución de sanciones penales.

c) Uso malintencionado de la inteligencia artificial

La inteligencia artificial permite ataques más complejos y mejorados como luego expondré, por ejemplo, de ingeniería social avanzada, *malware* impulsado por inteligencia artificial, generación y mantenimiento dirigido de cuentas falsas en redes sociales, modelos generativos profundos para crear datos falsos, descifrado de contraseñas o entornos cifrados mediante inteligencia artificial o ataques de denegación de servicio DDoS aumentados mediante inteligencia artificial. La inteligencia artificial tiene enorme potencial para mejorar los ataques tradicionales.

[258] Según el Instituto Nacional de Ciberseguridad de España -INCIBE-, se conoce como *fuzzing* o técnicas de *fuzzing* al conjunto de pruebas de caja negra que permiten descubrir errores en los programas o protocolos mediante la introducción de datos al azar, inválidos y malformados. El fin último de estas pruebas es provocar comportamientos inesperados, como fallos que lleguen a hacer que el dispositivo, aplicación o servicio, dejen de funcionar. Este tipo de pruebas permiten, de forma automática o semiautomática, detectar potenciales vulnerabilidades de una forma rápida. Recuperado de: https://www.incibe-cert.es/blog/fuzzing-y-testing-sistemas-control-industrial. Consultado el 5.02.2021.

d) La inteligencia artificial como objeto de ataques.

Los sistemas inteligentes pueden ser objeto de ataques por medios tradicionales por medios físicos y/o virtuales (ciberataques), así como mediante otros sistemas de inteligencia artificial.

Y a todo ello debo de añadir una dimensión adicional transversal, que es la inteligencia artificial para la seguridad nacional de un Estado, más allá de su utilidad para descubrir vulnerabilidades de *software* útiles para *hackear* sistemas informáticos y para dotar de capacidades de defensa a entidades privadas para detectar códigos maliciosos en sus redes. Sobre esta cuestión me permito significar las reflexiones de Buchanan en su informe publicado bajo el título *A National Security Research Agenda for Cybersecurity and Artificial Intelligence*[259].

Estas dimensiones están relacionadas con distintos usos que serán objeto de análisis por mi parte a continuación: La inteligencia artificial como medio o instrumento para la comisión de actos ilícitos o delictivos, como objeto de actos ilícitos o delictivos y como instrumento para la mejora de la ciberseguridad y contra usos maliciosos.

5.8.1. La inteligencia artificial como medio o instrumento para la comisión de actos ilícitos o delictivos.

El uso malicioso de la inteligencia artificial puede clasificarse de múltiples formas. En función del tipo de norma de la que se aleje o incumpla, dicho uso puede clasificarse como ilegítimo, ilícito o delictivo. No todo lo ilegítimo e inmoral es ilegal, ni todo lo ilegal es delictivo.

A modo de ejemplo, hasta fechas relativamente recientes el *hacking* no era un delito en España. En función de la seguridad y dimensión que se pueda ver afectada, estos usos pueden considerarse acciones o ataques contra la seguridad informática, lógica o digital

[259] BUCHANAN, B. (2020). *A National Security Research Agenda for Cybersecurity and Artificial Intelligence.* Center for Security and Emerging Technology -CSET-. Mayo 2020.

(ciberataques), contra la seguridad física, contra la seguridad psicológica o emocional, la seguridad jurídica o contra la propia seguridad política.

Sobre el uso de la inteligencia artificial con finalidades maliciosas, me permito significar el estudio denominado *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*[260], publicado en febrero de 2018.

El informe fue redactado por 26 autores pertenecientes al mundo académico, la sociedad civil y la industria, en el que participaron entidades como el *Future of Humanity Institute* de la Universidad de Oxford y el *Centre for the Study of Existential Risk* de la Universidad de Cambridge, entre otras entidades.

El informe analiza el panorama de las posibles amenazas a la seguridad derivadas de los usos maliciosos de las tecnologías de inteligencia artificial y propone formas de prever, prevenir y mitigar mejor estas amenazas, focalizándose en los tipos de ataques que probablemente sufriremos, si no se desarrollan mecanismos de defensa adecuados.

De entre sus recomendaciones, destacar las siguientes:

a) Los responsables políticos deberían colaborar estrechamente con los investigadores técnicos para investigar, prevenir y mitigar los posibles usos maliciosos de la inteligencia artificial;

b) Los investigadores e ingenieros en inteligencia artificial deberían tomarse en serio la naturaleza de doble uso de su trabajo, permitiendo que las consideraciones relacionadas con el uso indebido influyan en las prioridades y normas de investigación, y contactando de forma proactiva con los actores relevantes cuando se prevean aplicaciones dañinas;

c) Deben identificarse las mejores prácticas en áreas de investigación con métodos más maduros para abordar las preocupaciones de doble uso, como la seguridad

[260] BRUNDAGE, M. ET AL. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. Febrero 2018. Recuperado de: https://maliciousaireport.com/. Consultado el 18.02.2021.

informática, e importarse cuando sea aplicable al caso de la inteligencia artificial, y;

d) Ampliar el abanico de partes interesadas y expertos en la materia que participan en los debates sobre estos retos.

Del mismo modo, sus autores proponen la exploración de distintas cuestiones y posibles actuaciones consecuentes en distintos aspectos, en particular, aprender de y con la comunidad de la ciberseguridad, explorar diferentes modelos de apertura de la investigación -con nuevos marcos que incorporen análisis de riesgos previos a publicaciones, modelos de licencia o régimen de intercambio-, promover una cultura de la responsabilidad -destacando la importancia de la educación, las normas éticas, las normas y las expectativas- y desarrollar soluciones tecnológicas y políticas, y respuestas legislativas adecuadas.

El informe anticipa algunas de las principales amenazas en materia de seguridad, especialmente conforme las capacidades de la inteligencia artificial aumenten en calidad y cantidad y se generalice su uso.

De un lado, las amenazas ya existentes aumentarán y se expandirán. Los costes de los ataques se reducirán gracias a la escalabilidad de los sistemas e inteligencia artificial y es previsible, de manera consecuente, que aumente el número de actores que puedan llevar a cabo los ataques, el ritmo al que pueden realizarlos y el conjunto de objetivos potenciales.

De otro, se introducirán de nuevas amenazas, ya que es previsible que surjan nuevos ataques mediante el uso de sistemas de inteligencia artificial para realizar tareas que, de otro modo, resultarían imposibles de llevar a la práctica para los seres humanos sin estos medios. Además, los actores maliciosos podrán focalizar sus acciones contra sistemas también de inteligencia artificial para explotar sus vulnerabilidades.

Y, por último, se espera un cambio en las amenazas habituales, con mayor eficacia, con objetivos muy precisos, difíciles de imputar y que probablemente aprovechen las propias vulnerabilidades de los sistemas de inteligencia artificial.

En mi opinión, la inteligencia artificial va a incrementar tanto cuantitativa como cualitativamente los riesgos de ataque, aumentando su eficacia, probabilidad y su impacto.

En el ámbito de la seguridad digital, la inteligencia artificial permite la automatización de múltiples tareas asociadas a un ciberataque, aumentando su escala, eficacia y precisión. Un ejemplo de ello sería el *phishing* selectivo o "*spear phishing*" comentado anteriormente.

A modo de ejemplo, el *malware* denominado "DeelLocker" es capaz de identificar su objetivo utilizando indicadores como el reconocimiento facial y de voz y la geolocalización. Además, según Marc Stoecklin[261], director del grupo de investigación *Cognitive Cybersecutiry Intelligence* de IBM, podría integrase en *software* no malicioso siendo la inteligencia artificial la responsable de determinar la activación de su comportamiento malintencionado.

Del mismo modo, se producirán nuevos tipos de ataques dirigidos a la explotación de las vulnerabilidades propias de la persona (suplantación de voz mediante su síntesis), del *software* y de los sistemas, inteligentes o no, siendo previsible, como refleja el informe precitado, que se orienten a sistemas inteligentes mediante técnicas de ingeniería inversa o viciado de datos.

Un ejemplo de ello sería el *hackeo* y posterior manipulación de un coche autónomo o incluso su simple engaño mediante la mera manipulación de los datos que capta, como ya ha ocurrido en la práctica[262]. Este tipo de ataques ha sido ya descrito por la literatura científica[263]. En este caso, se trata de ataques de la dimensión física a la dimensión lógica.

---

[261] JANOFSKY, A. (2018). "AI Could Make Cyberattacks More Dangerous, Harder to Detect". *The Wall Street Journal*. 13.11.2018. Recuperado de https://www.wsj.com/articles/ai-could-make-cyberattacks-more-dangerous-harder-to-detect-1542128667?mod=article_inline. Consultado el 11.03.2021.

[262] En 2017, investigadores de la compañía de ciberseguridad McAfee consiguieron engañar a un algoritmo de reconocimiento de señales de tráfico utilizando técnicas de aprendizaje automático denominadas GANs. Alteraron con una pegatina una señal de tráfico imperceptible para el ojo humano pero no para la máquina, de modo que en lugar de identificar el límite de seguridad en 35 millas/hora, pasó a identificar 85 millas/hora. Benjamins, R. y Salazar, I. "*El mito del algoritmo*". Op.cit. Pos. 1164.

[263] EYKHOLT, KEVIN; EVTIMOV, IVAN ET ALT. (2018). "Robust Physical-World Attacks on Deep Learning Visual Classification*". CVPR* 10.04.2018. Resultados publicados por Ackerman, E. (2017). "Slight Street Sign Modifications Can Completely Fool Machine Learning Algorithms". Publicado en la revista *IEEE*

En el ámbito de la seguridad física, la inteligencia artificial puede conllevar el aumento de los ataques, especialmente utilizando drones o sistemas físicos como las armas autónomas, incluso, como he expuesto anteriormente, tomando el control de sistemas *ciberfísicos* haciendo que un vehículo autónomo se estrelle o sistemas físicos como un enjambre de *minidrones* que no pueden controlarse manualmente a distancia.

En el ámbito de la seguridad psicológica y emocional de la persona, pueden llevarse a cabo ataques con la estabilidad emocional de las personas provocando cambio de estados de ánimo hasta incluso situaciones que puedan poner en riesgo la vida, especialmente en el caso de sistemas inteligentes estrechamente relacionados con la persona, como los asistenciales y compañía, o con los sentimientos.

A modo de ejemplo, actualmente existen en el mercado aplicaciones que permiten darle vida a fotos de personas queridas ya no están con nosotros, o sistemas de realidad virtual con inteligencia artificial integrada, que permiten crear la ilusión de tener delante a una persona ya fallecida con la que podríamos incluso hablar o incluso tocarla, así como generar videos de la misma donde nos hablan y nos trasladan los mensajes que consideremos oportunos.

En el ámbito de la seguridad jurídica, los sistemas inteligentes podrían ser utilizados o alterados contra la seguridad jurídica de las personas, afectando a sus expectativas de derechos y facultades en el ámbito de las relaciones *off* y *online*, por ejemplo, en el ámbito de la concesión de un préstamo, una prestación, una subvención o la adopción de medidas preventivas o evaluación del riesgo de reincidencia en el ámbito penal.

Y, por último, en el ámbito de la seguridad personal y política, las capacidades de la inteligencia artificial para automatizar tareas relacionadas con la vigilancia -especialmente para el análisis de datos recogidos en masa-, la persuasión -por ejemplo mediante la creación de comunicaciones, mensajes y publicidad dirigida- y el engaño -

Spectrum el 04.08.2017. Disponible en: https://spectrum.ieee.org/cars-that-think/transportation/sensors/slight-street-sign-modifications-can-fool-machine-learning-algorithms. Consultado el 17.02.2021.

por ejemplo y como he referenciado anteriormente, los *deep fakes*-, permiten ampliar las amenazas relacionadas con la vulneración de la privacidad y la manipulación social.

Asimismo, y relacionado con lo anterior, su mayor capacidad para analizar los comportamientos, estados de ánimo y creencias de los seres humanos a partir de los datos disponibles, hace previsible el correlativo aumento de los ataques relacionados con las mismas, especialmente significativos en el contexto de los países autoritarios, pero que afectará a países más democráticos, como ya está sucediendo.

En conclusión, el desarrollo y aplicación de la inteligencia artificial comportará el aumento y expansión de las amenazas existentes, la irrupción de nuevas amenazas y la mejora y transformación de las existentes, perfeccionándolas, haciéndolas más personalizadas, masivas, más efectivas, aumentando de forma exponencial su probabilidad e impacto, por lo que debemos prepararnos ante todo ello, conforme advierte el informe precitado.

Y de todos estos actos ilícitos y, en su caso, delictivos, se exigirán no solamente la responsabilidades administrativas y penales que correspondan sino, en su caso, las responsabilidades civiles derivadas de la infracción administrativa o comisión de un delito.

Las primeras deberán depurarse en procedimiento aparte, las segundas, podrán ser depuradas en el procedimiento penal.

El *Libro Blanco sobre la inteligencia artificial*[264] de la Comisión Europea alude a los esfuerzos adicionales que se requerirán para evitar y combatir el uso abusivo de la inteligencia artificial con fines delictivos, si bien, se considera una cuestión independiente al mismo y no abordada.

Los usos que he comentado se hayan en su gran mayoría contemplados y prohibidos, en mayor o menor medida, en los distintos ordenamientos jurídicos, en muchas ocasiones no sólo como actos ilegítimos o ilícitos, sino como actos constitutivos de infracción

[264] COM (2020) 65 final

administrativa o delito, expresamente previstos en un ordenamiento jurídico-penal como tales.

No obstante, el Derecho Penal es reactivo, de modo que opera cuando ya la conducta se ha llevado a cabo y se ha podido producir la consecuencia y, en su caso, el daño.

Es por ello que se hace necesario potenciar los marcos regulativos de carácter proactivo que contemplen acciones y obligaciones preventivas y detectivas orientadas a evitar que todas estas acciones ocurran y, si ocurren, que lo hagan con la probabilidad más baja o menor impacto posible, para lo que la exigencia de la seguridad proactiva y gestionada, la gestión de riesgos y la seguridad en el diseño y por defecto se erigen en esenciales con esta finalidad.

Esta es la base sobre las que se han construido, entre otros, los marcos europeos y españoles en materia de seguridad de datos personales y de seguridad de operadores de servicios esenciales e infraestructuras críticas a los que he aludido anteriormente, y parece que es la fórmula por la que deberían apostar los nuevos marcos reguladores de la inteligencia artificial, si bien, como veremos, la reciente Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, regula estos aspectos en sus artículos 9 y 15, pero como requisitos para sistemas de inteligencia artificial de alto riesgo, no para el resto, como he manifestado anteriormente.

La mayoría de los ordenamientos jurídicos prohíben sabotear una infraestructura crítica, manipular procesos democráticos, *crackear* sistemas u obtener o tratar datos personales sin legitimación legal, entre otras acciones.

La cuestión es si la inteligencia artificial, por sus características singulares, debería disponer de un marco regulativo específico para evitar usos maliciosos.

En mi opinión, sobre este aspecto específico y considerando la inexistencia actual de una autonomía real, conciencia y libertad plena en los sistemas más avanzados, considero que actualmente no sería necesario un marco específico para evitar dichos usos, sin perjuicio de que el ordenamiento jurídico deba integrarse, dotarse de coherencia y adaptarse a la

realidad tecnológica y social en cada momento bajo un paradigma adaptativo, flexible, evolutivo y *responsive*, al que he aludido en anteriores apartados de esta investigación. Y todo ello, sin perjuicio de la regulación de la inteligencia artificial desde un enfoque de riesgos, entre los que la (ciber) seguridad constituye uno de los mismos.

A mi juicio, los futuros marcos reguladores de la inteligencia artificial deberían exigir análisis y gestión de riesgo, seguridad gestionada, proactiva en el diseño y por defecto a cualquier sistema de inteligencia artificial, cualquiera que sea su nivel de riesgo inicial o simplemente no clasificado, habida cuenta de las características y capacidades de las que puede estar dotado.

En consecuencia, debe llevarse a cabo una deseable armonización en el ámbito de la UE y los ordenamientos jurídicos nacionales, que deben ser integrados y complementados en aquello que sea necesario en determinados aspectos relacionados con los sistemas de inteligencia artificial, en especial en materia de obligaciones éticas, jurídicas, seguridad y responsabilidad asociada a los mismos, en la línea analizada en esta investigación.

5.8.2. La inteligencia artificial como objeto de actos ilícitos o delictivos

Los sistemas inteligentes pueden ser objeto de ataques provenientes del mundo físico, por ejemplo, mediante el engaño como he referido anteriormente, o del virtual, por ejemplo, mediante ciberataques.

Conforme a los distintos marcos éticos y jurídicos anteriormente referidos y que son objeto de análisis en esta investigación, la seguridad es una exigencia no sólo por razones éticas y jurídicas sino por cuestiones de política económica y de incentivos para la innovación y competitividad empresarial.

Los sistemas inteligentes deben ser robustos y fiables de modo que garanticen la seguridad y confiabilidad a la sociedad para propiciar su despegue y aplicación, lo que a su vez redunda en la seguridad para las iniciativas empresariales, la inversión y la innovación para su desarrollo y aplicación.

La consecución de estos objetivos requiere, a mi juicio, no sólo la definición de marcos éticos claros, sino marcos jurídicos exigentes en esta materia, que establezcan bases mínimas de seguridad y que se complementen con la promoción de la autorregulación sectorial complementaria para aplicaciones y sectores concretos que la mejoren y garanticen.

Y todo ello acompañado de la necesaria coordinación entre las estrategias gubernamentales de ciberseguridad, de transformación digital, de inteligencia artificial, de gobierno de los datos y de la tecnologías con este firme propósito, con el objetivo de crear técnicas, metodologías y herramientas que promuevan y garanticen el diseño, desarrollo, evaluación, validación, despliegue y aplicación de sistemas de inteligencia artificial sustentados en la seguridad por en el diseño y por defecto, esto es, la *Security by design and by default*, similar a la privacidad y seguridad exigida por marcos vigentes, como el Reglamento General de Protección de Datos Europeo.

Un sistema de inteligencia artificial no deja de ser un sistema informático o de información que se compone principalmente por *hardware*, *software* e información, bajo modelos y algoritmos de procesamiento. Estos sistemas pueden llevar a cabo distintas funciones, como la captación y procesamiento de datos, más rápidamente y con mayor precisión que cualquier humano.

No obstante, la ausencia de seguridad en su diseño y la mayor innovación de los *ciberatacantes* pueden causar brechas de seguridad explotadas por los mismos con enorme impacto para todos los agentes relacionados con el sistema afectado, que, además, deberán responder unos frente a otros, en función de elemento o componente que puede estar relacionado con la vulnerabilidad explotada y su mayor o menor control del riesgo.

Las consecuencias negativas pueden derivarse tanto para nuestra esfera personal, empresarial como administrativa. Los elementos, aplicaciones, sistemas, productos y servicios que utilizan la inteligencia artificial son comunes en nuestra vida cotidiana en cualquier tipo de ámbito, como he comentado en anteriores apartados.

La concesión de un crédito, el acceso a un empleo o un concreto tratamiento médico son procesos que conllevan la toma de decisiones que pueden verse influenciadas de manera determinante o ser adoptadas por sistemas dotados de inteligencia artificial.

Pero también la circulación por las calles de una *smartcity* donde el tráfico se haya regulado por dichos sistemas y/o en vehículos de transporte o conducción "autónoma" implica también decisiones automatizas en momentos determinados, en base a la información recabada por los sistemas, tratamiento de la misma en base a sus datos de entrenamiento e instrucciones asociadas y su ejecución física o digital, como puede ser poner un semáforo en rojo o en verde para parar o dar paso a la circulación, o acelerar o frenar un vehículo ante un hecho de la circulación. Todo ello comporta indudables riesgos para la seguridad de personas y cosas.

Si hablamos de sistemas de "defensa" o armamentísticos dotados de inteligencia artificial y dotados de autonomía total o parcial, resultan obvios sus riesgos y las posibles consecuencias de un error de programación, de valoración o defecto funcional.

En consecuencia, un sistema de inteligencia artificial, como cualquier otro sistemas informático, puede ser objeto de ataque con origen en vulnerabilidades no detectadas o en herramientas creadas por el atacante con el objetivo de entrar en el mismo, alterar los algoritmos, instrucciones, autonomía, capacidades, datos, condiciones, controles, restricciones, medidas de seguridad, sesgos, impredecibilidad, decisiones y/o acciones, que incluso podrían convertir al sistema en un *zombi* bajo control del *ciberdelicuente* o incluso en arma para causar daños de cualquier naturaleza, incluso físicos.

Pensemos en la posibilidad de *crackear* y tomar el control de un dron o un vehículo autónomo para llevar a cabo un atentado, provocar un accidente o facilitar un robo, o en la mera manipulación de las decisiones de un sistema.

Los ataques a sistemas de inteligencia artificial pueden ser de muy diversa naturaleza, tipología, origen, metodología y finalidad, y pueden llevarse a cabo tanto durante su

entrenamiento como durante su operación como recogen distintos estudios como el proyecto SPARTA[265].

Los ataques a sistemas comerciales de inteligencia artificial han afectado a todo tipo de organizaciones, desde las más pequeñas a las más grandes como Microsoft, Tesla, Google o Amazon, entre otras[266].

La exigencia de una inteligencia artificial segura, sólida y fiable, como referí anteriormente, constituye una exigencia ética y se halla prevista tanto en las recomendaciones del Parlamento Europeo de 20 de octubre de 2020 como en la reciente Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, si bien, este último, como he referido, lo regula como exigencia para sistemas inteligentes considerados de alto riesgo conforme al mismo.

5.8.3. La inteligencia artificial como instrumento contra usos maliciosos

La inteligencia artificial se ha convertido igualmente en una eficaz herramienta de ayuda a los profesionales de la seguridad y para gestionar la misma, especialmente ante la creciente complejidad de los sistemas y redes y tecnologías que integran.

Los gobiernos y entidades del sector público y privado están actualmente inmersos en múltiples iniciativas para utilizar la inteligencia artificial para su ciberseguridad, si bien, el interés no es nuevo en la medida que los gobiernos occidentales y la *Agencia de Proyectos de Investigación Avanzada de Defensa* -DARPA por sus siglas en inglés- llevan apoyando la investigación en materia de inteligencia artificial y redes informáticas de la década de 1960.

[265] Recuperado de https://www.sparta.eu.
[266] AYERBE, A. (2020). *La ciberseguridad y su relación con la inteligencia artificial*. ARI 128/2020. 10.11.2020. Real Instituto Elcano. 2020. P. 4.

La adopción de soluciones basadas en la inteligencia artificial para lograr objetivos *ciberdefensivos* y *ciberofensivos*, esto es, la denominada "IA para la ciberseguridad" o *AI for Cybersecurity*[267] es muy prometedora, pero también supone un gran reto.

La asociación entre inteligencia artificial y *Big data* se ha convertido en una eficaz herramienta para la lucha contra la delincuencia, cibernética o no.

A pesar de los posibles usos maliciosos de los sistemas inteligentes, como he referido anteriormente, la inteligencia artificial constituye una herramienta muy eficaz frente a todas estas nuevas amenazas que, si se sustenta en la ética, la seguridad y el cumplimiento regulatorio en el diseño, junto con la inclusión del aseguramiento de su calidad, análisis y gestión de riesgos, evaluaciones de impacto en su proceso de diseño y desarrollo, y la monitorización y revisión continua, podrá bloquear e incluso adelantarse a las mismas, prevenirlas y detectarlas y, en el peor de los casos, gestionarlas adecuadamente reduciendo su probabilidad e impacto.

La ciberseguridad se enfrenta a múltiples desafíos como la defensa proactiva, la predicción e identificación de comportamientos sospechosos, la detección de amenazas sofisticadas, la detección de intrusiones o la protección de la privacidad, por lo que la posibilidad de disponer de sistemas integrados hombre-máquina que faciliten la predicción y detección temprana, el análisis y la toma de decisiones en tiempo real y la adoptación de medidas de gestión inmediatas e incluso automatizadas, constituye un objetivo estratégico para la gestión de la ciberseguridad.

Los algoritmos de aprendizaje automático y la *ciberinteligencia* asociada están potenciando las predicciones de ataques cibernéticos que mejoran las tasas de detección y pueden ser un aspecto esencial para revertir la tendencia actual de crecimiento incesante de delitos cibernéticos y para potenciar la ciberseguridad.

[267] BONFANTI, M.E. Y KOHLER, K. (2020). "Artificial Intelligence for Cybersecurity". *CSS Analyses in Security Policy*. Nº 265. Center for Security Studies (CSS), ETH Zurich. 2020.

La inteligencia artificial puede ser utilizada en todas las etapas de una seguridad integral inteligente, esto es, en la prevención, la identificación, la detección, la protección, la gestión, la respuesta y la recuperación ante una brecha de seguridad.

La ciberseguridad es un nuevo campo o sector de aplicación de la inteligencia artificial y que, a mi juicio, debería ser reconocido de antemano como "de alto riesgo" en sí mismo y la investigación en su aplicación contra las *ciberamenazas* debería ser incesante a lo largo del todo el ciclo de vida de la gestión y mitigación de los riesgos de seguridad[268] y en paralelo el aumento constante a nivel cuantitativo y cualitativo de las mismas.

Estas previsiones están contempladas expresamente en los artículos 9 y 15 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, que analizaré con detalle en el capítulo IV.

No obstante, no podemos olvidar que un pilar esencial de la ciberseguridad, junto la dotación de estos sistemas, lo constituye la concienciación y educación, por lo que estos sistemas híbridos o totalmente automatizados deberán ser acompañados por la inversión en información, concienciación y formación en el seno de las organizaciones, pero también de la sociedad en general y cada vez más en edades más tempranas, asegurando la misma mediante la inclusión de la seguridad en los planes de estudio, lo que corresponderá a los gobiernos y responsables políticos.

A modo de ejemplo, en relación con una de las prácticas maliciosas más extendidas y comentada anteriormente, la información y concienciación en materia de *phishing* selectivo, a través o no de la inteligencia artificial, podría evitarse en buena medida mediante dicha información, la concienciación y la formación.

Por otra parte, la investigación para la creación de sistemas inteligentes para su aplicación en ciberseguridad es incesante.

---

[268] *Research Topics. From January 2019 to April 2020*. ENISA 2020. Recuperado de: https://www.enisa.europa.eu/topics/threat-risk-management/threats-and-trends/etl-review-folder/etl-2020-cybersecurity-research. Consultado el 02.01.2021.

A modo de ejemplo, IBM ya reorientó una nueva versión de su sistema de inteligencia artificial "Watson", del que he hablado en anteriores apartados, para la lucha contra el cibercrimen bajo el nombre *Watson for Cyber Security*[269]. Para ello, durante 2016 y hasta 2017 fue entrenado para conocer los sistemas de seguridad más avanzados.

El sistema permite la predicción inteligente y la gestión masiva de datos para anticiparse en segundos a rutinas de ataques incluso antes de que se materialicen, lo que era imposible sin este tipo de sistemas. De este modo puede anticiparse al delito y reducir al mínimo el tiempo entre la detección de la comisión de un ciberdelito y la articulación de los mecanismos de respuesta.

La consultora Accenture lanzó en marzo de 2016 una plataforma de *ciberinteligencia*[270] orientada a la detección y monitorización de *ciberamenazas* en tiempo real y que combina servicios de seguridad gestionada, inteligencia artificial, aprendizaje automático, análisis de datos y servicios *cloud*. La plataforma analiza la actividad en las redes para aprender, identificar y reportar comportamientos sospechosos atribuibles a ciberataques.

Otras aplicaciones como "Magnifier" se presentan como un sistema de comportamiento analítico que utiliza aprendizaje automático con datos estructurados y no estructurados para modelar el comportamiento de la red y mejorar la detección de amenazas, o "Chronicle" que ofrece servicios de plataforma inteligente de ciberseguridad[271].

"DeepCode" es una inteligencia artificial desarrollada por ingenieros de la Universidad de Boston que utiliza el *Machine Learning* para encontrar los errores en el código que usan los *hackers*.

[269] JOYANES, L. (2017). "Ciberseguridad: La colaboración público-privada en la era de la cuarta revolución industrial (Industria 4.0 versus ciberseguridad 4.0)". *Cuadernos de Estrategia 185. Ciberseguridad: La cooperación público-privada*. Instituto Español de Estudios Estratégicos. Departamento de Seguridad Nacional (DSN). Ministerio de Defensa. Marzo 2017. P. 44.
[270] Recuperado de: www.accenture.com/es-es/company-news-release-new-cybersecurity-platform. Consultado el 21.02.2021.
[271] LEÓN, G. (2019). "Situación y perspectivas de las tecnologías y aplicaciones de inteligencia artificial". En Documento de Seguridad y Defensa 79. *La inteligencia artificial aplicada a la defensa*. Instituto Español de Estudios Estratégicos. Marzo 2019. P. 57.

Por su parte, Microsoft y MITRE[272], en colaboración con IBM, NVIDIA, Airbus, Bosch, Deep Instinct, Two Six Labs, Cardiff University, University of Toronto, PricewaterhouseCoopers, Software Engineering Institute en Carnegie Mellon University, y el Berryville Institute of Machine Learning, han desarrollado "Adversarial ML Threat Matrix"[273], un marco abierto diseñado para ayudar a identificar, responder y remediar ataques dirigidos a sistemas de aprendizaje automático -*Machine Learning*-.

El objetivo es que cualquier organización pueda usar *Adversarial ML Threat Matrix* para probar la resistencia de sus modelos de inteligencia artificial mediante la simulación de escenarios de ataque realistas, y para ello utiliza una lista de tácticas para obtener acceso inicial al entorno, ejecutar modelos de aprendizaje automático inseguros, contaminar los datos de entrenamiento y extraer información confidencial, a través de ataques de robo de modelos.

La herramienta pretende permitir que los analistas de seguridad puedan orientarse en estas amenazas nuevas y futuras para detectar, responder y resolver las mismas.

Según Gartner[274], para el año 2022, el 30% de todos los ciberataques emplearán técnicas de *Data Poisining* (envenenamiento o contaminación de datos) o la sustracción de muestras y modelos de *Machine Learning* para atacar sistemas basados en inteligencia artificial. La seguridad de la inteligencia artificial tiene tres perspectivas claves para Gartner, según refleja en el estudio citado al pie:

a) La protección de sistemas impulsados por la inteligencia artificial (de los datos de entrenamiento, canales de entrenamiento y modelos de *Machine Learning*);

[272] The MITRE Corporation, conocida comúnmente como MITRE, es una organización estadounidense sin ánimo de lucro constituida en 1958 que provee ingeniería de sistemas, investigación y desarrollo, y soporte sobre tecnologías de la información al gobierno de EE.UU. y tiene entre sus misiones intentar resolver problemas que contribuyan a un mundo más seguro.
[273] Recuperado de https://github.com/mitre/advmlthreatmatrix. Consultado el 02.01.2021.
[274] "*Gartner Top 10 Strategic Technology Trends for 2020*". Gartner Inc. 21.10.2019. Recuperado de: https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020/. Consultado el 02.01.2021.

b) El aprovechamiento de la inteligencia artificial para mejorar la defensa de la seguridad, esto es, usar *Machine Learning* para comprender patrones, descubrir ataques y automatizar partes de los procesos de ciberseguridad, y;

c) Anticipar el uso malicioso de la inteligencia artificial por parte de los atacantes, esto es identificar ataques y defenderse de ellos.

La utilización de la inteligencia artificial para llevar a cabo actos maliciosos puede ir desde ataques menores de impacto leve o poco significativo a casos de *ciberespionaje*, ciberterrorismo o ciberguerra. La variedad de sujetos atacantes, finalidades, metodología, forma de operar y grado de sofisticación y complejidad es amplia y su análisis es apasionante, pero excede del objeto y alcance de esta investigación.

El uso de la misma puede ser inmediatamente ofensivo dirigido a infraestructuras críticas, acceso a sistemas, sustracción o manipulación de información o interceptación de comunicaciones, y también puede utilizarse como medio para romper y obtener contraseñas para que acceso lo lleven a cabo terceros, construir *malware* que evite la detección inmediata, su ocultación, adaptación a contramedidas de reacción, así como incluso para obtener automáticamente información utilizando métodos de procesamiento del lenguaje natural -*Natural Language Processing* o NLP por sus siglas en inglés- y la suplantación y generación de audios, vídeos y textos falsos como he referido anteriormente.

Los atacantes también están utilizando redes generativas antagónicas -*Generative Adversarial Networks* o GAN por sus siglas en inglés-, para imitar patrones de tráfico de comunicaciones normales con el objetivo de distraer la atención de un ataque y encontrar y extraer datos sensibles rápidamente[275].

Los últimos informes llevados a cabo a nivel internacional reflejan como la necesidad de reforzar la seguridad con inteligencia artificial se ha convertido en un imperativo para cualquier organización. El informe *Reinventing Cybersecurity with Artificial*

[275] Ayerbe, A. (2020). *La ciberseguridad y su relación con la inteligencia artificial*. ARI 128/2020. 10.11.2020. Real Instituto Elcano. 2020. P. 4.

Intelligence: The new frontier in digital security[276] de Capgemini Research Institute así lo concluye.

Entre otros aspectos, significa de nuevo el papel de la inteligencia artificial como herramienta de ayuda a los analistas de ciberseguridad y concluye que la respuesta a las *ciberamenazas* mediante inteligencia artificial es la nueva frontera de la ciberseguridad en la medida que "los hackers ya la están utilizando". De los participantes en este estudio, el 69% de las organizaciones consideraban que la inteligencia artificial será necesaria para responder a los ciberataques.

La adopción de la inteligencia artificial en ciberseguridad aumenta incesantemente como destaca el estudio precitado. "Casi una de cada cinco organizaciones utilizaba la IA antes de 2019" y casi dos de cada tres ya planeaban emplear IA para 2020. Entre sus ventajas destaca una capacidad de respuesta más rápida ante brechas, reduce el coste detección, y mejora de la precisión y eficiencia de los *ciberanalistas*.

Del informe precitado me permito destacar un elemento esencial para la creación de una hoja de ruta para la implementación de inteligencia artificial en la ciberseguridad, esto es, la gobernanza de la inteligencia artificial para garantizar los resultados esperados.

De nuevo, se significa la necesaria colaboración hombre-máquina para poder acometer los complejos retos del futuro y que se evidencia igualmente en otros informes internacionales.

Por ejemplo, en materia de ciberinteligencia, el informe *Artificial Intelligence and Cyber Intelligence. An Implementation Guide*[277] de la Carnegie Mellon University, destaca como las herramientas y técnicas de aprendizaje automático -*machine learning*- mejoran eficazmente los flujos de trabajo de la inteligencia cibernética en todo el entorno, la recopilación de datos, el análisis de amenazas, el análisis estratégico y la información a los responsables de la toma de decisiones, si bien, significa que el mero hecho de añadir *Machine Learning* a los conjuntos de herramientas organizativas existentes,

[276] *Reinventing Cybersecurity with Artificial Intelligence: The new frontier in digital security*. Capgemini Research Institute. 2019.
[277] GALYARDT, A. ET AL. (2019). *Artificial Intelligence and Cyber Intelligence. An Implementation Guide*. Carnegie Mellon University. Software Engineering Institute. 20.05.2019

procedimientos y flujos de trabajo existentes no resolverá todos los retos de la *ciberinteligencia*, en la medida que estas tecnologías funcionan con personal experimentado y cualificado que sabe cómo entender, integrar e incluso mejorar los procesos en el contexto de los retos de la *ciberinteligencia*.

En el mismo sentido, merece significar igualmente el informe *Machine Learning in Cybersecurity*[278], también de la Carnegie Mellon University, que se centra en cómo garantizar que las herramientas de *Machine Learning* (ML) sean útiles cuando se apliquen para abordar un problema de ciberseguridad.

Como he analizado anteriormente, el uso de la inteligencia artificial comporta riesgos de distinta naturaleza, tanto para personas, bienes como instalaciones, incluyendo los sistemas de información, especialmente cuando incorpora la toma de decisiones sin supervisión humana y, en cualquier caso, si va asociada o integrada en los denominados sistemas de "alto riesgo".

El marco de seguridad de los productos en la UE ya impone a las empresas la obligación de garantizar que únicamente se comercialicen productos seguros y conformes a los marcos vigentes, si bien, las nuevas aplicaciones que permite la inteligencia artificial, especialmente ante su capacidad de toma de decisiones automatizada y autoaprendizaje, comporta nuevos desafíos y no sólo en su concepción inicial, sino posteriormente a su comercialización ante su evolución, aprendizaje e interacción, que deben ser adecuadamente abordados por el marco regulativo.

Desde el punto de vista de protección del ciudadano en su calidad de consumidor, la UE estableció su posicionamiento en su Resolución de 12 de febrero de 2020[279] sobre los desafíos derivados del rápido desarrollo de la inteligencia y la toma automatizada de decisiones, en el que el Parlamento Europeo reclama un plan de evaluación de riesgos generales asociados a la IA y la toma automatizada de decisiones.

[278] SPRING, J.M.; FALLON. J. ET AL. (2019). *Machine Learning in Cybersecurity*. Carnegie Mellon University. Septiembre 2019

[279] Resolución del Parlamento Europeo, de 12 de febrero de 2020, sobre los procesos automatizados de toma de decisiones: garantizar la protección de los consumidores y la libre circulación de bienes y servicios (2019/2915(RSP))

Entre otras recomendaciones, sugiere la necesaria aclaración por parte de la Comisión Europea sobre cómo va a garantizar que los consumidores estén protegidos ante prácticas comerciales desleales y/o discriminatorias así como ante los riesgos que conlleva la inteligencia artificial, proponiendo el establecimiento de garantías de una mayor transparencia en los procesos involucrados o sustentados en inteligencia artificial y asegurar que únicamente se utilicen conjuntos de datos de alta calidad y sin sesgos en los sistemas de decisión algorítmica -ADS por sus siglas en inglés-.

Asimismo, el Parlamento Europeo instó a la Comisión a que presentara propuestas al objeto de adaptar la normativa de seguridad de la UE para los productos que esta regula mediante legislación específica al objeto de establecer requisitos armonizados, en particular, la Directiva relativa a las máquinas[280], la Directiva sobre la seguridad de los juguetes [281], la Directiva sobre equipos radioeléctricos[282] y la Directiva sobre baja tensión[283], así como en relación con los "productos no armonizados" de los que se ocupa la Directiva relativa a la seguridad general de los productos[284], a fin de velar por la adecuación de las nuevas normas a su función y por la protección de los usuarios y los consumidores contra posibles perjuicios, así como por que tengan claras tanto los fabricantes sus obligaciones como los usuarios la manera de utilizar los productos con capacidad de toma de decisiones automatizada.

En este sentido, la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, pretende dar cumplimiento, al menos

[280] Directiva 2006/42/CE del Parlamento Europeo y del Consejo, de 17 de mayo de 2006, relativa a las máquinas y por la que se modifica la Directiva 95/16/CE (versión refundida) (DO L 157 de 9.6.2006. P. 24)
[281] Directiva 2009/48/CE del Parlamento Europeo y del Consejo, de 18 de junio de 2009, sobre la seguridad de los juguetes (DO L 170 de 30.6.2009.P. 1).
[282] Directiva 2014/53/UE del Parlamento Europeo y del Consejo, de 16 de abril de 2014, relativa a la armonización de las legislaciones de los Estados miembros sobre la comercialización de equipos radioeléctricos, y por la que se deroga la Directiva 1999/5/CE (DO L 153 de 22.5.2014. P. 62).
[283] Directiva 2014/35/UE del Parlamento Europeo y del Consejo, de 26 de febrero de 2014, sobre la armonización de las legislaciones de los Estados miembros en materia de comercialización de material eléctrico destinado a utilizarse con determinados límites de tensión (DO L 96 de 29.3.2014. P. 357).
[284] Directiva 2001/95/CE del Parlamento Europeo y del Consejo, de 3 de diciembre de 2001, relativa a la seguridad general de los productos (DO L 11 de 15.1.2002. P. 4).

parcialmente, en relación con la legislación de necesaria armonización de la UE basada en el denominado "Nuevo Marco Legislativo".

El Parlamento Europeo también destaca en la precitada Resolución de 12 de febrero de 2020 la necesidad de un enfoque regulador basado en el riesgo, ante la variada naturaleza y complejidad de los retos planteados por los diferentes tipos y aplicaciones de inteligencia artificial y de los sistemas de toma de decisiones automatizadas.

Del mismo modo, el Parlamento Europeo también instó a la Comisión para que desarrollara un sistema de evaluación de riesgos para la IA y la toma de decisiones automatizadas a fin de garantizar un enfoque coherente de la aplicación de la legislación sobre seguridad de los productos en el mercado interior. Y en relación con todo ello, destacó que los Estados miembros deberán elaborar estrategias armonizadas de gestión de riesgos para la inteligencia artificial en el contexto de sus estrategias nacionales de vigilancia del mercado.

Y, por último, instó igualmente a la Comisión a que revise la Directiva sobre la responsabilidad por productos defectuosos[285] para adaptar los conceptos como "productos", "daños" y "defectos", así como las normas en materia de carga de la prueba, dado que la misma ha proporcionado seguridad a los consumidores contra los daños causados por productos defectuosos, pero necesita ser actualizada. Asimismo, instó a la Comisión a que actualice todos estos conceptos y normas de ser necesario, especialmente ante el reto que plantea determinar la responsabilidad en caso de que el perjuicio del consumidor resulte de procesos autónomos de toma de decisiones.

Este posicionamiento del Parlamento Europeo se haya alineado con el análisis, reflexiones, argumentos y conclusiones que expondré en el capítulo V en relación con el marco actual de responsabilidad en España y en la UE, y la necesidad de su revisión y adecuación.

[285] Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados miembros en materia de responsabilidad por los daños causados por productos defectuosos (DO L 210 de 7.8.1985. P. 29).

Durante el análisis de las propuestas regulatorias posteriores tanto del Parlamento como de la Comisión, expondré cuáles de estas recomendaciones han sido abordadas y cómo.

Por último, como he analizado en el anterior apartado con mayor detalle, la actividad legislativa en materia de ciberseguridad por parte de la UE ha sido prolífica durante los últimos años, especialmente mediante la Directiva (UE) 2016/1148 del Parlamento Europeo y del Consejo, de 6 de julio de 2016, relativa a las medidas destinadas a garantizar un elevado nivel común de seguridad de las redes y sistemas de información en la Unión (También conocida como "Directiva NIS").

Esta Directiva pretende ser la respuesta europea ante los problemas de seguridad de las redes y sistemas de información en la UE, mediante el establecimiento de unos requisitos comunes de seguridad para operadores de servicios esenciales y proveedores de servicios digitales, así como unas bases comunes para el desarrollo de capacidades y planificación, intercambio de información y cooperación en esta materia.

Posteriormente, se promulgó su Reglamento de Ejecución (UE) 2018/151 de la Comisión, de 30 de enero de 2018, por el que se establecen normas para la aplicación de la Directiva (UE) 2016/1148 del Parlamento Europeo y del Consejo en lo que respecta a la especificación de los elementos que han de tener en cuenta los proveedores de servicios digitales para gestionar los riesgos existentes para la seguridad de las redes y sistemas de información, así como de los parámetros para determinar si un incidente tiene un impacto significativo.

Actualmente se está preparando la denominada Directiva NIS 2 como he expuesto anteriormente.

Estos marcos establecen los requerimientos reguladores vinculantes en materia de seguridad para servicios esenciales y digitales, como he analizado anteriormente.

Por último, por lo que se refiere a la seguridad de datos personales, tal y como abordé anteriormente, la UE promulgó el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo

que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento General de Protección de Datos).

Además de ello, España también ha promulgado distintas normas reguladoras de la seguridad en sistemas informáticos. De un lado, el Real Decreto-ley 12/2018, de 7 de septiembre, de seguridad de las redes y sistemas de información, que pretende transponer la precitada Directiva NIS, ampliando los servicios incluidos, regulando nuevas obligaciones, así como estableciendo un régimen sancionador en caso de incumplimiento de sus disposiciones. De otro, en relación con los sistemas informáticos de las Administraciones Públicas o sus proveedores, el Real Decreto 3/2010, de 8 de enero, por el que se regula el Esquema Nacional de Seguridad en el ámbito de la Administración Electrónica. Y, por último, significar otras normas reguladoras de la confidencialidad de la información y exigencia de su seguridad, en particular, la Ley 1/2019, de 20 de febrero, de Secretos Empresariales, o reguladoras de la seguridad de cualquier información sino de naturaleza específica como los datos personales, entre otras, la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD).

La cuestión es determinar en qué punto nos encontramos en relación con la inteligencia artificial como sistema que debe ser seguro, así como la inteligencia artificial como instrumento o medio para el gobierno y gestión de la ciberseguridad y su posible utilización como medio o instrumento para la comisión de actos ilícitos o delictivos.

Durante la elaboración de esta investigación, en particular, el 22 de octubre de 2020, asistí a una conferencia sobre el futuro de la ciberseguridad organizada por la Agencia Española de Protección de Datos bajo el título "Innovación, Protección de Datos y Transformación Digital: El futuro de la ciberseguridad", dentro del Ciclo de Conferencias "Innovación y Protección de datos. Mujer y Ciencia".

La conferencia fue impartida por Soledad Antelada Toledano, experta en ciberseguridad con reconocimiento internacional, ingeniera informática del Departamento de Ciberseguridad del *Berkley Lab* y *Network Security Chair* del equipo de expertos que tiene a cargo la seguridad de "SCinet", la red de alto rendimiento de la Conferencia de

Supercomputación de EEUU, que sigue batiendo récords, especialmente de velocidad, llegando a los más de 4 Terabits por segundo.

Durante la misma se abordó como la inteligencia artificial está siendo utilizada para garantizar la ciberseguridad, especialmente en redes complejas. Se destacó el importante papel de la inteligencia artificial en este campo, todavía en fase experimental, que permite automatizar tareas, procesar datos y tomar decisiones mucho más rápido de lo que un humano es capaz.

Y al hablar sobre cómo se estaba utilizando la inteligencia artificial para garantizar la seguridad de una red como "SCinet", se ilustró con una imagen de la misma durante la que tuve un "*dejà-vu*" con la irrupción de un pensamiento y reflexión consecuente: "Esto es *Skynet*[286]". Y de manera correlativa me abordaron dos preguntas recurrentes en personas más alejadas de la visión técnica de la inteligencia artificial: ¿Podría tomar decisiones autónomas incluso de autoprotección de los intereses encomendados aún en contra del ser humano que la creó? ¿Podría considerar al ser humano -el eslabón más débil en ciberseguridad- como una amenaza para sus objetivos?

Reaccioné a tiempo: "*Keep Calm y Carry On*". Y al igual que es eslogan diseñado por el Ministerio de Información del gobierno británico en 1939 ante la inminente invasión alemana. No lo compartí ni lo hice público.

Entonces volví a mi sentido racional, hipotéticamente algo más cualificado que el de otras personas ajenas a lo jurídico y tecnológico por mi profundización y dedicación profesional a estos temas, y me autocontesté "No, inicialmente y por ahora", porque no puedo obviar algunas de las frases más célebres de personalidades del mundo científico, como Stephen Hawking, que he citado y citaré en esta investigación en diversas ocasiones, que el tiempo dirá si fueron o no proféticas -espero sinceramente que no lo sean-, pero que siguen resultando inquietantes y motivan mi autorreflexión constante

[286] Skynet es el nombre que recibe la inteligencia artificial que lidera al ejército de las máquinas en la saga de películas Terminator, con capacidad de aprendizaje automático y que toma decisiones autónomas para su autoprotección que, en el marco de ese autoaprendizaje, incluye la eliminación de decisiones humanas y la consideración del ser humano como una amenaza.

cuanto más trabajo y profundizo en inteligencia artificial, y no es una sensación aislada sino compartida por muchos expertos en la materia.

Entre otras citas de Hawking a las que hago referencia en mi investigación, me quedo con estas dos: "Los robots podrían llegar a tomar el control y se podrían rediseñar a sí mismos" y "La humanidad tiene un margen de mil años antes de autodestruirse a manos de sus avances científicos y tecnológicos".

Y como he comentado, no es el único. El científico de la computación Steve Omohundro[287], ha manifestado públicamente en diversas ocasiones que pueden llegarse a identificar hasta seis tipos diferentes de sistemas de inteligencia artificial "malignos", exponiendo tres formas de detenerlos:

a) Impedir que se creen sistemas de inteligencia artificial perjudiciales -lo que en mi opinión ya está sucediendo por parte de organizaciones cibercriminales en contra de los principios y normas éticas básicas de benevolencia-;

b) Detectar la inteligencia artificial maliciosa de manera precoz en su ciclo de vida, en mi opinión en el diseño y mediante la *Ethics by design* y la *Security by design*, y antes de que adquiera demasiados recursos. Ello requeriría monitorización del sistema y disponer de la facultad de apagarlo ante su funcionamiento inadecuado;

c) Identificar la inteligencia artificial maliciosa cuando ya ha adquirido muchos recursos, lo que nos situaría en algunos de los supuestos que ya nos ha mostrado la ciencia ficción, lo que nunca se debería producir, en mi opinión, si se ha construido un sistema inteligente sustentado en la ética, la seguridad y el cumplimiento normativo.

[287] Recuperado de: https://www.businessinsider.com/stephen-hawking-on-artificial-intelligence-2014-5. Consultado el 21.02.2021.

La literatura y el cine nos han mostrado repetidamente, no tanto un futuro, sino una mera ficción que el ser humano ha venido convirtiendo en deseo y objetivo, mostrando la grandeza del mismo al ser capaz de convertir en realidad lo que era una ficción irrealizable en el momento en que fue ideado, escrito o filmado. No obstante, esta conversión sustentada en la tecnología comporta enormes riesgos con origen, no tanto en la intencionalidad, sino en la potencialidad de la misma.

La inteligencia artificial será imprescindible para garantizar la ciberseguridad en el futuro, especialmente redes complejas, máxime ante la velocidad de envío y procesamiento y volumen de los datos gestionados, que evidencian cada vez más la insuficiencia de los algoritmos clásicos de ciberseguridad.

Las redes informáticas y de comunicaciones actuales requieren una monitorización continua y un esfuerzo constante en garantizar su seguridad, lo que es cada vez más difícil debido a su dimensión y complejidad. Para ello la inteligencia artificial se ha mostrado como un medio imprescindible, y no sólo para solucionar problemas de ciberseguridad sino para manejar de forma eficiente el Big data.

Y, del mismo modo que será utilizada para gestionar de manera eficiente la ciberseguridad, deberá ser utilizada para defensa ante ataques, preferiblemente para su previsión y detección precoz y, en caso de impacto para su mitigación o eliminación. Inexorablemente será utilizada como medio o instrumento de ataque, en una carrera incesante de anticipación entre ambos bandos de la contienda.

Actualmente se está investigando en distintas líneas y aplicando distintas técnicas en materia de ciberseguridad con apoyo en la inteligencia artificial y el aprendizaje automático en universidades y National Labs en EEUU.

En concreto, desarrollando métodos de planificación y razonamiento para, por ejemplo, ayudar a descubrir rutas de ataque no identificadas, utilizando aprendizaje automático para, por ejemplo, identificar *malware* o para la detección de intrusos, desarrollando algoritmos de aprendizaje automático con datos de diferentes dispositivos, información de tráfico de red o instancias de comportamiento malicioso, o usando supercomputadoras

y desarrollando algoritmos de aprendizaje automático para reconocer patrones específicos de eventos que conducen a ataques.

La investigación y profundización en las áreas de la inteligencia artificial y ciberseguridad en redes complejas únicamente podrá llevarse a cabo a través de supercomputadoras.

Para concluir, la seguridad y, consiguientemente la confiabilidad en la inteligencia artificial, a mi juicio, únicamente podrá alcanzarse si se contempla desde el diseño y por defecto -como igualmente consideran expertos como Ayerbe[288]-, y no sólo la seguridad en sus distintas dimensiones, sino también aspectos como la privacidad, la imparcialidad, la transparencia y la explicabilidad, adicionando a todo ello su evaluación posterior, seguimiento durante su operación y, en caso de brechas de seguridad, análisis y auditorías que documenten las mismas y propongan medidas correctivas, complementarias y recomendaciones.

Según Ayerbe, para diseñar, desarrollar, validar y desplegar sistemas de inteligencia artificial se debe velar por la calidad del dato, del modelo y del resultado, debiendo considerar los siguientes aspectos: Privacidad, equidad, trazabilidad, robustez, fiabilidad, causalidad, explicabilidad, transparencia y gobernanza.

Otros expertos como Bonfanti[289], anteriormente citado, destacan el papel de los gobiernos para abordar estos riesgos y oportunidades que supone la inteligencia artificial, y apuesta por una gobernanza positiva e inclusiva mediante la puesta en práctica de principios de alto nivel como los adoptados por la UE o la OCDE para una inteligencia artificial segura y confiable.

A las empresas les resulta cada vez más necesario reforzar la ciberseguridad en inteligencia artificial. Casi dos tercios no creen que puedan identificar amenazas críticas sin la misma.

[288] AYERBE, A. (2020). *La ciberseguridad y su relación con la inteligencia artificial*. ARI 128/2020. 10.11.2020. Real Instituto Elcano. 2020. P. 7.
[289] BONFANTI, M. E. (2020). *Artificial Intelligence and Cybersecurity: A Promising but Uncertain Future*. Op.cit. P. 9.

El ritmo de adopción de la inteligencia artificial en ciberseguridad se está acelerando: Según el informe *Reinventing Cybersecurity with Artificial Intelligence: The new frontier in digital security*[290] 2019 de Capgemini, el 51% de las empresas utilizan la inteligencia artificial para detectar ataques contra su seguridad y casi el 75% de las empresas están probando inteligencia artificial en casos de uso de ciberseguridad de algún modo.

6. Estrategias europeas de ciberseguridad ante los retos de la inteligencia artificial

La preocupación de la UE sobre los retos que plantea la inteligencia artificial en el marco de la seguridad y, específicamente, de la ciberseguridad, es creciente, como expondré a lo largo de esta investigación y que está presente en todo momento en informes, comunicaciones y propuestas regulatorias en materia ética y de responsabilidad.

La ciberseguridad es un pilar esencial para la fiabilidad y confianza deseable de los sistemas inteligentes.

En esta materia, la *Agencia de Ciberseguridad de la Unión Europea* -ENISA[291]-está desempeñando un papel fundamental en el ámbito de sus funciones y competencias para evaluar la inteligencia artificial en este contexto y ofrecer información esencial para las políticas futuras de la UE para la adecuada gestión de los retos que plantea.

ENISA creó a principios de 2020 el denominado *Grupo de Trabajo Ad Hoc sobre Ciberseguridad para inteligencia artificial*[292] para apoyar a la misma en el proceso de generación de conocimiento sobre ciberseguridad de la inteligencia artificial, en especial,

---

[290] TOLIDO, R. ET AL. (2019). *Reinventing Cybersecurity with Artificial Intelligence: The new frontier in digital security*. Capgemini 2019. Recuperado de: https://www.capgemini.com/es-es/wp-content/uploads/sites/16/2019/07/AI-in-Cybersecurity_Report_20190710_V05.pdf. Consultado el 12.02.2021.

[291] ENISA es la Agencia de la Unión Europea para la Ciberseguridad y tiene como misión garantizar un nivel de seguridad adecuado en todos los países de la UE. Creada en 2004, contribuye, entre otras cosas, a gestionar una política de ciberseguridad en la UE, garantizar la confianza de los productos tecnológicos, servicios y procesos con los esquemas de ciberseguridad, cooperar con los estados miembros y dar soporte a los desafíos que se pueden presentar. https://www.enisa.europa.eu/

[292] Recuperado de https://www.enisa.europa.eu/topics/iot-and-smart-infrastructures/artificial_intelligence/ad-hoc-working-group/adhoc_wg_calls

en asuntos relacionados con la ciberseguridad de la inteligencia artificial, en el desarrollo de un panorama de amenazas de la misma y en el suministro de pautas de ciberseguridad proporcionales al riesgo para la inteligencia artificial.

El grupo está integrado por 15 miembros de la *Dirección General de Redes de Comunicaciones, Contenidos y Tecnología de la Comisión Europea* (DG CONNECT), el *Comité Conjunto de Investigación de la Dirección General de la Comisión Europea* (DG JRC), *Europol*, la *Agencia Europea de Defensa* (EDA), la *Agencia de la Unión Europea para la gestión operativa de sistemas informáticos a gran escala en el espacio de libertad, seguridad y justicia* (eu-LISA), el *Instituto Europeo de Normas de Telecomunicaciones* (ETSI), así como académicos y expertos de la industria.

El 30 de septiembre de 2020 se organizó el primer taller de "Ciberseguridad para la inteligencia artificial (C4AI)" para analizar los retos de seguridad que plantea. ENISA destacó los nuevos métodos de manipulación y ataque que los sistemas inteligentes pueden permitir, así como los de privacidad. La principal conclusión del taller fue el papel esencial de la ciberseguridad para el despliegue seguro y confiable de la inteligencia artificial.

De las intervenciones durante el encuentro precitado, destacar la de la eurodiputada y presidenta del panel -Eva Kaili-[293], sobre el futuro de la ciencia y la tecnología en el Parlamento Europeo, durante la que destacó la confianza es uno de los factores más importantes para la adopción de nuevas tecnologías y la ciberseguridad es clave para inspirar confianza en la IA, por lo que los reguladores europeos deben asegurar una estrategia de ciberseguridad integral en Europa.

Por su parte, el Director Ejecutivo de la Agencia de la UE para la Ciberseguridad, Juhan Lepassaar, destacó durante el mismo que la ciberseguridad es la base de las soluciones confiables de inteligencia artificial y servirá como un trampolín para el despliegue seguro y generalizado de la inteligencia artificial en toda la UE.

[293] Recuperado de: https://www.enisa.europa.eu/events/cybersecurity-for-ai-c4ai (Incluye video del evento). Consultado el 25.03.2021.

6.1. Informe ENISA *AI Cybersecurity Challenges*

La *Agencia de Ciberseguridad de la Unión Europea* -ENISA- publicó en diciembre de 2020 un informe bajo el título *AI Cybersecurity Challenges*[294] -Retos de Ciberseguridad de la inteligencia artificial- que sitúa la ciberseguridad y la protección de datos como los pilares esenciales para la creación y sustento de un ecosistema de inteligencia artificial seguro y confiable.

Según ENISA, la ciberseguridad servirá de trampolín para el despliegue generalizado y seguro de la inteligencia artificial en toda la UE, si bien, sólo lo hará cuando exista una comprensión de contexto de amenazas relacionadas con la misma y de retos asociados a las mismas, a lo que pretende contribuir el informe citado.

A mi juicio, es necesario comprender que nos enfrentamos a una realidad compleja y a un contexto de amenazas amplio y dinámico, que evoluciona de manera incesante conjuntamente con las innovaciones que se van produciendo en el ámbito de la inteligencia artificial y su continua interrelación con otras tecnologías que generan efectos disruptivos en todos los campos y sectores, por lo que las soluciones a los retos y amenazas, es decir, las políticas, normas, procedimientos, análisis de riesgos, las evaluaciones de impacto, los controles y las medidas deberán ser igualmente dinámicas, evolutivas y adaptativas y estar en constante revisión y mejora durante todo su ciclo de vida.

El informe se enmarca en una serie de iniciativas europeas previas relacionadas con la inteligencia artificial que han sido y serán tratadas a lo largo de esta investigación, entre otras, el *Libro blanco sobre la inteligencia artificial* de la Comisión Europea que resalta la importancia de la inteligencia artificial en la sociedad y economía, el *Plan coordinado sobre inteligencia artificial* que busca armonizar y coordinar las diferentes iniciativas de inteligencia artificial en la UE, incluyendo los aspectos de seguridad, la política común y

[294] *AI Cybersecurity Challenges*. Diciembre 2020. Disponible en: https://www.enisa.europa.eu/publications/artificial-intelligence-cybersecurity-challenges. Consultado el 22.02.2021

las recomendaciones para crear una inteligencia artificial de confianza, los informes y guías éticas para una inteligencia artificial confiable emanadas del *Grupo de Expertos en inteligencia artificial de la UE,* que contemplan aspectos éticos, legales y de seguridad, así como las propuestas regulatorias de octubre de 2020 en materia ética y de responsabilidad civil de la inteligencia artificial.

Asimismo, destacar otras iniciativas europeas a las que aludiré posteriormente, en particular de la *Agencia de Defensa Europea (EDA),* que ha desarrollado una taxonomía para la inteligencia artificial en el ámbito de la defensa, el *Centro Común de Investigación de la UE (JRC),* que ha establecido un observatorio de la inteligencia artificial con el objetivo de monitorizar el desarrollo y el impacto de la UA y el *Grupo de Especificaciones Industriales para la Securización de la IA* (ISG SAI) del Instituto de Estándares de Telecomunicaciones Europeo (ETSI), cuyo objetivo es la creación de estándares para preservar y mejorar la seguridad de las nuevas tecnologías basadas en inteligencia artificial.

6.1.1. Objetivo

El objetivo del informe elaborado por ENISA es la identificación de los principales activos y de las amenazas asociadas a la inteligencia artificial a las que pueden verse expuestos aquellos desde un punto de vista global, así como los principales retos que todo ello supone, dando de este modo seguimiento a las prioridades definidas en el documento de trabajo previo denominado *WP2020 Output O.1.1.3 on Building Knowledge on Artificial Intelligence Security* y en el *Libro blanco sobre la inteligencia artificial* de la Comisión Europea[295].

El informe pretende ser global, horizontal y transversal, por lo que no aborda aplicaciones sectoriales, si bien, ya advierte de la necesidad de que en el futuro se aborden enfoques

[295] Recuperado de: https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en. Consultado el 22.02.2021

sectoriales para identificar los riesgos y evaluar la probabilidad y el impacto de las amenazas en aplicaciones específicas.

La ética de la inteligencia artificial queda formalmente fuera del ámbito de este informe, ya que fue una de las áreas de interés del *Grupo de Expertos de Alto Nivel de la Comisión Europea sobre IA*, si bien, de nuevo, la seguridad constituye un principio y norma ética esencial alrededor de la cual se ha elaborado este informe.

Por lo que se refiere a sus antecedentes inmediatos, significar que la UE aprobó su *Plan Coordinado sobre inteligencia artificial*[296] con el objetivo de armonizar y coordinar las iniciativas de la inteligencia artificial en toda la UE, incluyendo sus aspectos relacionados con la seguridad. Asimismo, la reciente *Estrategia de Seguridad de la Unión*[297] de julio de 2020 destacó la importancia de la inteligencia artificial en este contexto. Por su parte, el denominado *Grupo de Expertos de Alto Nivel sobre inteligencia artificial* -AI HLEG[298]- incluyó dentro de sus recomendaciones, las relativas a la fiabilidad y seguridad de los sistemas de inteligencia artificial.

6.1.2. Destinatarios

El informe está dirigido a los responsables de crear y aprobar futuras políticas para la implementación segura de la inteligencia artificial en organizaciones, así como a los técnicos expertos para sustentar la adecuada evaluación de sus riesgos, y a los organismos de estandarización para respaldar los futuros estándares de seguridad de la inteligencia artificial.

[296] Recuperado de: https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence. Consultado el 22.02.2021
[297] Recuperado de: https://ec.europa.eu/info/sites/info/files/communication-eu-security-union-strategy.pdf. Consultado el 22.02.2021
[298] Recuperado de: https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence. Consultado el 14.03.2021.

El informe pretende facilitar a las empresas la evaluación de riesgos de ciberseguridad de una forma uniforme y repetible para este tipo de aplicaciones y permitirles identificar cuáles son los controles de seguridad más adecuados.

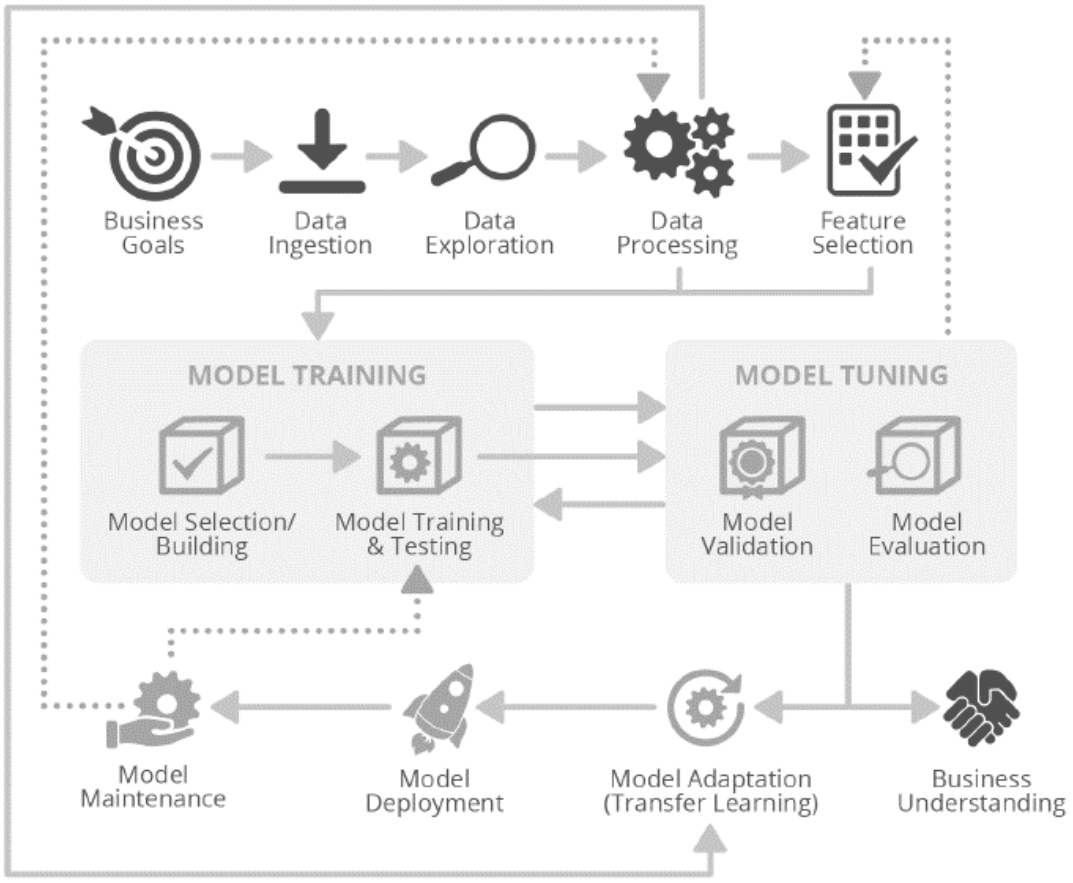
La metodología seguida se basa en las propias de ENISA para otros trabajos sobre amenazas en otros contextos y está alineada con las iniciativas europeas precitadas (ETSI, EDA, etc.), ISO 27005 y otros estándares internacionales relacionados con la ciberseguridad como NIST o MITRE.

6.1.3. Contenido

Por lo que se refiere a su contenido, el informe analiza las amenazas en todo el ciclo de vida de la inteligencia artificial, desde el análisis de requerimientos hasta su implementación y uso, identifica los activos que integran el ecosistema de inteligencia artificial objeto de protección y de seguridad, identifica las amenazas asociadas mediante un mapeo y taxonomía detalladas para identificar posibles vulnerabilidades que podrían ser explotadas y escenarios de ataque, clasifica las amenazas y lista los actores más relevantes relacionados con aquéllas, significa el impacto de las amenazas en las diferentes dimensiones de la seguridad e identifica los retos y oportunidades para la implementación de sistemas y servicios de inteligencia artificial seguros en la UE.

6.1.4. Ciclo de vida de un sistema de inteligencia artificial

El ciclo de vida de un sistema de inteligencia artificial incluye varias fases interdependientes que van desde su diseño y desarrollo (incluyendo subfases como el análisis de requisitos, la recopilación de datos, la formación, las pruebas y la integración), la instalación, el despliegue, el funcionamiento, el mantenimiento y la eliminación. El gráfico adjunto muestra el *Modelo de referencia genérico del ciclo de vida de la IA* que toma como referencia el informe precitado:



Fuente: ENISA. *AI Cybersecurity Challenges*. Diciembre 2020. P. 15

Conforme se puede apreciar en el ciclo de vida, los datos son los activos de más valor y están en constante transformación.

6.1.5. Agentes participantes

En todo el ciclo de vida participan distintos agentes:

a) Diseñadores y desarrolladores: Los diseñadores de inteligencia artificial -*AI designers*-, los diseñadores de aplicaciones de inteligencia artificial -*AI applications designers*- que participan en el diseño y la creación de sistemas inteligentes, los desarrolladores de inteligencia artificial -*AI developers*- que desarrollan y construyen software y los algoritmos utilizados en los sistemas de inteligencia artificial (que pueden igualmente trabajar para su perfeccionamiento

y mejora). El papel que desempeñan para el desarrollo de sistemas de IA seguros es fundamental, especialmente por su experiencia y capacidad.

b) Científicos de datos: Los diseñadores y desarrolladores de inteligencia artificial trabajan en estrecha colaboración con los científicos de datos -*data scientists*-. Las aportaciones de éstos pueden consistir en ayudar a diseñar a diseñar y desarrollar modelos de inteligencia artificial o pueden consistir en utiliza dichos modelos y analizar sus resultados. Los científicos de datos participan en la recogida e interpretación de datos, centrándose en la extracción de conocimientos e ideas de esos datos.

c) Ingenieros de datos: El trabajo de los ingenieros de datos -*data engineers*- consiste principalmente en extraer y cotejar datos de diferentes fuentes, y luego transformarlos, limpiarlos, normalizarlos y almacenarlos. Los ingenieros de datos se centran principalmente en el diseño, la gestión y la optimización del flujo de datos.

d) Propietarios/responsables de los datos: Los propietarios/responsables de los datos -*data owners/controllers*- son los responsables de los conjuntos de datos que se utilizan para entrenar/validar los sistemas de inteligencia artificial o que estos sistemas utilizan para realizar tareas. Suele tratarse de empresas que tienen sus propios conjuntos de datos vinculados a su negocio y que proporcionan a un sistema de inteligencia artificial para que realice una tarea en su nombre.

e) Proveedores de datos o *data brokers*. Se trata de terceros que monetizan los datos utilizados por los sistemas de sistemas de inteligencia artificial, ya sea con fines de formación o para realizar diversas tareas. Pueden ser *data brokers* comerciales,

que recogen, almacenan y venden diversos tipos de datos, de manera legal. No obstante, existen también *data brokers* que actúan desde la ilegalidad, recopilando datos sobre los usuarios sin que éstos sean conscientes de que sus datos personales se están recopilando, almacenando y vendiendo.

f) Proveedores de modelos -*model providers*-: Su aportación consiste en proporcionar modelos, así como implementaciones de los mismos, que ya han sido entrenados y puestos a punto. Algunos proveedores de modelos son proveedores de servicios de nube *o cloud providers,* que ofrecen los modelos como un servicio, en particular el uso de las capacidades de cálculo y análisis de datos basados en la inteligencia artificial en la nube.

g) Proveedores de terceros que suministran el software a éstos o *third-party providers*.

h) Usuarios finales o *end users*: Hacen uso de los sistemas de inteligencia artificial, incluidos los consumidores de servicios. Puede tratarse de empresas, consumidores y el público en general.

Todos estos actores están relacionados con los riesgos y las amenazas, por lo que deben ser considerados para generar un marco de seguridad, y su mayor o mejor control sobre el riesgo puede determinar su mayor o menor grado de responsabilidad contractual o extracontractual por daños derivados del sistema inteligente.

ENISA destaca también la especial importancia de la cadena de suministro relacionada con la inteligencia artificial, considerando que, para lograr un ecosistema seguro y confiable en la UE, debe considerarse y garantizarse la seguridad en los elementos que integran la cadena de suministro.

6.1.6. Ciberseguridad y privacidad para una IA confiable y segura

El informe también destaca que el ecosistema confiable y seguro en el ámbito de la UE debería situar la ciberseguridad y la protección de datos en primer plano y fomentar las

iniciativas pertinentes de innovación, creación de capacidades, concienciación, investigación y desarrollo.

Como he referido anteriormente, los servicios, sistemas y técnicas que hacen uso de la inteligencia artificial generan resultados inesperados y pueden igualmente ser objeto de manipulación para alcanzar éstos, como ocurre con el software que sustenta la inteligencia artificial que se basa frecuentemente en modelos de caja negra -*Black box*-. Además, estos sistemas pueden ser utilizados con finalidades maliciosas como medio para facilitar ataques o para realizarlos directamente.

En este sentido, el informe se sustenta en un enfoque de la gestión de riesgos para una inteligencia artificial segura, significando la necesidad de comprender lo que hay que asegurar y proteger -es decir, identificar los activos a proteger que pueden verse afectados por las amenazas y los modelos de los adversarios-, comprender los modelos de gobernanza de datos -diseño, evaluación, protección y proceso de entrenamiento de la inteligencia artificial-, gestionar los riesgos adecuadamente y desarrollar medidas y controles para garantizar la seguridad.

6.1.7. Sujetos activos

El informe analiza las amenazas identificando a sus actores que pueden tener la intención de dañar los sistemas inteligentes, las metodologías adoptadas para derivar las amenazas a los diferentes activos y finaliza con una descripción de las amenazas y su categorización. En relación con estos aspectos, me permitiré realizar una breve exposición de los agentes involucrados, para analizar con mayor detalle las amenazas y su categorización.

Por lo que se refiere a los actores de *ciberamenazas*, estos se pueden mover por múltiples motivos.

a) Ciberdelincuentes: El informe los denomina “ciberdelincuentes”, si bien, como he manifestado anteriormente, no todo acto ilegítimo es ilícito, ni todo acto ilícito es delictivo en función del ordenamiento jurídico aplicable. Los “ciberdelincuentes” utilizan la inteligencia artificial como herramienta para realizar ataques. pero

también para explotar vulnerabilidades de los sistemas, desde ataques de *ransomware* contra sistemas basados en inteligencia artificial utilizados para la gestión de la cadena de suministro y el almacenamiento como, por ejemplo, para intentar *crackear chatbots* mediante inteligencia artificial para robar datos de tarjeta de crédito.

b) Usuarios internos y externos de la organización: Personas con acceso a información privilegiada de la empresa, incluidos directivos, empleados y contratistas que tiene acceso a las redes de la misma, pueden perjudicar a la empresa de manera intencionada, por ejemplo sustrayendo o alternando los datos utilizados por los sistemas inteligentes, o no intencionada, por ejemplo, el denominado "ignorante accidental", esto es, el usuario de sistemas de información de la empresa que no ha sido suficientemente informado y formado sobre determinados usos o prácticas no permitidas que puedan causar enormes brechas de seguridad y corromper los datos utilizados por los sistemas inteligentes.

c) Actores pertenecientes o patrocinados por determinados Estados para atacar a otros países, incluyendo no sólo infraestructuras gubernamentales sino industrias e infraestructuras críticas.

d) Terroristas que suelen buscar los daños físicos, como por ejemplo que los que *hackean* vehículos autónomos para utilizarlos como arma.

e) *Hacktivistas* que suelen tener una motivación ideológica, también pueden tratar de *hackear* los sistemas de inteligencia artificial con el fin de demostrar que se puede hacer.

f) Competidores cuyo principal objetivo es conseguir una ventaja de mercado y/o perjudicar la posición competitiva de su competidor.

g) Otros: Por ejemplo, actores menos sofisticados como los *script kiddies* mediante el uso de *scripts* o programas preestablecidos para atacar sistemas dado que carecen de los conocimientos para desarrollos los suyos propios, y que pueden estar movidos por motivaciones criminales o ideológicas.

6.1.8. Clasificación de las amenazas

Respecto de las amenazas, se clasifican en siete categorías: a) Actividad/abuso malicioso; b) Escucha/interceptación/secuestro; c) Ataques físicos; d) Daños no intencionados; e) Fallo/interrupción; f) Desastre; g) Jurídico-legal. El gráfico que se específica a continuación sintetiza la taxonomía detallada de las amenazas de la inteligencia artificial:

Fuente: ENISA *AI Cybersecurity Challenges.* Diciembre 2020. P. 30

Respecto de las amenazas de ciberseguridad que pueden afectar la inteligencia artificial, conforme destaca Ballarín[299], ENISA aprovecha el trabajo realizado por el *Grupo de Expertos de la UE para la IA* -EC AI HLEG-, que identificó un conjunto de atributos que deben ser evaluados para medir la confianza de las aplicaciones de inteligencia artificial.

Estos atributos son la autenticidad, autorización, no repudio, así como otras que son más específicas al ámbito de la inteligencia artificial, como son la solidez, fiabilidad, safety, transparencia, explicabilidad, responsabilidad y protección de datos.

En la medida que estos atributos se consideran propiedades de la seguridad, ENISA recoge los impactos potenciales de las amenazas de ciberseguridad de manera asociada a soluciones de inteligencia artificial que pueden atentar contra la autenticidad, autorización, no repudio, solidez, fiabilidad, transparencia, explicabilidad, rendición de cuentas y protección de datos.

6.1.9. Conclusiones del informe

De las conclusiones del informe me permito destacar la necesidad de:

- Garantizar la ciberseguridad de la inteligencia artificial para asegurar que la misma y el conjunto de tecnologías asociadas sean dignas de confianza, fiables y robustas.

- Asegurar la comprensión común sobre las amenazas relevantes relacionadas con la ciberseguridad de la inteligencia artificial como aspecto clave para el despliegue y la aceptación generalizados de los sistemas y aplicaciones de la inteligencia artificial.

[299] BALLARÍN, P. (2021). *Desafíos de la ciberseguridad en la inteligencia artificial (IA). Análisis del informe de ENISA*. ODISEIA. Enero 2021. P.9.

- Acometer los retos de ciberseguridad asociados a la misma en torno a la complejidad, las cuestiones técnicas, la integridad, la confidencialidad y la privacidad que plantea.

- Desarrollar un conjunto de medidas concretas de mitigación para las amenazas de la inteligencia artificial identificadas sobre la base de evaluaciones de riesgo.

Las amenazas que refleja el informe pueden predicarse respecto de cualquier sistema/aplicación de inteligencia artificial (en función de su configuración y de las técnicas utilizadas).

La realización de evaluaciones de riesgo específicas de los sistemas de inteligencia artificial debe tener en cuenta el contexto de uso y sectores, en la medida que éstos presentan diferentes grados de riesgo que deben ser evaluados y, en consecuencia, deben establecerse diferentes medidas y controles de seguridad.

No obstante, a pesar de la necesidad de análisis verticales y sectoriales, la introducción de metodologías horizontales y mejores prácticas podría ser valiosa para establecer una línea de base común y promover así una capa común de ciberseguridad y confianza en todos los sectores.

Del panorama de las amenazas se desprende que hay que trabajar especialmente en el ámbito de la verificación y validación formal automática, la explicabilidad y la transparencia, y las nuevas técnicas de seguridad para contrarrestar las amenazas emergentes de la inteligencia artificial.

La investigación se considera esencial en diversos ámbitos para desarrollar algoritmos, sistemas y soluciones fiables de inteligencia artificial que mejoren las operaciones industriales y de seguridad y la competitividad de los mercados de la UE, desarrollando la marca "*IA made in Europe*" como sello de calidad para una inteligencia artificial ética, segura y de vanguardia que pueda convertirse en una referencia mundial.

Del mismo modo, la investigación también es necesaria para promover mejores sistemas y soluciones robustas, por lo que el trabajo de previsión por parte de los cuerpos de seguridad y autoridades, que deberían evaluar y predecir proactivamente el uso indebido

de la inteligencia artificial para mejorar la preparación, como el informe de Europol, TrendMicro y UNICRI sobre *Usos y abusos maliciosos de la inteligencia artificial*.[300]

La inteligencia artificial, como destaca el informe, puede apoyar las operaciones de ciberseguridad y ciberdelincuencia con técnicas que pueden ser utilizadas para aumentar y automatizar las operaciones de ciberseguridad, como los cortafuegos inteligentes.

También es esencial asegurar los diversos activos del ecosistema y el ciclo de vida de la inteligencia artificial, especialmente activos que residen en complejas cadenas de suministro y que implican relaciones transfronterizas e intersectoriales. La seguridad e integridad de la cadena de suministro de la inteligencia artificial es, por tanto, de vital importancia.

Con esta finalidad, es esencial el aprovechamiento de las asociaciones público-privadas y el fomento de la creación de grupos multidisciplinares de expertos en ciberseguridad de la IA, como el *Grupo de trabajo ad hoc de ENISA sobre ciberseguridad de la IA,* así como profundizar la normalización de la seguridad de la inteligencia artificial.

La complejidad y la amplitud del panorama de las amenazas a la ciberseguridad de la inteligencia artificial requieren el fomento de un ecosistema de la UE para una inteligencia artificial segura y digna de confianza, que incluya todos los elementos de la cadena de suministro de la inteligencia artificial.

El informe ENISA destaca que el ecosistema de inteligencia artificial segura de la UE debe situar la ciberseguridad y la protección de datos en primer plano y fomentar las iniciativas pertinentes de innovación, creación de capacidades, sensibilización e investigación y desarrollo.

Por último, el informe incorpora distintos anexos que conforman la fotografía más actual de los riesgos asociados a la inteligencia artificial para su gestión: Anexo A: Descripción de la taxonomía de activos; Anexo B: Descripción de la taxonomía de las amenazas; C:

[300] Recuperado de: https://www.europol.europa.eu/newsroom/news/new-report-finds-criminals-leverage-ai-for-malicious-use-%E2%80%93-and-it%E2%80%99s-not-just-deep-fakes. Consultado el 20.03.2021.

Asignación de activos al ciclo de vida de la inteligencia artificial; Anexo D: Mapa de amenazas del ciclo de vida de la inteligencia artificial.

6.2. Seguridad y privacidad como base para una IA segura y fiable.

Como he referido, el informe confiere a la seguridad y la privacidad un papel principal para alcanzar una inteligencia artificial segura y confiable.

La seguridad ha pasado de ser una mera disposición técnico-organizativa a un principio esencial, como así se regula por ejemplo en el Reglamento General de Protección de Datos (RGPD), en particular, en su artículo 5, que debe regir todo tratamiento de datos y cuya inexistencia comporta su ilegalidad, es decir, la seguridad no es una opción, sino una necesidad y una obligación.

El artículo 32 del RGPD regula las medidas de seguridad obligatorias que deben aplicarse de manera adecuada en función del contexto y de los riesgos asociados al tratamiento, que deberán ser debidamente analizados previamente para determinar su probabilidad y el impacto que podría tener para los intereses en juego.

El informe ENISA toma como referencia el RGPD para construir su enfoque de la seguridad de la inteligencia artificial basada en el análisis y gestión de riesgos, con identificación de activos, amenazas, probabilidad, impacto, medidas y controles.

Conforme destaca, el activo a proteger no es meramente la información sino el ejercicio sin restricciones de los derechos de las personas, por lo que la seguridad es una forma de reforzar los derechos y libertades de los individuos en su conjunto.

Y en relación con todo ello contiene una afirmación de extrema relevancia para el objeto de esta investigación y directamente relacionado con las cuestiones de responsabilidad ética y jurídica: “Los sistemas de IA son sistemas lógicos y, como tales, pueden no ser totalmente consistentes y completos, lo que significa que los humanos nunca podrán predecir, por adelantado durante la fase de diseño, todos los posibles factores contextuales que pueden perjudicar su funcionamiento”, lo que comporta sus riesgos inherentes de

obtener resultados inesperados cuando los resultados de un sistema inteligencia no están debidamente limitados. Los resultados deberían estar siempre limitados.

ENISA apuesta por un concepto de seguridad proactiva, en el diseño y por defecto en la línea recogida en el Reglamento General de Protección de Datos en su artículo 25, debiendo establecerse medidas técnicas y organizativas adecuadas a los riesgos y relacionadas con los principios de protección de datos, desde la minimización hasta la exactitud de los datos, con medidas como la seudonimización o el cifrado. Y en este sentido, considera que estos análisis y medidas consecuentes deben considerar el contexto de la inteligencia artificial de modo que, por ejemplo, los individuos no sean identificables por defecto por las máquinas a menos que lo deseen.

Del mismo modo, reitera la seguridad genera confianza, atrae inversión, retiene a usuarios y crea retroalimentación positiva para desarrollar nuevas aplicaciones.

En este sentido, propone como estrategias a reflexionar, la consideración de la seguridad como un elemento un funcional de los sistemas inteligentes y en el diseño (*Security by design*) y la posibilidad de certificación para promover una cultura de seguridad entre los agentes económicos.

Por último, destacar la descripción que efectúa ENISA de un conjunto de escenarios de ataque o fallo correspondientes a las principales etapas del ciclo de vida de la inteligencia artificial, a partir del cual se desprende con claridad el escenario actual en el que nos encontramos.

Durante 2021 se pretende intensificar el diálogo entre ENISA, la Comisión Europea y demás instituciones europeas sobre la ciberseguridad y las iniciativas reguladoras de la inteligencia artificial, objeto de análisis en esta investigación. ENISA ha anunciado la elaboración de controles para la mitigación de los riesgos identificados con publicación de las pautas pertinentes para facilitar la gestión de aquéllos.

6.3. Retos de ciberseguridad en la utilización de la IA en la conducción autónoma

Recientemente, ENISA y el *Joint Research Centre* de la Comisión Europea -JRC- publicaron el pasado 11.02.2021 un informe sobre los retos de la inteligencia artificial en la conducción autónoma bajo el título: *Cybersecurity challenges in the uptake of artificial intelligence in autonomous driving*[301]. Se trata de un informe sectorial orientado a esta aplicación de la inteligencia artificial.

El objetivo de este informe es abordar los desafíos de ciberseguridad específicamente relacionados con la adopción de técnicas de inteligencia artificial en la conducción autónoma, profundizando en los aspectos técnicos de la inteligencia artificial en el sector de la automoción con el objetivo de comprender mejor las preocupaciones tecnológicas y reflejar el nivel de integración de la inteligencia artificial en la conducción autónoma.

El informe identifica los principales retos y propone distintas recomendaciones para mejorar la seguridad de la inteligencia artificial en los vehículos autónomos y mitigar las amenazas y riesgos potenciales asociados a los mismos.

Este informe también es fruto de la creciente preocupación y concienciación de responsables políticos, organismos reguladores y de la industria de la automoción de las necesidades de seguridad de estos sistemas, especialmente a raíz de los ataques de los que se ha tenido conocimiento durante los últimos años en el contexto de la automoción en los últimos tres años.

El informe destaca los innegables beneficios de la conducción autónoma para muchos aspectos de nuestras sociedades, si bien, en paralelo, se ha planteado la seguridad de esta tecnología, que por definición está destinada a funcionar con una supervisión humana limitada.

Las técnicas de aprendizaje automático que constituyen la base de estos sistemas inteligentes según el informe, han evidenciado ser muy vulnerables a una amplia gama

[301] *Cybersecurity challenges in the uptake of artificial intelligence in autonomous driving*. UE 2021. Recuperado de: https://www.enisa.europa.eu/publications/enisa-jrc-cybersecurity-challenges-in-the-uptake-of-artificial-intelligence-in-autonomous-driving. Consultado el 22.02.2021.

de ataques que podrían comprometer el buen funcionamiento de los vehículos autónomos y plantear graves amenazas para la seguridad de las personas, tanto dentro como fuera de un vehículo.

ENISA aborda la cuestión desde un enfoque propio de seguridad de los sistemas de información y de sus principios esenciales, si bien, por su naturaleza, los componentes de estos sistemas inteligentes no obedecen las mismas reglas que el *software* tradicional. Las técnicas de *Machine Learning* se basan en reglas implícitas que se fundamentan en el análisis estadístico de grandes colecciones de datos lo cual, de un lado, permite que la automatización alcance capacidades cognitivas sin precedentes y, de otro, abre nuevas oportunidades para la explotación de la complejidad de estos sistemas por los *ciberatacantes* en su beneficio.

La seguridad de estos sistemas, modelos y técnicas exige tener en cuenta los riesgos tradicionales de ciberseguridad de sistemas digitales junto con las cuestiones específicas que plantea la adopción de técnicas de inteligencia artificial en vehículos autónomos, evaluándolos en toda la cadena de suministro de *hardware* y de *software* como en todo el ciclo de vida de estos sistemas, componentes y datos, esto es, desde su desarrollo hasta su integración con otros sistemas de automatización, aplicación y funcionamiento.

En este sentido hay múltiples iniciativas institucionales y privadas para la definición de los principios y normas que deben regir el desarrollo de vehículos autónomos, con directrices específicas en unos casos, o mediante la definición de conjuntos de prácticas que apoyen el desarrollo y despliegue de la inteligencia artificial y la ciberseguridad.

La UE ha llevado a cabos distintas iniciativas para el desarrollo de una inteligencia artificial fiable, como he referido anteriormente, en las que la seguridad o "ciberseguridad" con un elemento esencial.

El informe precitado destaca que los componentes de un *software* de inteligencia artificial no forman un sistema monolítico, sino que se integran una compleja combinación de grandes y variadas colecciones de datos obtenidos de múltiples sensores y un conjunto de metodologías de inteligencia artificial basadas en trabajos científicos en las áreas de la

estadística, la informática, las matemáticas y la robótica, a los que asocian las tareas cognitivas del sistema.

Sobre la seguridad de la inteligencia artificial, ENISA se remite en este informe al global emitido por la misma y anteriormente citado, que debe ser completado con este en el ámbito de la conducción autónoma.

Como he referido en distintos puntos de esta investigación, los ataques a estos sistemas pueden ser de distinta naturaleza. Por una parte, pueden tratarse de ataques llevados a cabo en el entorno donde circula el vehículo, es decir, ataques diseñados para el engaño de los sistemas mediante la difusión de patrones minuciosamente estudiados en el entorno para, de un lado, alterar el proceso de toma de decisiones e inducir a un comportamiento no esperado del vehículo, como, por ejemplo, pegatinas en una señal de tráfico para confundir o impedir su reconocimiento o la pintura en el asfalto, y, de otro, hacerlos imperceptibles para el ser humano.

Por otra parte, pueden tratarse de ataques al sistema interno del vehículo y sus componentes.

El informe indicado ilustra y documenta, tanto teórica como experimentalmente, distintos escenarios de ataque.

Por último, el informe define un conjunto de retos y recomendaciones para la mejora de seguridad de la inteligencia artificial en los vehículos autónomos y mitigar las posibles amenazas y riesgos, sustentándose en las metodologías propias de gobierno y gestión de la ciberseguridad en sistemas de información junto con las especificidades de estos sistemas y su aplicación para estos usos.

Las recomendaciones con las que concluye ENISA su informe son:

- Validación sistemática de la seguridad de los modelos y datos de la inteligencia artificial

- Seguridad de la cadena de suministro de la inteligencia artificial en la industria del automóvil

- Procesos y controles de ciberseguridad de las técnicas de inteligencia artificial en la conducción autónoma.

- Aumentar las capacidades de preparación y respuesta a incidentes.

- Aumentar la capacidad y la experiencia en ciberseguridad de la inteligencia artificial para los sistemas de automoción.

6.4. Otros estudios e informes

Sin perjuicio del informe de ENISA abordado, la Agencia Europea de Defensa -*European Defence Agency* o EDA por sus siglas en inglés- ha desarrollado también una taxonomía exhaustiva para la inteligencia artificial[302] en el ámbito de la defensa estructurada en tres líneas: Algoritmos, funciones realizadas por algoritmos y áreas de apoyo o relacionadas como la ética, la implementación de hardware o las técnicas de aprendizaje.

Por su parte, el Instituto Europeo de Normas de Telecomunicaciones -*European Telecommunication Standards Institute* o ETSI por sus siglas en inglés- ha creado un Grupo de Especificación Industrial sobre Seguridad de la inteligencia artificial[303] - Industry *Specification Group on Securing Artificial Intelligence* o ISG SAI por sus siglas en inglés- con el objetivo de crear normas para preservar y mejorar la seguridad de las nuevas tecnologías de inteligencia artificial.

Asimismo, el precitado Centro Común de Investigación de la Comisión Europea -*Joint Research Centre* o JRC por sus siglas en inglés, ha creado la iniciativa *AI Watch* como servicio de conocimiento para supervisar el desarrollo, la adopción y el impacto de la inteligencia artificial para Europa y supervisar la investigación en la misma.

[302] Recuperado de: https://www.eda.europa.eu/info-hub/press-centre/latest-news/2020/08/25/artificial-intelligence-joint-quest-for-future-defence-applications. Consultado el 14.03.2021.
[303] Recuperado de: https://www.etsi.org/technologies/securing-artificial-intelligence. Consultado el 14.03.2021

6.5. Estrategias de ciberseguridad y nuevas propuestas reguladoras

El 16 de diciembre de 2020 la Comisión Europea presentó su nueva *Estrategia de Ciberseguridad*[304] con la finalidad de reforzar la resiliencia colectiva de Europa contra las amenazas cibernéticas y garantizar un ciberespacio global, abierto, estable y seguro, basado en el estado de derecho, los derechos humanos, las libertades fundamentales y valores democráticos.

De la estrategia europea me permito destacar su principal objetivo estratégico: "La ciberseguridad debe integrarse en todas estas inversiones digitales, especialmente en las tecnologías clave como la inteligencia artificial (IA), el cifrado y la computación cuántica."

En paralelo, el legislador europeo ha dado un paso firme para la revisión de los marcos regulativos vigentes en materia de seguridad en la UE, que analicé anteriormente. En particular, la Comisión Europea presentó sendas propuestas para abordar la resiliencia física y cibernética de las entidades y redes críticas, en particular, la *Directiva sobre medidas para un alto nivel común de ciberseguridad en toda la Unión*[305], también conocida como "Directiva NIS revisada" o "NIS 2" y una nueva *Directiva sobre resiliencia de entidades críticas*[306].

Estos instrumentos jurídicos pretenden abordar tanto los riesgos actuales como futuros *off* y *online*, desde ciberataques hasta delitos o desastres naturales bajo un enfoque coherente y complementario a los marcos actuales.

La Directiva "NIS 2" pretende actualizar a su predecesora, analizada anteriormente, ante el panorama actual de amenazas y nuevos desafíos que requieren respuestas adaptadas e innovadoras. Con este propósito, la Directiva propuesta amplia el alcance de la actual Directiva incluyendo nuevos sectores por su relevancia para la economía y la sociedad,

[304] *The EU's Cybersecurity Strategy for the Digital Decade.* Joint Communication to the European Parliament and the Council.. Bruselas. 16.12.2020

[305] Proposal for a Directive of the European Parliament and of the Council on measures for a high common level of cybersecurity across the Union, repealing Directive (EU) 2016/1148. Bruselas. 16.12.2020. COM (2020) 823 final.

[306] Proposal for a Directive of the European Parliament and of the Council on the resilience of critical entities. Bruselas. 16.12.2020. COM (2020) 829 final.

pero orientada a mediana y grandes organizaciones, sin perjuicio de dejar cierta flexibilidad a los Estados miembros para identificar entidades más pequeñas, pero con un perfil de riesgo de seguridad elevado.

La propuesta también elimina la distinción entre operadores de servicios esenciales y proveedores de servicios digitales, clasificándose como categorías esenciales e importantes.

La propuesta refuerza los requerimientos de seguridad e impone un enfoque de gestión de riesgos con un listado de requerimientos básicos de seguridad que deben aplicarse, lo que aporta claridad y objetividad a la norma. También precisa algunos aspectos sobre notificación de brechas de seguridad, contenido de informes y plazos.

Siguiendo las inquietudes evidenciadas en el informe ENISA referido, se aborda igualmente la seguridad en la cadena de suministro de tecnología y relaciones con proveedores, de modo que se asegure la ciberseguridad por los participantes en la misma.

Además, contempla la realización de evaluaciones de riesgo coordinadas de las cadenas de suministro consideras críticas por parte de los Estados miembros -en cooperación con la Comisión Europea y ENISA-, basándose en el enfoque llevado a cabo con éxito en el contexto de la *Recomendación de la Comisión sobre la ciberseguridad de las redes 5G*[307].

La propuesta también introduce una supervisión más estricta de las autoridades nacionales, requisitos de ejecución más estrictos y pretende armonizar los regímenes sancionadores en los Estados miembros.

Por último, la propuesta establece un marco básico de actores clave responsables sobre la divulgación coordinada de vulnerabilidades, la creación de un registro de las mismas y refuerza el rol del denominado *Grupo de Cooperación NIS* -creado por la Directiva NIS para garantizar la cooperación estratégica y el intercambio de información entre los Estados miembros de la UE en ciberseguridad-, en la configuración de decisiones políticas estratégicas sobre tecnologías emergentes y nuevas tendencias, aumenta el

[307] Recuperado de: https://ec.europa.eu/digital-single-market/en/news/cybersecurity-5g-networks. Consultado el 30.03.2021.

intercambio de información y cooperación entre autoridades de los Estados miembros y mejora la cooperación operativa, incluyendo la gestión de crisis de ciberseguridad.

Por lo que se refiere a España, publicó su *Estrategia de Ciberseguridad*[308] en 2019, precedente al análisis europeo de las amenazas de ciberseguridad de la inteligencia artificial, y en la que no se abordó específicamente la inteligencia artificial, sin perjuicio de significar que sus implicaciones, al igual que la robótica, el *Big data*, el *Blockchain* y el *Internet de las Cosas*, van más allá de la dimensión tecnológica, se extienden hacia nuevos modelos sociales y se adentran en el campo de las relaciones personales y la ética.

Con posterioridad España aprobó su *Estrategia Española de I+D+I en inteligencia artificial en 2019*[309], que contempla como esencial la investigación, desarrollo y aplicación en las tecnologías de la inteligencia artificial dedicada a los sistemas de ciberseguridad para detectar y repeler amenazas, mediante las tecnologías del lenguaje, el análisis de imágenes y el aprendizaje automático, se plantea esencial.

Dentro de España, algunas comunidades autónomas como Cataluña[310] y la Comunidad Valenciana[311] han asumido cierto liderazgo regional y han creado y aprobado sus estrategias en materia de inteligencia artificial.

A nivel internacional son múltiples las organizaciones supranacionales y estados que han aprobado en los últimos tres años sus estrategias sobre inteligencia artificial[312], que se suman a sus estrategias de ciberseguridad, muchas de ellas sin un tratamiento específico de la inteligencia artificial[313]. No obstante, la mayoría de países como China, Canadá,

---

[308] *Estrategia Nacional de Ciberseguridad*. Ministerio de la Presidencia, relaciones con las Cortes e Igualdad. 2019

[309] *Estrategia Española de I+D+I en inteligencia artificial*. Ministerio de Ciencia, Innovación y Universidades. 2019. Disponible en: https://www.ciencia.gob.es/stfls/MICINN/Ciencia/Ficheros/Estrategia_Inteligencia_Artificial_IDI.pdf. Consultada el 02.02.2021.

[310] Recuperado de: https: //participa.gencat.cat/uploads/decidim/attachment/f ile/818/Document-Bases-Estrategia-IA-Catalunya.pdf. Consultado el 27.02.2021.

[311] Recuperado de: http://www.presidencia.gva.es/documents/80279719/169117420/Dossier_cas. pdf/88361b83-0e33-4b49-99c0-ad894ffc0f75. Consultado el 27.02.2021.

[312] Recuperado de: https://futureoflife.org/national-international-ai-strategies/. Consultado el 21.03.2021.

[313] A modo de ejemplo citar la *National Cyber Security Strategy 2016-2021* de Reino Unido. Disponible en: https://www.gov.uk/government/publications/national-cyber-security-strategy-2016-to-2021. Consultada el 02.01.2021.

India, Japón y Francia contemplan en sus estrategias nacionales de inteligencia artificial la ciberseguridad como un campo de aplicación.

A nivel europeo, en la fecha de finalización de esta investigación, 20 de sus Estados miembros han adoptado estrategias nacionales en materia de inteligencia artificial, en particular, Alemania, Bulgaria, Chequia, Chipre, Dinamarca, Eslovaquia, Estonia, Finlandia, Francia, Hungría, Letonia, Lituania, Luxemburgo, Malta, Países Bajos, Portugal, Suecia, España, Polonia y Noruega.

Distintos países están complementando estas estrategias con otros instrumentos políticos y técnicos que abordan las aplicaciones sectoriales de la inteligencia artificial, algunos de ellos con instrumentos normativos como la UE, conforme citaré en el capítulo IV.

Las políticas gubernamentales y sus medidas de aplicación persiguen el triple objetivo de fomentar la adopción de la inteligencia artificial, maximizar sus beneficios y minimizar los riesgos asociados.

En resumen, las iniciativas e informes precitados evidencian que se va en la dirección correcta. No obstante, la velocidad de cambio constante en el que nos hallamos inmersos requiere soluciones ágiles a los problemas inminentes, por lo que no puede haber lugar a la relajación para poder anticiparnos a los ciberdelincuentes y elaborar estrategias de protección con medidas y controles adecuados y eficaces.

La investigación e innovación en ciberseguridad debe ser incesante, y sus estrategias deben ser revisadas de manera continua para adaptarse a las nuevas amenazas resultantes no sólo de la innovación y el desarrollo, sino de las nuevas aplicaciones de las tecnologías ya existentes en todos los ámbitos y sectores de nuestra de vida, teniendo en cuenta las singularidades de sectores y aplicaciones específicos.

7. Seguridad en el diseño.

De los riesgos, retos, propuestas, informes y marcos normativos analizados debe concluirse la necesidad urgente de consolidar la exigencia de una "*Security by design*" en

los sistemas de inteligencia artificial para garantizar una inteligencia artificial robusta, confiable y segura, que garantice desde su diseño y concepción la seguridad física y psíquica de las personas, de sus datos, de sus bienes, de sus derechos, su seguridad lógica o virtual y la seguridad jurídica.

En mi opinión, la seguridad debe constituir una norma ética y jurídica vinculante a todo sistema de inteligencia artificial, cualquiera que sea su nivel de riesgo inicial, por las características y capacidades de las que puede estar dotado, por lo que debe ser un requisito y exigencia para los mismos en los futuros marcos reguladores de la inteligencia artificial.

En este sentido el Parlamento Europeo, en su Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial[314], considera que, además de la actualización del marco de establecido por la Directiva 85/374/CEE en materia de responsabilidad por los daños causados por productos defectuosos[315], debe acompañarse de la actualización de la Directiva 2001/95/CE del Parlamento Europeo y del Consejo, de 3 de diciembre de 2001, relativa a la seguridad general de los productos[316], a fin de garantizar que los sistemas de inteligencia artificial incorporen los principios de seguridad y protección desde el diseño, esto es, la *Security by design*, en congruencia con su exigencia ética y también jurídica en la Propuesta de Reglamento incorporada en la coetánea Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas[317].

Como he referido, esta exigencia también se haya regulada en los artículos 9, 15, siguientes y concordantes de la Propuesta de Reglamento del Parlamento Europeo y del

[314] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))
[315] Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos. OJ L 210, 7.8.1985. Pp. 29-33.
[316] Directiva 2001/95/CE del Parlamento Europeo y del Consejo, de 3 de diciembre de 2001, relativa a la seguridad general de los productos. OJ L 11, 15.1.2002. Pp. 4-17.
[317] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))

Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, que será objeto de análisis detallado en el capítulo IV.

8. Seguridad, responsabilidad civil y penal

Si la ética exige seguridad, conforme he expuesto y abordaré con mayor profundidad con posterioridad al analizar los distintos marcos éticos y jurídicos propuestos en el ámbito de la UE, también la ética exige responsabilidad jurídica dimanante del incumplimiento de sus principios y normas, así como por los daños que pueda sufrir la persona afectada por los mismos.

Del mismo modo, la seguridad exige responsabilidad, que puede derivarse tanto de actos lesivos, constituyan o no ilícitos civiles o penales.

9. Relación e interacción con otros sistemas y tecnologías

9.1. Aspectos generales

El incesante avance tecnológico y el crecimiento exponencial del desarrollo y aplicación de la inteligencia artificial nos está llevando a una aceleración de la fusión entre el mundo físico y virtual, una interrelación obligada hombre-máquina cada vez más estrecha, así como a la convergencia de diferentes tecnologías, como Internet de las Cosas -IoT por sus siglas en inglés-, *Blockchain*, la robótica o las tecnologías de sensores, así como a la creciente cantidad y variedad de datos, así como a sus nuevas características -por ejemplo, los datos distribuidos-, para emplear la inteligencia artificial a gran escala.

De nuevo, esa relación e interacción comporta un aumento del poder y potencial de la inteligencia artificial y sus aplicaciones, pero tanto en sentido positivo como también negativo, como he expuesto anteriormente, aumentando cualitativa y cuantitativamente los riesgos y sus impactos, entre otros, los de ciberseguridad.

La necesaria comunicación e interoperabilidad con otros sistemas y tecnologías es esencial para explotar las ventajas y oportunidades que supone la inteligencia artificial, máxime en la medida que la inteligencia artificial no es una sola tecnología, sino una diversidad de tecnologías sustentadas en *software* pero integradas por otros elementos imprescindibles como los datos, y que pueden utilizarse de distintas formas en un número potencialmente ilimitado de aplicaciones, en casi todos los sectores.

Y para ello la estandarización es esencial con este propósito. Ya se está trabajando en ello a nivel internacional.

En 2018, el comité técnico conjunto *ISO/IEC JTC 1, Tecnologías de la Información* de ISO y la Comisión Electrotécnica Internacional (IEC) fundó el subcomité SC 42 sobre inteligencia artificial[318], constituyendo el primer ecosistema de la inteligencia artificial.

Tarek Besold[319], es uno de los miembros clave de este comité y en relación a la responsabilidad que comporta el potencial poder de la inteligencia artificial, ha venido significando la necesidad de definir la situación actual y adoptar definiciones sensatas sobre los mecanismos y tecnologías de la inteligencia artificial, en particular, el desarrollo de normas y estándares, que es una ingente tarea, pero la interoperabilidad es vital ante el enorme alcance de la inteligencia artificial.

A continuación, me permito simplemente significar algunas de estas relaciones e interacciones con todo lo que las mismas suponen, tanto positivo como negativo, potenciando algunas de sus bondades y poder, pero también acrecentando algunos de sus desafíos y aumentando algunos de sus riesgos y posibles impactos.

---

[318] Recuperado de: ISOfocus Noviembre-diciembre 2019. P. 2. Disponible en: https://www.iso.org/files/live/sites/isoorg/files/news/magazine/ISOfocus%20(2013-NOW)/sp/ISOfocus_137_sp.pdf. Consultado el 05.02.2021.

[319] Recuperado de: ISOfocus Noviembre-diciembre 2019. P. 2. Disponible en: https://www.iso.org/files/live/sites/isoorg/files/news/magazine/ISOfocus%20(2013-NOW)/sp/ISOfocus_137_sp.pdf. Consultado el 05.02.2021.P. 9

9.2. Big data

El *Big data* podemos definirlo como un conjunto inmenso de datos, estructurados o no estructurados, cuyo crecimiento es exponencial y cuyo tratamiento no es posible a través de las aplicaciones informáticas de procesamiento de datos tradicionales o convencionales.

El tratamiento y análisis de este volumen de datos es imposible si no se utilizan tecnologías especiales para ello, en especial, la inteligencia artificial.

La disponibilidad y acumulación masiva de datos exige capacidad de alojamiento y procesamiento, así como la disposición de la tecnología y recursos necesarios para captarlos, analizarlos y procesarlos en tiempo razonable utilizando distintos métodos, para facilitar su gestión, generar conocimiento y tomar decisiones, para lo que la inteligencia artificial constituye un instrumento idóneo.

Y no sólo la inteligencia artificial es un instrumento en este sentido, sino que los propios macrodatos se han convertido en el recurso esencial para el desarrollo y, sobre todo, la aplicación de la inteligencia artificial.

Como expuse anteriormente, la mayor disponibilidad de datos está ayudando a determinados países a liderar la aplicación de la inteligencia artificial.

Del mismo modo, se están planteando cuestiones éticas y jurídicas relacionadas con ese acceso creciente a datos para el bien común, como igualmente expuse.

La inteligencia artificial precisa para su despliegue y aplicación grandes cantidades de datos para poder entrenar los algoritmos. Su eficiencia y eficacia depende principalmente de la cantidad y calidad de los datos sobre los que opere.

No obstante, como he indicado al analizar los retos y riesgos, todo ello puede tener un enorme impacto en las personas afectadas por el tratamiento de los mismos, especialmente en materia de privacidad y protección de datos y no discriminación.

El procesamiento masivo de datos para la extracción de información sobre comportamientos colectivos es uno de los paradigmas y tecnologías que ha permitido impulsar el *neuromarketing* por parte de la inteligencia artificial, entendido como la ciencia de la lectura de la mente de los consumidores para medir sus reacciones a los estímulos de marketing.

9.3. Cloud

De nuevo, no estamos ante una tecnología en sí misma ni desde luego nueva, sino más bien un servicio. Conforme he referido anteriormente, el concepto se atribuye igualmente a John McCarthy, cuyo origen se sitúa a mediados del siglo pasado.

El *cloud computing* constituye un modelo de servicio, que no tecnología, que igualmente potencia las capacidades y usos de la inteligencia artificial, así como sus retos y riesgos, especialmente ante la ubicación y tratamiento de los sistemas y los datos en infraestructuras de terceros, y que pueden hallarse en territorio de la UE o fuera del mismo, con todas las implicaciones que ello conlleva, por ejemplo, en materia de protección de datos y cumplimiento regulativo, especialmente ante la exigencia de mecanismos adecuados de transferencia de datos en caso de proveedores ubicados fuera del territorio de la misma o asimilados.

No existe un marco legal específico regulador de estos servicios, si bien, la UE ha identificado los mismos como un objetivo estratégico en el marco de su estrategia digital y de gobierno de datos.

Las tendencias de este paradigma de servicios se orientan, según Gartner[320], hacia la denominada la nube distribuida, esto es, la distribución de servicios de nube pública a

[320] Gartner Top 10 Strategic Technology Trends for 2020". 21.10.2019. Disponible en: https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020/. Consultado el 02.01.2021

ubicaciones fuera de los centros de datos físicos del proveedor de nube, pero controladas por el proveedor.

Los centros de datos están ubicados en cualquier lugar lo que resuelve, de un lado, problemas técnicos como la latencia y, de otro, según Gartner, los desafíos regulativos como la soberanía de los datos. Este modelo ofrecería los beneficios de un servicio de nube pública junto con los beneficios de una nube local privada.

En este tipo de nube, el proveedor de la misma es responsable de todos los aspectos de la arquitectura, la entrega, las operaciones, la gobernanza y las actualizaciones del servicio en la nube. Según esta consultora de referencia internacional, la evolución de la nube pública centralizada a la nube pública distribuida marca el comienzo de una nueva era de la computación en la nube.

9.4. Blockchain

Tampoco estamos hablando de una tecnología especialmente novedosa, en la medida que su origen se sitúa en 2008, como tecnología desarrollada para generar *bitcoins*.

Blockchain no es ni más (ni menos) que un registro de transacciones descentralizado. Es una tecnología que surgió a raíz de la creación del bitcoin por parte de una persona o grupo de personas anónima/o, que se autodenominó Satoshi Nakamoto.

De manera muy básica, podría definirse como una tecnología que permite gestionar un registro descentralizado de transacciones de todo tipo, es decir, llevar un libro mayor a través de internet, encriptado y que genera confianza entre las partes, ya que se puede verificar cualquier información dentro de la cadena.

Blockchain registra todas las transacciones en bases de datos distribuidas entre sus participantes, prescindiendo de una base centralizada, por lo que no está gestionada ni custodiada por ninguna entidad pública -gobiernos- ni privada -por ejemplo, empresas o bancos-.

De este modo, *Blockchain* permite a las empresas realizar y rastrear una transacción sin necesidad de intermediarios, lo que permite minimizar costes y los tiempos de liquidación de transacciones, así como mejorar el flujo de efectivo.

No existe un marco legal específico regulador de la misma a nivel europeo y español, si bien, Liechtenstein fue el primer país en aprobar el primer marco específico sobre *Blockchain*.

La combinación de *Blockchain* con otras tecnologías como la realidad virtual o aumentada puede transformar radicalmente la forma en que las personas ven la televisión, asisten a un concierto o juegan a videojuegos a través de un entorno inmersivo e interactivo donde se producen multitud de eventos simultáneamente como recibir publicidad, contratar un producto o contenido o realizar un pago.

Actualmente, todavía es una tecnología objeto de investigación y experimentación por múltiples de consorcios internacionales para su aplicación futura a todo tipo de sectores.

En mi opinión, es una tecnología que está revolucionando y revolucionará los negocios y los modelos actuales, creando modelos disruptivos, y contribuirá a la democratización de la Sociedad Digital.

Lo que parece evidente es que la denominada "*Blockchain* completa" transformará las industrias y sectores completos, e incluso la propia economía, máxime ante su integración con tecnologías complementarias como la inteligencia artificial o el *Internet de las Cosas* -IoT-, con posibilidad de que las máquinas y sistemas de inteligencia artificial formen parte de las redes *Blockchain* y participen en el intercambio y contratación de todo tipo de activos. Por ejemplo, como expone Gartner en el informe precitado[321], la posibilidad de que un coche negocie los precios del seguro directamente con la compañía de seguros em función de los datos recopilados por sus sensores.

[321] Gartner Top 10 Strategic Technology Trends for 2020". 21.10.2019. Disponible en: https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020/. Consultado el 02.01.2021

Sin perjuicio de todo ello, los expertos atribuyen a la tecnología un *Blockchain* un carácter instrumental en el marco de la inteligencia artificial, reflejado como ejemplo en los *smart contracts*, que algunos autores, entre otros, Gallego Sánchez[322], lo consideran una manifestación de la inteligencia artificial, cuya eficacia ha sido mejorada por el *Blockchain*.

No comparto completamente estas calificaciones en la medida que, de nuevo, corremos el riesgo de asociar el atributo de inteligencia artificial a todo producto tecnológico o codificado con origen en la inteligencia humana. En el posterior apartado abordo sucintamente el concepto de *smart contract*.

Gartner[323] prevé que a partir de 2025 la tecnología *Blockchain* incorpore tecnologías complementarias como el *Internet de las Cosas* (IoT), inteligencia artificial y la denominada identidad autónoma descentralizada (SSI), lo que permitirá su eclosión en múltiples sectores. Además, prevé en dicho informe cambios en los modelos de negocio sustentados en esta tecnología, conforme los sistemas autónomos adquieran la capacidad de interactuar comercialmente y operar independientemente de un ser humano, lo que considera un hecho que sucederá, que choca frontalmente con la posición de los órganos legislativos y ejecutivos de la UE en relación con la posible personalidad jurídica y capacidad de obrar de estos sistemas y el necesario control y supervisión humana.

9.5. smart contracts

Un *smart contract* es un protocolo informático compuesto por un conjunto de instrucciones para su gestión y ejecución automáticamente conforme a las mismas.

De este modo puede ser utilizada para ejecutar un contrato de forma automática, sin necesidad de un intermediario físico, es decir, gestiona y ejecuta acuerdos establecidos

[322] GALLEGO, E. (2019). "La patentabilidad de la inteligencia artificial. La compatibilidad con otros sistemas de protección". *La Ley Mercantil*. Nº 59, de 1 de junio 2019. Wolters Kluwer 2019. P. 6.
[323] Gartner Top 10 Strategic Technology Trends for 2020". 21.10.2019. Disponible en: https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020/. Consultado el 02.01.2021

entre dos o más partes haciendo que ciertas acciones sucedan como resultado del cumplimiento de una serie de condiciones específicas y operaciones programadas. Es decir, que permite ejecutar automáticamente lo que previamente se ha definido por el programador y/o por el abogado cualificado para ello, pero sin realmente conocimiento ni consciencia de lo que se está ejecutando automatizadamente.

Una vez más, no estamos ante un instrumento especialmente novedoso. El origen del término se sitúa en 1993 y se asocia al criptógrafo Nick Szabo, a pesar de la inexistencia entonces de la tecnología que ha permitido su desarrollo: *Blockchain*.

Los *smart contracts* garantizan de una manera segura y sin intervención de terceros la ejecución y cumplimiento de un contrato.

Las partes programan las condiciones, firman en prueba conformidad y la tecnología *Blockchain* se encarga de que no se modifique, así como de su ejecución.

La legislación del Estado de Arizona[324] los define como "*an event-driven program, with state, that run on a distributed, decentralized, shared and replicated ledger and that can take custody over and instruct transfer of assets on that ledger".*

En resumen, los *smart contracts* pueden definirse, en general, como un conjunto de protocolos informáticos que permite a un dispositivo por sí mismo procesarlos y ejecutarlos de forma autónoma, sin necesidad de intervención humana.

Entre sus principales ventajas, su objetividad -no hay lugar a la interpretación, si se cumplen las condiciones objetivas el contrato se ejecutará automáticamente en tiempo real, sin intermediarios y sin intervención humana-, seguridad en el cumplimiento de los contratos, ahorro de tiempo y costes, y prevención de fraudes e impagos.

[324] Bill HB 2417 31.03.2017. Definición traducida libremente por el autor: "*Un programa basado en eventos, con estado, que se ejecuta en un libro de contabilidad distribuido, descentralizado, compartido y replicado, y que puede asumir la custodia e instruir la transferencia de activos en ese libro de contabilidad*".

Entre sus principales desventajas, necesidad de enfoque global, la seguridad y la necesidad de auditorías y *due diligence*, especialmente en materia de gobernanza.

La utilidad de la inteligencia artificial para su elaboración como para su gobernanza es evidente.

9.6. Internet de las Cosas (IoT)

El denominado *Internet de las Cosas -Internet of Things* o IoT por sus siglas en inglés- tampoco es un concepto nuevo. Fue introducido en 1999 por Kevin Ashton del MIT[325], y se refiere al mecanismo por el que sensores y transmisores están conectados por redes a sistemas informáticos y se comunican entre sí, algo que ha sido posible gracias al abaratamiento de los sensores y del acceso a internet.

La denominada "revolución 4.0" o "cuarta revolución industrial" sustentada en la tecnología y, en especial en la inteligencia artificial y en IoT, comporta un incremento de los retos y riesgos en materia de seguridad.

Este concepto se consolidó definitivamente en la *Reunión Anual del Foro Económico Mundial* de 2016, en el que se abordó también el crecimiento exponencial de la vulnerabilidad de las empresas antes los ciberataques, especialmente en la medida que el 84% de sus activos son ya intangibles. Además, se significó que la gran mayoría de las brechas de seguridad -entre en un 75% y un 90%- están causadas por errores humanos y fugas de información, destacando la necesidad de la educación empresarial en materia de seguridad y de concienciación a las personas sobre su importancia, especialmente ante una digitalización total o el denominado *Internet del todo* -IOE por sus siglas en inglés o *Internet of everything*-.

[325] DE LA TORRE, I. (2018). "La disrupción tecnológica ya está aquí. Cómo afecta a las personas, los gobiernos y las empresas". En Cuadernos de Estrategia 199. *Gobernanza futura: Hiperglobalización, mundo multipolar y Estados menguantes*. Instituto Español de Estudios Estratégicos. Ministerio de Defensa. Diciembre 2018. P. 40.

La utilización del internet de las cosas con redes neuronales para predecir el futuro conforme como el aprendizaje profundo. Esta tecnología, que existe desde los años setenta, se ha desarrollado aceleradamente a medida que aumentaba la capacidad de procesar datos

Las redes neuronales son sistemas de cálculo basados en el comportamiento observado en cerebros biológicos, a través de neuronas individuales y sus conexiones, activando o inhibiendo las adyacentes. Los sistemas acaban aprendiendo y formándose a sí mismos y son muy útiles para ofrecer soluciones donde la programación convencional no es capaz de llegar.

9.7. Smartcities

Las ciudades inteligentes son una realidad cada vez más próxima, si bien, en relación con determinados servicios y dotaciones.

Las ciudades inteligentes pretender disponer de áreas comerciales, residenciales e industriales diseñadas utilizando marcos inteligentes de ecosistemas urbanos y donde la inteligencia artificial puede controlar parte de su vida diaria como, por ejemplo, la regulación del tráfico rodado.

La inteligencia artificial permite soluciones urbanas inteligentes aportando innumerables ventajas, entre otras, una gestión más eficiente de la energía, el agua y los residuos, la reducción de la contaminación, el ruido y las congestiones de tráfico.

No obstante, también comporta importantes retos para las autoridades locales, en especial, la disponibilidad y fiabilidad de la tecnología y los datos, la dependencia de terceros privados y la falta de competencias, los retos éticos para el uso de la inteligencia artificial y la ausencia de sesgo o la dificultad de regular las infraestructuras y los datos interdependientes, respectivamente.

Un informe en preparación en la fecha de cierre de esta investigación, solicitado por el Comité Especial AIDA[326], recoge algunas recomendaciones frente a estos retos, como el apoyo en toda la UE a la infraestructura y la gobernanza en materia de digitalización para mejorar la eficiencia y garantizar al mismo tiempo la recopilación imparcial de datos, la inclusión de la inteligencia artificial urbana en los programas de investigación de la UE, armonización de las políticas relacionadas con la inteligencia artificial en la UE y la adoptación de procedimientos de contratación pública innovadores, que impliquen requisitos para una inteligencia artificial técnica y éticamente responsable. Esto último, es otra de las posibles vías para la exigencia de la ética ante la ausencia de regulación legal, mediante su definición y exigencia, no como requerimiento regulatorio, sino como requerimiento contractual (público).

Los retos y riesgos expuestos de distinta naturaleza son evidentes y pueden plantear aspectos singulares en materia de privacidad, ante sistemas de reconocimiento facial para acceder a determinados servicios o instalaciones, o de responsabilidad.

Por ejemplo, ante la interacción de los sistemas inteligentes gestores del tráfico rodado que interaccionan con vehículos autónomos que circulan por las vías destinadas a los mismos, en caso de accidente en un cruce supuestamente gestionado automatizadamente ¿Quién sería el responsable? ¿El vehículo autónomo? ¿El propietario? ¿El sistema inteligente gestor del tráfico rodado? ¿El Ayuntamiento y a través de la responsabilidad patrimonial de las Administraciones públicas?

El diseño y desarrollo de las ciudades inteligentes requerirán la utilización de distintas tecnologías y servicios como *Blockchain*, IoT, inteligencia artificial, *Big data* o *Cloud*, lo que potenciará los retos y riesgos asociados.

[326] *Special Committee on Artificial Intelligence in a Digital Age* -AIDA por sus siglas en inglés, creado por el Parlamento Europeo en su sesión plenaria de 18 de junio de 2020.

9.8. Realidad aumentada, virtual y extendida

Se trata de distintas tecnologías con distinto alcance e impacto en la persona y los negocios, que la inteligencia artificial puede potenciar a niveles inimaginables.

La Realidad Virtual -VR por sus siglas en inglés- saca visualmente al usuario de su entorno del mundo real y lo sumerge en un entorno completamente virtual para navegar por el mismo.

La Realidad Aumentada -AR por sus siglas en inglés- superpone objetos digitales (información, gráficos, sonidos) en el mundo real, lo que permite al usuario experimentar la relación entre el mundo digital y el físico.

Por su parte la Realidad Mixta -MR por sus siglas en inglés- superpone objetos digitales en el mundo real y ancla los objetos virtuales y reales entre sí, lo que permite al usuario interactuar con objetos virtuales/reales combinados.

Y, por último, la Realidad Extendida -XR por sus siglas en inglés hace referencia a la gama completa de experiencias inmersivas que permiten la interacción humana entre los mundos físico y digital (o virtual). Esto incluye la realidad aumentada, la realidad virtual y la realidad mixta, así como técnicas más amplias que usan y mejoran los sentidos humanos como hápticos, hologramas y más.

Los estudios internacionales evidencian la utilidad y ventajas de estas tecnologías y no sólo en los sectores popularmente más asociados a estas tecnologías, como el del ocio interactivo, turismo o cultura, sino que de manera exponencial su impacto es cada vez mayor en la industria, educación o sanidad, como así lo muestran distintos informes[327].

[327] Informe "The App Date y Oarsis" Fundación Telefónica. 2018 o Informe "*Realidad virtual y aumentada. Entendiendo la carrera por la próxima plataforma informática*" de Goldman Sachs Group.

De hecho, la realidad extendida es una de las tecnologías incluidas en las denominadas "DARQ" junto con el *Blockchain*, la inteligencia artificial y la computación cuántica, consideradas de mayor impacto.

Entre sus ventajas en distintos sectores, destaca la mejora de la formación, el entrenamiento y de la experiencia del usuario en sectores como el de la salud, la generación de nuevas terapias inmersivas o la mejora de la productividad de los trabajadores incluso en los sectores industriales mediante la colaboración hombre-máquina[328].

Según el informe *A responsible future for immersive technologies* 2019 de Accenture referenciado al pie, el crecimiento de inversión y patentes en estas tecnologías se está disparando.

No obstante, como toda tecnología, supone distintos retos y riesgos, en especial contra la privacidad o la *Deep privacy,* ante el uso y tratamiento de datos que están profundamente conectados con la identidad personal, los comportamientos y pensamientos íntimos, y el riesgo de su posible uso indebido, no autorizado o sustracción de datos relacionados con sentimientos, comportamientos, juicios y semejanzas físicas entre personas.

También el riesgo de experiencias falsas o *fakes experiences* ante la dificultad de identificar la realidad de la falsedad, pudiendo influir profundamente en los comportamientos, opiniones y decisiones.

También en materia de ciberseguridad que han sido evidenciados por investigadores del del MIT y Harvard[329], tanto específicos como generales que pueden aplicar a cualquier sistema, como el *ransomware*.

También riesgos importantes relacionados con la salud física y psicológica, ante el posible alejamiento del mundo real, así como de impacto mental de alcance desconocido

[328] Según el Informe "*A responsible future for immersive technologies*" de Accenture, de 2019, el uso de la realidad extendida podría aumentar el 21% del tiempo de trabajo, y podría llegar al 30% en servicios de salud y sociales, fabricación y construcción.
[329] FINLAYSON, S. G., BOWERS, J. ET AL. (2019). "Adversarial attacks on medical machine learning". *Science Journal*. Vol. 363, Issue 6433, pp. 1287-1289. March 22, 2019. DOI: 10.1126/science.aaw4399.

en función de uso y contexto. Estas tecnologías implican conexiones directas con nuestra mente y percepción de la realidad que aún no se comprenden completamente.

De hecho, generan impacto en las emociones, son empleadas para el *neuromarketing* comentado anteriormente, permiten las relaciones virtuales hiperrealistas con el consiguiente riesgo de crear una realidad sintética o *ciberrealidad* distinta a la física adulterando o sacando al individuo de ésta.

Y también, la dificultad de reversión de errores en su uso y funcionamiento, especialmente por su interconexión, descentralización y velocidad de distribución.

Si asociamos la inteligencia artificial a todas estas tecnologías el potencial benefactor o perjudicial se dispara exponencialmente.

Las tecnologías inmersivas no sólo contribuirán a cambiar el modo en el que los usuarios interactúan con el mundo sino la forma también en la que lo perciben y cada vez más será más difícil diferenciar entre lo real y lo virtual.

El mercado de la realidad virtual ha pasado de 2.600 millones de dólares en 2015 a 20.000 millones de dólares en 2019. Se estima que alcanzará 66.680 millones dólares en 2022 y algunas consultoras, como DigiCapital que elevan esta cifra a 90.000 millones[330].

9.9. Impresión 3D y 4D

A pesar de los avances que tenemos ya hoy, la industria de la impresión 3D está todavía muy lejos de alcanzar su madurez, pero desde luego es uno de los factores de una nueva revolución industrial en la que ya estamos.

[330] VIDAL, M. (2019). *La era de la humanidad*. Posición 2569 Versión Kindle. Ediciones Deusto. Barcelona. 2019.

Las máquinas y técnicas de fabricación para crear objetos basados en modelos digitales utilizan ya todo tipo de materiales para la creación de objeto o incluso órganos vitales utilizando desde plástico, metal o gel hasta tejido vivo.

Los principales factores de disrupción de esta tecnología son la alteración de la cadena de suministro, reducción de los costes y de los tiempos de fabricación y transporte, la personalización de los productos (especialmente en el sector salud y ortopédico), contribución al desarrollo sostenible.

Esta tecnología ha evolucionado todavía más adicionando una cuarta dimensión, que sería el tiempo, en la denominada impresión 4D, de modo que los objetos cambiarán de forma con autonomía o podrán autoensamblarse.

No obstante, de nuevo, junto a sus ventajas y bondades se suscitan algunos retos, especialmente en el ámbito de la propiedad industrial (patentes) y la necesidad de licencias obligatorias para imprimir determinadas invenciones y diseños industriales, como se ha planteado en la actualidad en el marco de la pandemia mundial ante la ausencia de respiradores en hospitales.

La inteligencia artificial potencia las ventajas de esta tecnología y la precisa, en la medida que los sistemas inteligentes orientados, por ejemplo, a la generación de resultados supuestamente fruto de la creatividad o la invención, precisarán integrar estas tecnologías para su materialización, planteando las cuestiones sobre la protección del resultado y la posible aplicación de los marcos reguladores de la propiedad intelectual e industrial, conforme igualmente abordó al final de esta investigación.

9.10. Redes e interfaces neuronales. Biotecnología, neurotecnología y otras

La posibilidad de interactuar con una máquina por medio de la actividad cerebral activada por un sujeto es una realidad.

Su aplicación en el ámbito de la salud y para facilitar la vida a personas con movilidad reducida es infinito.

La simbiosis entre inteligencia artificial y el Internet de las Cosas en su aplicación puede tener infinitas aplicaciones, desde la conducción de un vehículo, participar en un videojuego, practicar una cirugía con implicaciones neurológicas con mayor precisión, tratamiento del dolor, manejar una prótesis, un ordenador o gobernar un exosqueleto.

Incluso ya hoy, se ha demostrado la transferencia de pensamiento muy básica en condiciones de laboratorio, lo que podría ser posible llevarlo a entornos reales en unas décadas conforme recoge el informe de *The Royal Society* británica al que aludiré posteriormente. No obstante, es evidente que igualmente comporta evidentes riesgos, especialmente en materia de seguridad física.

En este sentido, significar su aplicación en el sector de la salud o la dependencia.

Podríamos encontrarnos igualmente ante sistemas que podrían integrar inteligencia artificial de ayuda o soporte a la humana o, incluso, prescindiendo de ésta o de consciencia alguna en casos graves de afectación física e intelectual de "grandes dependientes" o personas con grados de discapacidad elevada y ausencia absoluta de movilidad, lo que de nuevo nos lleva a hablar de los importantes riesgos que ello puede comportar y de la necesidad de integrar en su diseño y funcionamiento la ética y todos los valores humanos que representa y la seguridad.

La tecnología neuronal podría traer un cambio aún más profundo que el que plantea la propia inteligencia artificial en sí misma, en la medida que vincula y asocia el poder cognitivo del cerebro a la potencia de procesamiento del aprendizaje automático y la supercomputación.

Las interfaces neuronales conectan el cerebro o el sistema nervioso con equipos, normalmente dispositivos digitales o sistemas informáticos. Algunas actúan para registrar la actividad fisiológica, como las señales o los movimientos del cerebro, otras las estimulan y otras efectúan ambas actividades.

Las interfaces ofrecen beneficios que son inimaginables actualmente y permitirán mejorar al ser humano, su salud, su memoria o su concentración. Actualmente ya se ofrecen terapias para personas con enfermedades como el ictus, epilepsia, parálisis o depresión, y

pueden ofrecer oportunidades de mejora de la concentración, la toma de decisiones y la colaboración, y también pueden contribuir a mejorar la salud, forma física y el bienestar de las personas.

Pero también comporta riesgos de alto impacto, para empresas gobiernos y ciudadanos, especialmente en materia de privacidad y otros derechos humanos, desigualdad social, así como riesgos de seguridad.

En este sentido destacar el informe *iHuman. Blurring lines between mind and machine*[331], elaborado por la *The Royal Society* británica, que advierte de la necesidad de que los responsables políticos, los empresarios y los ciudadanos se preparen para esta esta nueva ola de tecnología neuronal, creando las estructuras y los sistemas necesarios para aprovechar las oportunidades, gestionar los riesgos -especialmente de seguridad y éticos- y abordar adecuadamente las cuestiones fundamentales.

El informe se acompaña con una serie de principios y buenas prácticas a las que adiciona en sus conclusiones la necesidad de crear un marco regulativo de esta tecnología en el ámbito de la tecnología más proactivo y anticipatorio en la línea en la que ya Reino Unido está trabajando a través de del denominado *Ministerial Working Group for Future Regulation*[332], Nesta[333] y el Department for Business, Energy and Industrial Strategy.[334]

En relación directa con los interfaces precitados, la biotecnología, neurotecnología o la nanotecnología van a tener un impacto cada vez mayor en nuestras vidas. Se trata de tecnologías que aplican tecnología para intervenir en el cuerpo o cerebro con el objetivo de entender, mitigar, predecir o curar enfermedades.

---

[331] *iHuman: blurring lines between mind and machine*. The Royal Society. Septiembre 2019. Recuperado de: https://royalsociety.org/-/media/policy/projects/ihuman/report-neural-interfaces.pdf. Consultado el 04.03.2021.

[332] *Business Secretary hosts first crossgovernment working group on future regulation*. UK Government. 2018. Recuperado de: https:// www.gov.uk/government/news/business-secretary-hostsfirst-cross-government-working-group-on-future-regulation. Consultado el 02.02.2021.

[333] *2019 Anticipatory regulation*. NESTA. Recuperado de: https://www.nesta. org.uk/feature/innovation-methods/anticipatory-regulation/. Consultado el 08.02.2021.

[334] *2019 Regulation for the Fourth Industrial Revolution*. Department for Business, Energy and Industrial Strategy (BEIS). Recuperado de: https://www.gov.uk/government/publications/regulation-forthe-fourth-industrial-revolution. Consultado el 02.02.2021.

El desarrollo y uso de estas tecnologías plantea cuestiones similares a la inteligencia artificial, no tanto desde el punto de vista de la tecnología, sino del humano mejorado a través de la misma.

Si a las mismas adicionamos la inteligencia artificial su potencial y riesgos asociados aumentan de manera exponencial, por lo que todavía hace más necesario la definición de marcos regulativos sólidos y efectivas en materia ética, de seguridad y de responsabilidad en estas áreas específicas, pero en especial, en materia de inteligencia artificial y su aplicación en el sector médico y con estas finalidades.

Desde hace tiempo son muchas las voces científicas más autorizadas que avisan del peligro de estas tecnologías, incluso para la pérdida de la privacidad de nuestra esfera más interna, la de nuestros pensamientos, como, por ejemplo, Rafael Yuste, Catedrático de Universidad de Columbia en EE.UU. e impulsor de la mayor iniciativa para conocer el cerebro humano bajo el título "BRAIN".

Según este experto "a corto plazo, el peligro más inminente es la pérdida de privacidad mental". Incluso estas tecnologías podrían servir para manipular los pensamientos a través de señales eléctricas[335]. Este experto propone una regulación desde dos dimensiones: Una autorregulación para ingenieros, informáticos y otros especialistas, y un marco legal que contemple los denominados *neuroderechos* dentro de los derechos humanos básicos y que evite los abusos.

Chile será el primer país del mundo en regular en una ley que la identidad mental no es manipulable, en una iniciativa liderada por el precitado experto.

Las diademas inteligentes actuales permiten registrar la actividad cerebral de los usuarios para el control mental de dispositivos, drones o coches o medir el nivel de concentración o estrés de las personas. En China se están utilizando para medir el nivel de estrés en conductores públicos y en escolares para comprobar su nivel de concentración.

[335] YUSTE, R. ET AL. (2017). "Four ethical priorities for neurotechnologies and AI". *Nature.* Vol. 551. Nº 7679. Noviembre 2017.

Las tecnologías actuales permiten acceder a lo que piensa una persona y "descifrar" los *neurodatos*. Incluso se baraja la posibilidad de implantes cerebrales que supongan mejora del ser humano. En relación con todo ello el grupo de investigación que lidera el experto precitado significan su preocupación en relación con los nuevos "neuroderechos": Derecho a la identidad personal, al libre albedrío, a la privacidad mental, al acceso equitativo a las tecnologías de mejora humana y a la protección contra sesgos y discriminación.

Los gigantes tecnológicos ya están invirtiendo y trabajando en conectar nuestro cerebro con las máquinas, como Facebook, Microsoft (Neuralink) o Google. Y superpotencias como China están trabajando en la fusión entre la inteligencia artificial y neurotecnología.

9.11. Computación cuántica

Las tecnologías exigen cada vez mayor capacidad de computación ante el crecimiento exponencial del volumen de datos sobre los que operan, mayor capacidad de almacenamiento y mayor eficiencia en su procesamiento.

La computación cuántica supondrá resolver complejos cálculos científicos imposibles de procesar con las técnicas actuales.

Quizás la manera gráfica de resumir lo que supone la computación cuántica y la diferencia entre computadoras tradicionales y cuánticas la ha propuesto Gartner. Si imaginamos una biblioteca gigante de libros, la computadora clásica leería todos los libros de forma lineal, mientras que la cuántica los leería todos simultáneamente, gracias a su capacidad para realizar millones de cálculos al mismo tiempo.

Durante estos últimos dos años, hemos visto distintos anuncios de algunos de los gigantes tecnológicos a través de medios de comunicación, en los que se anunciaba haber alcanzado la denominada "supremacía cuántica".

A modo de ejemplo, Google aseguró en un artículo científico publicado a finales de 2019[336] que había conseguido que un ordenador cuántico realizara en apenas 3 minutos una operación para calcular números aleatorios que a la supercomputadora más potente del mundo actual le hubiera llevado al menos 10.000 años. Esta publicación fue rebatida inmediatamente por IBM[337] que argumentó que el cálculo propuesto, en teoría, podría ejecutarse en una computadora actual en menos de dos días y medio.

Del mismo modo, IBM también anunció en fechas coetáneas su primer ordenador cuántico para uso comercial denominado "System One", que combina computación cuántica con computación tradicional.

El concepto de "supremacía cuántica" se atribuye a John Preskill, investigador y profesor del Instituto Tecnológico de California, que pretende identificar el momento en el que seamos capaces de construir un procesador cuántico capar de realizar una determinada tarea no pueda ser ejecutada por un ordenador clásico en un tiempo razonable.

No obstante, esto no es una realidad ni lo será en breve, dado que quedan muchos aspectos técnicos que resolver, tanto de hardware (ordenadores cuánticos) como de software (lenguaje y algoritmos cuánticos), para el desarrollo y aplicación de esta tecnología, y debería ser exclusivamente regulado y usado para finalidades concretas en el marco de resolver problemas o necesidades del ser humano, mejorar nuestra vida o nuestro mundo.

No obstante, parece indudable que llegará y previsiblemente de la mano de las grandes tecnológicas como Google, IBM o Intel. La inversión privada también se ha orientado hacia la misma.

La computación cuántica supondría realizar en segundos operaciones que miles de ordenadores tradicionales tardarían años en hacer. Sin embargo, la reflexión política, ética y jurídica debería empezar a hacerse ya, en la medida que esta tecnología promovida y

---

[336] ARUTE, F., ARYA, K., BABBUSH, R. ET AL. (2019). "Quantum supremacy using a programmable superconducting processor". Publicado el 23.10.2019 en *Nature*, Nº 574. Pp. 505–510. https://doi.org/10.1038/s41586-019-1666-5

[337] YAFFE-BELLANY, D. (2019). "Computación cuántica explicada en unos minutos". Publicado en *The New York Times*, el 24.10.2019. Recuperado de: https://www.nytimes.com/es/2019/10/24/espanol/ciencia-y-tecnologia/computacion-cuantica-google.html. Consultado el 24.02.2021.

creada por el sector privado, supondría dejar en sus manos capacidades abrumadoras y un poder que, inadecuadamente usado, permitirían atacar y llegar a hundir países enteros, gobiernos, economías y empresas.

A modo de ejemplo, ¿cuánto costaría cifrar o descifrar una red encriptada mediante el uso de esta tecnología? Las ventajas y bondades del Blockchain se ha construido sobre la criptografía garantizando el cifrado y la seguridad consecuente de toda la cadena. ¿La computación cuántica podría suponer el final del Blockchain que conocemos actualmente? Ante este escenario, distintas entidades están ya investigando desde hace años en todo ello, como la *Agencia de Seguridad Nacional* estadounidense -NSA por sus siglas en inglés-, que pretende construir una computadora cuántica que pueda descifrar la mayoría de los tipos de cifrado.

Como toda tecnología, el problema no es la misma sino el uso que se haga de ella. La computación cuática podría impulsar avances absolutamente disruptivos en inteligencia artificial, la generación de nuevas técnicas de seguridad, perfeccionamiento de la criptografía, crear nuevos materiales o nuevos medicamentes, pero también puede permitir romper el cifrado de redes o atentar contra la seguridad de plataformas de comercio electrónico o contra la propia seguridad nacional de cualquier estado. Va a cambiar el marco actual de ciberseguridad.

Prueba de su potencial y de los riesgos que plantea para los gobiernos de EE.UU. y China, como lo debería ser para el resto del mundo, la computación cuántica es una prioridad nacional. China está invirtiendo en un laboratorio cuántico nacional y ha presentado en los últimos años más del doble de patentes cuánticas que EE.UU. Por su parte, el gobierno estadounidense lanzó su Iniciativa Cuántica Nacional.

En diciembre de 2020, un grupo de investigación de la Universidad de Ciencia y Tecnología de China y la Universidad Tsinghua de Pekín, dirigido por Jian-Wei Pan, publicó un artículo científico en *Science*[338] en el que explicaba cómo había logrado resolver utilizando un sistema cuántico, en poco más de tres minutos, un problema en el

[338] ZHONG, HAN-SEN; WANG, HUI ET AL. (2020). "Quantum computational advantage using photons". Publicado en *Science* el 18 de diciembre de 2020. Vol. 370, Issue 6523. Pp. 1460-1463. DOI: 10.1126/science.abe8770

que los superordenadores clásicos más potentes del planeta habrían invertido 600 millones de años.

A nivel científico se está trabajando tanto en su desarrollo como en la creación de nuevas técnicas de seguridad para contrarrestar su capacidad para descifrar códigos.

La futura irrupción de la computación cuántica permitirá a los sistemas inteligentes aprender y tomar decisiones con una rapidez desconocida hasta la fecha, mejorando en paralelo la calidad, predicción, capacidad de entendimiento y respuesta, etc. Esta asociación entre inteligencia artificial y computación cuántica se denomina *Quantum Machine Learning* -QML por sus siglas en inglés-.

Según un reciente estudio de Goldman Sachs, la computación cuántica podría aplicarse a algunos de los cálculos más complejos de los mercados financieros en un plazo máximo de 5 años.

Asimismo, una reciente investigación llevada a cabo por Cambridge Quantum Computing (CQC), bajo el título *QNLP in Practice: Running Compositional Models of Meaning on a Quantum Computer*[339]y publicada en marzo de 2021, ha desarrollado una implementación experimental, la más grande llevada hasta la fecha según se ha presentado, de las tareas de procesamiento del lenguaje natural (PNL) en una computadora cuántica.

Según esta investigación constituye una prueba de concepto de que el procesamiento cuántico del lenguaje natural está "al alcance de la mano" y que los ordenadores cuánticos pueden aprender a razonar en condiciones de incertidumbre, como lo hacemos los seres los humanos. El Procesamiento del Lenguaje Natural (PLN) es una rama de la inteligencia artificial que transforma el lenguaje natural en un lenguaje formal que los ordenadores pueden procesar.

La computación cuántica, así como la computación biológica y la computación *neuromórfica* son algunas de las disciplinas con mayor potencial en los próximos años,

[339] Recuperado de: https://cambridgequantum.com/cambridge-quantum-announces-largest-ever-natural-language-processing-implementation-on-a-quantum-computer/. Consultado el 21.03.2021.

especialmente en relación con la inteligencia artificial, en lo que están trabajando las grandes tecnológicas, universidades y centros de investigación como el MIT, Stanford o el IMEC.

La computación neuromórfica propone emular el comportamiento del sistema nervioso animal en general, y el del cerebro en particular, con aplicaciones tan relevantes como la aplicación de patrones, el aprendizaje automático, búsqueda de la mejor solución y satisfacción de requisitos predefinidos.

La UE ha hecho público su posicionamiento frente a la computación cuántica en el precitado *Libro blanco sobre la inteligencia artificial*, considerando que puede situarse a la vanguardia de esta tecnología gracias a su fortaleza académica en computación cuántica y la posición de la industria europea en materia de simuladores cuánticos y entornos de programación para la computación cuántica.

10. Consideraciones finales

De inicio, me remito a mis consideraciones particulares y conclusiones específicas que he efectuado en relación con los distintos aspectos abordados durante este capítulo.

El uso de la inteligencia artificial aporta tantísimas ventajas al ser humano que es impensable cuestionarse su valor.

No obstante, como he reflejado en este capítulo, comporta igualmente retos y riesgos que necesariamente deben identificarse, analizarse y gestionarse adecuadamente con el objeto de evitarlos y, en caso de imposibilidad, mitigarlos reduciendo al máximo su probabilidad e impacto, y/o trasladándolos a terceros en caso de materializarse, por ejemplo, mediante seguros, conforme se recoge ya en las iniciativas legislativas europeas de regulación de la inteligencia artificial objeto de análisis en esta investigación.

Citando a Paul Virilio[340] "Cuando inventas el barco, también inventas el naufragio, cuando inventas el avión, también inventas el accidente aéreo; y cuando se inventa la electricidad, se inventa la electrocución... Cada tecnología tiene su propia negatividad, que se inventa al mismo tiempo que el progreso técnico". Y en relación con todo ello, aspectos como el mero error o pérdidas de conectividad asociadas la inteligencia artificial deben ser esperados.

Como destaca Featherstone[341], la teoría del accidente de Virilio sugiere que cuando uno crea tecnología también diseña las fallas y errores que plagan la máquina, considerando que este autor "muestra cómo la tecnología y el accidente están atrapados en una relación dinámica".

La relación e interacción de la inteligencia artificial con determinadas tecnologías y servicios basados en las mismas, potencia sus ventajes y utilidades, pero también sus retos, riesgos y su potencial lesivo a nivel cualitativo y cuantitativo.

La gestión y tratamiento de algunos de estos riesgos, en base a criterios éticos y de la relevancia de los bienes jurídico protegidos en juego, no debería contemplar de antemano la mitigación o su traslado a terceros, sino directamente su erradicación de inicio, prohibiendo el despliegue y aplicación de determinados sistemas dotados de inteligencia artificial en atención a su autonomía, grado de impredecibilidad, sector donde opere, usos previstos y riesgos asociados considerando su probabilidad e impacto. En definitiva, sistemas de riesgo inadmisible, prohibidos por esta razón.

Por ejemplo, sistemas armamentísticos sustentados en inteligencia artificial con plena o cierta autonomía, donde el principal valor jurídico protegido es la vida humana.

En otros riesgos, debe velarse por su adecuada gestión desde el diseño y por defecto, es decir, debe obligarse a fabricantes, diseñadores, desarrolladores, integradores y operadores que efectúen análisis de riesgos previos sobre los elementos, sistemas, aplicaciones, productos y servicios que generen, con información clara y transparente

[340] VIRILIO, P. (1999). *Politics of the Very Worst*. New York. Semiotext (e). 1999. P.89.
[341] FEATHERSTONE, M. (2000). "Velocidad y violencia: sacrificio en Virilio, Derrida y Girard". Publicado en *Anthropoethics: The Journal of Generative Anthropology*. Antropoética VI. Nº 2. Disponible en: http://anthropoetics.ucla.edu/ap0602/virilio/. Consultado el 20.12.2020.

sobre los mismos y su uso, con las advertencias oportunas para que sus posteriores operadores y usuarios adopten las precauciones necesarias en su utilización, especialmente en relación con la calidad de los datos y conocimientos con los que proveer al sistema.

Del mismo modo, los operadores, gestores y usuarios de los dichos elementos, sistemas, aplicaciones, productos y servicios, deberían llevar a cabo una evaluación de impacto de los mismos en su organización, personas y cosas de manera previa a su habilitación y uso, conforme ya es actualmente obligatorio[342] respecto de aquellos sistemas dotados de inteligencia artificial que puedan conllevar un alto riesgo y tener un impacto en los derechos y libertades de las personas relacionados con su privacidad y la protección de datos de carácter personal, en particular, cuando se traten datos de especial protección, como salud, vida sexual, origen racial, datos genéticos o biométricos.

La ética, la seguridad y una regulación específica son esenciales para acometer los mismos adecuadamente, por lo que los nuevos marcos reguladores de la inteligencia artificial deben construirse sobre estos pilares.

La Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*- , de 21 de abril de 2021, se construye desde un enfoque de riesgos y sobre las bases indicadas, esto es, prohibiendo una serie de sistemas que considera, de inicio, de riesgo inadmisible y permitiendo el resto, si bien, en particular, sujetando los sistemas inteligentes de alto riesgo a un detallado conjunto de requisitos y obligaciones de origen ético y jurídico que, sorpresivamente, al menos para mí, no son exigidos al resto de sistemas inteligentes, clasificados o no clasificados, de riesgo medio o bajo o, simplemente no clasificados.

Considero que estos aspectos, entre otros, deberán ser abordados por el Parlamento Europeo y la Comisión para su revisión, en la medida que considero que aquellos

[342] Artículo 35 del Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento General de Protección de Datos o RGPD/GDPR).

principios y normas éticas y jurídicas básicas y esenciales deben ser exigibles a cualquier sistema inteligente por las propias características y capacidades de las que puede estar dotado.

Capítulo III

Marco ético

1. Introducción

La inteligencia artificial, como he referido anteriormente, está contribuyendo de manera indudable a la mejora de la asistencia médica y social, a mejorar la calidad de vida de las personas, a mejorar y personalizar servicios, a mejorar los procesos de fabricación, logística y transporte, a mejorar la recogida y tratamiento de productos agrícolas, a mejorar la sostenibilidad y la responsabilidad social, a reducir costes, a aumentar la producción, a ser más competitivos, entre otros.

Según el McKinsey Global Institute[343], el impacto económico de las distintas aplicaciones de inteligencia artificial para el 2025 se estima entre los 6,5 y los 12 trillones de euros anuales.

La inteligencia artificial es definida en el *Libro Blanco sobre la inteligencia artificial*[344] como "una tecnología estratégica que ofrece numerosas ventajas a los ciudadanos, las empresas y la sociedad en su conjunto *siempre que sea antropocéntrica, ética y sostenible y respete los derechos y valores fundamentales*".

Sin embargo, parece que esta exigencia se ha obviado parcialmente por la propia Comisión Europea en la última Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, en el que únicamente exige el

[343] Informe *Disruptive technologies: Advances that will transform life, business, and the global economy*. Mckinsey Global Institute 2013. Recuperado de: www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Disruptive%20technologies/MGI_Disruptive_technologies_Full_report_May2013.ashx. Consultado el 05.03.2021.

[344] *Libro Blanco sobre la inteligencia artificial*. Comisión Europea. 19.02.2020. COM (2020) 65 Final

cumplimiento de algunos principios y normas éticas esenciales, no todos, a los sistemas considerados conforme al mismo de alto riesgo, pero no respecto del resto.

Conforme he abordado en el capítulo anterior, la inteligencia artificial comporta importantes riesgos potenciales de origen ético en base a sus propia naturaleza, características y capacidades, que recogen los distintos informes y análisis publicados a nivel internacional y, en particular, tal y como será objeto de análisis en este capítulo.

Entre otros informes, me permito destacar el realizado por el FHI de la Universidad de Oxford en 2018[345], en colaboración con otros centros de investigación, en el que analiza los riesgos actuales de la inteligencia artificial y su profunda relación con la seguridad digital, la seguridad física y seguridad política, apostando por el desarrollo de una inteligencia artificial segura y justa.

Algunos de los riesgos asociados a la inteligencia artificial que mayor debate están generando en la actualidad desde un enfoque ético son su seguridad, falibilidad, continuidad, resiliencia, manipulación y vulnerabilidad, la imparcialidad, sesgo y discriminación, rendición de cuentas (*accountability*), responsabilidad, explicabilidad, vulneración de la intimidad, tratamiento ilícito/ilegítimo de datos personales, la destrucción del empleo, transformación de las relaciones sociales, o erosión de la sociedad civil, sistemas democráticos y gobiernos.

Muchos de ellos fueron ya abordados en el capítulo anterior.

Frente a todos estos riesgos se pretende consensuar a nivel internacional un conjunto de principios y requisitos éticos sobre los que se deberían diseñar y funcionar los sistemas dotados de inteligencia artificial, como la rendición de cuentas, la explicabilidad, la imparcialidad, la privacidad y la protección de datos.

La reciente literatura científica ha documentado múltiples de ellos en relación con distintas aplicaciones actuales de los sistemas de inteligencia artificial, sean o no

[345] BRUNDAGE, M.; AVIN, S.; CLARK, J.; TONER, H.;, ECKERSLEY, P.; GARFINKEL, B.; DAFOE, A. ET AL. (2018). *El uso malicioso de la inteligencia artificial: pronóstico, prevención y mitigación*. Oxford University 2018. https://doi.org/10.17863/CAM.22520.

calificables como de alto riesgo, a diferencia de las propuestas regulatorias europeas de 2020 y 2021, en la medida que la primera de 20 de octubre de 2020 si incluyó algunos principios éticos para todo sistema inteligente, mientras que la reciente de 21 de abril de 2021 únicamente contempla como requisitos y obligaciones las relativas a los sistemas de inteligencia artificial de alto riesgo. Entre otros, significar su falta de explicabilidad (Bostrom y Yudkowsky[346]), el registro de conversaciones sin advertencia por asistentes virtuales (Fussell[347] y Shulevitz[348]), el sesgo en su diseño o datos de entrada (Mittelstadt, Allo, Taddeo, Watcher y Floridi[349], Buolamwini y Gebru[350], así como Barlett, Morse, Stanton y Wallace[351]).

En definitiva, la inteligencia artificial plantea importantes riesgos y diversos retos desde una óptica ética que serán abordados en este capítulo, los cuales se hayan en directa relación con los nuevos marcos jurídicos de responsabilidad.

No obstante, de manera previa a su análisis, considero necesario realizar una primera aproximación y reflexión sobre el concepto de "ética" y la complejidad de su integración en y aplicación por sistemas inteligentes.

2. Definición de ética. Integración y aplicación por sistemas inteligentes.

Sin ánimo de entrar en un análisis profundo del concepto y su alcance, podemos definir la ética como el conjunto de normas morales que rigen la conducta de la persona en

[346] BOSTROM, N. Y YUDKOWSKY, E. (2014). *The ethics of artificial intelligence*. The Cambridge Handbook of Artificial Intelligence. Machine Intelligence Research Institute, Cambridge University Press. Junio 2014. Pp. 316-334.
[347] FUSSELL, S. (2019). "Consumer Surveillance Enters Its Bargaining Phase". *The Atlantic*. 2019 www.theatlantic.com/technology/archive/2019/06/alexa-googleincognito-mode-not-real-privacy/590734. Consultado el 2.03.2021.
[348] SHULEVITZ, J. (2018). "Alexa, Should We Trust You?" *The Atlantic*. 2018. Recuperado de: www.theatlantic.com/magazine/archive/2018/11/alexa-how-will-you-change-us/570844.
[349] MITTELSTADT, B.; D., ALLO, P.; TADDEO, M.; WATCHER, S. Y FLORIDI, L. (2016). *The ethics of algorithms: Mapping the debate. Big Data & Society*, 2016. P.7.
[350] BUOLAMWINI, J. Y GEBRU, T. (2018). *Gender shades: Intersectional accuracy disparities in commercial gender classification*. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, 2018. PMLR 81, Pp. 77-91.
[351] BARLETT, R., MORSE, A., STANTON, R. Y WALLACE, N. (2019). *Consumer-Lending Discrimination in the FinTech Era*. National Bureau of Economic Research. 2019.

cualquier ámbito de la vida o, tomando una definición más amplia de la ética como disciplina, como el estudio y discusión de los bienes, las normas y las conductas que contribuyen al desarrollo y florecimiento de la vida humana[352].

Algunos autores la definen como la ciencia filosófica que se encarga de reflexionar sobre los comportamientos morales del ser humano con el objetivo de realizar valoraciones genéricas que puedan ser universalizables[353].

Los bienes incluidos dentro de la misma incluyen la protección de la vida humana, la libertad o la dignidad y, sin lugar a dudas, la inteligencia artificial puede impactar en los mismos en sentido positivo como negativo, especialmente ante su desarrollo, despliegue y aplicación incesante en todos los ámbitos de nuestra vida.

El análisis de los riesgos éticos e impacto de cada una de las infinitas aplicaciones de la inteligencia artificial sería inabordable y máxime en el marco de esta investigación, especialmente ante las distintas tipologías de inteligencia artificial, capacidades y supuesto grado de "autonomía" en la toma de decisiones, si bien, considero necesario realizar una reflexión horizontal y global de los principales beneficios potenciales que supone, de sus principales riesgos y retos que comporta desde un punto de vista ético, de la relevancia de la ética para la construcción de los futuros marcos jurídicos reguladores de la inteligencia artificial y para la responsabilidad por daños causados por o mediante sistemas inteligentes, así como analizar las propuestas regulatorias formuladas hasta la fecha en el ámbito de la UE en el ámbito ético.

De manera consecuente a todo lo anterior y como primera reflexión obligada sobre la inteligencia artificial desde un punto de vista ético, deberíamos cuestionarnos cuál es el objetivo principal y actual de la inteligencia artificial: ¿Ser un medio para satisfacer necesidades, resolver problemas o mejorar la vida del ser humano? ¿O replicar con la mayor fidelidad y profundidad la inteligencia humana en sistemas de información y

[352] STERBA, J. P. (2009). *Ethics: The Big Questions*. Reino Unido: John Wiley & Sons. 2009

[353] CASTRILLÓN, O.D.; RODRÍGUEZ, M. Y LEYTON, J.D. (2008). "Ética e inteligencia artificial ¿Necesidad o urgencia?". International Institute of Informatics and Systemics. Recuperado de: http://www.iiis.org/CDs2008/CD2008CSC/CISCI2008/PapersPdf/C054TM.pdf. Consultado el 2.03.2021.

máquinas? ¿Y con qué rasgos, capacidades y atributos? ¿Cuáles serían los adecuados? ¿Quién los determina?

Pretender replicar la inteligencia humana, exigiría descomponer los procesos cognitivos para expresarlos en lenguaje lógico en forma de algoritmos y, en función de las facultades cognitivas que se busca emular y programar en los sistemas, nos encontraríamos con distintas tipologías de inteligencia artificial.

Como segunda reflexión obligada desde una perspectiva ética, deberíamos cuestionarnos el término inteligencia artificial y su evolución.

La inteligencia y la supuesta autonomía inherente y asociada actualmente al concepto podrían comportar encontrarnos ante entes capaces de tomar decisiones y actuar de forma supuestamente racional, siguiendo instrucciones propias distintas a las predefinidas por el ser humano en su concepción.

El término "autonomía" asociado a la inteligencia artificial, desde una óptica ética y al margen del concepto jurídico analizado en el capítulo I, designaría la capacidad de escoger un curso de acción de forma libre, la cual, tradicionalmente, se ha identificado como un rasgo distintivo y exclusivo de los seres humanos[354]. Este concepto determina el enlace entre ética y responsabilidad.

Esta capacidad de escoger, desde un punto de vista ético, llevaría asociada la capacidad de hacerse cargo de los motivos que guían su conducta y acción, es decir, la responsabilidad sobre los motivos y razones y por los actos y decisiones adoptadas[355], lo que desde un punto de vista ético nos enlazaría con los conceptos de culpabilidad, imputabilidad y responsabilidad jurídica que serán objeto de análisis en el los capítulos V y VI.

Si los sistemas de inteligencia artificial actuales no disponen de esa capacidad de responder de sus propias conductas y decisiones de la misma manera que las personas

[354] *Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems*. European Group on Ethics in Science and New Technologies. 2018. P. 9.
[355] MARÍN, S. (2019). *Ética e inteligencia artificial*. Cuadernos de la Cátedra CaixaBank de Responsabilidad Social Corporativa-IESE Business School. Nº 42. 2019.

dan cuenta de sus propios actos y decisiones, en consecuencia, desde un punto de vista ético, no sería del todo correcto hablar generalizadamente de aplicaciones, sistemas y máquinas dotados de inteligencia artificial, capaces de operar de forma autónoma conforme es concebida esta capacidad del ser humano.

Por todo ello, desde un punto de vista ético, posiblemente sería más correcto hablar de sistemas o aplicaciones capaces de operar sin supervisión humana o automáticamente, no autónomamente, lo que de nuevo refrenda desde un punto de vista ético mi posicionamiento sobre esta cuestión expuesta en el capítulo I.

La autonomía constituiría un rasgo exclusivo de los seres humanos desde un punto de vista estrictamente ético, en la medida que sólo el ser humano es capaz de escoger con libertad y guiar sus acciones, a diferencia de los animales que actúan por instinto. Y en congruencia con ello, no debería confundirse "autonomía" con "automaticidad", conforme significan distintos autores y será analizado más adelante.

Es más, el primer principio y norma ética esencial debería ser el control y la supervisión humana durante todo el ciclo de vida de un sistema inteligente, lo que de inicio impide la atribución de dicha autonomía al sistema, cuanto menos total. Como referí anteriormente al abordar la definición de inteligencia artificial y autonomía, algunas de las incorporadas en las propuestas reguladoras objeto de análisis en esta investigación, incorporan como rasgo inherente la autonomía y, en algún caso, la posibilidad de operar sin tener que circunscribirse a las instrucciones incorporadas en la concepción del sistema inteligente, lo que comporta sistemas más avanzados que los construidos sobre la base de una inteligencia artificial "débil", en los términos definidos en los capítulos anteriores.

Los sistemas inteligentes actuales actúan de forma "automática" en la medida que operan y actúan siempre conforme a su programación y dentro de unas instrucciones y parámetros predefinidos, bajo una capacidad de decisión aparentemente autónoma que realmente responde a un cálculo de probabilidades y de impacto, realizado en base a distintos parámetros previamente predefinidos en su diseño.

En consecuencia, la responsabilidad ética derivada del funcionamiento de los sistemas dotados de esta inteligencia artificial, inicialmente y en base a ello, debería situarse inicial y necesariamente sobre las partes involucradas en su diseño y desarrollo

No obstante, la previsión y futura existencia de sistemas inteligentes más avanzados, con cierto grado de autonomía efectiva e impredecibilidad, supuesta capacidad de interpretación, razonamiento lógico-informático, capacidad de autoaprendizaje y hasta incluso dotados de una posible pseudo-conciencia, plantea la duda sobre si solo los seres humanos serán capaces de actuar de forma autónoma, la cual debe abordarse, no sólo por sus implicaciones éticas sino jurídicas: ¿Quién deberá ser el responsable de los razonamientos y las decisiones de estos sistemas y aplicaciones? ¿Las personas encargadas de su diseño o los propios sistemas?

En este sentido, desde hace algunos años, distintos autores empezaron a distinguir entre la denominada inteligencia artificial "débil" e inteligencia artificial "fuerte" desde la introducción de estos conceptos por el filósofo John Searle, como expuse en el capítulo I, en particular López de Mántaras[356], investigador del CSIC y director del Instituto de Investigación de inteligencia artificial.

La inteligencia artificial débil se circunscribiría al diseño y programación de sistemas capaces de realizar tareas de forma inteligente, es decir, sistemas con inteligencia especializada, carentes de autonomía en sentido estricto y, en consecuencia, sería incapaces de actuar de forma racional por lo que su actividad no sería éticamente imputable (sistemas que juegan al ajedrez o diagnostican enfermedades).

La responsabilidad ética derivada de su funcionamiento y "conducta" debería recaer inicialmente y en su totalidad sobre las personas encargadas de su diseño y funcionamiento, al igual que, inicialmente, la jurídica, conforme se analizará en el capítulo V, es decir, la responsabilidad ética en esta tipología de inteligencia artificial se centraría en la programación y en el diseño de los sistemas.

[356] LÓPEZ DE MÁNTARAS, R. Y MESEGUER, P. (2017). *Inteligencia artificial*. Madrid, CSIC 2017.

Sin embargo, la inteligencia artificial fuerte permitiría, supuestamente, replicar la inteligencia humana en sistemas y máquinas, creando sistemas que podrían actuar de forma autónoma y no sólo automática.

La responsabilidad ética recaería en su totalidad sobre las personas encargadas de su diseño y funcionamiento, así como, inicialmente, la jurídica. No obstante, dado el mayor grado de automatización, la reflexión ética debería ir más allá y no debería centrarse exclusivamente en la programación y en el diseño de los sistemas, sino también en la explicabilidad, la rendición de cuentas, la trazabilidad, sus capacidades o la supervisión humana. La pregunta obligada a formularnos: ¿Es o será posible crear sistemas o máquinas con esta inteligencia? Según López de Mántaras, anteriormente citado, la complejidad del cerebro humano sería prácticamente imposible de replicar en el futuro.

No obstante, se investiga de manera incesante en esa línea. Un reciente estudio científico publicado en la revista *Nature*[357], llevado a cabo por investigadores de la Northwestern University en Estados Unidos y la Universidad de Hong Kong, asegura la creación del primer cerebro electrónico que sería capaz de aprender por sí mismo como cualquier ser humano.

Si en la actualidad no es posible crear sistemas dotados de inteligencia y sentido común, quizás las propuestas actuales sobre la denominada "singularidad tecnológica" no tienen demasiado sentido, todavía.

Actualmente disponemos de sistemas dotados con inteligencia artificial con capacidad para tomar decisiones y escoger entre diversas opciones de actuación, por ejemplo, los denominados vehículos "autónomos" que, en base a lo expuesto, quizás sería más adecuado denominarlos "automáticos" o de "conducción automática".

Según distintos autores, en estos sistemas la capacidad de tomar decisiones es aparente, dado que están programados para realizar un cálculo de probabilidades teniendo en cuenta el contexto y, dentro de éste, una serie de factores y circunstancias que determinarán la

[357] JI, X.; PAULSEN, B.D.; CHIK, G.K.K.ET AL. (2021). "Mimicking associative learning using an ion-trapping non-volatile synaptic organic electrochemical transistor". *Nature Communications*. 12, 2480. 2021. https://doi.org/10.1038/s41467-021-22680-5

acción, pero dado un escenario concreto y un conjunto de datos limitado, el curso de acción considerado por el sistema como óptimo será único, aunque para el cálculo de probabilidades integre factores de relevancia ética.

De nuevo, como he referido anteriormente, la responsabilidad ética en estos supuestos recaería inicialmente en las personas encargadas de su diseño y funcionamiento, si bien, como también he comentado, debiendo considerar aspectos más específicos como la explicabilidad, la trazabilidad, la rendición de cuentas, las capacidades de las que disponga o la supervisión humana.

Sin embargo, la ética no es algo único o universal, sino que puede variar especialmente en función del contexto cultural.

El proyecto denominado *The Moral Machina* desarrollado por el MIT[358], pretende reunir opiniones de distintas regiones del mundo sobre la mejor manera y más ética de actuar en situaciones extremas por parte de los vehículos autónomos con el objetivo de alcanzar un consenso.

Las principales cuestiones éticas que se pueden plantear en relación con todo ello fueron ya evidenciadas por el denominado *Dilema del tranvía*, cuya primera versión presentó la filósofa Philippa Foot[359] en un artículo de 1967, que fue analizado posteriormente con profundidad por otros filósofos y, en particular, por Judith Jarvis Thomson[360], que planteó una variable adicional sobre el mismo, y ha sido posteriormente tomado como referencia para múltiples estudios y variables.

El dilema del tranvía planteado por Foot planteaba una situación en la que un tranvía sin frenos se dirige hacia cinco trabajadores que están en la vía, sin poder avisarles y sin poder parar el tranvía, pero si podríamos accionar una palanca que lo desviaría hacia otra vía, donde hay un solo trabajador.

---

[358] *Moral Machine*. Massachusetts Institute of Technology (MIT). MIT Media Lab, 2016. Recuperado de: https://www.moralmachine.net/. Consultado el 30.10.2020.
[359] FOOT, P. (1967). "The Problem of Abortion and the Doctrine of the Double Effect". Oxford Review, N. 5, 1967. Recuperado de: http://pitt.edu/~mthompso/readings/foot.pdf.
[360] JARVIS THOMSON, J. (1985). "The Trolley Problem". Publicado en *The Yale Law Journal*. Vol. 94, No. 6, May, 1985. The Yale Law Journal, Inc. Pp. 1395-1415.

Las preguntas inmediatas relacionadas con el objeto de esta investigación son algunas como éstas: ¿Debemos accionar la palanca? ¿Quién decide quién vive y quién muere? ¿Quién asume los daños colaterales? Respecto de la primera, la gran mayoría de las personas a las que se plantea esta cuestión responden que accionarían la palanca.

Por su parte, la precitada Jarvis Thomson propuso una variable en la que nos encontraríamos en una pasarela y vemos como el tranvía se dirige hacia esos cinco trabajadores, pero como expertos en este tipo de vehículos, nos apercibimos que sólo habría una fórmula de detenerlo empujando a un "hombre gordo" que tenemos al lado, él morirá, pero los otros cinco salvarán su vida. En este otro contexto, la gran mayoría contesta que no sería admisible empujar a una persona, a pesar de que también estamos hablando de sacrificar una vida para salvar otras cinco.

Se han realizado múltiples estudios en relación con este dilema y sobre la pregunta qué y cómo decidir entre lo correcto e incorrecto.

En el denominado *Test de Sentido Moral* de la Universidad de Harvard[361] contestaron más 200.000 personas. Según recoge David Edmonds en su libro *Would You Kill The Fat Man?*, el 90% de las personas que contestaron a este test accionaría la palanca, pero el 90% se niega a empujar al "hombre gordo", como lo describía Jarvis Thomson.

Estudios posteriores elaborados por neurocientíficos como Joshua Greene, sustituyeron al hombre gordo por un hombre con una mochila al detectar que el calificativo utilizado en el ejemplo, generaba un cierto rechazo preliminar en los encuestados.

Según los estudios sobre estos aspectos, el ser humano tiende a censurar acciones dañinas que suponen la aplicación de fuerza de modo personal. Por ejemplo, vemos peor empujar a una persona que accionar una palanca o trampilla para que caiga a la vía.

[361] Recuperado de: https://www.moralsensetest.com/

Para los denominados utilitaristas, ambos supuestos son equivalentes, en la medida que se estarían salvando cinco vidas a cambio de una. El enfoque opuesto sería el deontológico o de la ética de los deberes.

Cuando la precitada Jarvis Thomson habla de la diferencia entre el escenario de la palanca y del hombre "gordo", hace referencia al imperativo categórico de Kant como requisito moral sin excepciones que, en su formulación relacionada con el dilema planteado, vendría a decir que "Obra de tal modo que uses a la humanidad, tanto en tu persona como en la persona de cualquier otro, siempre como fin y nunca solo como medio". Siguiendo la misma, un "kantiano" no empujaría a nadie, pero tampoco accionaría la palanca.

La misma aborda si los derechos tienen más valor que la utilidad. Según Jarvis Thomson, para muchos, la diferencia entre ambos escenarios está en que "los derechos superan a las utilidades". Es decir, al accionar la palanca, estaríamos redirigiendo una amenaza ya existente, mientras al empujar a alguien, estaríamos infringiendo sus derechos y creando una nueva amenaza. Y tampoco considera que esta diferencia sea tan clara, en la medida que también infringimos los derechos del trabajador solitario cuando accionamos la palanca, dado que no se ha presentado voluntario a sacrificar su vida por otras cinco, la persona que acciona la palanca ha decidido por él.

Jarvis Thomson publicó en 2008 que el cuasi consenso respecto al primer escenario era incorrecto, a pesar de que ella también estaba de acuerdo y lo fundamenta en que muy pocos accionarían la palanca si eso supusiera que el tranvía nos arrollaría a nosotros. No tendríamos pues derecho a desviar el tren.

Otros filósofos como Thomas Cathcart en su libro *The Trolley Problem*, significa que una cosa es lo que muchos haríamos y otra diferente es lo que deberíamos hacer.

A la vista de todo lo expuesto y su contextualización en el ámbito de los sistemas inteligentes, tiene más actualidad que nunca, especialmente en su diseño y en relación con la supuesta "toma de decisiones" o "ejecución contextual de instrucciones o eventos"

por sistemas inteligentes y los coches autónomos. Y en este sentido, debo destacar la variante denominada *El túnel*, planteada por el ingeniero y filósofo Jason Millar[362].

El *Dilema del túnel* plantea que un contexto en el que viajamos por una carretera de un solo sentido en un coche sin conductor y al acercarnos a un túnel muy estrecho, justo cuando estamos a punto de entrar, una niña intenta cruzar, pero tropieza y cae, bloqueando la entrada al túnel. No hay tiempo para frenar y el coche solo tiene dos opciones, arrollar a la niña o girar la dirección y estrellar el coche contra el muro. ¿Qué debería hacer el coche? ¿Debería ser el marco regulador el que lo definiera? ¿Lo debería definir el diseñador o programador?

Un artículo publicado en la revista *Science*[363] bajo el título "El dilema social de los vehículos autónomos*"*, sugiere que la regulación puede no contribuir a su resolución, en la medida que las encuestas mostraron que una regulación que exigiera evitar el atropello y estrellar el vehículo constituiría un autosacrificio que no sería apoyado por la mayoría de las personas y que comportaría que las mismas evitarían comprar coches sin conductor. Todo ello evidencia la relevancia de la ética no sólo para la responsabilidad jurídica, sino incluso para los mercados y competitividad.

Como indica Millar[364], un humano reaccionaría por instinto, pero los coches "autónomos" o automáticos como los que está desarrollando Google, estarán diseñados y programados para responder a este tipo de situaciones, tomando una decisión ¿O más que tomando una decisión, estaría ejecutando contextualmente una instrucción? ¿Quién será el responsable de la misma y de los daños causados? Parece que desde un punto de vista ético todo apunta de nuevo hacia el diseñador o programador del sistema.

Millar propone que la responsabilidad de responder a esta pregunta no debería ser de los ingenieros, sino de los conductores, en consecuencia, del propietario/usuario, dado que el coche podría estar diseñado para mostrar ciertas preferencias, como, por ejemplo, intentar

[362] MILLAR, J. (2014). "An ethical dilemma: When robot cars must kill, who should pick the victim?". Publicado en *Robohub* el 11.06.2014. Recuperado de https://robohub,org/and-ethical-dilemma-when-robot-cars-must-kill-who-should-pick-the-victim/. Consultado el 27.01.2021.
[363] BONNEFON, J.F.; SHARIFF A.; RAHWAN, I. (2016). "El dilema social de los vehículos autónomos", en revista *Science*. Publicado el 24 de junio de 2016. Vol. 352, Nº 6293. Pp. 1573-1576.
[364] Recuperado de: https://robohub.org/an-ethical-dilemma-when-robot-cars-must-kill-who-should-pick-the-victim/. Consultado el 27.01.2021. Op.cit.

salvar el máximo de vidas, lo que significaría tener en cuenta cuánta gente va en el coche y a cuánta podría arrollar, pero podría dejarse margen a los conductores. Por ejemplo, quienes viajen con sus hijos podrían preferir atropellar a otras personas antes que poner en peligro sus vidas.

¿Qué ocurriría si en el coche viaja una sola persona de 80 años? ¿Qué ocurriría si viajan dos con dicha edad? ¿Y si optar por girar a la izquierda supondría un 50% de probabilidad de matar a dos peatones y girar a la derecha supondría un 100% de probabilidad de matar a una? Y si viajo sólo en el coche, o mi mujer, y ante la posibilidad de atropellar a tres personas, ¿el sistema debería optar por sacrificarme a mí y salvar la vida a tres personas? Es posible, que en ese caso el conductor con tiempo para pensar y decidir, pudiera tomar una decisión distinta al sistema.

En definitiva, la tecnología ya está disponible para ser aplicada pero los conflictos éticos que supone la toma de decisiones por parte de un sistema inteligente, especialmente sobre la vida o la muerte, no están resueltos, no es una cuestión pacífica a nivel de expertos y, desde luego, no es algo aceptado por la sociedad.

El proyecto *Moral Machine* precitado pretende consensuar a nivel internacional unos principios éticos para su integración en el sistema inteligente, para que determinen su objetividad en determinadas situaciones.

La pregunta fundamental planteada en el seno de este proyecto es la relativa a qué decisión debería tomar un coche autónomo si se encuentra en una situación comprometida como la de atropellar a unas personas u otras en función de parámetros como número de personas, sexo, edad, raza o, incluso, estado de salud.

Las respuestas de más de dos millones de personas de más de 200 países no han sido ni categóricas ni concluyentes, dado que su cultura, situación geográfica, circunstancias familiares o personales influyen significativamente en su concepto de ética. Ello evidencia la complejidad de consensuar un marco ético general para cualquier sistema específico de inteligencia artificial y para sus distintas aplicaciones.

En resumen, los retos y responsabilidades éticas de la inteligencia artificial son complejos y están relacionados con sus niveles de automatización, supuesta "autonomía", relativa impredecibilidad y el resto de capacidades de las que esté dotada, como la interpretación, el racionamiento lógico-informático, el autoaprendizaje o su capacidad operativa, y de incuestionable relevancia para la responsabilidad por daños.

3. Principios y normas éticas básicas de la IA

La necesidad de definir un marco ético para el desarrollo, despliegue y aplicación de la inteligencia artificial con el objetivo de garantizar una inteligencia artificial segura y fiable ha motivado que tanto instituciones y organizaciones internacionales como gobiernos, empresas, asociaciones profesionales e investigadores se hayan puesto a trabajar intensamente en su definición, especialmente en los últimos tres años, dando como resultado la definición y relativo consenso de distintos marcos.

De hecho, como a continuación expondré, no existe un consenso pleno en la actualidad.

Existe coincidencia en la necesidad de inteligencia artificial ética invocando principios presentes en la mayoría de propuestas como transparencia, equidad, justicia, no maleficencia, responsabilidad, privacidad, libertad y autonomía, dignidad, solidaridad o sostenibilidad, como destaca Jobin[365]. Sin embargo, se plantean distintos enfoques y discrepancias sobre el contenido de estos principios, su relevancia o las exigencias específicas que debería conllevar su aplicación.

Algunos autores como Fjeld[366], agrupan estos principios e identifican un conjunto de áreas comunes en las distintas declaraciones y manifiestos éticos, en particular, privacidad, seguridad, transparencia y explicabilidad, rendición de cuentas,

[365] JOBIN, A., IENCA, M AND VAYENA, E. (2019). "The global landscape of AI ethics guidelines". *Nature Machine Intelligence*. 1, 2019. Pp. 389-399.
[366] FJELD, J. ET ALT. (2020). *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI*. Berkman Klein Center for Internet & Society, 2020

responsabilidad profesional, equidad y no discriminación, promoción de los valores humanos y control humano de la tecnología.

Oliver[367], una absoluta eminencia internacional en inteligencia artificial, hace referencia agrupada a las mismas bajo el acrónimo en inglés *FATEN*, en el que la "F" hace referencia a *fairness* (justicia), la "A" a autonomía, atribución de responsabilidad y aumento de inteligencia, la "T" a *trust* (confianza) y transparencia, -entendiendo que para que exista confianza se tienen que cumplir tres condiciones, esto es, de un lado la competencia como habilidad para realizar con solvencia la tarea comprometida, de otro, la fiabilidad entendida como la competencia sostenida en el tiempo y, por último, la honestidad y transparencia-, la "E" a educación, efecto beneficioso y equidad, y la "N" a la no maleficencia.

La UE tomó la iniciativa de generar un marco común y encomendó a un grupo de expertos su definición. Las recomendaciones y principios emanados de este grupo y de las distintas instituciones y organismos de la UE serán objeto de análisis específico en los siguientes apartados.

Las iniciativas y directrices europeas en esta materia se produjeron en un contexto de debate internacional sobre la ética en materia de inteligencia artificial, especialmente prolífico desde la presidencia japonesa del G7[368].

En este sentido, desde un inicio la UE se mostró especialmente colaborativa con otros países con el objetivo de alcanzar la deseable convergencia ante los distintos proyectos de directrices éticas de países como Japón, Canadá o Singapur, mediante el *Instrumento de Colaboración para la cooperación con terceros países*[369].

---

[367] OLIVER, N. (2019). "Governance in the Era of Data-driven Decisionmaking Algorithms". *Women Shaping Global Economic Governance*. CEPR. Julio 2019.

[368] Foro económico formado por Canadá, Estados Unidos, Francia, Alemania, Italia, Japón y Reino Unido cuyo principal objetivo es aunar posturas para coordinarse en temas como la economía, el empleo, la seguridad o la política comercial y diseñar respuestas políticas a los retos mundiales. Recuperado de: https://ec.europa.eu/info/food-farming-fisheries/farming/international-cooperation/international-organisations/g7_es. Consultado el 18.01.2021

[369] Reglamento (UE) n.º 234/2014 del Parlamento Europeo y del Consejo, de 11 de marzo de 2014, por el que se establece un Instrumento de Colaboración para la cooperación con terceros países (DO L 77 de 15.3.2014. P. 77).

El proyecto lanzado en este contexto bajo el título "*Una alianza internacional para un enfoque centrado en el ser humano para la inteligencia artificial*", se presentó como el marco para la adopción de iniciativas conjuntas con socios de ideas afines, con el fin de promover unas directrices éticas y adoptar principios comunes y conclusiones operativas. Esta cooperación se consideró que posibilitaría realizar un seguimiento del despliegue de la inteligencia artificial a nivel mundial. El proyecto previó desde su definición la organización de actividades de diplomacia durante actos internacionales, entre otros, del G7, G20 y la Organización para la Cooperación y el Desarrollo Económicos.

La Comisión Europea ha venido desarrollando un destacado y activo papel en las conversaciones e iniciativas internacionales, con contribuciones en el G7 y el G20, y contribuyendo a actividades de normalización en organizaciones para el desarrollo de normas internacionales basadas en la inteligencia artificial centrada en el ser humano, promoviendo el intercambio de visiones, trabajando conjuntamente con organizaciones internacionales de referencia en esta materia y promoviendo y participando en reuniones con países no pertenecientes a la UE para alcanzar un consenso sobre este principio esencial.

A nivel internacional, existen en la actualidad distintos marcos de referencia en materia de ética de la inteligencia artificial, como las *Recomendaciones del Foro Económico Mundial para un buen uso de la IA por parte de empresas y gobiernos*, los informes publicados por la UNESCO y la UNICRI Centre for AI and Robotics de las Naciones Unidas, los estándares formulados por la IEEE -*Institute of Electrical and Electronics Engineers*-, los *Principios de Asilomar* sobre IA propuestos por el Future of Life Institute o la declaración para el desarrollo responsable de la IA elaborada por el *Forum on the Socially Responsible Development of Artificial Intelligence* celebrado en la Universidad de Montreal en 2017.

Como haré referencia con más detalle en el capítulo IV de esta investigación, los 37 países miembros de la Organización para la Cooperación y el Desarrollo Tecnológico (OCDE)[370] han suscrito los principios sobre inteligencia artificial alineados con valores

[370] Australia, Austria, Bélgica, Canadá, Chile, Colombia, República Checa, Dinamarca, Estonia, Finlandia, Francia, Alemania, Grecia, Hungría, Islandia, Irlanda, Israel, Italia, Japón, Corea, Letonia, Luxemburgo, México, Holanda, Nueva Zelanda, Noruega, Polonia, Portugal, República de Eslovaquia, Eslovenia,

humanos propuestos por EE.UU., supeditando el uso de esta tecnología al respeto del Estado de derecho, los derechos humanos, la democracia, la transparencia y la rendición de cuentas.

De todas estas iniciativas, me permito destacar la *Conferencia de Asilomar* organizada por el Future of Life Institute, en la que se establecieron 23 principios o recomendaciones conocidos como los *Principios de Asilomar*, cuyo objetivo es minimizar los riesgos que puede generar la inteligencia artificial y lograr aprovechar su increíble potencial para beneficiar a la humanidad y no para dañarla.

De los mismos, significo los de naturaleza ética, en especial, seguridad, transparencia, responsabilidad, alienación de valores con los del ser humano, compatibilidad con los mismos, la dignidad, la diversidad cultural y la tolerancia racial, la privacidad, la libertad, el beneficio y la prosperidad común, el control humano, no subversión de los procesos cívicos y sociales, uso no armamentístico y acabando por el "bien común", de modo que cualquier súper inteligencia debe desarrollarse únicamente al servicio del bienestar del mayor número posible de personas y para el beneficio de toda la humanidad y no para el bien individual de un estado u organización. Es decir, se basan igualmente en el necesario alineamiento de la tecnología y la inteligencia artificial con los objetivos de la humanidad.

Del mismo modo, la Universidad de Harvard llevó a cabo una investigación en 2019 sobre los principios éticos de 32 organizaciones que hasta ese momento había publicado principios éticos de inteligencia artificial[371], incluyendo organizaciones civiles, gobiernos, organizaciones intergubernamentales y compañías privadas.

El estudio agrupó los principios en nueve dimensiones: Derechos humanos, valores humanos, responsabilidad profesional, control humano de la tecnología, justicia y no

España, Suecia, Suiza, Turquía, Reino Unido y Estados Unidos. Recuperado de: https://www.oecd.org/acerca/miembros-y-socios/. Consultado el 02.12.2020.

[371] FJELD, J.; HILLIGOSS, H; ACHTEN, N.; DANIEL, M.L. ET AL. (2019). *Principled artificial intelligence: a Map of Ethical and Rights-Based Approaches*. Berkman Klein Center, 2019. Recuperado de: https://ai-hr.cyber.harvard.edu/images/primp-viz.pdf. Consultado el 14.02.2021.

discriminación, transparencia y explicabilidad, seguridad, rendición de cuentas y privacidad.

Por su parte, la Escuela Politécnica Federal de Zúrich -ETH por sus siglas en alemán-, analizó [372]los principios de la inteligencia artificial de 84 organizaciones de todo el mundo y los clasificó en las siguientes dimensiones, de mayor a menor uso: Transparencia, justicia, no maleficencia, responsabilidad, privacidad, beneficencia, libertad y autonomía, confianza, sostenibilidad, dignidad y solidaridad.

En España, la Universidad de Huelva llevó a cabo un estudio sobre el uso de la inteligencia artificial en las empresas españolas que cotizan en el IBEX 35[373] basándose en sus informes anuales de gestión.

Y, por último, Algorithm Watch, organización sin ánimo de lucro para la observación y evaluación de los procesos algorítmicos de toma de decisiones que tienen un impacto social, mantiene un inventario global[374] con 83 organizaciones que han publicado principios éticos de inteligencia artificial entre 2018 y abril de 2020.

No obstante, los principios reflejados en todos estos informes y declaraciones no son nuevos. De hecho, de nuevo, debemos remontarnos algunas décadas atrás para encontrarnos con muchos de esos principios proclamados en un contexto notablemente diferente del actual, pero en el que se abordaron aspectos de indudable actualidad en el presente.

Y, es más, si nos remontamos a mediados del siglo pasado y tomando como referencia, una vez más, la ciencia ficción, debo significar las leyes de la robótica de Isaac Asimov[375], orientadas a regir la relación entre los hombres y las máquinas con el fin de garantizar el

[372] JOBIN, A.; IENCA, M; Y VAYENA, E. (2019). *Artificial Intelligence: the global landscape of ethics guidelines.* 2019. https://arxiv.org/ftp/arxiv/papers/1906/1906.11668.pdf. Consultado el 14.02.2021.
[373] *Análisis comparativo de la información sobre IA en los informes anuales de las empresas Ibex 35 y OMX Helsinki 25*. AECA 27.11.2019.
[374] *AI Ethics Guidelines Global Inventory*. Algorithm Watch, 2019. Recuperado de: https://algorithmwatch.org/en/project/ai-ethics-guidelines-global-inventory/. Consultado el 14.02.2021.
[375] ASIMOV, I. (1942). *Runaround*. Astounding Science Fiction, 1942. N.29 (1). Pp. 94-103.

sometimiento de la tecnología al ser humano y evitar que el desarrollo tecnológico pueda perjudicar a la humanidad, y que me permito transcribir a continuación:

- Un robot no puede hacer daño a un ser humano o, por inacción, permitir que un ser humano sufra daño.

- Un robot debe obedecer las órdenes dadas por los seres humanos, excepto si estas órdenes entrasen en conflicto con la primera ley.

- Un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con la primera o la segunda ley.

Se trata de leyes creadas en un contexto de ciencia ficción entonces, pero que hoy tienen más actualidad que nunca.

Los retos éticos que plantean los sistemas dotados de inteligencia artificial y la pretensión de garantizar el uso y aplicación de manera segura y fiable de la misma, requieren la definición y aplicación de principios y normas éticas esenciales desde su diseño, centradas en el ser humano, que a su vez deben ser la base de los futuros marcos reguladores de la misma.

Si analizamos los principios propuestos por las instituciones y organizaciones precitadas anteriormente como por parte de algunos de los expertos precitados, podemos extraer una serie de principios éticos generales comunes que considero, aglutinan y concretan la capa ética que debe regir el diseño, desarrollo, despliegue y uso de la inteligencia artificial y que me permito compilar y sintetizar:

- Respeto de la libertad, dignidad y autonomía humana, de la vida y de los derechos humanos (fundamentales o no)

- Transparencia y explicabilidad, de fácil comprensión y acceso.

- Responsabilidad y rendición de cuentas, permitiendo la asignación de responsabilidades ante los posibles daños y perjuicios que los sistemas de inteligencia artificial puedan causar.

- Confianza, robustez, fiabilidad y seguridad. Los algoritmos deben operar de manera precisa y segura. Pero ¿quién debe resolver los errores de diseño o programación, defectos funcionales o incoherencias durante todas las fases del ciclo de vida del sistema? ¿Quién debe supervisar de forma continua los mismos, realizar mantenimientos preventivos, detectivos, correctivos y evolutivos? Los sistemas deben diseñarse desde la seguridad, la resiliencia y la continuidad, es decir, mediante la denominada "*Security by design*".

- Solidaridad, igualdad, justicia y no discriminación.

- Beneficencia

- Sostenibilidad

- Supervisión y control humano

Y a los mismos, me permito adicionar otros imprescindibles como el principio de precaución para anticiparnos sobre los impactos de los sistemas de inteligencia artificial, la reversibilidad para posibilitar su control y asegurar un comportamiento fiable y seguro (que podrían entenderse incluida en la precitada supervisión y control humano), la accesibilidad, así como los de protección de la intimidad y de los datos personales, esto es, la "*Privacy by design*", y el cumplimiento regulativo "*Compliance by design*".

Los principios, requisitos y normas éticas esenciales, éstas u otros, deben contemplarse desde el propio diseño, es decir, bajo un método basado en la denominada "*Ethics by design*", pero también deben ser considerados en la fase de desarrollo, explotación, despliegue, integración, aplicación y uso de la inteligencia artificial, en definitiva, en todo su ciclo de vida.

La cuestión es si estos principios y normas éticas deberían exigirse a cualquier sistema de inteligencia artificial, cualquiera que sea el riesgo que pueda presentar para la vida, la salud, la seguridad, cualquier otro derecho fundamental y otros derechos o bienes en juego, o si deberían únicamente exigirse a sistemas calificados como del alto riesgo para todos o para concretos bienes y derechos.

En mi opinión, de todas estas principios y normas éticas, hay algunos que deben ser considerados esenciales e imperativos a nivel ético, de modo que no debería concebirse un sistema de inteligencia artificial, cualquiera que sea su nivel de riesgo inicial, sin que los integre en su diseño y en su funcionamiento durante todo su ciclo de vida, por ejemplo, la seguridad o la supervisión y control humano.

Esta es la línea seguida por la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas[376], que incorpora una Propuesta de Reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas.

Sin embargo, como he anticipado en los apartados precedentes, la Propuesta de Reglamento de inteligencia artificial del Parlamento Europeo y del Consejo, de 21 de abril, únicamente exige algunos de esos requisitos y normas éticas esenciales en base a su consenso mayoritario a nivel internacional, a los sistemas clasificados conforme al mismo como de alto riesgo, no al resto de sistemas, sean de nivel medio o bajo o simplemente no clasificados, a excepción de la obligación de transparencia para sistemas concretos.

De los estudios realizados se evidencia que son pocas las compañías que tienen definidos corporativamente todos estos principios.

Llevar a la práctica los mismos requiere la aplicación de un conjunto de métodos y herramientas desde un enfoque global, considerando tanto las relativas al proceso de diseño como a su posterior puesta en funcionamiento y uso, que involucra principalmente a ingenieros, programadores y proveedores como otras complementarias, como la creación de marcos regulativos especiales, autorregulación, certificaciones, evaluación por laboratorios u organismos independientes, investigación, concienciación y formación.

[376] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))

Los parámetros éticos deberían ser considerados desde el diseño de las tecnologías[377] como he comentado y durante todo su ciclo de vida, pero no son suficientes para garantizar que los sistemas de inteligencia artificial sean seguros y sean usados de manera responsable.

Como he referido en los capítulos anteriores, la inteligencia artificial está siendo liderada a nivel mundial por las grandes tecnológicas y algunos gobiernos que la han identificado como un medio para alcanzar diversos objetivos de distinta naturaleza, y los grandes proyectos no nacen y están liderados bajo un enfoque *down-top* sino al *top-down*.

La tecnología se desarrolla por personas y organizaciones que definen o representan unos objetivos, pero también unas normas y unos valores propios de la cultura a la que pertenecen[378]. De este modo, algunos de los usos por parte de buscadores o redes sociales pueden responder a objetivos económicos y comerciales que pueden estar muy alejados de la ética y del denominado bien común.

La tecnología forma parte de nuestra vida e influye en nuestra manera de vivir, de relacionarnos y de interactuar, y sus riesgos ponen en juego valores y bienes como la propia vida humana o los derechos fundamentales de las personas, por lo que no puede dejarse la responsabilidad sobre la misma exclusivamente en manos de intereses privados empresariales y, en especial, de sus diseñadores.

En consecuencia, en mi opinión, deben construirse marcos éticos sólidos que sean reconocidos y aceptados por la industria generadora de sistemas de inteligencia artificial y formar parte de su propia autorregulación. Pero además deben ser acompañados, de un lado, de mecanismos de participación que permitan debatir e incorporar principios y normas de funcionamiento a las distintas partes involucradas como investigadores, empresas, ciudadanos, autoridades, reguladores o gobiernos y, de otro, de marcos regulativos que establezcan su carácter obligatorio y vinculante, para evitar usos contrarios a las normas éticas más esenciales, en ocasiones guiados por el liderazgo

[377] BUCHHOLZ, R. A. Y ROSENTHAL, S. B. (2002). "Technology and Business: Rethinking the Moral Dilemma*"*. *Journal of Business Ethics* 2002. N. 41(1-2) Pp. 45-50.
[378] BIJKER, W., HUGHES, T. Y PINCH, T. (1987). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. Cambridge: MIT Press. 1987. Pp. 40-42.

tecnológico, la competitividad, el avance o la innovación sin objetivos, el ánimo de lucro u otros, totalmente alejados de las necesidades del ser humano como base para su construcción.

En este sentido, la UE creó la *European AI Alliance*[379], una comunidad que se ha convertido en un punto de referencia en los debates impulsados por las distintas partes interesadas sobre la política de inteligencia artificial, que está contribuyendo al debate europeo sobre la inteligencia artificial y a la formulación de políticas de la Comisión Europea en esta materia.

Al cierre de esta investigación hay más de 117 iniciativas de ética de la inteligencia artificial a nivel mundial y parece que todo ello puede desembocar en el lanzamiento en los próximos meses de la plataforma Globalpolicy.AI en base a la colaboración de distintas organizaciones intergubernamentales, entre otras, la Comisión Europea, la OCDE, las Naciones Unidas y el Banco Mundial, al objeto de definir los principios para las futuras aplicaciones de la inteligencia artificial y fomentar el desarrollo y el uso responsable de una inteligencia artificial ética y fiable para la consecución de los objetivos de desarrollo sostenible y su alineación con el respeto de los derechos humanos y los valores democráticos.

El pasado mes de mayo de 2021 se publicó el Informe final sobre el proyecto de Recomendación sobre la Ética de la inteligencia artificial, adoptado en la reunión intergubernamental de expertos de la UNESCO.

De los documentos publicados más recientes con este propósito, destacar el publicado por el Consejo de Europa y The Alan Turing Institute bajo el título *Artificial Intelligence, Human Rights, Demacracy and the Rule of Law: A primer*[380], en el que se destaca el papel de la regulación en relación con los derechos, principios y valores precitados y cuáles serían los aspectos clave a abordar de los marcos legales.

[379] Recuperado de: https://ec.europa.eu/digital-single-market/en/european-ai-alliance

[380] LESLIE, D., BURR, C., AITKEN, M., COWLS, J., KATELL, M., AND BRIGGS, M. (2021). *Artificial intelligence, human rights, democracy, and the rule of law: a primer*. The Council of Europe

No obstante, es necesario encontrar el deseable y necesario equilibrio entre la regulación, la normalización, la seguridad, la innovación, la competitividad, el consenso internacional y el compromiso de la industria, dado que, de otro modo, se podría provocar una disyuntiva ética en caso de encontrarnos ante sistemas de inteligencia artificial de última generación y con beneficios incuestionables para el ser humano, que incluso podrían estar salvando vida diariamente, pero con origen en países con un marco normativo menos proteccionista o restrictivo que no podrían ser desplegados y aplicados en otros países debido a sus marcos más proteccionistas o restrictivos. ¿Qué priorizar en este contexto desde un punto de vista ético?

De un lado, podríamos estar privando a los ciudadanos de tecnologías que, en función de su aplicación, podrían estar salvando vidas y, de otro, podríamos estar perjudicando la innovación y la competitividad empresarial frente a marcos regulativos más permisivos.

4. Marco ético en Europa

La UE pretende liderar el consenso internacional en materia ética de la inteligencia artificial.

4.1. Aspectos generales

La UE ha venido desarrollando un intenso trabajo sobre la ética de la inteligencia artificial, especialmente significativo durante los últimos años y que tiene su reflejo, principalmente y entre otros, en los siguientes documentos:

- Resolución de 16 de febrero de 2017, en la que incorporó sus recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica[381].

[381] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))

- Declaración sobre inteligencia artificial, robótica y sistemas autónomos[382] de 9 de marzo 2018, del Grupo Europeo sobre Ética de la Ciencia y las Nuevas Tecnologías.

- Proyecto de directrices éticas para una inteligencia artificial Confiable de 18 de diciembre de 2018, del Grupo de expertos de alto nivel en inteligencia artificial de la Comisión Europea, revisado en marzo de 2019.

- Informe, de 8 de abril de 2019, del Grupo de expertos de alto nivel sobre inteligencia artificial creado por la Comisión, titulado "*Directrices éticas para una IA fiable*"[383].

- Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social europeo y al Comité de las Regiones, "Generar confianza en la inteligencia artificial centrada en el ser humano"[384], de 8 de abril de 2019.

- Recomendación del Consejo de la OCDE sobre la inteligencia artificial, aprobada el 22 de mayo de 2019.

- *Estudio de evaluación del valor añadido europeo*[385] realizado por el Servicio de Estudios Parlamentarios (EPRS) titulado "*European framework on ethical aspects of artificial intelligence, robotics and related technologies: European added value assessment*" (Marco europeo de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas: valor añadido europeo).

- Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, de 19 de febrero de 2020,

[382] Recuperado de: http://www.bioeticayderecho.ub.edu/archivos/pdf/EGE_inteligencia-artificial.pdf. Consultado el 08.02.2021
[383] Recuperado de: https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1. Consultado el 08.02.2021
[384] Recuperado de: https://ec.europa.eu/transparency/regdoc/rep/1/2019/ES/COM-2019-168-F1-ES-MAIN-PART-1.PDF. Consultado el 08.02.2021.
[385] https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2 020)654179

titulada *Libro blanco sobre la inteligencia artificial: Un enfoque europeo orientado a la excelencia y la confianza*[386].

- Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas[387].

Asimismo, el marco ético de nuestro acervo europeo se complementa con los marcos definidos por otras instituciones relacionadas, especialmente en las comunicaciones y los estudios elaborados a petición del *Grupo de Expertos sobre el Futuro de la Ciencia y la Tecnología (STOA)*, dirigido por la Unidad de Prospectiva Científica del Servicio de Estudios del Parlamento Europeo, titulados *What if algorithms could abide by ethical principles?* -¿Y si los algoritmos obedeciesen a principios éticos?-, *Artificial Intelligence ante portas: Legal & ethical reflections* -inteligencia artificial ante portas: reflexiones legales y éticas-, *A governance framework for algorithmic accountability and transparency* -Un marco de gobernanza para la rendición de cuentas y la transparencia de los algoritmos-, *Should we fear artificial intelligence?* -¿Debemos temer a la inteligencia artificial?- y *The ethics of artificial intelligence: Issues and initiatives* -La ética de la inteligencia artificial: problemas e iniciativas-.

En Reino Unido, el 13 de mayo de 2021 se ha publicado su nueva guía sobre el marco de ética, transparencia y rendición de cuentas para la toma de decisiones automatizadas en relación sector público, bajo el título *Ethics, Transparency and Accountability Framework for Automated Decision-Making*.

Y en las mismas fechas, la actividad de múltiples organizaciones supranacionales ha sido ingente.

[386] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, de 19 de febrero de 2020, titulada "Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza".(COM(2020)0065)
[387] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))

La OMS ha publicado una nueva guía sobre ética y gobernanza de la inteligencia artificial bajo el título *Ethics and Governance of Artificial Intelligence for Health*, que sitúa la ética y los derechos en el centro del diseño, implementación y uso de la misma.

La Autoridad Europea de Seguros (EIOPA) acaba de publicar un marco de gobernanza y principios bajo el título: *Artificial Intelligence Governance Principles: Towards Ethical and Trustworthy Artificial Intelligence in the European Insurance Sector*, con el objetivo de poner en el mercado asegurador soluciones de inteligencia artificial de forma confiable y segura.

A nivel internacional, la U.S. Government Accountability Office ha publicado en la fecha conclusión de esta investigación su informe *Artificial Intelligence: An Accountability new Framework for Federal Agencies and Other Entities*[388], en la que identifica algunas prácticas clave para la rendición de cuentas o accountability. Previamente había presentado su marco de gobernanza algorítmica en el sector público en un informe bajo el título *Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies*.

4.2. Resolución del Parlamento Europeo de 16 de febrero de 2017, sobre normas de Derecho civil sobre robótica[389].

El Parlamento Europeo incorporó en esta Resolución sus inquietudes, reflexiones, retos y posibles soluciones en materia ética y de responsabilidad civil, recogiendo algunas de las bases sobre las que se han construido las sendas Resoluciones y correlativas propuestas de Reglamento del Parlamento Europeo, de 20 de octubre de 2020.

Esta Resolución será objeto de análisis más detallado en el capítulo VII de esta investigación, si bien, por lo que se refiere a los aspectos contemplados en materia ética

[388] *Artificial Intelligence: An Accountability new Framework for Federal Agencies and Other Entities*. U.S. Government Accountability Office. 30.06.2021. Recuperado de: https://www.gao.gov/products/gao-21-519sp. Consultado el 30.06.2021.
[389] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))

y su indudable aportación, considero necesario destacar aquí sus aspectos clave en materia ética.

El Parlamento Europeo propuso la creación de un marco ético focalizado en robots y sistemas de inteligencia artificial avanzada que sirviera de orientación a los ingenieros para el diseño y producción de los mismos, con la finalidad de afrontar sus riesgos.

El Parlamento Europeo recogió en la misma los principales riesgos contra la seguridad, la salud, la libertad, la intimidad, la integridad física o psíquica, la dignidad, la autonomía, la autodeterminación del individuo, la no discriminación y la protección de los datos personales, considerando que el marco jurídico actual en el ámbito de la UE debía actualizarse y completarse mediante directrices éticas que, de un lado, contemplaran la complejidad de la robótica y sus numerosas implicaciones sociales, médicas y bioéticas y, de otro, que orientaran el diseño, desarrollo, producción, uso y modificación de los robots dotados de inteligencia artificial fuerte.

El Parlamento Europeo consideró que estas orientaciones éticas debían basarse en los principios de beneficencia, autonomía y justicia, así como en los principios consagrados en la *Carta de los Derechos Fundamentales de la Unión Europea*, como la dignidad humana, la igualdad, la justicia y la equidad, la no discriminación, el consentimiento informado, la vida privada y familiar y la protección de datos, así como en otros principios y valores inherentes al Derecho de la UE, como la no estigmatización, la transparencia, la autonomía, la responsabilidad individual, y la responsabilidad social, sin dejar a un lado las prácticas y códigos éticos ya existentes.

Asimismo, la Resolución destacó la necesidad transparencia, de modo que siempre fuera posible justificar cualquier decisión que se haya adoptado con ayuda de la inteligencia artificial y que pudiera tener un impacto significativo sobre la vida de una o varias personas.

La Resolución incorporó una *Carta sobre Robótica* que se acompañó con un código de conducta ética para los ingenieros en robótica, un código deontológico para los comités de ética de la investigación y de sendos modelos de licencia para diseñadores y para usuarios.

Ante la relevancia e interés de estos códigos y normas propuestos para garantizar la precitada "*Ethics by design*" en el diseño y desarrollo de sistemas inteligentes, me permito abordar su contenido.

4.2.1. Carta sobre Robótica

Una de las principales aportaciones de la Resolución precitada es la *Carta sobre Robótica*[390] que incorpora como anexo a la misma y que recoge los principios que deberían ser considerados para la elaboración de futuras propuestas legislativas.

Se trata de un código de conducta que recoge los principios éticos fundamentales que deberían cumplirse desde el diseño de un sistema inteligente, y que, como el mismo refleja, pretende ser una referencia y complemento para abordar los retos jurídicos que plantea la inteligencia artificial.

Además, no sólo pretende establecer el marco de normas éticas a cumplir por investigadores, profesionales, usuarios y diseñadores, sino también introducir un procedimiento para la resolución de los dilemas éticos y permitir que estos sistemas puedan funcionar de una manera éticamente responsable.

4.2.2. Código de Conducta Ética para los ingenieros en robótica.

Se propone como un código de naturaleza voluntaria, si bien, establece los principios que deberían respetar los ingenieros en robótica, en particular, los siguientes:

- Beneficencia: Los robots deben actuar en beneficio del hombre.

[390] https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_ES.html#title2

- Principio de no perjuicio o maleficencia: Los robots no deberían dañar o perjudicar a las personas.

- Autonomía: La capacidad de tomar una decisión con conocimiento de causa e independiente sobre los términos de interacción con los robots.

- Justicia: La distribución justa de los beneficios asociados a la robótica y la asequibilidad de los robots utilizados en el ámbito de la asistencia sanitaria a domicilio y de los cuidados sanitarios en particular.

Asimismo, el Código de Conducta adiciona a los mismos un conjunto de normas a respetar durante el desarrollo de actividades de investigación:

- Derechos fundamentales: Las actividades de investigación en materia de robótica deben respetar los derechos fundamentales. Las actividades de concepción, ejecución, difusión y explotación han de estar al servicio del bienestar y la autodeterminación de las personas y de la sociedad en general. La dignidad y la autonomía humanas -tanto físicas como psicológicas- siempre tienen que respetarse.

- Precaución: Las actividades de investigación en el ámbito de la robótica deben llevarse a cabo de conformidad con el principio de precaución, anticipándose a los posibles impactos de sus resultados sobre la seguridad y adoptando las precauciones debidas, en función del nivel de protección, al tiempo que se fomenta el progreso en beneficio de la sociedad y del medio ambiente.

- Participación: Garantizar la transparencia y el respeto al derecho legítimo de acceso a la información de todas las partes interesadas. La integración permite la participación en los procesos de toma de decisiones de todas las partes interesadas o afectadas por las actividades de investigación en el ámbito de la robótica.

- Rendición de cuentas: Los ingenieros en robótica deben rendir cuentas de las consecuencias sociales y medioambientales y el impacto sobre la salud humana que la robótica puede conllevar.

- Seguridad: Los diseñadores de robots han de tener en cuenta y respetar la integridad física, la seguridad, la salud y los derechos de las personas. Un ingeniero en robótica debe preservar el bienestar sin dejar de respetar los derechos humanos, y divulgar con prontitud los factores susceptibles de poner en peligro a la población o al medio ambiente.

- Reversibilidad: Condición necesaria de la posibilidad de control y concepto fundamental en la programación de robots para que se comporten de manera segura y fiable. Un modelo de reversibilidad indica al robot qué acciones son reversibles y, en su caso, el modo de invertirlas. La posibilidad de deshacer la última acción o secuencia de acciones, permite al usuario anular las acciones no deseadas.

- Privacidad: El derecho a la intimidad debe siempre respetarse. Un ingeniero en robótica debe garantizar que la información privada se conservará en total seguridad y solo se utilizará de forma adecuada. Por otra parte, el ingeniero en robótica ha de garantizar que los individuos no son personalmente identificables, salvo en circunstancias excepcionales, y únicamente en caso de consentimiento claro, consciente e inequívoco. El consentimiento consciente de la persona tiene que solicitarse y recabarse con anterioridad a cualquier interacción hombre-máquina. A tal efecto, los diseñadores en robótica tienen la responsabilidad de desarrollar y aplicar procedimientos para garantizar el consentimiento válido, la confidencialidad, el anonimato, el trato justo y el respecto de la legalidad. Los diseñadores llevarán a cabo todas las solicitudes de destrucción de los datos relacionados y de eliminación de las bases de datos. Se trata de una manifestación de la privacidad en el diseño y por defecto -*Privacy by design* y *by default*-.

- Maximización de beneficios y reducción al mínimo los daños: Los investigadores deben intentar maximizar los beneficios de su actividad en todas las fases, desde su concepción hasta su difusión. Se debe evitar cualquier daño a los participantes o a los seres humanos que participen en los experimentos, ensayos o estudios en el ámbito de la investigación. En caso de aparición de riesgos inevitables que formen parte de un elemento integrante de la investigación, sería necesario llevar a cabo una evaluación sólida de los riesgos, desarrollar protocolos de gestión y adecuarse a los mismos. Normalmente, los riesgos a un daño no deberían ser superior a los

existentes en la vida cotidiana, es decir, las personas no han de estar expuestas a riesgos mayores o adicionales a aquellos a los que están expuestos en su vida cotidiana. La explotación de un sistema de robótica debería basarse siempre en una profunda evaluación de los riesgos, y reposar en los principios de proporcionalidad y de precaución.

4.2.3. Código Deontológico para los comités de ética de la investigación

El código propuesto recoge a los principios éticos a los que deben estar sujetos estos comités, en particular, independencia, competencia, transparencia y rendición de cuentas.

4.2.4. Modelo de licencia para diseñadores

Los términos y condiciones propuestos para ese modelo de licencia, consolidan la denominada "*Ethics by design*", consistente en un conjunto de principios y normas que los diseñadores de un sistema deberán considerar y que me permito transcribir[391] a continuación:

- Los diseñadores deberán tener en cuenta los valores europeos de dignidad, autonomía y autodeterminación, libertad y justicia, antes, durante y después del proceso de concepción, desarrollo y de aplicación de esas tecnologías, incluida la necesidad de no perjudicar, herir, engañar o explorar a los usuarios (vulnerables).

- Los diseñadores deberán introducir principios de diseño de sistemas fiables en todos los aspectos del funcionamiento de un robot, tanto para la concepción del material y de programas informáticos, como para el tratamiento de datos dentro o fuera de la plataforma a efectos de seguridad.

[391] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))

- Los diseñadores deberán introducir dispositivos concebidos para asegurar que las informaciones privadas se conservan con total seguridad y solo se utilizan de manera adecuada.

- Los diseñadores deberán integrar mecanismos de salida evidentes (teclas de interrupción de urgencia) que deberán ser coherentes con los objetivos de diseño razonables.

- Los diseñadores deberán garantizar que un robot funciona de modo conforme a los principios éticos y jurídicos a nivel local, nacional e internacional.

- Los diseñadores deberán asegurarse de que las etapas de toma de decisión del robot puedan ser objeto de reconstrucción y trazabilidad.

- Los diseñadores deberán asegurarse de que es conveniente una transparencia máxima en la programación de los sistemas robóticos, así como la previsibilidad del comportamiento de los robots.

- Los diseñadores deberán analizar la previsibilidad de un sistema humano-robot teniendo en cuenta la incertidumbre en la interpretación y en la acción, así como los posibles fallos de los robots o del hombre.

- Los diseñadores deberán desarrollar instrumentos de rastreo en la fase de concepción del robot. Estos instrumentos permitirán tener en cuenta y explicar los comportamientos de los robots, aunque sea de forma limitada, en los distintos niveles previstos para los expertos, los operadores y los usuarios.

- Los diseñadores deberán elaborar protocolos de concepción y evaluación, y colaborar con los usuarios y las partes interesadas potenciales para evaluar las ventajas y los riesgos de la robótica, incluido a nivel cognitivo, psicológico y medioambiental.

- Los diseñadores deberán asegurarse de que los robots son identificables como tales al relacionarse con seres humanos.

- Los diseñadores deberán salvaguardar la seguridad y la salud de las personas que interactúan y entran en contacto con los robots, teniendo en cuenta que estos, como productos, deberán elaborarse utilizando procesos que garantizan su seguridad y protección. Un ingeniero en robótica ha de preservar el bienestar humano, al tiempo que respeta los derechos humanos, y no podrá accionar un robot sin garantizar la seguridad, la eficacia y la reversibilidad del funcionamiento del sistema.

- Los diseñadores deberán obtener el dictamen favorable de un comité de ética de la investigación antes de probar un robot en un entorno real o implicando a seres humanos en los procedimientos de concepción y desarrollo.

Muchos de estos aspectos son considerados en las propuestas reguladoras posteriores, y algunas de ellas son indudable relevancia a efectos de responsabilidad, especialmente respecto de la previsibilidad de los robots.

4.2.5. Modelo de licencia para los usuarios

Los términos y condiciones de la licencia recogen un conjunto de derechos y obligaciones que los usuarios de un sistema deberán considerar y que me permito igualmente transcribir[392] a continuación:

- Los usuarios estarán autorizados a hacer uso de un robot sin miedo de perjuicio físico o psicológico.

- Los usuarios deben tener derecho a esperar que un robot efectúe las tareas para las que haya sido expresamente concebido.

- Los usuarios deben ser consciente de que los robots pueden tener límites de percepción, límites cognitivos y límites de accionamiento.

[392] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))

- Los usuarios deberán respetar la fragilidad humana, tanto física como psicológica, así como las necesidades emocionales de los seres humanos.

- Los usuarios deben tener en cuenta el derecho a la vida privada de las personas, incluida la desactivación de videomonitores durante procedimientos íntimos.

- Los usuarios no están autorizados a recoger, utilizar o divulgar información personal sin el consentimiento explícito de la persona concernida.

- Los usuarios no están autorizados a utilizar un robot de modo contrario a los principios y normas éticas o jurídicas.

- Los usuarios no están autorizados a modificar los robots para utilizarlos como armas.

Estos códigos y modelo de licencia para diseñadores y usuarios no han sido contemplados en las propuestas posteriores.

4.3. Declaración sobre la inteligencia artificial, la robótica y los sistemas autónomos.

El *Grupo Europeo sobre Ética de la Ciencia y las Nuevas Tecnologías* publicó el 9 de marzo 2018 sus reflexiones y conclusiones en materia ética, en la denominada *Declaración sobre inteligencia artificial, robótica y sistemas autónomos*[393], en la que se recogían los principios éticos que deben inspirar la futura regulación de los mismos en la Unión.

El documento pretende constituir las bases para la construcción de un marco ético y legal común e internacionalmente reconocido para el diseño, producción, uso y gobernanza de la inteligencia artificial, la robótica y los sistemas "autónomos", proponiendo un conjunto de principios éticos fundamentales que pueden servir de guía para el desarrollo de este

[393] Recuperado de: http://www.bioeticayderecho.ub.edu/archivos/pdf/EGE_inteligencia-artificial.pdf. Consultado el 08.02.2021

marco ético y legal, basados en los valores establecidos en los Tratados de la UE y en la Carta de Derechos Fundamentales de la UE.

Los principios éticos que recoge son los siguientes:

a) Dignidad Humana

El principio de la dignidad humana, el reconocimiento de la condición inherente del ser humano que lo hace digno de respeto, no debe ser violado por las tecnologías "autónomas". Esto implica, por ejemplo, que la toma de decisiones y la clasificación de individuos hechas por algoritmos y sistemas "autónomos" debe ser regulada, especialmente cuando los involucrados ignoran estas prácticas. También implica que tienen que existir límites (legales) para evitar que se le haga creer a las personas que están tratando con seres humanos, cuando en realidad están tratando con algoritmos y máquinas inteligentes. En este contexto es valioso adoptar una concepción relacional de la dignidad humana, que es aquella que se define según nuestras relaciones sociales. De acuerdo a esta concepción, es necesario que conozcamos si estamos interactuando con una máquina u otro ser humano y cuándo ocurre. Además, esta concepción de dignidad requiere que nos reservemos el derecho de decidir si asignamos determinadas tareas a humanos o máquinas.

b) Autonomía

El principio de autonomía implica la libertad del ser humano y se halla directamente relacionado directamente con la transparencia y previsibilidad de los sistemas. Esto se traduce en responsabilidad humana. Para evitar que los sistemas "autónomos" menoscaben la libertad de los seres humanos de establecer sus propios estándares y normas, y de poder vivir de acuerdo con ellos, es necesario tener control y conocimiento sobre ellos. Por consiguiente, todas las tecnologías "autónomas" deben respetar la capacidad humana de elegir si delegarles determinadas decisiones o acciones, cuándo y cómo hacerlo. Esto requiere que los sistemas "autónomos" sean transparentes y previsibles, características sin las cuales sería imposible para los usuarios intervenir o detenerlos cuando lo así lo consideren moralmente necesario.

c) Responsabilidad

El principio de responsabilidad debe ser fundamental en la investigación e implementación de la inteligencia artificial. Los sistemas "autónomos" sólo deberían desarrollarse y aplicarse si sirven al bienestar social y ambiental a nivel global. Establecer dicho bienestar requiere procesos democráticos deliberativos. En otras palabras, los sistemas "autónomos" deben ser diseñados de manera que sus impactos respeten la pluralidad de valores y derechos humanos fundamentales. En vista del gran desafío que supone el potencial abuso de tecnologías "autónomas", es crucial ser conscientes de los riesgos y adoptar el principio de precaución.

Desde este punto de vista, las aplicaciones de la inteligencia artificial y la robótica no deben entrañar riesgos inaceptables para los seres humanos. Tampoco deben comprometer la libertad o la autonomía humana al reducir ilegítima y subrepticiamente las opciones o el conocimiento de los ciudadanos. Al contrario, el desarrollo y el uso de estas aplicaciones deberían dirigirse a incrementar el acceso al conocimiento y a las oportunidades para los individuos.

La investigación, el diseño y el desarrollo de la inteligencia artificial, la robótica y los sistemas "autónomos" deben ser guiados por un auténtico interés en la ética de la investigación, en la responsabilidad social de los programadores y en la cooperación académica mundial para proteger derechos y valores humanos fundamentales.

Además, estas tecnologías deben ser diseñadas de forma que promuevan esos derechos y valores, evitando a toda costa que más bien los deterioren.

d) Justicia, equidad y solidaridad

La inteligencia artificial debería contribuir a la justicia global y facilitar la igualdad de acceso a los beneficios y ventajas de la misma, la robótica y los sistemas "autónomos". Los sesgos discriminatorios en los conjuntos de datos utilizados para entrenar y ejecutar los sistemas de inteligencia artificial deben evitarse y, en caso de imposibilidad, detectarse, notificarse y neutralizarse en la etapa más temprana

del proceso. Además, se destaca la necesidad de prestar especial atención a los impactos negativos de la acumulación creciente y masiva de datos personales.

e) Democracia

Las decisiones clave sobre la regulación de la inteligencia artificial, específicamente sobre su desarrollo y aplicaciones, deben ser el resultado de procesos de debate democrático y participación ciudadana, para lo que se requiere un espíritu de cooperación global y procesos de diálogo público. Del mismo modo, debe garantizarse el derecho a la educación y a la información sobre las nuevas tecnologías y sus implicaciones éticas para permitir que la sociedad en general comprenda los riesgos y las oportunidades. También debe permitirse a la sociedad participar en los procesos de toma de decisiones que son cruciales para construir su futuro y considera que el derecho de los seres humanos a la autodeterminación a través de medios democráticos es central a la dignidad humana y a la autonomía. Del mismo, pone en valor el pluralismo como valor, la diversidad y la incorporación de las diferentes concepciones de lo que es tener una buena vida para los sistemas políticos democráticos. Las nuevas tecnologías no deben poner en peligro a los ciudadanos, despojarlos de sus derechos o de su individualidad. Tampoco deben inhibir o influir en la toma de decisiones políticas, infringir la libertad de expresión y el derecho a recibir y difundir información sin interferencia. Al contrario, estas tecnologías deberían ser herramientas para beneficiarnos de la inteligencia colectiva, y para apoyar y mejorar los procesos cívicos de los que dependen nuestras sociedades democráticas.

f) Estado de derecho y rendición de cuentas

Este principio se halla asociado a la exigencia de responsabilidad de la inteligencia artificial.

El estado de derecho, el acceso a la justicia y el derecho de recibir una compensación y un juicio justo, proporcionan el marco necesario para garantizar la observancia de las normas de derechos humanos. Asimismo, estos proveen los mecanismos para el desarrollo de eventuales regulaciones específicas para la

inteligencia artificial. Esto incluye la protección contra la violación de los derechos humanos por parte de los sistemas "autónomos", por ejemplo, la seguridad o la privacidad.

Este principio ético fundamental se halla asociado a la responsabilidad derivada de los daños causados por los sistemas inteligentes. Conforme recoge el documento, los desafíos legales prácticos deben abordarse con un esfuerzo oportuno para desarrollar soluciones sólidas que asignen responsabilidades de manera clara y justa y para establecer una legislación vinculante eficiente. En este sentido, los gobiernos y las organizaciones internacionales deben incrementar sus esfuerzos para establecer en quién recae la responsabilidad de los daños causados por el desempeño no deseado de los sistemas "autónomos". Asimismo, deben instituirse sistemas efectivos de mitigación de daños.

g) Seguridad, protección, e integridad física y mental

La seguridad y la protección de los sistemas "autónomos" se concretan en: (1) la seguridad externa, que se ofrece al entorno y a los usuarios, (2) la confiabilidad y la robustez interna, por ejemplo, contra la piratería y (3) la seguridad emocional, que se refiere a la interacción humano-máquina. Estas tres dimensiones de la seguridad y la protección deben ser tomadas en cuenta por los desarrolladores de inteligencia artificial y deben ser estrictamente evaluadas antes del lanzamiento de cualquier sistema "autónomo". Esto con el fin de garantizar que estos sistemas no infrinjan el derecho de los seres humanos a la integridad física y mental, y a un entorno seguro. Se debe prestar especial atención a aquellas personas en posiciones vulnerables, así como al posible doble uso y a la militarización de la inteligencia artificial. Por ejemplo, en los campos de ciberseguridad, finanzas, infraestructura y conflictos armados.

h) Protección de datos y privacidad

En una era de recopilación generalizada y masiva de datos a través de tecnologías digitales de la comunicación, el derecho a la protección de la información personal y el derecho a la privacidad están siendo decisivamente cuestionados. En efecto, la

inteligencia artificial debe respetar las regulaciones de protección de datos. Esto tanto si la inteligencia artificial se concreta en la forma de robots físicos, aquellos que forman parte del internet de las cosas, o en la forma de *softbots*, aquellos que operan a través de la red informática mundial. Ni robots ni *bots* deben recopilar o difundir datos, ni ser ejecutados en conjuntos de datos para los que estas actividades no han sido consentidas. Los sistemas "autónomos" no deben interferir en el derecho a la vida privada. Esto incluye el derecho a estar libres de tecnologías que influyan en las opiniones y el desarrollo personal, el derecho a establecer y desarrollar relaciones con otros seres humanos, y el derecho a estar libres de vigilancia. También se deben definir criterios precisos y establecer los mecanismos apropiados para asegurar que el desarrollo y la aplicación de los sistemas "autónomos" sean éticamente correctos. Los posibles efectos de los sistemas "autónomos" en la vida privada y la privacidad generan gran preocupación. A la luz de tales inquietudes, es valioso considerar el debate actual sobre la introducción de dos nuevos derechos: el derecho al contacto humano significativo y el derecho a no ser perfilado, medido, analizado, aconsejado -en inglés *coached*- o provocado -en inglés *nudged*.

i) Sostenibilidad

La tecnología de inteligencia artificial debe responder a la responsabilidad humana de garantizar los prerrequisitos fundamentales para la vida en nuestro planeta, la continua prosperidad de la humanidad y la conservación del medioambiente para las generaciones futuras. Las estrategias para evitar que las futuras tecnologías afecten negativamente la vida humana y la naturaleza necesitan fundamentarse en políticas que prioricen la protección del medio ambiente y la sostenibilidad.

La declaración considera que todos estos principios no son un obstáculo sino un estímulo y oportunidad para la innovación, por lo que propone a la Comisión Europea investigar sobre los actuales y nuevos instrumentos normativos y de gobernanza que plantean estas tecnologías, y propone iniciar el camino hacia un marco ético y legal común e internacional para el diseño, la producción, el uso y la gobernanza de la inteligencia artificial, la robótica y los sistemas "autónomos".

Las Resoluciones del Parlamento Europeo de 20 de octubre de 2020, objeto de análisis en esta investigación, constituyen un paso firme del mismo con este propósito.

4.4. Directrices europeas para una IA fiable

La Comisión Europea consideró necesario definir unas directrices éticas para el desarrollo y utilización de la IA y ello, sobre la base del sólido marco regulador actual construido sobre un conjunto de valores fundamentales, la novedad de esta tecnología que, como he analizado en anteriores apartados, no es tal realmente, y los retos que plantea, especialmente dado su despliegue, aplicación y uso incesante en todo tipo de actividades y sectores.

La UE construyó este nuevo enfoque ético con el objetivo de reforzar la confianza de los ciudadanos en el desarrollo digital y, en particular, en la inteligencia artificial, así como proporcionar también seguridad y una ventaja competitiva para las empresas europeas, con el objetivo de mantener la reputación de seguridad y calidad asociadas en la actualidad a los productos europeos, en la medida que la Comisión Europea venía considerando que la inteligencia artificial únicamente podría considerarse fiable si se desarrolla y utiliza de forma que respete unos valores éticos ampliamente aceptados y compartidos.

La consecución de los objetivos precitados requería, a juicio de la Comisión Europea, que las aplicaciones de inteligencia artificial no solo se ajustaran al marco regulador, sino que también respeten unos principios éticos que garanticen que su implementación evite daños involuntarios. Con esta finalidad, la Comisión consideró que debía garantizarse la diversidad de género, origen racial, étnico, religión, creencias, discapacidad o edad en todas las fases de desarrollo de la inteligencia artificial, es decir, que no se permitiera ningún tipo de sesgo o discriminación, y que se debía empoderar a los ciudadanos y garantizar el respeto de sus derechos fundamentales.

En base a todo ello, la Comisión consideró necesario establecer unas directrices éticas basadas en el marco regulador existente para su aplicación por desarrolladores, proveedores y usuarios de inteligencia artificial en la UE, estableciendo igualmente condiciones de competencia ética en todos los Estados miembros.

Con esta finalidad la Comisión creó un *Grupo de expertos de alto nivel sobre IA*[394] al que encomendó la elaboración de dichas directrices éticas, así como de una serie de recomendaciones para la definición de una política más amplia en esta materia. Simultáneamente se creó la *Alianza europea de la IA*[395], una plataforma multilateral abierta para contribuir a la labor del precitado grupo de expertos.

El grupo de expertos publicó un primer borrador de las directrices éticas en diciembre de 2018 y, tras consultas a las partes interesadas[396] en las que participaron 511 organizaciones, incluyendo asociaciones, empresas, institutos de investigación, particulares y otros, y tras las reuniones con representantes de los Estados miembros, el grupo de expertos sobre inteligencia artificial presentó el documento revisado a la Comisión en marzo de 2019.

El *Grupo independiente de expertos de alto nivel sobre IA*, nombrado por la Comisión Europea en junio de 2018, publicó el 8 de abril de 2019 las *Directrices éticas para una IA fiable*[397] con el objeto de conseguir una inteligencia artificial en la que se pueda confiar.

Las directrices elaboradas por el *Grupo de expertos de alto nivel sobre inteligencia artificial* para alcanzar una inteligencia artificial fiable se sustentaron en tres pilares fundamentales: Licitud, ética y seguridad, que básicamente son las tres dimensiones

---

394 Recuperado de: https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence. Consultado el 15.12.2020.

395 Recuperado de: https://ec.europa.eu/digital-single-market/en/european-ai-alliance. Consultado el 15.12.2020

396 A la consulta respondieron 511 organizaciones, asociaciones, empresas, institutos de investigación, particulares y otros. Puede consultarse un resumen de las respuestas en: https://ec.europa.eu/futurium/en/system/files/ged/consultation_feedback_on_draft_ai_ethics_guidelines_4.pdf

397 *Directrices éticas para una IA fiable*. Recuperado de: https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-es/format-PDF. Consultado el 20.12.2020.

desde las que he abordado esta investigación para llegar a los aspectos de la responsabilidad por daños causados por o mediante la inteligencia artificial.

Las directrices pretenden una inteligencia artificial basada en la conformidad regulatoria, respeto de los principios éticos y la solidez desde un punto de vista ético y social.

Las directrices se construyen con el objetivo de crear una inteligencia artificial confiable o *Trustworthy AI*, que debe sustentarse en el respeto de los derechos fundamentales, el marco legal aplicable y sus principios básicos y estar alineada con la ética. Según el mismo, únicamente cuando la tecnología sea confiable, la humanidad podrá disfrutar de sus beneficios con plena libertad.

El grupo de expertos no realizó valoraciones o recomendaciones de carácter regulativo, sino que se centró exclusivamente en los principios éticos que la inteligencia artificial debería respetar y en las posibles fórmulas para alcanzar los objetivos pretendidos.

Las directrices se sustentan a su vez en cuatro grandes principios que aglutinan las normas éticas esenciales, como son el respeto a la autonomía humana, la prevención del daño, la equidad y la explicabilidad.

La ética y sus valores actúa como medida preventiva, la seguridad también. El derecho tradicionalmente lo hace como medida correctiva y reparadora y depuradora de las responsabilidades de distinta naturaleza. De ahí, esa necesaria interrelación entre las tres dimensiones desde las que se ha abordado esta investigación y nuevo enfoque legislativo, hacia un derecho más preventivo y detectivo, mediante la exigencia de evaluaciones y certificaciones previas a la puesta en uso de un sistema inteligente.

Del mismo modo, las directrices recogen siete requisitos esenciales que, a juicio del grupo de expertos, deben respetar las aplicaciones y sistemas de inteligencia artificial desde su diseño, significando que no son los únicos. El orden en que se relacionan y tratan sigue el de los principios y derechos de la Carta de la UE con los que se relacionan.

Estos requisitos afectan a los distintos actores involucrados en el diseño, desarrollo, despliegue y aplicación de los sistemas de inteligencia artificial, por lo que deben ser considerados por diseñadores y desarrolladores en su concepción.

Según estas directrices, sus principales destinatarios son las organizaciones públicas o privadas que recurren a los mismos para prestar sus servicios o para sus procesos internos, así como los responsables de su despliegue, a los que corresponderá asegurar que se cumplen. Y serán los usuarios finales y la sociedad en general quienes podrán exigir su cumplimiento.

La Comisión Europea avaló el trabajo realizado por el grupo de expertos y emitió su Comunicación[398] "*Generar confianza en la inteligencia artificial centrada en el ser humano*" en la que abordó la necesidad de generar confianza en la inteligencia artificial centrada en el ser humano, recogiendo un argumento que sustenta mi opinión manifestada en los capítulos precedentes y posicionamiento sobre la naturaleza y objetivo de la tecnología y, en particular, de la inteligencia artificial, al considerar que la inteligencia artificial no es un fin en sí mismo, sino un medio que debe servir a las personas para aumentar su bienestar.

Según las directrices y Comunicación precitadas, los requisitos esenciales para una inteligencia artificial fiable son los siguientes:

a) Intervención y supervisión humanas

   La inteligencia artificial no debe limitar la autonomía humana. La supervisión humana debe garantizar que una aplicación o sistema de inteligencia artificial no afecte a la autonomía humana, apoyando su intervención y la protección de los derechos fundamentales. Como recoge la *Comunicación de la Comisión*[399] citada, "El bienestar global del usuario debe ser primordial en la funcionalidad del sistema".

   Las aplicaciones y sistemas de inteligencia artificial deben garantizar el grado adecuado de control, incluida la adaptabilidad, la exactitud y la explicabilidad de

[398] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. "*Generar confianza en la inteligencia artificial centrada en el ser humano*". 8 de abril de 2019.
[399] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. "*Generar confianza en la inteligencia artificial centrada en el ser humano*". 8 de abril de 2019. COM/2019/168 final

los mismos, lo que además constituye no sólo un requisito de la misma sino que, en algunos contextos, es una obligación legal vinculante conforme a lo previsto en el RGPD, que regula en su artículo 22 el derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado de datos, cuando produzca efectos jurídicos en los usuarios o les afecte significativamente de modo similar.

A juicio del comité de expertos y de la Comisión, la precitada supervisión debe lograrse mediante la gobernanza de las aplicaciones y sistemas de inteligencia artificial, con lo que obviamente no puedo estar más de acuerdo, en particular, sustentando la misma en un enfoque de supervisión humana o "*Human on the loop*" (HOTL), de participación humana o "*Human in the loop*" (HITL) y/o de control humano o "*Human in command*" (HIC).

No obstante, la participación humana, es decir, la intervención humana en cada ciclo de decisión del sistema, en muchos casos y según el *Grupo de expertos de alto nivel sobre inteligencia artificial*, ni es posible ni deseable, como expondré.

La supervisión humana, de un lado, hace referencia, a juicio del comité de expertos, a la capacidad de la intervención humana durante el ciclo de diseño del sistema y a la supervisión del funcionamiento del sistema. Y el control humano, de otro, hace referencia a la capacidad de supervisar la actividad global del sistema de inteligencia artificial (incluido su impacto más amplio económico, social, jurídico y ético) y a la capacidad de decidir cuándo y cómo utilizar el sistema en cada situación determinada. Esto puede incluir la decisión de no utilizar un sistema de inteligencia artificial en una situación concreta, establecer niveles de discreción humana durante el uso del sistema o garantizar la capacidad de imponerse a una decisión tomada por el sistema.

El control humano en el diseño se materializaría en los análisis de riesgos y evaluaciones de impacto de los sistemas inteligentes.

La Comunicación precitada apuesta por garantizar que las autoridades públicas tengan la capacidad de ejercer sus competencias de supervisión, por lo idénticos contextos, cuando menor sea la supervisión que puede ejercer un ser humano sobre

un sistema o aplicación de inteligencia artificial, más extensas tendrán que ser las pruebas sobre el mismo y más estructurada su gobernanza.

b) Solidez y seguridad técnicas

Las aplicaciones y sistemas de inteligencia artificial deben ser seguros, fiables y resilientes, es decir, resistir ataques y evitar usos malintencionados, incluyendo ataques de fuerza bruta o actuación frente a la manipulación de datos o algoritmos.

Los sistemas deben integrar medidas de seguridad adecuadas a los riesgos previamente identificados para su gestión.

Los programas informáticos y algoritmos sobre los que se sustentan deben ser seguros, confiables y sólidos. Asimismo, deben garantizar de planes de contingencia, la reproductibilidad y me permitiría adicionar de manera consecuente, planes de recuperación ante desastres. Un ejemplo, un sistema de gestión de tráfico rodado en una *smart city.*

El testeo y aseguramiento de la calidad son esenciales para la consecución de estos objetivos si bien, no se puede garantizar de manera plena la inexistencia de posibles errores o defectos funcionales, riesgos inherentes o residuales que deberán ser adecuadamente gestionados mediante medidas alternativas o compensatorias de prevención y detección, así como de contingencia, mitigación y reparación, sin perjuicio de contemplar en paralelo, de manera recomendable, el traslado del riesgo a terceros, mediante seguros que aseguren el resarcimiento al perjudicado.

Las decisiones deberán ser adecuadas y acertadas o, cuanto menos, reflejar su nivel de acierto, lo que de nuevo supone analizar y gestionar adecuadamente los riesgos asociados a la toma de decisiones teniendo en cuenta su probabilidad e impacto. Y los resultados de la aplicación o sistema de inteligencia artificial reproducibles.

El grupo de expertos apuesta por integración de la seguridad desde el diseño -*Security by design*- y no estaría de más adicionar "proactiva y por defecto" -*Security by design*-, como así lo exige el RGPD en el ámbito de la privacidad, debiendo poder acreditar dicha seguridad en cada fase de su diseño y desarrollo. Y no sólo su

seguridad física, sino también la que denomina como "psicológica" de cualquier afectado.

Con este propósito, el grupo de expertos destaca la necesidad de analizar y evaluar los riesgos potenciales asociados al uso de las aplicaciones y sistemas de inteligencia artificial y su gestión, incluyendo mecanismos de minimización y, cuando sea posible, reversibilidad de resultados no deseadas o errores de funcionamiento.

c) Privacidad y gobernanza de datos

Los sistemas y aplicaciones de inteligencia artificial y su uso deben garantizar la privacidad y la protección de datos en todas las fases del ciclo de vida de los sistemas, de modo que las personas interesadas tengan el control sobre sus datos y no pueda utilizarse para fines distintos ni para usos perjudiciales ni discriminatorios. Además, esto ya es una exigencia legal de conformidad con lo previsto en el RGPD.

El potencial de determinados sistemas de inteligencia artificial para el registro digital y análisis de comportamientos humanos, puede derivar en que no solamente se infieran aspectos de las personas como edad, sexo o preferencias, sino orientación sexual, religión, creencias o ideología política. A modo de ejemplo, pensemos en el potencial de los sistemas de inteligencia artificial asociado al uso de una red social desde el mismo momento en que el usuario navega por la misma e interacciona ante determinados estímulos e indicadores a través de las ventanas que visualiza.

Los tratamientos de datos mediante sistemas de inteligencia artificial deben cumplir el marco regulativo en materia de privacidad y protección de datos y, por ende, sus principios y obligaciones.

En relación con todo ello, la Comunicación citada recoge expresamente algunos de los principios y obligaciones que deben garantizarse en relación el tratamiento de datos contemplados ya en el RGPD, en particular, los siguientes:

- Calidad de los datos que recogen, almacenan y utilizan los sistemas dotados de inteligencia artificial -especialmente por los sesgos sociales que puedan contener, inexactitudes y errores-, lo que deberá abordarse antes de entrenar un sistema de inteligencia artificial con datos reales, y me permito adicionar, sean personales o no;

- Integridad de los datos, sometiendo a prueba y documentando los procesos y conjuntos de datos en cada fase desde la planificación, el entrenamiento, el ensayo y el despliegue;

- Acceso a los datos.

En definitiva, garantizar la privacidad en el suministro o captación de datos, su tratamiento y resultados del mismo, tanto en el diseño, en el despliegue y como en el funcionamiento del sistema, esto es, en todo su ciclo de vida.

d) Transparencia

La transparencia debe ser tanto del sistema, de los datos como del modelo de negocio u organizativo. Los sistemas y aplicaciones de inteligencia artificial deben tener garantizada su trazabilidad, explicabilidad y adecuada comunicación de sus capacidades y limitaciones a las distintas partes interesadas.

De un lado, la trazabilidad exige que se registren y documenten tanto las decisiones tomadas por los sistemas como el conjunto de datos y la totalidad de los procesos sobre los que opera el sistema -recogida, etiquetado y algoritmo utilizado-, es decir, garantizar su explicabilidad -*Explainable AI*-.

De otro, el sistema debe diseñarse de modo que sea posible explicar tanto sus procesos técnicos como las decisiones humanas asociadas, de manera comprensible por las personas, es decir, explicar el proceso de toma de decisiones y de manera adaptada a las personas afectadas. Las explicaciones deberían alcanzar el grado en que un sistema de inteligencia artificial puede influir, condicionar y configurar el

proceso organizativo de toma de decisiones, las opciones de diseño del sistema y la justificación de su despliegue que deben estar formalizadas y disponibles.

Y, por último, los sistemas de inteligencia artificial deben ser identificables como tales al objeto de garantizar que los usuarios puedan saber que está interactuando con un sistema de inteligencia artificial y quienes son los responsables del mismo.

e) Diversidad, no discriminación, equidad y accesibilidad

Los sistemas de inteligencia artificial deben estar orientados y construidos sobre la diversidad, la no discriminación, la equidad y la accesibilidad.

No obstante, no debemos olvidar que los algoritmos discriminan por diseño y constituye su función principal legítima, por lo que debe evitarse usar datos o atributos especialmente sensibles que puedan generar una discriminación ilegítima o contraria al ordenamiento jurídico y los derechos fundamentales.

Por ende, los diseñadores deben velar por una discriminación funcional legítima, lícita y respetuosa con los derechos fundamentales.

Además de todo ello, ya existen herramientas para gestionar y mitigar los riesgos de sesgo en los datos o en los algoritmos para que el resultado no sea discriminatorio[400], detectando posibles datos o variables sensibles en las bases de datos de la que se nutre el sistema como, por ejemplo, género, raza o religión. No obstante, no son eficaces para todos los supuestos. En el capítulo anterior ya expuse algunas de las herramientas para prevenir, detectar, mitigar y erradicar el sesgo, incluso utilizando la inteligencia artificial.

[400] BENJAMINS, R.; BARBADO, A Y SIERRA, D. (2019). "Responsible AI by Design in Practice*"*. *Proceedings of the Human-Centered AI: Trustworthiness of AI Models & Data (HAI) track at AAAI Fall Symposium*, DC, 2019.

Los sesgos pueden ser de distintos tipos, especialmente los que tienen su origen en los datos y contexto tomado como referencia, y en los sesgos humanos.

Los sesgos más habituales se producen en los datos de muestra de los que se alimenta el sistema inteligente, de modo que no representen adecuadamente el entorno en o con el que deberá operar el sistema. Los datos que se seleccionen para el entrenamiento del sistema deben ser lo suficientemente representativos y adecuados del contexto real donde operará el sistema. Algunos sesgos se producen cuando se omiten o eliminan características o variables relevantes por inconsciencia o percepción de irrelevancia, de modo que la muestra es sesgada. Asimismo, también se pueden producir sesgos en la propia medición y recogida de datos en función del medio de recogida.

Los datos usados para entrenar el sistema inteligente pueden influenciar significativamente su comportamiento.

En relación con este asunto, me permito significar un estudio llevado a cabo los investigadores del *Massachusetts Institute of Technology* (MIT) para crear la primera inteligencia artificial psicópata[401]: "Norman" -nombre inspirado en el protagonista de la película *Psicosis* (1960), Norman Bites-.

El estudio evidencia el comportamiento obtenido de dos sistemas al entrenar ambos algoritmos utilizando una metodología de *Deep Learning,* pero con diferentes conjuntos de datos, unos con sesgos integrados.

La conclusión del estudio evidencia que los datos que se utilizan para enseñar un algoritmo de aprendizaje automático pueden influir significativamente en su comportamiento, por lo que cuando se dice que el sistema o el algoritmo de inteligencia artificial es sesgado e injusto, el culpable a menudo no es el algoritmo en sí, sino los datos sesgados que se le enviaron.

[401] Recuperado de: http://norman-ai.mit.edu/. Consultado el 14.02.2021.

La misma metodología de aprendizaje puede ver cosas muy diferentes en una imagen, incluso cosas enfermas, si fue entrenada con el conjunto de datos adecuado o inadecuado.

El estudio sometió a "Norman" a una exposición prolongada a algunos de los contenidos más oscuros relacionados con la muerte del sitio web www.reddit.com.

El experimento representa un estudio de caso sobre los peligros de la inteligencia artificial y su resultado evidencia las peligrosas consecuencias cuando se utilizan datos sesgados en algoritmos de aprendizaje automático.

Además de estos sesgos, se pueden producir sesgos humanos con origen en la persona que diseñó el sistema o creó el modelo de relación e interacción, incluso de forma absolutamente inconsciente. Los sesgos que pueda tener el diseñador o programador pueden proyectarse al sistema, ya sea de manera consciente o inconsciente, especialmente si tiene su origen en la cultura, país, costumbres o convicciones ideológicas o religiosas del creador del sistema, especialmente por su origen o ubicación.

Los sesgos pueden conllevar una responsabilidad ética, no del sistema inteligente, ante su ausencia de consciencia, personalidad e imputabilidad cualquiera que sea el grado de autonomía conferido, sino al creador del sistema que pudo intervenir sin la diligencia debida, dolo o mera negligencia que posibilitó el sesgo en el funcionamiento del sistema.

Los sistemas deben ser diseñados y desarrollados desde estos principios en la medida que los conjuntos de datos pueden verse afectados por sesgos históricos, culturales, por no estar completos o por aplicar modelos de gobernanza inadecuados. El mantenimiento de estos sesgos puede conllevar una discriminación ilegítima, ilícita o causar daños de distinta naturaleza al facilitar su explotación intencionada de sesgos o por competencia desleal. La forma en que se desarrollan y programan los sistemas pueden conllevar también dichos sesgos.

Las directrices del grupo de expertos recomiendan la utilización de equipos de diseño diversificado y mecanismos de participación ciudadana en el desarrollo de la inteligencia artificial para evitar estos problemas.

Por último, este conjunto de requisitos recoge la accesibilidad, de modo que los sistemas de inteligencia artificial deberían tener en cuenta toda la gama de capacidades, habilidades y necesidades humanas para garantizar la accesibilidad de un enfoque de diseño universal con independencia de las capacidades de cada persona.

f) Bienestar social y responsabilidad medioambiental

La fiabilidad en los sistemas de inteligencia artificial debe ir acompañada de sostenibilidad y responsabilidad medioambiental, generalmente ya prevista tanto en la normativa de la Unión como en los distintos ordenamientos jurídicos de los estados miembros.

Los sistemas de inteligencia artificial pueden tener impacto tanto medioambiental y ecológico, pero también social, por lo que se introduce en este conjunto de requisitos la necesidad de prestar especial atención a contextos relacionados con procesos democráticos, en especial, la formación de la opinión, toma de decisiones políticas y procesos electorales.

g) Rendición de cuentas.

Los sistemas de inteligencia artificial deben garantizar la responsabilidad y rendición de cuentas antes y después de su implementación, es decir, del sistema y sus resultados, considerando fundamental la posibilidad de auditar, esto es, la evaluación de los sistemas, los datos y los procesos por auditores internos y externos que, sin duda, contribuye en gran medida a la fiabilidad pretendida.

El concepto responsabilidad de un sistema inteligente desde un punto de vista ético ha evolucionado para situarse en la obligación de justificar sus decisiones y la posibilidad de enfrentarse a sanciones si dicha justificación resulta inadecuada. Y

como tal se colocado por encima de otros correlacionados como la equidad, la transparencia y la explicabilidad.

Para la Comisión Europea, conforme recoge en la Comunicación precitada, *"*la posibilidad de realizar auditorías externas debe garantizarse especialmente en aplicaciones que afecten a los derechos fundamentales, por ejemplo, las aplicaciones críticas para la seguridad*"*.

En este sentido, la Agencia Española de Protección de Datos publicó el pasado mes de enero de 2021 un documento de *Requisitos para Auditorías de Tratamientos que incluyan IA*[402]. Asimismo, la Comisión Europea aboga por la realización de evaluaciones de impacto, como un mecanismo idóneo para identificar riesgos, prevenir y mitigar sus efectos, y que constituye un requisito en determinados contextos por el RGPD en relación con el tratamiento de datos personales.

La Comisión también establece la necesidad de disponer de mecanismos de reparación adecuada cuando se produzcan efectos negativos injustos.

La rendición de cuentas debe fundamentar la responsabilidad por los daños causados por los sistemas inteligentes y constituye la base para su construcción jurídica, con relación con otros principios éticos como transparencia, explicabilidad y auditabilidad. De nuevo, se evidencia la necesaria vinculación entre ética y Derecho para construir los marcos regulatorios especiales que deben regular una realidad tan compleja.

Conforme analizaré en el capítulo V, los marcos actuales de responsabilidad no son suficientes para abordar adecuadamente toda la casuística que pueden envolver a la inteligencia artificial, por lo que tanto los marcos específicos que regulen la responsabilidad civil derivada de la inteligencia artificial como su seguridad deberán partir de la ética, como base y fundamento de su exigencia, dado que todos

[402] Recuperado de: https://www.aepd.es/es/documento/requisitos-auditorias-tratamientos-incluyan-ia.pdf. Consultado el 15.03.2021.

de estos principios, solo algunos se hallan expresamente regulados por los marcos vigentes, por ejemplo, el de privacidad.

Las directrices éticas analizadas coinciden en gran medida con las recogidas en otros marcos previos promovidos por otras instituciones a los que he hecho mención anteriormente.

Las directrices expuestas ya contemplan que estos requisitos deben ser considerados en cada contexto específico de aplicación para su implementación concreta y proporcionada bajo un enfoque basado en el impacto, conforme recoge expresamente la Comunicación de la Comisión Europea objeto de análisis**.**

En mi opinión, más que un enfoque basado en el impacto debería ser un enfoque basado en riesgos, de modo que no sólo tenga en consideración el impacto, sino también la probabilidad.

Las directrices elaboradas por el grupo de expertos no son vinculantes, por lo que no crean nuevas obligaciones legales, sin perjuicio de que algunas disposiciones legales vigentes del Derecho de la Unión como del Derecho interno de los distintos Estados, como he referido, contemplen ya la exigencia legal de algunos de estos requisitos esenciales a nivel ético como, por ejemplo, los marcos regulatorios de seguridad, privacidad y protección de datos personales o de protección medioambiental.

El mayor reto que plantean desde su concepción es su efectiva adopción y cumplimiento, para lo que disponemos de distintas vías que considero no excluyentes sino más bien complementarias y necesarias, como su inclusión en metodologías de desarrollo y de verificación de sistemas inteligentes, códigos de conducta y otros instrumentos de autorregulación, sistemas de evaluación y certificación, normalización, creación de marcos de gobernanza y, por supuesto, la regulación legal.

La Comisión Europea expresa en su Comunicación citada su apoyo a los requisitos esenciales relacionados propuestos por el Grupo de expertos, en la medida que están basados en los valores europeos, animando a las partes interesadas a aplicarlos para generar el pretendido marco de confianza y seguridad.

Sin embargo, como ya he apuntado y analizaré con más detalle más adelante, en mi opinión, la definición de estos principios y requisitos éticos de carácter no vinculante no son suficientes sin la necesaria concienciación y colaboración de la industria para su autorregulación y sin el necesario acompañamiento legal que complemente el marco jurídico de UE y de cada Estado miembro para promover y garantizar su observancia, previendo los mecanismos disuasorios y conminatorios para su cumplimiento, así como las consecuencias derivadas de su incumplimiento, generando así un marco integral y de seguridad jurídica para todas las partes involucradas.

Durante los últimos años hemos visto distintas iniciativas sectoriales como códigos de buenas prácticas y directrices en distintos ámbitos y sectores, por ejemplo de seguridad en dispositivos IoT, si bien, la inobservancia de las mismas a nivel global puede generar inseguridad para todas las partes involucradas e incluso problemas de competitividad para aquéllos que las siguen y que se ven obligados a incurrir en mayores costes de producción, precios más altos en consecuencia y mayores complejidades de uso para el usuario.

A modo de ejemplo, el Internet de las Cosas (IoT) ha desembocado en los últimos años en una "tormenta perfecta", dado que el número de dispositivos conectados a Internet aumenta a un ritmo vertiginoso y exponencial.

Según CISCO, la cantidad de dispositivos conectados en Internet superará los 50 mil millones este mismo año, es decir, la cantidad de dispositivos IoT será tres veces mayor que la población mundial para 2021. Y para 2022, 1 billón de sensores en red estarán integrados en el mundo que nos rodea, con hasta 45 billones en 20 años[403].

Si al número de dispositivos conectados (smartwatch, robot-aspiradora, cámaras IP, oso de peluche, etc..) adicionamos su configuración inicial de seguridad -en algunos inexistentes, en otros "de fábrica", e inexistencia de obligaciones de *securización* personalizada al usuario del dispositivo en el primer acceso- y la falta de concienciación

[403] *2019 Cybersecurity Almanac: 100 Facts, Figures, Predictions and Statistics*. Cybersecurity Ventures, February 2019.

y capacitación del usuario sobre sus riesgos asociados a uso, tenemos ciberataques como "Mirai Attack"[404].

Por todo ello, algunos países, como Reino Unido, han considerado necesario ir más allá de los meros códigos de buenas prácticas de seguridad en IoT, mediante iniciativas legislativas para la aprobación de nuevos marcos reguladores de seguridad para estos dispositivos, es decir, para su establecer su exigencia legal y carácter vinculante.

En definitiva, estos todos esos marcos éticos son imprescindibles y, aunque no sean vinculantes en la actualidad, en mi opinión deben guiar al legislador y conformar la base esencial sobre la que deben diseñarse y construirse los futuros marcos jurídicos y requisitos regulatorios de la inteligencia artificial -cualquiera que sea su nivel de riesgo y sin perjuicio de considerar cada contexto particular-.

Además, ya contribuyen -y más si son aceptados y exigidos por la propia industria-, a definir la denominada diligencia debida, contractual o extracontractual, de cómo debe ser, qué seguridad y qué comportamiento esperar de una aplicación o sistema de inteligencia artificial que se ponga en el mercado, especialmente en caso de derivarse daños de su despliegue, aplicación o uso.

La propuesta de directrices éticas objeto de análisis durante los párrafos anteriores fue revisada posteriormente por el *Grupo de expertos de alto nivel sobre inteligencia artificial* a la vista de las aportaciones y reacciones recibidas tras su comunicación para su actualización y, tras su debate, se llegó a la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas[405], que integra la Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre los

[404] El ataque se produjo en octubre de 2016 y fue llevado a cabo por una red de robots o botnet Mirai que, mediante la infectación de dispositivos del Internet de las Cosas -*IoT* por sus siglas en inglés-, principalmente cámaras IP y *routers* domésticos, conformó una red de dispositivos zombis y lanzó un ataque masivo de denegación de servicio distribuido (DDoS) contra la infraestructura de DNS del proveedor Dyn, afectando a usuarios de empresas como Twitter, Amazon, Reddit, Spotify, PayPal o Netflix, denegándoles el acceso.

[405] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))

principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, que será objeto de análisis en los posteriores apartados.

Como expondré más adelante, la UE dio un paso más en este sentido para dotarlos de carácter vinculante mediante la Propuesta de Reglamento precitada, promocionando la autorregulación, pero asegurando su observancia mediante la regulación y exigencia legal de determinados principios y normas éticas y obligaciones jurídicas asociadas, las esenciales a todo sistema inteligente cualquiera que sea su nivel de riesgo y otras exclusivamente a los sistemas inteligentes considerados de alto riesgo conforme al mismo. No es esta la postura que evidencia la reciente Propuesta de Reglamento regulador de la inteligencia artificial, del Parlamento Europeo y del Consejo, de 21 de abril de 2021, conforme he anticipado anteriormente y abordaré con mayor detalle en el capítulo IV, en la medida que los contempla exclusivamente respecto de sistemas inteligentes de alto riesgo.

4.5. Inteligencia artificial centrada en el ser humano

Como he referido y prosiguiendo con mi análisis cronológico, la Comisión Europea acogió favorablemente los requisitos esenciales de la inteligencia artificial contemplados en las directrices precitadas del *Grupo de expertos de alto nivel sobre inteligencia artificial*, recogiendo los principios y requisitos éticos de la inteligencia artificial en su Comunicación[406] de 08.04.2019, siguiendo su estrategia en esta materia que sitúa a la persona en el centro del desarrollo de la inteligencia artificial, es decir, una inteligencia artificial centrada en el ser humano.

[406] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Bruselas, 08.04.2019 COM (2019) 168 final.

4.6. Libro Blanco sobre la inteligencia artificial: Un enfoque europeo orientado a la excelencia y la confianza

Las directrices y Comunicación analizadas en los apartados anteriores se centraron en los aspectos éticos, obviando las cuestiones jurídicas dimanantes y la referencia a los futuros marcos reguladores. Esta cuestión fue abordada en el denominado *Libro sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza*, que es objeto de análisis y comentario en este apartado.

La UE es muy consciente de lo que califica como "feroz competencia mundial"[407] por liderar la inteligencia artificial.

Dentro de sus acciones y sobre la base de la *Estrategia Europa para la IA*[408] de 2018 para la promoción de su desarrollo y adopción, la UE apostó por promover el avance científico en esta materia con el objetivo de preservar el supuesto liderazgo tecnológico que considera atesora ya la UE, garantizando la disponibilidad y seguridad de la inteligencia artificial para todos los ciudadanos.

En este sentido, el enfoque europeo durante los últimos dos años se ha construido sobre las implicaciones éticas y humanas de la inteligencia artificial y en la mejora de la utilización de los macrodatos para la innovación, y ha apostado por su regulación y la inversión para promover la adopción de la misma, abordar adecuadamente los retos que supone y los riesgos potenciales asociados a la misma.

La Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, de 19 de febrero de 2020,

[407] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, de 19 de febrero de 2020, titulada "Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza"
[408] inteligencia artificial para Europa [COM(2018) 237 final].

incorporó el denominado *Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza*[409].

El objetivo de este documento es formular las alternativas políticas para alcanzar los objetivos precitados, entre las que una adecuada regulación cobra especial protagonismo, siendo abordada de manera clara y directa sobre un conjunto de reflexiones que pretendo exponer y analizar.

La inteligencia artificial confiable ha sido identificada por la UE como uno de los pilares de la denominada *economía de los datos*, así como una oportunidad y una ventaja competitiva para la misma, sustentada en la capacidad consolidada de la UE para la creación de productos seguros, fiables y vanguardistas en todos los sectores.

El *Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza* acompaña a la denominada *Estrategia Europea de Datos* que tiene el ambicioso objetivo de conseguir que la UE se convierta en la economía más atractiva, segura y dinámica del mundo en el manejo de datos, siguiendo el modelo europeo proteccionista por el apuestan sus órganos rectores, que caracteriza las iniciativas reguladoras de los últimos años.

La UE apuesta por un modelo de calidad y seguridad con una regulación basada en sus valores y en la protección de los derechos fundamentales de las personas, con el objetivo de convertirse en líder mundial de la innovación en la economía de los datos y sus aplicaciones, siguiendo su *Estrategia Europea de Datos*[410].

La estrategia europea parte de un enfoque común alrededor de la inteligencia artificial que evite un contexto no armonizado y la fragmentación del mercado único, en la medida que dejarla en manos de iniciativas nacionales comportan riesgos contra la necesaria

[409] *On Artificial Intelligence - A European approach to excellence and trust*. European Commision, 19 de febrero de 2020. Disponible en: https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en. Consultado el 14.02.2021.

[410] COM (2020) 66 final.

seguridad jurídica, la confianza de los ciudadanos y la consolidación de la industria europea.

El documento destaca que el impacto de la inteligencia artificial no puede considerarse exclusivamente desde una perspectiva individual, sino desde una perspectiva de la sociedad en su conjunto, lo que evidencia el compromiso y calidad del enfoque europeo, que es ya una referencia internacional.

El Libro Blanco ofrece medidas políticas para facilitar el desarrollo de una inteligencia artificial segura y fiable, respetuosa con los valores y los derechos de los ciudadanos en la UE, y todo ello sustentado en un marco político que favorezca la creación de un ecosistema de excelencia y en las bases de un marco normativo que genere un ecosistema de confianza.

Según este documento, el marco normativo regulador de la inteligencia artificial debe velar por el cumplimiento de las normas de la UE, especialmente en materia de protección de derechos fundamentales y derechos de los consumidores, y debe respaldar la seguridad a ciudadanos, empresas y Administraciones públicas, bajo el enfoque antropocéntrico conforme a lo recogido en la Comunicación anteriormente citada bajo el título *Generar confianza en la inteligencia artificial centrada en el ser humano*[411].

Para la creación de ese pretendido ecosistema de excelencia, el *Libro Blanco* establece un plan de actuación basado en las siguientes acciones: La necesaria colaboración de todos los Estados miembros, mejorar la investigación e innovación -especialmente mediante la creación de centros de experiencia y pruebas-, promover la capacitación, cualificación y habilidades en este campo ante la escasez de competencias, garantizar que las PYMES puedan acceder y utilizar la inteligencia artificial, involucrar al sector privado en la agenda de investigación e innovación y que coinvierta, promover la adopción de la inteligencia artificial en el sector público, asegurar el acceso a datos e infraestructuras informáticas conforme al marco de la *Estrategia Europea de Datos* y, finalmente, en

[411] COM (2019) 168.

relación con la ética, liderar mundialmente el uso ético de la inteligencia de la inteligencia artificial, creando alianzas internacionales en torno a valores compartidos.

En relación con estos últimos aspectos, como he expuesto, la UE ha mostrado su liderazgo internacional en materia ética durante los últimos años, habiendo colaborado estrechamente en la elaboración de los principios éticos de la OCDE en materia de inteligencia artificial los cuáles, posteriormente, fueron suscritos por el G20 en su *Declaración Ministerial sobre Comercio y Economía Digital* de junio de 2019.

Asimismo, la UE está asumiendo un protagonismo en materia ética en distintos foros multilaterales, como el Consejo de Europa, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura -UNESCO-, la precitada Organización para la Cooperación y el Desarrollo Económicos (OCDE), la Organización Mundial del Comercio, la Unión Internacional de Telecomunicaciones (UIT) y Naciones Unidas.

La Comisión Europea recoge expresamente en el documento cuáles deben ser los principios que deben regir esa cooperación internacional a promover basados en el respeto de los derechos fundamentales, la dignidad, el pluralismo, la inclusión, la ausencia de discriminación o sesgo y la protección de la privacidad y de los datos personales, considerando en que estos valores deben ser exportados al resto del mundo.

Por otra parte, para la creación del pretendido ecosistema de confianza, el *Libro blanco* establece igualmente un plan de acción focalizado en la regulación.

No obstante, conforme he expuesto, la confianza había sido ya identificada como un objetivo estratégico por la UE en el que focalizar esfuerzos. Junto con la *Estrategia sobre inteligencia artificial*[412] de 25 de abril de 2018, se aprobó el denominado *Plan coordinado* con los Estados miembros para armonizar estrategias y creó también el precitado *Grupo de expertos de alto nivel sobre inteligencia artificial* que publicó las precitadas *Directrices éticas para una IA fiable* seguida de la *Comunicación de la Comisión Europea*[413] que acogió favorablemente los siete requisitos esenciales de la inteligencia

[412] COM (2018) 237
[413] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Bruselas, 08.04.2019 COM (2019) 168 final.

artificial recogidos en dichas directrices: Acción y supervisión humanas, solidez técnica y seguridad, gestión de la privacidad y de los datos, transparencia, diversidad, no discriminación y equidad, bienestar social y medioambiental, y rendición de cuentas. Como expuse anteriormente, tras el período de consultas posterior a la publicación de estas directrices, entre otros aspectos, se detectó que muchos de estos requisitos esenciales se hallaban ya contemplados en los ordenamientos jurídicos, por ejemplo, privacidad y protección de datos, pero otros no se hallan contemplados específicamente para distintos sectores, como los relativos a transparencia, seguimiento o supervisión humana.

Actualmente, los desarrolladores e implementadores de sistemas de inteligencia artificial se hallan sujetos a los marcos reguladores europeos y nacionales en materia de derechos fundamentales como, por ejemplo, la normativa de protección de datos, protección de la intimidad, no discriminación o igualdad, entre otros, así como a los marcos reguladores en materia de protección de los consumidores, normas sobre la seguridad de los productos y responsabilidad civil.

Sin embargo, como he referido, todas estas directrices éticas carecían de eficacia vinculante y de hecho se comprobó que no se hallaban contempladas por todos los marcos jurídicos vigentes, lo que evidenció la necesidad de establecer un marco regulador claro en el seno de la UE y todo ello alineado con la precitada estrategia europea, en la medida que el mismo, conforme recoge expresamente el *Libro Blanco*, no sólo generaría la necesaria confianza en consumidores y empresas en la inteligencia artificial -lo que permitiría su adopción-, sino que además contribuiría a la innovación y a la competitividad y garantizaría su conformidad con la legislación, los principios y valores de la UE.

Por todo ello y ante las características específicas de la inteligencia artificial, sus retos, riesgos y la regulación actual a nivel europeo, el *Libro Blanco* planteó dos cuestiones fundamentales que enlazan con el objeto de esta investigación:

De un lado, la necesidad de analizar si los marcos reguladores vigentes pueden hacer frente a los riesgos de la inteligencia artificial o si es necesario adaptarlos o crear nueva legislación en torno a todo ello.

De otro, la posible necesidad de armonizar y unificar estos marcos en el seno de la UE, ante la inexistencia de un marco común europeo, pero todo ello bajo un enfoque que considero necesario destacar: "El marco regulador debe dejar margen para abordar su desarrollo en el futuro". Es decir, un marco evolutivo, adaptativo y flexible.

De otro modo, el riesgo de fragmentación del mercado interior podría materializarse, lo que a su vez podría poner en peligro los objetivos de confianza y la seguridad pretendidos, así como, de manera consecuente, los de despliegue y aplicación de la inteligencia artificial en la UE. En ese sentido, me permito significar a continuación algunos ejemplos de la variedad de estrategias reguladoras internas sobre la inteligencia artificial.

El Comité alemán sobre ética en materia de datos propuso un sistema de regulación de cinco niveles sustentado en un enfoque basado en el riesgo, que contempla desde la ausencia absoluta de regulación en el caso de los sistemas inteligentes más inocuos hasta la prohibición absoluta en el caso de los más peligrosos, como a mi juicio no puede ser de otra forma, conforme referiré al abordar distintos aspectos en esta investigación.

Otros países como Dinamarca han creado un sello de ética y Malta ha creado un sistema voluntario de certificación[414].

Otra de las importantes aportaciones del *Libro blanco* sobre la inteligencia artificial es su análisis de los principales riesgos que conlleva la misma, conforme expuse al abordar sus retos y riesgos, que deben motivar también la actividad regulatoria pretendida.

La inteligencia artificial, conforme expuse en el capítulo anterior, comporta riesgos y puede causar daños de distinta naturaleza, tanto materiales que pueden afectar a la seguridad y la salud de las personas y que pueden conllevar desde el menoscabo de su patrimonio a su muerte, como inmateriales que pueden afectar a la dignidad humana, causar discriminación, impactar en su intimidad y privacidad o conllevar limitaciones del derecho a la libertad de expresión, entre otros, pero no olvidemos que también a empresas, al mercado, a las Administraciones públicas, gobiernos, estados o al medio ambiente.

[414] *Libro blanco sobre la inteligencia artificial*. Comisión Europea. 19.02.2020. COM (2020) 65 final. P. 12.

En consecuencia, el futuro marco regulador debe focalizarse no sólo en los aspectos éticos precitados anteriormente, objeto de reflexión en este capítulo, sino también en cómo minimizar o evitar los diversos riesgos de la misma.

Los principales riesgos relacionados con el uso de la inteligencia artificial analizados en el capítulo II de esta investigación se hayan relacionados con el objeto de protección de muchos de los marcos reguladores ya vigentes, concebidos para la protección de los derechos fundamentales -como la protección de datos y la no discriminación-, la seguridad -ciberseguridad, infraestructuras críticas o uso malintencionado-, o la responsabilidad civil, conforme destaca expresamente el propio *Libro blanco sobre inteligencia artificial*.

Y dichos riesgos tienen principalmente su origen en su diseño, en el uso de los datos y en su aplicación. Es por ello que el nuevo marco armonizado deberá considerar estos aspectos, abordar los aspectos no contemplados y complementar los existentes.

5. La propuesta regulatoria europea: Una inteligencia artificial centrada en el ser humano, ética y confiable. Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas

5.1. Aspectos generales

La UE dio un paso más hacia la regulación de la inteligencia artificial mediante la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas[415], que incorporó como recomendación de aquél a la Comisión, la Propuesta de Reglamento del Parlamento europeo y del Consejo sobre los

[415] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))

principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas.

De inicio, su propio título es ya significativo, en la medida que supone la regulación de unos principios éticos en una norma de alcance general y de eficacia directa como lo es un Reglamento, lo que no es habitual.

El contenido de la propuesta deriva de la visión regulatoria de la Comisión Europea objeto de análisis en los anteriores apartados -especialmente el *Libro blanco sobre inteligencia artificial*- y sustentada también en el trabajo del *Grupo de Expertos de Alto Nivel sobre inteligencia artificial* -AI HLEG- y las *Directrices éticas para una inteligencia artificial fiable*.

El documento supone una primera propuesta regulatoria a nivel europeo desde la ética y para la ética, al objeto de garantizar, de un lado, un conjunto de principios y normas éticas esenciales en todo sistema de inteligencia artificial, y otros adicionales y específicos para los considerados de alto riesgo y, de otro, un conjunto de obligaciones legales relacionadas que aseguren su observancia en su desarrollo, despliegue y uso.

De manera previa a su análisis, el futuro marco que contemple los principios y normas éticas convirtiéndolos en vinculantes debería ser un instrumento que evite las estrategias de *Ethics washing* o de ética meramente anunciada o definida pero no implementada y gestionada, que algunas multinacionales han evidenciado en los últimos años.

Este instrumento supone un avance firme para la conversión de los principios y valores éticos esenciales en principios exigibles y en obligaciones jurídicas vinculantes.

No pretende ser un marco jurídico que se limite meramente a regular su comercialización o puesta en uso, sino que pretende tener eficacia preventiva *ex ante*, exigiendo ya en su concepción, diseño y desarrollo la observancia de determinados principios y normas, regulando parcialmente la propia tecnología, es decir, lo que algunos podrían considerar una injerencia gubernamental y posible afectación del principio de neutralidad tecnológica, inevitable ante los riesgos y retos de un conjunto de tecnologías tan complejas.

Como he referido anteriormente, ni el diseño, ni la aplicación ni el uso de la tecnología en cualquier sentido no la califica como neutral.

Los algoritmos pueden mentir, manipular, engañar o ser engañados, con consecuencias gravemente lesivas para personas y cosas, lo que exige un marco regulador que cuanto menos contribuya a identificar a los responsables en los contextos en que se vean involucrados sistemas inteligentes con mayor o mejor autonomía y a garantizar un resarcimiento efectivo de las personas afectadas, sin perjuicio de que dicho marco pueda ir más allá y no sólo contemplar aspectos reactivos o correctivos sino proactivos, preventivos y detectivos en todas la fases del ciclo de vida de un sistema inteligente, desde su diseño, hasta su despliegue y uso.

Una realidad tan compleja y con tal alto impacto en aspectos como la vida, la saludo y los derechos fundamentales requiere una intervención legislativa.

Este era uno de los principales objetivos del Reglamento propuesto objeto del presente análisis.

La Resolución que integra esta propuesta fue aprobada el 20 de octubre de 2020 por el Pleno de Parlamento Europeo y acometió los principios y requisitos éticos para el desarrollo, despliegue y uso de la inteligencia artificial, la cual se haya acompañada por dos Resoluciones más coetáneas del Parlamento Europeo: La Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial[416], que contiene una Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial, y la Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial[417].

[416] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))
[417] Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial (2020/2015(INI))

De inicio, destacar la clara determinación del Parlamento Europeo y la solvencia de este paquete de propuestas a la Comisión que van más allá de meros propósitos y recomendaciones, incorporando ya sendas propuestas legislativas motivadas como "reglamento", lo que comporta que serían directamente aplicables, una vez tramitadas y aprobadas, sin necesidad de transposiciones individuales por cada Estado de la Unión, lo que obedece, entre otras razones, al objetivo de evitar la fragmentación normativa en una estrategia europea única en materia de inteligencia artificial.

De este modo, al igual que en materia de protección de datos y ciberseguridad, la UE y sus Estados miembros apuestan inicialmente por una regulación específica y directamente vinculante para alcanzar los objetivos estratégicos de liderazgo europeo en inteligencia artificial y construidos con el objetivo de garantizar la precitada seguridad y confiabilidad tanto en la ciudadanía como en los operadores, así como la innovación y la competencia frente a otros países y regiones.

El Parlamento Europeo, como he anticipado, parte de las líneas definidas en materia regulatoria en el *Libro Blanco sobre la inteligencia artificial.*

El instrumento normativo propuesto entraría en vigor a los veinte (20) días de su supuesta publicación en el Diario Oficial de la Unión Europea, si bien, es previsible un largo debate para su aprobación de éste y otros instrumentos posteriores propuestos, previendo ya una fecha de exigencia y aplicación, muy posterior a su entrada en vigor, al igual que ocurrió con el RGPD.

La propuesta legislativa era previsible que se tramitara en el primer trimestre de 2021 en el marco de la estrategia europea de datos e inteligencia artificial, si bien, se fueron produciendo nuevos debates en el seno de la UE que llevaron a la posterior Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[418].

[418] COM (2021) 206 final 2021/0106 (COD)

A continuación, se analizará la Propuesta reglamento del Parlamento europeo y del Consejo sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, de 20 de octubre de 2020, por la importancia de los aspectos que fueron objeto de tratamiento en el mismo, y el cambio que supone la nueva propuesta precitada que será analizada posteriormente.

5.1.1. Proporcionalidad de la intervención reguladora propuesta

La intervención reguladora se justifica principalmente en garantizar el desarrollo, la innovación y la competitividad con la seguridad y la confianza, así como en la aplicación homogénea de principios y normas comunes a los sistemas calificados de alto riesgo.

La propuesta se construyó hipotéticamente bajo un principio de proporcionalidad y adecuación del marco regulador de estas tecnologías recogido en el precitado *Libro blanco sobre inteligencia artificial*, de modo que el nivel de intervención reguladora se ajusta al nivel de riesgo que representan determinados aplicaciones y usos de la inteligencia artificial, intentando salvaguardar el necesario y complicado equilibrio entre la innovación, el desarrollo y la competitividad con la seguridad y la confianza.

Esta proporcionalidad se proyecta en dos ámbitos, de un lado, en la regulación de la tecnología y sus usos y, de otro, en mantener el equilibrio precitado.

Respecto del primero de ellos, alineada con el principio de neutralidad tecnológica, la propuesta no pretende inicialmente regular la tecnología en sí misma, sino que se pretende focalizarse en su funcionamiento, aplicación y uso de la inteligencia artificial, si bien, algunas de sus exigencias obviamente afectan a su concepción, diseño y desarrollo, de modo que sí está regulando dicha tecnología aunque no de manera plena, para garantizar la "*Ethics by design*" y, por ende, la ética en su funcionamiento, aplicación y uso posterior, esto es, durante todo su ciclo de vida.

De este modo, el tenor literal del propio título de la Propuesta de Reglamento es claro en este sentido, en la medida que alude tanto al desarrollo, el despliegue y el uso de la inteligencia artificial, de modo que sí está regulando la tecnología, al menos, una parte

importante de la misma para garantizar la ética, la seguridad, el cumplimiento regulatorio y los derechos fundamentales, máxime cuando muchos de requisitos y obligaciones afectan ya a su diseño y concepción.

De hecho, la propia Propuesta contempla y define en su artículo 4 las distintas fases objeto del marco ético y obligacional que propone: El desarrollo, el despliegue y el uso.

El desarrollo lo concibe de una manera amplia, en particular, como la construcción y el diseño de algoritmos, la escritura y el diseño de programas informáticos o la recopilación, el almacenamiento y la gestión de datos con el fin de crear o entrenar la inteligencia artificial, la robótica y las tecnologías conexas o de crear una nueva aplicación para la inteligencia artificial, la robótica y las tecnologías conexas existentes.

El despliegue lo concibe como el funcionamiento y la gestión de la inteligencia artificial, la robótica y las tecnologías conexas, así como su comercialización o cualquier otra forma de puesta a disposición de los usuarios.

El uso lo define como toda acción relacionada con la inteligencia artificial, la robótica y las tecnologías conexas que no sea desarrollo o despliegue.

Y en el mismo precepto citado, define la "buena gobernanza" como la manera de garantizar que los desarrolladores, los desplegadores y los usuarios adoptan y cumplen unas normas y unos protocolos de conducta adecuados y razonables sobre la base de un conjunto formal de normas, procedimientos y valores que les permitan tratar adecuadamente las cuestiones éticas cuando se plantean o antes de que se planteen, es decir, esta buena gobernanza se exige también a los desarrolladores.

En definitiva, la propuesta pretende regular el desarrollo, función, gestión, comercialización y uso de la inteligencia artificial bajo un enfoque ético, de seguridad, de responsabilidad, de cumplimiento y de respeto de los derechos fundamentales.

La propuesta no entra en que sistemas de inteligencia artificial puede ser o no desarrollados ni en las capacidades de las que pueden estar dotados los mismos, incluyendo sistemas que puedan emular la consciencia o los sentimientos humanos. De este modo, se pretende una intervención mínima necesaria.

Del mismo modo, esa proporcionalidad también se proyecta en la necesidad de salvaguardar el desarrollo, la innovación y la capacidad competitiva empresarial con la seguridad y confianza en la tecnología, evitando cargas y barreras, pero garantizando una inteligencia artificial segura y confiable en todo la UE en base a unos principios y normas comunes.

5.1.2. Principios exigibles

La propuesta se sustenta en un conjunto de principios éticos pretendidamente vinculantes que deben regir la inteligencia artificial, la robótica y las tecnologías conexas, esto es, su carácter antropocéntrico, antropogénico, control y supervisión por el ser humano, evaluación de conformidad obligatoria, seguridad, transparencia, rendición de cuentas, prohibición del sesgo y la discriminación, respeto de la intimidad, restricciones del uso del reconocimiento biométrico, derecho de resarcimiento, responsabilidad social e igualdad de género, sostenibilidad medioambiental y gobernanza adecuada, tanto de la tecnología como de los datos utilizados o producidos por la misma. Posteriormente, en los apartados 5.5 y 5.6 analizaré los mismos con profundidad.

La Propuesta de Reglamento, como he anticipado, se construyó sobre la base y con el objetivo de asegurar una inteligencia artificial antropocéntrica (el ser humano como eje) y antropogénica (con origen en el ser humano).

El desarrollo, despliegue y uso de la inteligencia artificial, la robótica y las tecnologías conexas, según la Resolución, deberá respetar los siguientes principios éticos: Antropocéntricas, antropogénicas y controladas por el ser humano, sustentadas en marcos adecuados de gobernanza, seguridad, transparencia, rendición de cuentas, respeto de intimidad y la privacidad de las personas, sometidas a evaluación de conformidad previa y obligatoria (cuando se trate de tecnologías de alto riesgo), disponibilidad de salvaguardas y vías de recurso con el sesgo y la discriminación, resarcimiento de los daños causados, comprometidas con la responsabilidad social y el respecto a la igualdad de género, ambientalmente sostenibles y, respetuosas con los derechos fundamentales en cuanto a su uso para el reconocimiento biométrico.

De antemano, este es uno de los aspectos en los que esta propuesta difiere profundamente de lo regulado en la nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, dado que todos estos principios éticos esenciales exigibles a cualquier sistema y, algunos en particular, a los sistemas considerados de alto riesgo, no se contemplan en su integridad en esta última propuesta, lo que supone a mi juicio, una de sus grades carencias, en la medida que se limita a regular con detalle algunos de estos principios y normas pero exclusivamente exigible a los sistemas considerados conforme al mismo de alto riesgo y no al resto, conforme expondré en su posterior análisis.

Los principios éticos contemplados en la propuesta de 20 de octubre de 2020 son congruentes con la línea de trabajo de la Comisión y estaban alineados con los previamente definidos en marcos como las *Directrices éticas para una IA fiable* analizados anteriormente, si bien, no se recogían expresamente algunos especialmente relevantes como, por ejemplo, la reversibilidad.

Además, considero que algunos de estos principios sobre los que se debe diseñar, construir, sustentar, desplegar, explotar, aplicar y usar la inteligencia artificial y, por ende, el marco regulador de todo ello, deberían ir más allá.

A modo de ejemplo, la evaluación de conformidad debería tener un objeto y alcance más amplio y no circunscribirse meramente a la conformidad regulatoria sino al nivel de riesgo que pueda comportar de modo efectivo un sistema y evaluar el impacto que los sistemas de dotados de inteligencia artificial puedan tener en la persona, sus bienes y derechos, así como al propio mercado.

5.1.3. Conversión de normas éticas en normas jurídicas vinculantes

La propuesta plantea de inicio una fusión entre enfoques y técnicas legislativas distintas, esto es, la inclusión de la ética y el denominado *soft law* en un instrumento normativo de alcance general y eficacia directa como es un Reglamento, bajo pretensión de conformar

un *hard law*, pero sin acompañarse de los instrumentos coercitivos y sancionadores asociados en caso de incumplimiento.

5.1.4. Enfoque basado en el riesgo

La propuesta se sustenta en un enfoque basado en el riesgo que ya se recogía en el *Libro blanco* sobre inteligencia artificial, y al que también se hacía referencia, aunque de manera residual, en la *Directrices éticas para una inteligencia artificial fiable* del *Grupo de Expertos de Alto Nivel sobre inteligencia artificial* -AI HLEG-.

El enfoque normativo basado en el riesgo no es novedoso, dado que se haya presente en otras regulaciones europeas como el Reglamento General de Protección de Datos[419] (RGPD) -especialmente visualizable en la exigencia de una evaluación de impacto de protección de datos en función del riesgo del tratamiento en su artículo 35-, o en el Reglamento de Productos Sanitarios[420] (MDR) -en relación con las reglas de clasificación de los productos sanitarios contenidos en su artículo 51 y anexos-, y todo ello en relación con la *accountability* exigida en los mismos.

5.1.5. Elaboración de la propuesta y aprobación

La iniciativa fue elaborada por el parlamentario Ibán García del Blanco (S&D, España) y fue aprobada por 559 votos a favor, 44 en contra y 88 abstenciones.

[419] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos).

[420] Reglamento (UE) 2017/745 del Parlamento Europeo y del Consejo, de 5 de abril de 2017, sobre los productos sanitarios, por el que se modifican la Directiva 2001/83/CE, el Reglamento (CE) n.° 178/2002 y el Reglamento (CE) n.° 1223/2009 y por el que se derogan las Directivas 90/385/CEE y 93/42/CEE del Consejo.

5.1.6. Estructura del Reglamento propuesto

La Propuesta del Reglamento consta de veinticuatro (24) artículos agrupados en tres (3) capítulos, precedidos de una extensa sede expositiva con cincuenta y cinco (55) Considerandos, siguiendo la estela de Reglamentos previos como el Reglamento General de Protección de Datos.

5.1.7. Objetivos generales

Los principales objetivos de la propuesta, conforme recoge la misma, son:

a) Generar confianza y seguridad en la inteligencia artificial, la robótica y las tecnologías conexas por todas las partes implicadas, ciudadanos, empresas, Administraciones públicas y gobiernos;

b) Garantizar transparencia en su uso y aplicación, mejorando el flujo de información entre ciudadanos y operadores que diseñan, desarrollan, despliegan, integran o utilizan la inteligencia artificial, la robótica y las tecnologías conexas, y todo ello con la finalidad de que estas tecnologías respeten el marco europeo, así como los principios éticos, derechos y valores recogidos en la Propuesta de Reglamento formulada;

c) Promover y apoyar el desarrollo de la inteligencia artificial, la robótica y las tecnologías conexas de manera legal y segura, adecuándose a los marcos regulativos presentes y futuros -*compliance*- que les afecten, así como identificando y gestionando los riesgos asociados, y ello desde el diseño y durante todo el proceso de desarrollo, explotación y posterior uso por empresas, profesionales o particulares;

d) Promover y apoyar el posterior despliegue de la inteligencia artificial, la robótica y las tecnologías conexas de manera legal y segura, proporcionando un marco regulador general, adecuado y proporcionado sobre el que deberá sustentarse, con

la finalidad de garantizar, de un lado, la seguridad jurídica y la innovación, y de otro, el respecto de los derechos fundamentales y la protección de los consumidores;

e) Apoyar el uso de la inteligencia artificial, la robótica y las tecnologías conexas en la UE, garantizando que sean desarrolladas, desplegadas y utilizadas de manera ética, con respecto de los principios esenciales regulados en la propuesta.

5.1.8. Definiciones a considerar

De manera previa a abordar otros aspectos, alterando el orden de su tratamiento y sin perjuicio de su tratamiento más profundo en el apartado 5.4, considero igualmente oportuno anticipar previamente algunos de los conceptos jurídicos que regula la Propuesta para determinar su objeto y alcance.

En primer lugar, la Propuesta define jurídicamente en su artículo 4 el concepto de "Inteligencia artificial" como "un sistema basado en programas informáticos o incorporado en dispositivos físicos que manifiesta un comportamiento inteligente al ser capaz, entre otras cosas, de recopilar y tratar datos, analizar e interpretar su entorno y pasar a la acción, con cierto grado de autonomía, con el fin de alcanzar objetivos específicos".

La cuestión es que introduce igualmente una definición de la precitada autonomía que parece acotar la definición de sistema inteligente y el objeto del Reglamento propuesto a sistemas más avanzados y dotados de una inteligencia artificial superior a la "débil", conforme expuse al analizar esta cuestión en el capítulo I de esta investigación, en la medida que concibe su funcionamiento conforme a unas instrucciones predeterminadas, pero sin tener que limitarse a las mismas. Considero que no era ésta la intención del Parlamento Europeo en su propuesta.

Se trata de una nueva definición jurídica de inteligencia artificial adaptada de la definición efectuada en la Comunicación de la Comisión Europea bajo el título "Inteligencia

artificial para Europa" de 25 de abril de 2018[421], y distinta a la nueva definición que recoge la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021.

De la misma, reiterar que se concibe jurídicamente como un sistema informático, que manifiesta un comportamiento inteligente por tener determinadas capacidades con cierto grado de autonomía, lo que definitivamente asocia dicha "Inteligencia" al concepto "autonomía".

Conforme expuse al analizar el concepto de inteligencia artificial en el capítulo I, considerando las distintas atribuciones asociadas a estas propuestas a un sistema de inteligencia artificial, quizás una definición más avanzada e integradora de la misma podría definirla como "Sistema de información que integra un subsistema de gobierno y tratamiento de datos así como la capacidad asociada de tomar decisiones y ejecutar acciones automatizadas derivadas de su gestión, con cierto grado de autonomía, físicas y/o virtuales, pero con supervisión y control humano durante su ciclo de vida".

No obstante, el objeto y alcance la propuesta es más ambiciosa e incluye la robótica y las tecnologías conexas, por lo que aborda igualmente estos conceptos.

La Propuesta define la "Robótica" como las tecnologías que permiten que las máquinas controladas automáticamente, reprogramables y multifuncionales realicen en el mundo físico acciones tradicionalmente realizadas o iniciadas por los seres humanos, en particular, mediante la inteligencia artificial o las tecnologías conexas.

Es decir, las concibe como aquellas tecnologías que permiten que el *hardware* ejecute acciones físicas, ya sean activas (Ej. Impedir el paso) o pasivas (Ej. Dejar pasar) mediante la inteligencia artificial. En consecuencia, en la medida que la definición se aparta claramente del concepto físico de máquina, aparato o robot, es decir, del *hardware*, para concebirlo como "tecnologías", quizás podría haber sido más adecuado hablar de técnica

[421] Comunicación de la Comisión al Parlamento Europeo, al Consejo Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Inteligencia artificial para Europa. 25.04.2018. COM (2018) 237

informática de diseño y utilización de aparatos físicos para la realización de aquellas acciones incluidas en su definición mediante la inteligencia artificial o las tecnologías conexas.

No obstante, debo significar que la definición parece reducir la robótica únicamente a la realización de acciones en el mundo físico, ¿qué ocurre entonces con las acciones en el mundo virtual o digital? Conforme a dicha definición no encajarían en el concepto las acciones llevadas a cabo mediante *software* o máquinas en el mundo virtual, lo que considero una omisión que debe ser considerada en la elaboración de los futuros marcos.

Por último, la Propuesta aborda un concepto que considero sumamente importante a mi modo de ver, que son las "tecnologías conexas", concibiéndolas como aquellas que permiten que los programas informáticos controlen, con un grado de autonomía parcial o total, un proceso físico o virtual, las tecnologías capaces de detectar los datos biométricos, genéticos o de otro tipo, así como las tecnologías que copian o utilizan de otro modo características humanas.

Y a todo ello deberá aplicarse este futuro marco ético y jurídico, como iré exponiendo a lo largo de su análisis.

5.2. Objeto

El Capítulo I del Reglamento propuesto recoge las "Disposiciones generales" y está integrado por cinco (5) artículos que regulan su objeto, ámbito de aplicación, ámbito geográfico, definiciones y los principios éticos de la inteligencia artificial, la robótica y las tecnologías conexas.

El artículo 1 de la Propuesta define su objeto consistente en el establecimiento de un marco regulador de la UE, global y con visión de futuro, de principios éticos y obligaciones jurídicas para el desarrollo, despliegue y uso de la inteligencia artificial, la robótica y las tecnologías conexas en la UE.

Al hilo de lo indicado anteriormente, el objeto de la propuesta plasma esa necesidad de regular desde la globalidad y con visión de futuro, algo que considero necesario especialmente en tecnología, y pretende ir más allá de la ética, regulando no sólo principios y normas éticas para establecer su necesario cumplimiento y carácter vinculante, sino también obligaciones jurídicas para el desarrollo, despliegue y uso de la tecnología, incluyendo inteligencia artificial, robótica y tecnologías conexas.

5.3. Ámbito objetivo y subjetivo de aplicación

Por lo que se refiere a su alcance regulado en su artículo 2, considero acertadísimo su amplio enfoque, esto es, la inteligencia artificial, la robótica y las tecnologías conexas, incluyendo los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías, desarrolladas, desplegadas o utilizadas en la UE. No obstante, la definición de inteligencia artificial y autonomía asociada, deben ser revisadas para evitar equívocos respecto de su objeto y alcance, para la inclusión de todo tipo de sistema inteligente cualquiera que sea su grado de autonomía, incluyendo la inteligencia artificial "débil".

El marco ético y jurídico propuesto resultaría aplicable y exigible no sólo a la tecnología en sí misma, sino también a los programas, los algoritmos y los datos que utilicen o que generen estas tecnologías, ya sean elementos, aplicaciones, sistemas, tecnologías, productos o servicios desarrollados en la UE o fuera de la misma, pero que sean desplegados o utilizados en ésta por operadores o usuarios ubicados o no en ésta.

En este sentido, considero que se podría haber utilizado un concepto más genérico que aglutine todo lo pretendido, evitando así la confusión que pueden generar las distintas categorías a las que se alude en referencia a la inteligencia artificial.

Quizás y siguiendo lo expuesto por mi parte anteriormente, podría ser más adecuado hablar de sistemas de información dotados de inteligencia artificial (o con inteligencia artificial integrada) o directamente sistemas de inteligencia artificial, como comenté al abordar su definición jurídica, los cuales integran *hardware* -servidores, equipos,

dispositivos, máquinas, vehículos, robots, etc.-, *software* (programas informáticos), algoritmos, redes, datos, etc.

Asimismo, conforme a su tenor literal, resultaría igualmente de aplicación a los elementos, aplicaciones, sistemas, tecnologías, productos o servicios desarrollados en la UE, pero desplegados o utilizados fuera de la misma, y ello alienado con su exigencia e integración preceptiva desde el diseño y con el objetivo de mantener el prestigio, la calidad y la seguridad de la tecnología europea.

Su ámbito geográfico de aplicación, conforme a su artículo 3, establece que se aplicará este marco igualmente "con independencia de que los programas informáticos, los algoritmos o los datos utilizados o producidos por dichas tecnologías estén localizados fuera de la UE o no tengan una ubicación geográfica específica", cuando una parte de ésta se desarrolle, despliegue o utilice en la UE.

En definitiva, el Reglamento se aplicaría a los sistemas de inteligencia artificial, robótica y tecnologías conexas que puedan ser desarrolladas en la UE, pero también a las desarrolladas fuera pero que se desplieguen o utilicen en la UE.

Este extenso ámbito de aplicación permite, de un lado, garantizar el cumplimiento y la seguridad de los sistemas desarrollados, desplegados o utilizados en la UE, cualquiera que sea su origen y, de otro, garantizar una competencia leal y ética por parte de desarrolladores y fabricantes de terceros países, que se verán sujetos a este marco para desplegar o utilizar su tecnología en la UE.

Una de las cuestiones que se suscitan en la actualidad es la relativa a sistemas de inteligencia artificial virtualizados o basados en *software*, de modo que el usuario europeo accedería a los mismos desde aquí a una ubicación lógica (informática) en países terceros, y que pueden ser ofrecidos al mismo de manera integrada por proveedores de servicios digitales o servicios de sociedad de la información en los productos y soluciones que están ya ofreciendo a los usuarios europeos, por ejemplo redes sociales, o ser ofrecidos de manera diferenciada al usuario por cualquier proveedor de este tipo de sistemas para disfrutar de concretas ventajas o beneficios.

En este sentido, el Reglamento propuesto es claro en su tenor literal, en especial, cuando los programas, algoritmos o datos utilizados o producidos por estas tecnologías estén localizados fuera de la UE o no tengan una ubicación geográfica específica, como he indicado.

En cualquier caso, considero que cualquier valoración o interpretación relacionada con los previsibles conflictos que pudieran plantearse respecto de ámbito de aplicación, debería ser resuelta conforme a su propio tenor literal y una interpretación sistemática, lógica, integradora y garantista conforme a su objeto y finalidad, entendiendo que al dirigirse a o acceder los usuarios desde la UE, éstos deben considerase protegidos por este marco jurídico, considerando que su utilización se está produciendo desde y en la UE, conforme al criterio sobre el que han construido otras normas europeas, como el Reglamento General de Protección de Datos (RGPD)[422], que resultará igualmente de aplicación en los aspectos contemplados en el mismo.

En este sentido y como he expuesto, la interpretación de las normas jurídicas debe ser fruto no sólo de su tenor literal, sino teniendo en cuenta su sentido lógico y su ponderación sistemática, que obliga a considerar el ordenamiento jurídico como un todo orgánico, en la medida que la norma debe responder principalmente al fin supremo de la justicia.

5.4. Conceptos jurídicos

El artículo 4 de la Propuesta regula los distintos conceptos jurídicos clave para determinar la aplicación del futuro marco ético y jurídico, entre otros, los conceptos de inteligencia artificial, robótica y tecnologías conexas que he analizado anteriormente, con remisión a mis comentarios precedentes.

[422] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos)

Pero entre ellos, también aborda el concepto de tecnologías de alto riesgo, que son aquellas a las que resultará de aplicación el conjunto específico de obligaciones reguladas en el capítulo II del Reglamento propuesto.

La Propuesta de Reglamento define el concepto de "alto riesgo" en su artículo 4.I.e), considerando tecnologías de alto riesgo aquéllas cuyo desarrollo, despliegue y uso entrañen un riesgo significativo de causar lesiones o daños a particulares o a la sociedad, vulnerando los derechos fundamentales y las normas de seguridad establecidas en el Derecho de la Unión, a cuyos efectos deben tenerse en cuenta el sector en el que se desarrollan, despliegan o utilizan, su uso o finalidad específica y la gravedad de la lesión o daño que cabe esperar que se produzca.

Significar que la definición y calificación como tal no recoge un aspecto esencial en análisis y gestión de riesgos: La probabilidad. Cuestión que puede considerarse razonable desde un enfoque proteccionista y objetivo, y más de derechos fundamentales, si bien, abordaré esta cuestión más adelante.

Esta definición introduce un concepto indeterminado, como lo es "riesgo significativo", que abordaré más adelante.

Por último, define conceptos como "desarrollo", "desarrollador", "despliegue", "desplegador", "uso", "usuario", "discriminación", "lesión o daño", "sesgo" y "buena gobernanza.

5.5. Principios éticos generales de la IA, la robótica y las tecnologías conexas

La Propuesta de Reglamento incorpora en su artículo 5 los principios éticos exigibles a la inteligencia artificial, la robótica y las tecnologías conexas, sean o no de alto riesgo, a diferencia de las normas y obligaciones previstas en los posteriores artículos 6 a 16 que conforman su capítulo II, que únicamente serán aplicables a los sistemas considerados de alto riesgo.

Sobre este particular, reiterar que la nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, sigue esta técnica, si bien, contempla un conjunto de principios éticos exclusivamente exigibles a los sistemas inteligentes considerados conforme al mismo de alto riesgo y no al resto -a excepción de algunos aspectos relacionados con la transparencia para determinados sistemas-, regulando adicionalmente las obligaciones relativas a aquellos sistemas de alto riesgo.

Los principios éticos recogidos en la propuesta objeto de análisis en este apartado, exigibles en su desarrollo, despliegue y utilización, serían los siguientes:

- Respeto de la dignidad, la autonomía y la seguridad humanas, así como otros derechos fundamentales establecidos en la Carta de los Derechos Fundamentales de la Unión Europea.

- Protección de la intimidad y la privacidad, incluyendo el tratamiento de datos personales derivados de datos no personales y de datos biométricos.

  Protección que debe garantizarse tanto en la recogida inicial como los generados posteriormente en la interacción del sistema inteligente con su entorno y usuarios.

  En este sentido merece destacarse como el reiteradamente referenciado Reglamento General de Protección de Datos europeo (RGPD)[423] no sólo se ha erigido en referencia internacional de instrumento normativo en materia de protección de datos personales -como expresamente lo ha reconocido consultoras internacionales como Gartner[424]-, sino que incluso ha influenciado no sólo en esta propuesta normativa, sino en otras propuestas internacionales para limitar el riesgo de las aplicaciones de la inteligencia

[423] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento General de Protección de Datos o RGPD/GDPR).

[424] *Gartner Top 10 Strategic Technology Trends for 2020*". 21.10.2019. Recuperado de: https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020/. Consultado el 02.01.2021

artificial como, por ejemplo, las elaboradas por el *AI Now Institute*[425] de la Universidad de Nueva York, dedicado a la investigación interdisciplinar para comprender las implicaciones sociales de la inteligencia artificial.

Y lo ha hecho, en especial, en aspectos como la privacidad desde el diseño y por defecto, transparencia, responsabilidad proactiva, derecho de acceso, la necesidad de análisis previo de riesgos o la exigencia de evaluación de impacto en protección de datos -EIPD-.

Además de ello, las aportaciones del RGPD en materia de análisis, gestión y mitigación de riesgos son muy relevantes.

El RGPD regula en su artículo 13.2.f)[426] las obligaciones de transparencia y explicabilidad en el tratamiento de datos personales y, en particular, la obligación la información, en su caso, sobre la existencia de decisiones automatizadas, incluida la elaboración de perfiles, incluyendo información significativa sobre la lógica aplicada, así como la importancia y las consecuencias previstas de dicho tratamiento para el interesado.

Del mismo modo, en su artículo 22[427] regula concretos derechos y correlativas obligaciones para el responsable del tratamiento de datos personales en materia de

[425] CRAWFORD, K. ET AL. (2019). *AI Now Report 2019*. AI Now Institute, New York University. Diciembre 2019. Recuperado de: https://ainowinstitute.org/AI_Now_2019_Report.pdf. Consultado el 04.03.2021. P. 31-32.

[426] RGPD. Artículo 13.2.f) La existencia de decisiones automatizas, incluida la elaboración de perfiles, a que se refiere el artículo 22, apartados 1 y 4, y, al menos en tales casos, información significativa sobre la lógica aplicada, así como la importancia y las consecuencias previstas de dicho tratamiento para el interesado.

[427] RGPD. Artículo 22 Decisiones individuales automatizadas, incluida la elaboración de perfiles.
1. Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar.
2. El apartado 1 no se aplicará si la decisión:
a) es necesaria para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento;
b) está autorizada por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento y que establezca asimismo medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, o
c) se basa en el consentimiento explícito del interesado.
3. En los casos a que se refiere el apartado 2, letras a) y c), el responsable del tratamiento adoptará las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado,

explicabilidad y de intervención humana, en particular, de un lado, el derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en la persona o que le afecte significativamente de modo similar y, de otro, la obligación del responsable del tratamiento de adoptar las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos de la persona afectada, en concreto y como mínimo, el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión.

Asimismo, el RGPD regula otros derechos para las personas afectadas sobre sus datos que permitirían mitigar los riesgos asociados a la inteligencia artificial relacionados con la privacidad y la protección de datos, en particular, el derecho de acceso, rectificación, supresión, portabilidad, limitación del tratamiento u oposición.

El RGPD constituye un eficaz instrumento de protección de los derechos e intereses de las personas cuyos datos sean tratados por sistemas de inteligencia artificial, complementando el futuro marco ético europeo. No obstante, desde otros prismas también puede verse como un obstáculo para el desarrollo, despliegue y aplicación de los sistemas de inteligencia artificial en Europa frente a otros países ante las restricciones que comporta para el acceso y tratamiento de datos, lo que requiere mayor profundización por parte de las autoridades de protección de datos en esta materia y que podría motivar futuras reformas legislativas relaciones con las iniciativas actuales de gobernanza y tratamiento datos para el bien común, a lo que me refiero con mayor profundidad en otros apartados de esta investigación.

En este sentido, la Agencia Española de Protección de Datos está liderando algunas acciones orientadas en esta línea, mediante el documento publicado por la misma bajo el título *Adecuación al RGPD de tratamientos que incorporan inteligencia artificial.*

como mínimo el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión.
4. Las decisiones a que se refiere el apartado 2 no se basarán en las categorías especiales de datos personales contempladas en el artículo 9, apartado 1, salvo que se aplique el artículo 9, apartado 2, letra a) o g), y se hayan tomado medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado.

Una introducción"[428] y la reciente guía publicada por la misma publicada bajo el título "*Requisitos para Auditorías de Tratamientos que incluyan IA*"[429].

Los principios que regula el RGPD y que constituyen las grandes obligaciones para las empresas, exigen entre otras cosas, la limitación de la finalidad, la minimización de los datos y la limitación del plazo de conservación, junto con las obligaciones de licitud, lealtad, transparencia e información.

No obstante, los riesgos asociados a la recogida, tratamiento y explotación de datos mediante sistemas de inteligencia artificial no se circunscriben exclusivamente a la vulneración del derecho a la privacidad y protección de datos, sino que alcanzan a otros derechos y libertades, tal y como analizaba en el apartado 4 del capítulo II de esta investigación, como el derecho a la igualdad, libertad de expresión o de educación, entre otros, como ha puesto de manifiesto autores como Mantelero o Peguera[430].

- Conformidad regulatoria especial (RGDP y Directiva 2002/58/CE).

  Los sistemas de inteligencia artificial deberán ser respetuosos con el marco regulativo europeo y, en especial, como he analizado anteriormente, con el RGPD y demás normas nacionales que les resulten de aplicación.

- Fomento de los proyectos de investigación orientados a promover la inclusión social, la democracia, la pluralidad, la solidaridad, la equidad, la igualdad y la cooperación.

En resumen, se trata de un conjunto de principios básicos exigibles a cualquier sistema de inteligencia artificial, cualquiera que sea su nivel de riesgo, si bien, de inicio, parece una relación algo exigua y "de mínimos" a la vista de los marcos éticos actualmente preexistentes a nivel internacional, sin perjuicio de que la Propuesta de Reglamento

[428] *Adecuación al RGPD de tratamientos que incorporan inteligencia artificial. Una introducción*. Agencia Española de Protección de Datos. Madrid. Febrero 2020. Recuperado de: https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia.pdf. Consultado el 04.03.2021.
[429] *Requisitos para Auditorías de Tratamientos que incluyan IA*. Agencia Española de Protección de Datos. Madrid. Enero 2021. Recuperado de: https://www.aepd.es/sites/default/files/2021-01/requisitos-auditorias-tratamientos-incluyan-ia.pdf. Consultado el 04.03.2021.
[430] PEGUERA, M. (2020). "En búsqueda de un marco normativo para la inteligencia artificial" en Cerrillo i Martínez, A. y Pequera Poch, M. (Coord). *Retos jurídicos de la inteligencia artificial*. Aranzadi S.A.U. (Thomson Reuters). Navarra 2020. P. 46.

complete este marco de principios generales con un marco de principios éticos vinculantes y obligaciones jurídicas, aunque éstas exclusivas de los sistemas considerados de alto riesgo, conforme se analiza a continuación.

Los principios éticos precitados, en caso de aprobarse con sus disposiciones actuales, serán obligatorios y vinculantes para el desarrollo, despliegue y utilización de sistemas de inteligencia artificial, robótica y tecnologías conexas de alto riesgo, así como para los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías, con independencia de su clasificación como de riesgo alto o significativo, y todo ello conforme a una interpretación literal, sistemática y lógica de la norma propuesta.

5.6. Principios y obligaciones para los sistemas de inteligencia artificial de alto riesgo

El Capítulo II del texto propuesto regula las obligaciones que deberán cumplir exclusivamente las tecnologías de alto riesgo consideradas como tal conforme al mismo, esto es, inteligencia artificial, robótica y tecnologías conexas, lo que deja fuera inicialmente todas aquellas que no tengan dicha consideración.

En consecuencia, todo ello exige calificar previamente el nivel de riesgo asociado a la tecnología implicada, mediante la realización de un análisis previo de riesgos que permita determinar la concurrencia de un "riesgo significativo" de causar lesiones o daños a personas o a la sociedad, vulnerando los derechos fundamentales y las normas de seguridad de la UE, derivado de su desarrollo, despliegue y uso, en base al contexto de uso o finalidad específica, sector en el que se desarrolle, despliegue o uso y la gravedad de las lesiones o daños previsibles.

Sin embargo, como he referido anteriormente, no se define que debemos considerar "riesgo significativo".

La calificación y valoración de riesgos, en general, se determina a partir de la estimación de la probabilidad de que se materialice una determinada amenaza y de la valoración del impacto.

Según las metodologías y criterios de análisis de riesgos generales en tecnología, cumplimiento y seguridad, éstos pueden ser clasificados de distintas formas relacionando su probabilidad e impacto, en base, también de las metodologías, estándares y buenas prácticas tomadas como referencia. Particularmente, en mi actividad profesional suelo utilizar distintos niveles de clasificación tanto de la probabilidad como del impacto, esto es: Despreciable, Bajo, Limitado, Significativo o Máximo, lo que a su vez determina la clasificación de cada riesgo del mismo modo.

El *Libro blanco sobre inteligencia artificial* estableció un conjunto de criterios cumulativos para determinar si una determinada aplicación puede considerarse de alto riesgo, en primer lugar, que la actividad habitual del sector al que pertenezca la aplicación sea previsible que existan riesgos significativos (sanidad, transporte, etc,) y, en segundo lugar, que del uso específico de la aplicación puedan surgir del mismo modo riesgos significativos, siendo necesario la evaluación de las posibles repercusiones de este uso.

De inicio, estos criterios introducen el precitado concepto jurídico indeterminado de "riesgo significativo", que se mantiene en la propuesta objeto del presente análisis, como expondré a continuación y tampoco ofrece un marco claro de seguridad jurídica para la aplicación de este enfoque.

La Propuesta de Reglamento desarrolla el enfoque basado en el riesgo y, en particular, respecto del concepto "alto riesgo" que lo vincula y asocia al concepto "riesgo significativo" y que lo define, como he referido anteriormente, como el riesgo de "causar lesiones o daños a las personas o a la sociedad, vulnerando los derechos fundamentales y las normas de seguridad establecidas en el Derecho de la Unión, teniendo en cuenta su uso o finalidad específicos, el sector en el que se desarrollan, despliegan o usan y la gravedad de las lesiones o daños que cabe esperar que se produzcan".

El Considerando 11 de la propuesta, se establece que una inteligencia artificial de alto riesgo es aquella que contiene un "riesgo significativo de causar lesiones o daños a particulares o a la sociedad, vulnerando los derechos fundamentales y las normas de seguridad establecidas en el Derecho de la Unión".

No obstante, considero que la definición propuesta incorpora dicho concepto jurídico relativamente indeterminado de "riesgo significativo", en la medida que parece referirse más al hecho de que se materialice que al impacto que pueda tener, y dependiendo todo ello de las distintas variables del contexto, eso sí, siempre que se vean comprometidos derechos fundamentales y normas de seguridad. Es más, no integra criterios claros para determinar la "gravedad" y sería conveniente su inclusión en aras de la seguridad jurídica.

En este sentido, comparto con Lazcoz Moratinos[431] la posibilidad de integrar el concepto a nivel interpretativo, tomando como referencia la definición de "lesión o daño" que la Propuesta de Reglamento incorpora en su artículo 4.I.n) en relación con su Considerando 11 como referencia interpretativa, el cual establece que "El grado de gravedad debe determinarse sobre la base de la magnitud de la lesión o daño potencial, el número de personas afectadas, el valor total del perjuicio ocasionado y el daño a la sociedad en su conjunto".

El Considerando precitado incorpora algunos ejemplos de distintos tipos de lesiones y daños graves, entre otros, las violaciones de los derechos de los niños, los consumidores o los trabajadores que, debido a su alcance, el número de niños, consumidores o trabajadores afectados o su impacto en la sociedad en su conjunto, entrañen un riesgo significativo de vulneración de los derechos fundamentales y las normas de seguridad establecidas en el Derecho de la Unión.

De este modo, a mi juicio, podría considerarse "alto riesgo" o "riesgo significativo", a los efectos de esta propuesta, aquel riesgo susceptible de causar lesiones o daños graves a las personas o a la sociedad -en base a la magnitud de la lesión o daño potencial, el número de personas afectadas, el valor total del perjuicio ocasionado y el daño a la sociedad en su conjunto-, concurriendo igualmente una vulneración de derechos fundamentales y las normas de seguridad de la UE, y valorando todo ello en función del contexto de uso o finalidad específica y el sector en el que se desarrolle, despliegue o use.

[431] LACOZ MORATINOS, G. (2020). "Análisis de la propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas". *Ius et Scientia*. Vol. 6. Nº 2. 2020. P. 31.

No obstante, este concepto se aparta de los criterios de praxis profesional en materia de análisis de riesgos y evaluaciones de impacto en base a su objetivación, al no considerar el grado de probabilidad de que la amenaza se materialice -sea improbable o muy alto-, considerando meramente aspectos como el sector, uso, finalidad, contexto e impacto.

La evaluación y calificación de este riesgo, como pone de manifiesto el precitado Lacoz Moratinos[432], exige pues considerar tres factores, a los que se hace referencia el artículo 14 de la Propuesta de Reglamento en relación con la evaluación de riesgos. De un lado, el sector al que pertenece, de otro el uso o finalidad específica al que se dedica y, por último, la gravedad de daño que cabe esperar que se produzca.

No obstante, en mi opinión, se exigen sendos factores o elementos adicionales objetivos para su calificación como tal, esto es, que concurra una vulneración de derechos fundamentales y de las normas de seguridad de la UE, y ambos conjuntamente, de modo que podrían producirse los factores precitados y la vulneración de normas de seguridad de la UE, pero si, hipotéticamente, no existe vulneración de derechos fundamentales, no debería calificar como alto riesgo conforme a estas consideraciones basadas en el tenor literal del precepto.

Quizás hubiera sido más oportuno utilizar, en lugar de la conjunción copulativa "y", la conjunción disyuntiva "o". A continuación, profundizaré más sobre esta cuestión.

En cualquier caso, inicialmente, parece que la norma parte exclusivamente de un criterio proteccionista y del alto nivel del impacto para clasificar el riesgo como significativo prescindiendo de su probabilidad, a la que no hace referencia en su definición, lo que podría llevarnos a interpretar que el riesgo será en todo caso significativo si su impacto es grave o alto con independencia de su probabilidad, conforme al contexto de uso o finalidad específica, sector en el que se desarrolle, despliegue o uso y la gravedad previsible de las lesiones o daños.

[432] LACOZ MORATINOS, G. (2020). "Análisis de la propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas". Op.cit. Pp. 26-41.

Según el análisis efectuado conforme a su tenor literal y en base a una interpretación integradora en base a los Considerandos, en caso de que concurra un impacto alto, pero no se vulneraran derechos fundamentales o normas de seguridad de la UE -como factores de necesaria concurrencia para su consideración-, ¿no podría clasificarse el sistema como de riesgo significativo?

Sin perjuicio de la revisión de esta cuestión en futuras propuestas, una posible solución sería mantener una interpretación sistemática y lógica, es decir, que considere el contexto y la realidad social que pretende abordar, en base a la cual no debería ser necesaria la concurrencia de estas circunstancias adicionales para su consideración como "significativo" siempre que concurra ese alto nivel de impacto, y sin perjuicio de que pueda igualmente concurrir un alto grado probabilidad.

De otro modo, en mi opinión, se conduciría una solución no adecuada al contenido y filosofía que inspira la norma que lo contiene, dado su tenor literal. Y todo ello valorado conforme al contexto de uso o finalidad específica, sector en el que se desarrolle, despliegue o uso y la gravedad de las lesiones o daños previsibles.

Por todo ello, considero que la propuesta debería contemplar una definición más precisa de "alto riesgo" y, en especial, de "gravedad" con inclusión en sus preceptos de los criterios objetivos que deben sustentar su clasificación como tal.

En cualquier caso, la propuesta opta finalmente por objetivar el riesgo y su ámbito de aplicación, y clasificar de manera exhaustiva y acumulativa los sectores, usos y fines considerados *per se* de alto riesgo y que conllevan, a juicio del Parlamento Europeo, un riesgo de violación de los derechos fundamentales y las normas de seguridad. Y con este objetivo, incorpora un anexo en el que tipifica con efectos jurídicos la relación de sectores, usos y fines de alto riesgo, que sintetizo en la siguiente tabla:

SECTORES, USOS Y FINES DE ALTO RIESGO	
Sectores de alto riesgo	Empleo Educación Asistencia sanitaria Transporte Energía Sector público (asilo, migración, controles fronterizos, sistema judicial y servicios de seguridad social) Seguridad y defensa Finanzas, bancos, seguros
Usos o fines de alto riesgo	Contratación Clasificación y evaluación de estudiantes Asignación de fondos públicos Concesión de préstamos Comercio, corretaje, fiscalidad, etc. Tratamientos y procedimientos médicos Procesos electorales y campañas políticas Decisiones del sector público que tienen un impacto significativo y directo en los derechos y las obligaciones de las personas físicas o jurídicas Conducción automatizada Gestión del tráfico Sistemas militares autónomos Producción y distribución de energía Gestión de residuos Control de emisiones

Se trata de un listado inicialmente cerrado de sectores, usos y finalidades de alto riesgo, aunque susceptible de ampliación.

De la relación incluida en la versión propuesta, se echan en falta algunos sectores o usos de riesgo especialmente significativo como, por ejemplo, la asistencia social, los sistemas de movilidad personal y comunicación para personas con discapacidad, la asistencia visual, dispositivos médicos *intracorporales* o la prevención, detección precoz, respuesta e investigación de delitos. En relación con los usos o fines de alto riesgo se evidencia también una omisión muy relevante relacionada con el objeto de esta investigación: La

ciberseguridad. No obstante, se adicionó la seguridad y defensa en general dentro de los sectores incluidos.

Como he expuesto, estemos o no de acuerdo con el listado, a muy juicio, lejos de ser exhaustivo como así se presenta, aporta cierta seguridad jurídica para determinar los criterios objetivos para calificar el riesgo asociado a un sistema inteligente, si bien, otros aspectos como los criterios para determinar la gravedad se alejan de dicha seguridad jurídica como he expuesto.

En cualquier caso, para determinar la calificación del sistema inteligente como de alto riesgo, se deberá realizar una evaluación de riesgos previa conforme a lo dispuesto en el artículo 14 de la Propuesta de Reglamento y conforme al listado "exhaustivo y acumulativo" de tecnologías del alto riesgo que incorpora la misma como anexo.

La evaluación de riesgos prevista será analizada con mayor detalle en apartados posteriores, si bien, destacar que se trata de una evaluación previa, imparcial, regulada y externa en base a criterios concretos y definidos, y sobre la base del listado de sectores de alto riesgo y de usos o fines de alto riesgo que he citado anteriormente, que la propuesta califica formalmente como "exhaustivo y acumulativo".

Los principios y normas éticas obligatorias para los sistemas considerados de alto riesgo son los siguientes:

5.6.1. Control y supervisión humana

La Propuesta de Reglamento establece en su artículo 7 que las tecnologías de inteligencia artificial de alto riesgo, incluidos los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías, deberán desarrollarse, desplegarse y utilizarse de forma que se garantice en todo momento una supervisión humana integral y que se pueda restablecer en todo momento el control humano cuando sea necesario, incluso mediante la alteración o la desactivación de dichas tecnologías.

El precepto se centra en la regulación de la supervisión y control humano como mecanismo de gobernanza, y no aborda todo lo que comporta el carácter antropocéntrico y antropogénico de la inteligencia artificial al que alude el texto de la resolución, el cual incluye dicha supervisión y control, entre otros aspectos.

Se trata de un principio ético básico que integra los marcos éticos internacionales más avanzados y consensuados sobre inteligencia artificial y que supone, tal y como está redactada, una supervisión y control humano integral, tanto en su desarrollo como en su posterior despliegue y utilización, y de manera previa, coetánea y posterior, de modo que se pueda restablecer el control humano en cualquier momento cuando sea necesario, incluso mediante el uso del denominado "botón rojo" o *kill switch*, es decir, con posibilidad de alterar o desactivar dichas tecnologías en todo momento.

La supervisión humana se hallaba incluida como uno de los requerimientos para el desarrollo de inteligencia artificial fiable incorporados en las *Directrices éticas para una inteligencia artificial fiable* del *Grupo de Expertos de Alto Nivel sobre inteligencia artificial -AI HLEG-*, así como uno de los requisitos de obligado cumplimiento para las aplicaciones del alto riesgo en el *Libro blanco sobre inteligencia artificial.*

Esta supervisión se presenta además como un mecanismo necesario de gobierno de los sistemas inteligentes.

La exigencia de esta supervisión y control abarca la fase de desarrollo como las posteriores de despliegue y uso, es decir, todo el ciclo de vida de un sistema inteligente de alto riesgo, de modo que constituye un requerimiento regulatorio de diseño que, de nuevo, evidencia el objetivo de dicha propuesta de no focalizarse en la utilización de la tecnología sino ya en su etapa de diseño y concepción, con el establecimiento de un marco regulador con origen y pretensión de implementar la *Ethics by design*.

La propuesta enlaza las fases precitadas con las funciones asociadas del ser humano en relación con la tecnología, esto es, desarrollador, desplegador o usuario. La participación humana se produce en todas estas fases y la entidad de la misma no tiene por qué ser más garantista de cumplimiento de requerimientos en función de la fase que se produzca.

La intervención humana se haya igualmente relacionada con los distintos grados de autonomía, aspecto este último de indudable relevancia a los efectos de atribución de la responsabilidad ética y jurídica, pero que, sorpresivamente, a diferencia de la esta propuesta objeto de análisis, la nueva Propuesta de Reglamento regulador de la inteligencia artificial del Parlamento Europeo y del Consejo, de 21 abril de 2021, se aleja de este concepto en su definición.

La propuesta objeto ahora de análisis no regula los particulares mecanismos de supervisión humana integral en particular, exigiendo que se garantice de manera integral en todo el ciclo de vida de la inteligencia artificial.

No obstante en el Considerando 10 de la Propuesta de Reglamento refiere que "las decisiones adoptadas por la inteligencia artificial, la robótica y las tecnologías conexas o basadas en ellas deben seguir siendo objeto de revisión, evaluación, intervención y control humanos significativos", es decir, que la supervisión, control y participación humana debe materializarse en cuatro acciones principales, esto es, la revisión, la evaluación, la intervención y el control de las decisiones, y de manera "significativa", concepto que no define y sobre el que me detendré, pero que entendemos asociado a su carácter relevante respecto de la efectividad de esta supervisión.

Además, este Considerando, conforme recoge posteriormente su artículo 7.2, incorpora algunas acciones más específicas propias de esta supervisión específica, en particular, que el "desplegador o el usuario pueda, como mínimo, proceder a su desconexión segura, alterar o desactivar su funcionamiento o volver a un estado anterior que restaure las funcionalidades seguras en los casos en que esté en peligro la conformidad con el Derecho de la Unión y los principios éticos y obligaciones jurídicas establecidos en el presente Reglamento", es decir, la precitada desconexión, modificación, desactivación de su funcionamiento o la restauración a punto seguro.

En relación con ello, coincido con Lacoz Moratinos[433] que podría ser relevante adicionar al artículo 7 de la propuesta el atributo "significativo", de manera que se requiera

[433] LACOZ MORATINOS, G. (2020). "Análisis de la propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas". Op.cit. P. 35.

garantizar una supervisión humana integral y significativa en todas sus fases, si bien, en cualquier caso y por razones de seguridad jurídica y congruencia, se debería desarrollar y recoger este concepto para su operatividad.

A diferencia del requerimiento general de garantizar en todo momento una supervisión humana integral, el apartado 2º del artículo 7 de la propuesta objeto de análisis, contempla distintos mecanismos de gobernanza específicos, quizás y siendo purista a nivel técnico, más que mecanismos, acciones específicas derivadas de ésta, esto es, la posibilidad restablecimiento del control humano en cualquier momento, lo que deberá ser considerado en su desarrollo y aplicable en su despliegue y uso: "Las tecnologías a que se refiere el apartado 1 se desarrollarán, desplegarán y utilizarán de forma que se pueda restablecer en todo momento el control humano cuando sea necesario, incluso mediante la alteración o la desactivación de dichas tecnologías".

El precepto contempla, conforme recoge igualmente el Considerando 10 precitado, que los sistemas inteligentes deberán disponer de herramientas y procesos asociados en su concepción que permitan su desactivación, su modificación y restablecimiento del control humano sobre la decisión o proceso.

El cumplimiento de estas exigencias, como he referido, debe garantizarse desde el diseño como en la supervisión y vigilancia posterior en tiempo real por sus operadores en las fases de despliegue y uso.

En este sentido, el precitado *Libro blanco sobre inteligencia artificial* precitado recoge distintos ejemplos muy clarificadores en relación con vehículos sin conductor: Un botón de desactivación en las situaciones que la persona determine que no es seguro o restricciones funcionales en su diseño de modo que deje de funcionar en determinadas situaciones o lo haga de manera distinta en contextos determinados como, por ejemplo, visibilidad reducida.

El cumplimiento de esta exigencia, a mi juicio y como he expuesto en anteriores capítulos, impide considerar la existencia de una inteligencia artificial con plena autonomía, libertad e independencia, en el caso de que la tecnología actual o futura lo permitiese, en la medida que se trata de una capacidad de procesamiento, decisión y actuación restringida y

supeditada en todo momento al ser humano, como no puede ser de otra manera por razones éticas, de seguridad y de cumplimiento regulativo, en caso de aprobación y entrada en vigor de estos marcos.

Ello debería impedir que algunas predicciones de la ciencia-ficción e incluso de científicos de la talla de Stephen Hawking, a las que he hecho referencia anteriormente, se cumplan en el futuro.

No obstante, como aspecto menos positivo, esta supervisión y control podría adicionar al funcionamiento, despliegue y uso de los sistemas de inteligencia artificial determinados factores asociados al ser humano, como la subjetividad, imparcialidad, riesgos de seguridad -la persona posiblemente sea el elemento más débil en la seguridad de un sistema de información o para la evitación del sesgo- e incluso sesgos, alejándose de la objetividad, automatismo e imparcialidad propias de aquellos.

Esta norma ética de carácter vinculante pasaría a constituir una exigencia en su propio diseño y concepción, en su despliegue y en su utilización, y por defecto.

5.6.2. Seguridad, transparencia, trazabilidad y otras exigencias

La Propuesta de Reglamento regula en su artículo 8 la exigencia de que los sistemas de inteligencia artificial, la robótica y las tecnologías conexas de alto riesgo garanticen la seguridad, la resiliencia, la fiabilidad operativa, la exactitud, la explicabilidad, la transparencia, la reproductibilidad, la trazabilidad, la auditabilidad y su consecuente conformidad regulatoria, tanto en su desarrollo, como en su despliegue y utilización. De nuevo, se incluyen también los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías.

Se trata de un conjunto de principios y normas éticas considerados esenciales para los principales marcos éticos más avanzados y consensuados a nivel internacional durante los últimos años y que, mediante esta norma, en caso de aprobación, resultarían vinculantes desde un punto de vista jurídico.

Por lo que se refiere a la seguridad, el texto propuesto exige que se garanticen la resiliencia y salvaguardas en los sistemas, que incluyan soluciones alternativas y medidas frente a los riesgos para la seguridad o la protección.

En particular, exige un nivel de seguridad adecuado al riesgo identificado, con la finalidad de evitar que se exploten vulnerabilidades que puedan tener un impacto en los bienes jurídicos protegidos, así como su resiliencia, entendida como la capacidad de sistema de seguir operando ante cualquier alteración, es decir, de soportar y recuperarse ante desastres o perturbaciones. Se trata de una regulación general de la seguridad exigida con el objetivo de exigir la adecuada al contexto y riesgos en la línea del RGPD.

Esta exigencia ética se verá complementada con los marcos reguladores vigentes y futuros en materia de seguridad y privacidad a nivel europeo, a los que he hecho referencia en el capítulo II.

En cuanto a la exigencia de fiabilidad, se pretende garantizar un funcionamiento correcto, un rendimiento fiable conforme a la expectativa razonable del usuario para alcanzar los objetivos y realizar las actividades para las que inicialmente fueron concebidos los sistemas, y además asegurando la reproductibilidad de las operaciones.

En relación con su exactitud y correcto funcionamiento, el texto propuesto lo exige respecto de la consecución de los objetivos y realización de las actividades propias de dichos sistemas y, en caso de no fuera posible garantizar la misma plenamente, el sistema deberá informar sobre la probabilidad de que se produzcan errores e inexactitudes. Esta exigencia se halla igualmente relacionada con la obligación de transparencia y consecuente de información y, entiendo, no tanto, en relación con la posible elusión de determinadas responsabilidades para el fabricante por haber informado previamente al usuario de los posibles errores o inexactitudes en su funcionamiento, a la que me referiré al abordar su responsabilidad en el capítulo V.

Por lo que se refiere a su explicabilidad, la regulación propuesta pretende asegurar que los procesos técnicos de captación, análisis, decisión y acción de los sistemas puedan ser revisados, lo que está en directa relación con otros principios y exigencias como su auditabilidad. Ambas exigencias éticas y, con su inclusión en este instrumento normativo,

jurídicas, están directamente relacionadas con la responsabilidad derivada de los daños causados por o mediante sistemas de inteligencia artificial.

Esta exigencia ética, en caso de aprobación, ser hallará complementada con lo regulado en el precitado RGPD que, como he comentado anteriormente, recoge expresamente el derecho a la explicabilidad en el contexto de tratamiento de datos personales y decisiones individuales automatizadas.

Asimismo, el informe precitado anteriormente por mi parte del *AI Now Institute* de la Universidad de Nueva York pone precisamente de relieve la discusión doctrinal internacional sobre la posible construcción de una explicabilidad sobre la base del RGPD.

En cuanto a transparencia y obligación de información sobre la condición de "artificial" y "no humana" del sistema, así como de sus capacidades, exactitud y limitaciones, se configura como una obligación no sólo frente a usuarios sino frente a desarrolladores y "*desplegadores*" de estos sistemas.

En caso de incumplimiento de las exigencias precedentes de seguridad, resiliencia, fiabilidad, exactitud, explicabilidad, reproductibilidad, transparencia e información, el precepto exige que se garantice en cualquier caso la posibilidad de desactivar temporalmente las funcionalidades afectadas por dicho incumplimiento y volver a un estado anterior, es decir, lo que en la propuesta ética del Parlamento Europeo en su Resolución de 16 de febrero de 2017, sobre normas de Derecho civil sobre robótica, denominaba reversibilidad, que igualmente se recoge en las *Directrices éticas para una IA fiable*.

Por lo que se refiere a transparencia y trazabilidad, el precepto exige la transparencia para posibilitar su revisión, esto es, que se documenten sus elementos, procesos y fases conforme a las normas más estrictas posibles y de modo que permitan su evaluación por las autoridades competentes conforme recoge la Propuesta de Reglamento.

La inteligencia artificial, la robótica y las tecnologías conexas de alto riesgo, incluidos los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías deberán desarrollarse, desplegarse y utilizarse de manera transparente y

rastreable, de modo que sus elementos, procesos y fases estén documentados con arreglo a las normas aplicables más estrictas posibles y que las autoridades nacionales de control puedan evaluar que dichas tecnologías cumplen las obligaciones establecidas en el futuro Reglamento.

Uno de los principales retos de la inteligencia artificial, como analicé en el apartado 4 del capítulo II de esta investigación, es su opacidad. Conforme a esta exigencia ética se acomete pues este reto, exigiendo que los sistemas deberán documentar sus operaciones y permitir su evaluación posterior. Asimismo, facilitaría la posterior depuración de las responsabilidades dimanantes de los daños causado por o mediante sistemas inteligentes (de alto riesgo, obviamente).

La transparencia, tal y como está regulada en la Propuesta de Reglamento, se circunscribiría a elementos, procesos y fases, lo que podría conllevar el acceso a algoritmos, aunque exclusivamente por parte de "autoridades competentes", lo que en cualquier caso podría afectar a derechos de propiedad intelectual/industrial -en caso de que estuvieren protegidos por dichos derechos conforme analizaré con detalle en el capítulo VIII de esta investigación- y a los secretos empresariales de sus titulares.

El problema práctico será la dificultad de una transparencia y explicabilidad plena, especialmente ante la complejidad de algunos sistemas inteligentes, y particularmente aquellos con capacidad de autoaprendizaje, por lo que es posible que ni tan siquiera su diseñador o fabricante pueda estar en condiciones de explicarlo. En este sentido corresponderá a éstos integrar mecanismos para garantizar de forma efectiva dicha transparencia y explicabilidad.

Actualmente hay distintos proyectos e iniciativas para la mejora de la transparencia, entre otros, la propuesta del *Institute of Electrical and Electronics Engineers* -IEEE por sus siglas en inglés-[434] para desarrollar estándares que identifiquen niveles medibles y verificables de transparencia en sistemas inteligentes para su evaluación.

[434] Recuperado de https://standards.ieee.org/industry-connections/ec/autonomous-systems.html. Consultado el 03.03.2021.

Asimismo, el precepto concreta la obligación tanto de desarrolladores, *desplegadores* como usuarios de garantizar la seguridad proactiva y la conformidad regulatoria en esta materia en relación con esta Propuesta de Reglamento, esto es, no sólo garantizar su conformidad con las características de seguridad previstas en la norma sino, además, poder demostrarlo, en definitiva, la seguridad proactiva, efectiva y gestionada, conforme es igualmente requerida por otros marcos europeos, por ejemplo, en materia de privacidad, por RGPD reiteradamente citado.

No obstante, debo significar que esta obligación se hace igualmente recaer en el usuario -además de desarrolladores y desplegadores- y ello, a mi juicio, dependerá del concepto "usuario" y quién ocupa dicha posición en el uso y aplicación de cada sistema de inteligencia artificial específico, si bien, partiendo de la propia definición de "usuario" del artículo 4.I.k) del propio Reglamento propuesto, su posición está inicialmente alejada de la posibilidad de garantizar dicha seguridad y cumplimiento. Más bien es ajena. No tendría sentido establecer dicha obligación de garante para el "usuario" tetrapléjico de un exoesqueleto inteligente de movilidad o la persona invidente que utiliza un sistema inteligente de visión artificial.

En relación con la transparencia, en caso de aprobación, esta exigencia se verá complementada con lo dispuesto en el RGPD en relación con el tratamiento de datos personales y la toma de decisiones automatizadas, como he referido anteriormente.

En relación con las exigencias éticas anteriores, se exige igualmente la reproductibilidad.

Por último, el texto propuesto exige la rendición de cuentas (accountability) y auditabilidad en directa relación con la transparencia y explicabilidad y, como he referido, con la responsabilidad jurídica dimanante de las decisiones y conductas de los sistemas inteligentes.

Su exigencia se hallaba ya prevista en las propuestas éticas previas que he analizado anteriormente y en el propio *Libro Blanco sobre la inteligencia artificial* de la Comisión Europea.

La propuesta europea objeto de análisis los recoge a nivel ético y como principio obligatorio, pero la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial, objeto de posterior análisis en esta investigación, los contempla en relación con la responsabilidad civil derivada de la inteligencia artificial, en especial, respecto de la necesidad de transparencia del sistema y la necesaria trazabilidad y auditabilidad del mismo, para posibilitar la acreditación de los daños y perjuicios en caso de daños y perjuicios causados por el sistema, cuando se exija la misma.

El precepto citado establece la obligación de desarrolladores, *desplegadores* y, de nuevo, usuarios, de velar que las medidas adoptadas para cumplir los requerimientos de seguridad exigidos en el mismo puedan ser auditadas por las autoridades nacionales de control previstas en la Propuesta de Reglamento o, en su caso, por otros órganos de control sectorial nacionales o europeo.

5.6.3. Ausencia de sesgo y de discriminación

El artículo 9 del Reglamento propuesto exige la imparcialidad, igualdad de trato y no discriminación de los programas informáticos, algoritmos y los datos utilizados o producidos (de entrada o de salida) por los sistemas de inteligencia artificial, la robótica y las tecnologías conexas desarrolladas, desplegadas o utilizadas en la UE.

La ausencia de sesgo y la no discriminación se consideran un requerimiento ético de obligado cumplimiento en la propuesta analizada, así como causa de daños y perjuicios así recogido expresamente en el artículo 4.I.n).

El sesgo es definido en el artículo 4.I.l) como toda percepción personal o social prejuiciosa de una persona o de un grupo de personas sobre la base de sus características personales. Por su parte, la discriminación es definida en el apartado m) del precepto citado como todo trato diferenciado de una persona o de un grupo de personas basado en un motivo que no tiene justificación objetiva o razonable alguna y que, por tanto, está prohibido por el Derecho de la Unión.

La cuestión radica en que la propia función de los sistemas inteligentes y, en particular, de los sistemas de aprendizaje automático, es discriminar entre los datos con los que se genera ese aprendizaje.

En este sentido, el *Libro blanco sobre inteligencia artificial* analizado exigía la necesidad de adoptar medidas razonables como la utilización de conjuntos de datos suficientemente significativos o la necesidad de conservar la documentación sobre metodologías de programación y entrenamiento.

Se prohíbe expresamente y de manera amplia la discriminación por motivos de raza, sexo, orientación sexual, embarazo, discapacidad, características físicas o genéticas, edad, minoría nacional, origen étnico o social, lengua, religión o creencias, opiniones políticas o participación cívica, nacionalidad, estado civil o económico, educación o antecedentes penales.

Sin duda, una lista exhaustiva, entiendo que con el ánimo de cubrir los supuestos y categorías afectadas por el sesgo y la discriminación, especialmente ante las características propias de los sistemas inteligentes y, en especial, de los sistemas de aprendizaje automático, ante el análisis de las inferencias obtenidas por los mismos de los datos y el contexto.

Como destacan algunos autores como Zuiderveen[435] y Lacoz Moratinos[436], los marcos actuales que contemplan normas para evitar la discriminación pueden no ser suficientes para evitar que se produzcan dichos efectos derivados de las inferencias algorítmicas, por lo que proponen nuevas normas que aborden los efectos discriminatorios particulares desde una perspectiva sectorial.

Únicamente se podrá justificar un trato diferenciado entre personas o grupos de personas cuando exista una finalidad objetiva, razonable y legítima que sea proporcionada y

[435] ZUIDERVEEN, F.J. (2020). "Strengthening Legal Protection against Discrimination by Algorithms and Artificial Intelligence". *The International Journal of Human Rights*. Vol. 24, N.10, 2020. P.15.
[436] LACOZ MORATINOS, G. (2020). "Análisis de la propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas". Op.cit. P. 37.

necesaria, es decir, cuando no exista otra alternativa con mejor impacto en el principio de igualdad de trato.

5.6.4. Responsabilidad social, igualdad de género y otros aspectos

La Propuesta de Reglamento regula en su artículo 10, bajo el título de "responsabilidad social e igualdad de género" toda una serie de obligaciones adicionales en materias como el respeto de la Ley, los derechos laborales, la educación, el derecho de información, la igualdad de género o la propiedad intelectual.

El precepto indicado establece la exigencia de que los sistemas de inteligencia artificial, la robótica y las tecnologías conexas de alto riesgo, se desarrollarán, desplegarán y utilizarán de conformidad con la legislación (legalidad), los principios y los valores de la UE, de manera que no interfieran en elecciones ni contribuyan a la desinformación (veracidad y no interferencia), que respeten los derechos de los trabajadores, que promuevan una educación de calidad y la alfabetización digital, no aumenten la brecha de género impidiendo la igualdad de oportunidades para todos y no vulneren los derechos de propiedad intelectual o cualquiera de sus limitaciones o excepciones. De nuevo, se incluyen también los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías.

5.6.5. Sostenibilidad medioambiental

El artículo 11 del Reglamento propuesto establece la exigencia de que los sistemas de inteligencia artificial, la robótica y las tecnologías conexas de alto riesgo deben ser sostenibles medioambientalmente y ser evaluados por las autoridades nacionales de control previstas en la Propuesta de Reglamento o, en su caso, por otros órganos de control sectorial o europeos, a los que además el precepto les impone la función/obligación de velar por la adopción de medidas para mitigar y remediar su impacto medioambiental, para garantizar el cumplimiento legal a nivel nacional o de la UE y los compromisos de protección medioambiental contraídos por ésta, en especial su

impacto en los recursos naturales, el consumo de energía, la producción de residuos, la huella de carbono, la emergencia climática y la degradación del medio ambiente.

De nuevo, se incluyen también los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías.

5.6.6. Tratamiento de datos biométricos con finalidad de identificación

La Propuesta de Reglamento regula en su artículo 12 el uso y la recogida de datos biométricos con fines de identificación remota en zonas públicas, entre otros, el reconocimiento fácil o biométrico, estableciendo que únicamente deben ser desplegados o utilizados por las autoridades públicas de los Estados miembros para fines de interés público esencial, dado los riesgos que suponen para los derechos fundamentales.

Se trata de un marco complementario y específico para este tipo tratamientos en relación con el marco general establecido por el Reglamento General de Protección de Datos y la Directiva 2002/58/CE[437].

El texto propuesto establece un conjunto de obligaciones cuyos destinatarios, en este caso no son los desarrolladores, *desplegadores* o usuarios, sino las autoridades de los Estados miembros. En particular, se exige que el tratamiento sea público, proporcionado, específico, restringido a una finalidad y ubicación concreta y limitado en el tiempo, considerando la dignidad, la autonomía humana y los derechos fundamentales, en especial, el respeto a la intimidad y a la protección.

En este sentido es evidente la preocupación del legislador europeo por los riesgos de vigilancia y monitorización.

[437] Directiva 2002/58 / CE del Parlamento Europeo y del Consejo, de 12 de julio de 2002, relativa al tratamiento de los datos personales ya la protección de la intimidad en el sector de las comunicaciones electrónicas (Directiva sobre la privacidad y las comunicaciones electrónicas). DO L 201 de 31.7.2002. Pp. 37-47.

Sin embargo, la Propuesta de Reglamento de 21 de abril de 2021, supone un cambio drástico en la regulación de esta aplicación, especialmente sustentada en el posicionamiento de las autoridades europeas de protección de datos, como expondré en su análisis en el capítulo IV, partiendo de su prohibición en determinados contextos por considerarlos de riesgo inaceptable.

5.6.7. Derecho de resarcimiento

La Propuesta de Reglamento regula en su capítulo II no sólo distintas obligaciones para algunas de las partes implicadas, sino también derechos, en particular, para las personas que sufran daños con origen en los sistemas de inteligencia artificial. La nueva Propuesta de Reglamento de 21 de abril de 2021, no lo contempla, como será objeto de análisis posterior en el capítulo IV.

El artículo 13 de la propuesta regula el derecho de resarcimiento por parte de toda persona física o jurídica por las lesiones o los daños sufridos causados por el desarrollo, despliegue y/o el uso de la inteligencia artificial, la robótica y las tecnologías conexas de alto riesgo, así como los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías, si bien, conforme su tenor literal matiza, en el contexto de infracciones de Derecho de la UE y las obligaciones establecidas en el Reglamento propuesto.

De aprobarse esta exigencia ética en los futuro marcos, se verá complementada por la Propuesta de Reglamento sobre responsabilidad civil de la inteligencia artificial que será objeto de análisis en el capítulo V de esta investigación.

5.6.8. Evaluación de riesgos

La Propuesta de Reglamento regula en su artículo 14 la obligación de someter los sistemas de inteligencia artificial, la robótica y las tecnologías conexas, incluyendo los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías, a

análisis de riesgos en base, principalmente, a su uso o finalidad específicos, el sector donde se desarrollen, desplieguen o utilicen, y la gravedad de los posibles daños o lesiones causados.

Esta evaluación se regula como previa, imparcial y externa por parte de las autoridades nacionales de control en base a criterios concretos y definidos, y sobre la base del listado de sectores de alto riesgo y de usos o fines de alto riesgo que recoja el futuro Reglamento en sus anexos.

Es decir, esta obligación recaerá en los sistemas incluidos en la lista de tecnologías de alto riesgo que definitivamente recoja el anexo del futuro Reglamento. Este listado deberá ser actualizado posteriormente por la Comisión mediante actos delegados, sin perjuicio de que la misma elabore una lista común de tecnologías de alto riesgo, que deberá actualizar periódicamente.

Si se determina que su desarrollo, despliegue o uso entrañan un riesgo significativo de causar las lesiones o daños previsibles conforme a dicho análisis a las personas o a la sociedad, vulnerando los derechos fundamentales y las normas de seguridad establecidas en el Derecho de la UE, se considerarán tecnologías de alto riesgo, conforme son definidas en dicha Propuesta de Reglamento, con las consecuencias correspondientes conforme al mismo.

La evaluación de riesgos se deberá llevar a cabo por las autoridades nacionales de control, bajo supervisión de la Comisión Europea o la entidad u organismos que pueda designarse con esta finalidad en el marco de la cooperación.

5.6.9. Evaluación de conformidad

El artículo 15 de la Propuesta de Reglamento establece una obligación adicional de someter a un análisis de cumplimiento o evaluación de conformidad de los sistemas de inteligencia artificial, la robótica y las tecnologías conexas, en particular, de las obligaciones previamente analizadas y reguladas en sus artículos 6 a 12.

Esta evaluación, así como su seguimiento, deberán ser llevados a cabo por las autoridades nacionales de control indicadas en el Reglamento propuesto. Sin embargo, el texto actual no indica las metodologías o directrices para llevar a cabo la misma, limitándose a indicar que se elaborarán directrices vinculantes sobre la metodología a seguir por las autoridades nacionales de control para su realización, como muy tarde, en la fecha de entrada en vigor del Reglamento propuesto.

Los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías de alto riesgo que hayan sido evaluadas y calificadas como conformes, se considerarán que cumplen igualmente las precitadas obligaciones, salvo que la autoridad nacional de control decida realizar una evaluación por iniciativa propia o a petición del desarrollador, desplegador o usuario.

5.6.10. Certificado europeo de conformidad ética

La Propuesta de Reglamento regula en su artículo 16 la exigencia de un certificado europeo de conformidad ética, obligatorio para los sistemas de alto riesgo y facultativo para el resto. En este sentido, como he anticipado, varios países europeos han creado ya sus propios certificados.

El certificado se emitirá cuando los sistemas hayan sido sometidos a la evaluación de conformidad precitada y la misma haya resultado positiva.

Según regula el apartado segundo del precepto citado, los desarrolladores, desplegadores o usuarios de sistemas de inteligencia artificial, robótica y tecnologías conexas que no se consideren de alto riesgo -incluyendo los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías-, en consecuencia, no sujetos a las obligaciones precedentes, podrán solicitar que se certifique el cumplimiento de las obligaciones precitadas o parte de ellas cuando así lo justifique la naturaleza de la tecnología en cuestión, conforme a las decisiones que adopten en este sentido las autoridades nacionales de control. Se contempla pues como un certificado facultativo para sistemas no considerados de alto riesgo.

En cualquier caso, al igual que para los sistemas de alto riesgo, este certificado sólo se emitirá cuando hayan sido sometidos a la evaluación de conformidad precitada por la misma y ésta haya resultado positiva.

El texto propuesto no contempla el procedimiento de solicitud a seguir, remitiéndose a su futura elaboración a instancias de la Comisión.

5.7. Supervisión institucional

El texto propuesto por el Parlamento Europeo regula en su Capítulo III la supervisión de las instituciones de la inteligencia artificial, especialmente a través de la gobernanza, la creación de autoridades de control y la coordinación, que deberá ser objeto de desarrollo.

5.7.1. Gobernanza de los sistemas inteligentes

La Propuesta recoge uno de los aspectos que considero más importante para asegurar la ética y la consecuente seguridad y confianza en los sistemas inteligentes y su desarrollo, que no es otro que la gobernanza.

Conforme recoge en su Considerando 39, la gobernanza basada en las normas pertinentes refuerza la seguridad y fomenta una mayor confianza de los ciudadanos en el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas.

En este sentido, la Propuesta refiere en su Considerando 41 algunas de las normas de gobernanza ya existentes, algunas tratadas en otros apartados de esta investigación, como las *Directrices éticas para una inteligencia artificial fiable*, elaboradas por el *Grupo de expertos de alto nivel sobre inteligencia artificial* creado por la Comisión Europea, así como otras normas técnicas como las adoptadas por el *Comité Europeo de Normalización* (CEN), el *Comité Europeo de Normalización Electrotécnica* (Cenelec) y el *Instituto Europeo de Normas de Telecomunicaciones* (ETSI) a escala europea, y por la

Organización Internacional de Normalización (ISO) y el *Instituto de Ingenieros Eléctricos y Electrónicos* (IEEE) a escala internacional.

La gobernanza permite alinear la tecnología (el medio) con los objetivos perseguidos (el fin), al objeto de asegurar su consecución, y esencial para asegurar el cumplimiento ético y regulativo y garantizar sistemas inteligentes y seguros. Ello evita la irresistible tendencia -en ocasiones- a considerar la tecnología como un fin y no como medio o instrumento para la consecución de los objetivos pretendidos.

Y no es la tecnología, sino su gobernanza, lo que permitirá conseguir los objetivos finales pretendidos con la aplicación y uso de los sistemas inteligentes.

La Propuesta prevé en su artículo 17 el establecimiento de normas de gobernanza por las autoridades nacionales de control que igualmente regula en su artículo 18, bajo la coordinación de la Comisión Europea o entidad u organismo que se designe, contando para ello con la consulta de las partes interesadas.

Estas normas de gobernanza, según el texto propuesto, deberán incluir directrices de gobernanza y aplicación no vinculantes sobre la metodología que deberán seguir los desarrolladores, *desplegadores* y usuarios para adecuarse al nuevo marco regulador, que deberán publicarse de manera previa o coetánea a su entrada en vigor. En este sentido, se trataría de *soft law* de acompañamiento a los nuevos marcos regulativos.

Asimismo, el precepto citado regula específicamente la gobernanza de los datos utilizados o producidos por los sistemas de inteligencia artificial en su desarrollo, despliegue o utilización, los cuales deberán ser gestionados por desarrolladores, *desplegadores* y usuarios bajo los correspondiente protocolos industriales y comerciales y, en cualquier caso, de conformidad con las reglas y normas de gobernanza nacionales, de la UE y otras organizaciones europeas e internacionales -cuando resulten aplicables-.

Con este propósito, el texto propuesto exige a los desarrolladores y *desplegadores*, en particular y siempre y cuando sea viable, la realización de controles de calidad de las fuentes externas de los datos utilizados por los sistemas de inteligencia artificial y el

establecimiento de mecanismos de supervisión para su recopilación, almacenamiento, tratamiento y uso.

El apartado 4º del artículo citado, recoge expresamente la sujeción de la recopilación, el almacenamiento, el tratamiento, el intercambio y el acceso a los datos utilizados o producidos por los sistemas de inteligencia artificial desarrollados, desplegados o utilizados en la UE a las reglas y normas de gobernanza tanto nacionales, de la UE, de otra organizaciones europeas e internacionales que resulten de aplicación, así como a los correspondientes protocolos industriales y comerciales.

Con este objetivo, el texto propuesto exige a los desarrolladores y *desplegadores* velar por la aplicación de dichos protocolos durante el desarrollo y despliegue de la inteligencia artificial, con la exigencia de que definan con claridad los requisitos para el tratamiento y la concesión de acceso a los datos utilizados o producidos por estos sistemas, así como la finalidad, ámbito de aplicación, destinatarios del tratamiento y la concesión de acceso a dichos datos, en congruencia con lo exigido por el Reglamento General de Protección de Datos, lo cuales además, deberán ser susceptibles de auditoría y trazabilidad en todo momento.

5.7.2. Creación de autoridades nacionales de control

La Propuesta de Reglamento propone la creación de una autoridad pública de control estatal para garantizar la aplicación del marco ético y jurídico establecido en el mismo, que coopere con las respectivas autoridades del resto de Estados, con otras autoridades, instituciones, organismos competentes en estos aspectos y con la propia Comisión Europea. Esta autoridad deberá ser independiente conforme prevé expresamente la propuesta.

Sin duda, constituye una novedosa y necesaria figura para asegurar el cumplimiento del marco propuesto, sin embargo, se le deberá dotar de los recursos, herramientas, financiación y facultades necesarias para poder ejercer sus funciones y competencias, cuando, de inicio, se le priva de una de las principales herramientas para asegurar los

objetivos pretendidos, que no la única, como lo es el establecimiento de un régimen sancionador en caso de incumplimiento, como ha evidenciado la aplicación de los nuevos marcos normativos europeo en materia de privacidad.

Las principales funciones de esta autoridad estatal se hallan reguladas en el artículo 18 y concordantes de la Propuesta de Reglamento y consistirían principalmente en la evaluación y supervisión de la conformidad de los sistemas de inteligencia artificial con los principios éticos y las obligaciones jurídicas del Reglamento propuesto.

Sus funciones incluirán el control de la aplicación coherente del mismo en toda la UE mediante su cooperación, realizar las evaluaciones de riesgos y de conformidad precitadas, así como la emisión de certificados europeos de conformidad con todo ello.

La autoridad de control será la primera línea en caso de incumplimiento de los principios éticos y obligaciones jurídicas establecidas en la Propuesta de Reglamento.

Asimismo, será responsable de supervisar la aplicación de las reglas y normas de gobernanza nacionales, europeas e internacionales, colaborando con el mayor número posible de partes interesadas y constituyéndose en un foro de intercambio con y entre las partes interesadas en el ámbito académico, civil, industrial y de investigación.

Y, por último, también deberá proporcionar orientación y apoyo profesional y administrativo a organizaciones de investigación y desarrollo, pequeñas y medianas empresas o *startups* sobre la aplicación general de la legislación de la UE aplicable a la inteligencia artificial, la robótica y las tecnologías conexas, y sobre los principios éticos establecidos en el Reglamento propuesto objeto de análisis.

En definitiva, sus funciones principales serán las de control proactivo inicial, preventivo y detectivo de la aplicación de los principios éticos y cumplimiento de las obligaciones jurídicas, así como de evaluación y supervisión posterior, previa denuncia o actuación inspectora, junto con la emisión del certificado europeo de conformidad de los sistemas con los principios éticos y obligaciones jurídicas exigibles conforme a lo previsto en la Propuesta de Reglamento analizada.

Además de sus funciones proactivas, también desempeñaría funciones reactivas para la emisión de certificados a petición de cualquier desarrollador, operador o usuario para sistemas que no sean calificados de alto riesgo o de evaluación y supervisión en caso de denuncia previa de incumplimiento.

Sobre este aspecto, en mi opinión, considero que las facultades de esta autoridad de control merecen una reflexión más profunda, al objeto de considerar si debería tener unas atribuciones no sólo de policía sino más de difusión y de concienciación, que podría incluir la elaboración de guías, directrices, estándares generales o sectoriales o informes, aproximándose al papel que desempeñan en materia de protección de datos las autoridades nacionales, y que han demostrado su eficacia, entre otras, en sus labores de difusión y concienciación, así como de interpretación e integración.

5.7.3. Cooperación adicional de los Estados miembros

El texto propuesto exige cooperación adicional a los Estados miembros para garantizar el cumplimiento de los principios éticos y normas jurídicas establecidos en el Reglamento propuesto objeto de análisis.

En su artículo 18.7 exige que los Estados miembros deberán, de un lado, adoptar todas las medidas necesarias para garantizar la aplicación de los principios éticos y normas jurídicas establecidas en el Reglamento propuesto y, de otro, apoyar a las partes interesadas y a la sociedad civil para propiciar una respuesta ética adecuada a los nuevos retos y oportunidades que suponen los sistemas de inteligencia artificial, especialmente de carácter transfronterizo.

5.8. Infracciones

El Reglamento propuesto no prevé un régimen de infracciones y sanciones asociado a posibles incumplimientos ni un procedimiento relacionado, como considero sería deseable como incentivo adicional para su cumplimiento, si bien, establece en su artículo

19, la aplicación de la Directiva (UE) 2019/1937, de Parlamento Europeo y del Consejo[438], a las denuncias de infracción del mismo que se produzcan y para la protección de las personas denunciantes, proponiendo en su artículo 22 sendas enmiendas al artículo 2 y anexo de la Directiva precitada, incorporando el desarrollo, despliegue y uso de la inteligencia artificial, la robótica y las tecnologías conexas en su ámbito de aplicación material.

5.9. Otras cuestiones adicionales.

La propuesta del Parlamento Europeo también establece una serie de tareas que deberá llevar a cabo la Comisión Europea, en particular, la elaboración y posterior actualización de una lista común de las denominadas "tecnologías de alto riesgo".

Y, por último, la Propuesta de Reglamento establece que la Comisión Europea deberá examinar periódicamente el desarrollo de la inteligencia artificial, la robótica y las tecnologías conexas, incluidos los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías, para presentar un informe trienal sobre la aplicación del Reglamento propuesto, "*incluida una evaluación sobre la posible modificación del ámbito de aplicación del mismo*".

Es decir, la propuesta se ha construido, en mi opinión, sobre ese concepto adaptativo, evolutivo y "*responsive*" a la realidad social y tecnológica, necesario para dar soluciones, aportar eficacia y dar respuesta a los problemas de hoy o mañana, garantizando un marco ético y jurídico de toda esta tecnología que responda a sus riesgos y retos cambiantes y en constante aumento cualitativo y cuantitativo, especialmente ante su asociación con otras tecnologías.

[438] Directiva (UE) 2019/1937 del Parlamento Europeo y del Consejo, de 23 de octubre de 2019, relativa a la protección de las personas que informen sobre infracciones del Derecho de la Unión (DO L 305 de 26.11.2019. P. 17).

5.10. Recapitulación

La Propuesta de Reglamento contempla un conjunto de principios y normas consideradas esenciales para los principales marcos éticos elaborados a nivel internacional durante los últimos años para su consenso y que, mediante esta norma, en caso de aprobación, resultaría vinculantes desde un punto de vista jurídico.

La propuesta incluye una serie de principios y normas éticas generales exigibles a cualquier sistema de inteligencia artificial, robótica y tecnologías conexas -siempre y cuando se revise el concepto de sistema inteligente y autonomía asociada que incorpora la propuesta, en los términos analizados-, así como a los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías, así como otros específicos únicamente exigibles legalmente a los clasificados como de alto riesgo.

A lo largo del mismo se abordan conceptos como "alto riesgo", "supervisión humana", "sesgos" o "discriminación" que constituyen la base de los principales retos reguladores que plantea la inteligencia artificial en todas sus fases, tanto en su diseño como en su despliegue y uso.

En la tabla que incorporo a continuación, resumo los principales requerimientos éticos y jurídicos expresamente previstos en la propuesta analizada y exigibles a la inteligencia artificial:

Marco ético y jurídico	Sistemas de inteligencia artificial	
Requerimiento	General	Alto riesgo
Centrada en el ser humano	*	✓
Origen en el ser humano	*	✓
Supervisión y control humano	*	✓
Conformidad regulatoria (UE)	✓	✓
Respeto de la dignidad, autonomía y seguridad humana	✓	✓
Respeto de los demás derechos fundamentales	✓	✓
Protección de datos personales y cumplimiento normativo en esta materia	✓	✓
Seguridad proactiva	*	✓
Resiliencia	*	✓
Fiabilidad operativa	*	✓
Exactitud	*	✓
Explicabilidad	*	✓

Transparencia-Información	*	✓
Reproductibilidad	*	✓
Trazabilidad	*	✓
Auditabilidad-Rendición de cuentas	*	✓
Conformidad regulatoria	*	✓
Ausencia de sesgo y discriminación	*	✓
Responsabilidad social	*	✓
Legalidad	*	✓
Moralidad (principios y valores UE)	*	✓
No interferencia o distorsión informativa	*	✓
Respecto derechos de los trabajadores	*	✓
Promoción de la educación de calidad	*	✓
Promoción de la alfabetización digital	*	✓
Respeto propiedad intelectual	*	✓
Sostenibilidad medioambiental	*	✓
Intimidad y protección de datos	*	✓
Derecho de resarcimiento	*	✓
Evaluación de riesgos	*	✓
Evaluación de conformidad	*	✓
Certificado europeo de conformidad ética	*	✓

(*) Facultativos

Fuente: Elaboración propia.

Además de los precitados, como he comentado anteriormente, la Propuesta de Reglamento establece el deber de la UE y los Estados miembros de fomentar los proyectos de investigación dirigidos a ofrecer soluciones de inteligencia artificial que promuevan la inclusión social, la democracia, la pluralidad, la solidaridad, la equidad, la igualdad y la cooperación.

Sin perjuicio de las consideraciones particulares que he realizado al analizar cada uno de los aspectos abordados por el texto propuesto, debo significar algunas cuestiones para la reflexión futura en relación con el mismo.

En primer lugar, la necesidad de que se acompañe de un régimen sancionador como herramienta de actuación para las autoridades de control en caso de incumplimiento de sus disposiciones y contenido obligacional imperativo, sin perjuicio de que en paralelo se les dote del estatuto y recursos adecuados y necesarios para desarrollar sus funciones y competencias, y todo ello con el objetivo de establecer mecanismos de gobernanza y control efectivos y, a su vez, velar por un cumplimiento efectivo de los nuevos marcos regulatorios.

En segundo lugar, sería deseable la concreción del concepto de "riesgo significativo" conforme ha sido analizado anteriormente por mi parte, adicionar los criterios para determinar el mismo y que se abordan, al menos parcialmente en los Considerandos, de modo que se proporcione seguridad jurídica.

En tercer lugar, respecto a la "supervisión humana integral" en todo momento y en todas sus fases en los casos de sistemas de alto riesgo, debería adicionar "relevante" o "significativa", con un mayor detalle qué alcance debería tener dicha supervisión en los términos analizados, con el objetivo de proporcionar seguridad jurídica.

En cuarto lugar, la inclusión expresa de los requerimientos como el control y la supervisión humana en cualquier sistema de inteligencia artificial, cualquiera que sea su nivel de riesgo inicial, dado que su ausencia podría determinar de inicio su propia clasificación en altos niveles de riesgo. Del mismo modo, incluir de manera expresa para cualquier sistema inteligente la exigencia de que sea una inteligencia artificial centrada y con origen en el ser humano.

En quinto lugar, la revisión de la definición de sistema inteligente y de autonomía en los términos expuestos para evitar equívocos respecto de su aplicación a cualquier sistema inteligente.

En sexto lugar, la posible conveniencia de un análisis de riesgos preceptivo de cualquier sistema inteligente antes de su puesta en el mercado para determinar el nivel de riesgo y, en su caso, su posible sujeción a los requerimientos y obligaciones jurídicas establecidas por los futuros marcos reguladores, incluso la exigencia de una autodeclaración responsable que pueda ser requerida por la autoridad competente que se determine en los mismos.

Por último, algunos autores como el precitado Lacoz Moratinos, consideran que la definición de sesgo resulta muy limitada, proponiendo una definición que integre la dimensión ética y estadística de los sesgos, tomando como referencia a Hildebrandt[439],

[439] LACOZ MORATINOS, G. "Análisis de la propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas". Op.cit. P.40.

considerando sesgo toda desviación o representación estadística que reconfigura la distribución de los bienes, servicios, riesgos y oportunidades de una o varias personas físicas o jurídicas.

6. Ética y responsabilidad

Uno de los principales objetivos de esta investigación es reflexionar sobre los retos y riesgos de la inteligencia artificial desde una perspectiva ética, jurídica y de seguridad, para abordar desde la globalidad, entre otros aspectos, las consecuencias derivadas de la materialización de aquellos en el orden civil principalmente y, en especial, en relación con la responsabilidad por daños causados o derivados del funcionamiento o uso de los sistemas inteligentes, con una correlativa reflexión sobre la adecuación de los marcos actuales para dar soluciones efectivas en las distintas situaciones y contextos y el análisis de los nuevos marcos en el ámbito europeo.

Como he evidenciado durante el análisis realizado hasta este momento y se evidenciará en los posteriores, existe una relación directa entre todas estas dimensiones entre sí desde las que se ha abordado esta investigación. Entre la ética y la responsabilidad, la ética y la seguridad, la ética y el cumplimiento regulativo, así como entre la seguridad y el cumplimiento legal, y entre la seguridad y la responsabilidad.

Las normas éticas que conforman la mayoría de los marcos éticos internacionales así como las que incorpora la Propuesta de Reglamento del Parlamento Europeo y del Consejo, sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, exigen la seguridad, la privacidad, el control y la supervisión humana, el cumplimiento regulativo, la rendición de cuentas, la trazabilidad, la explicabilidad, la auditabilidad y la responsabilidad, como principios y normas éticas esenciales que deben respetar los sistemas inteligentes en su diseño, desarrollo, funcionamiento y uso.

En la medida que una propuesta de estas características sea aprobada en el futuro bajo su actual tenor literal y entre en vigor, estas normas éticas serán vinculantes y exigibles

jurídicamente, con independencia de su regulación adicional o complementaria en otras propuestas de tramitación simultánea o que se promuevan en lo sucesivo.

Desde la dimensión ética, los sistemas de inteligencia artificial deben debe ser seguros y hallarse sujetos a control y supervisión humana, y deberían serlo con independencia de su nivel de riesgo inicial conforme he referido.

Asimismo, estos marcos éticos, con independencia de que finalmente se conviertan en un requerimiento regulador, conforme han ido consensuándose a nivel internacional o regional y han sido promovidos por la industria, ayudan a conformar la denominada "diligencia debida" y su eficacia jurídica en algunos de los contextos y regímenes de responsabilidad contractual y extracontractual vigentes, especialmente en aquellos basados en la culpa, y la consideración de la ausencia de seguridad en sí misma constituiría un defecto en su producción.

La responsabilidad jurídica en el ámbito civil de la inteligencia artificial se tratará en el en el capítulo V de esta investigación, si bien, considero necesario expresar algunas sucintas reflexiones adicionales sobre la relación entre la responsabilidad ética y la jurídica.

Durante los apartados precedentes he abordado algunos de los retos y riesgos que plantea la inteligencia artificial a nivel ético y las responsabilidades a este nivel que se pueden derivar en atención a su tipología, capacidades, contexto, finalidad y aplicación.

La normalización o, más bien, la normativización de los principios y normas éticas que supondría una propuesta como la analizada anteriormente, aportará seguridad y transparencia a todas las partes implicadas, especialmente para desarrolladores, *desplegadores*, operadores y usuarios de los sistemas de inteligencia artificial.

Estos principios y normas ya contribuyen y contribuirán en mayor medida a conformar la precitada "diligencia debida" que comentaba a los efectos de la aplicación de determinados regímenes de responsabilidad.

No obstante, por lo que se refiere a la imputabilidad de una conducta y la relación causal con su resultado puede diferir en el ámbito ético y jurídico y, dentro de éste, si estamos en el ámbito civil, penal, administrativo o en otras disciplinas del Derecho.

Si analizamos algunos de los riesgos comúnmente asociados a los sistemas de inteligencia artificial y sus usos, considero que ayudarán mejor a reflejar esa relación entre responsabilidad ética y jurídica.

Si nos focalizamos en el contexto de sistemas dotados de inteligencia artificial que interaccionan con personas y su entorno, por ejemplo, en servicios asistenciales o cuidado de personas, en caso de error o toma de una decisión autónoma de la que se deriven perjuicios para la persona, ¿quién respondería de los daños?

Los requerimientos de "rendición de cuentas" y "resarcimiento" nos deberían dar la solución.

Inicialmente, desde un punto de vista ético, parece que cualquier daño causado por un sistema inteligente sería responsabilidad de las personas encargadas de su diseño y programación. El problema se plantea conforme se aumenta su autonomía y sus capacidades de autoaprendizaje y de decisión y, en paralelo, su supuesta y relativa impredecibilidad. Y más en casos de formación y entrenamiento posterior.

En estos contextos se empieza a producir una disociación entre el responsable ético y el jurídico, especialmente ante regímenes *cuasi objetivos*, objetivos o absolutos de responsabilidad.

A modo de ejemplo, si un dron o un vehículo autónomo decide colisionar contra un edificio, en el primero de los casos, para evitar su colisión con un helicóptero comercial y, en el segundo, para evitar un accidente mortal, ¿quién es responsable de los daños producidos al propietario del edificio? ¿el conductor del vehículo/piloto que participe en la acción activa o pasivamente, el fabricante, el diseñador/desarrollador, el propio vehículo/dron?

Si hablamos de vehículos autónomos o, como decía, "automáticos", sin duda, contribuirán en el futuro a reducir los accidentes de tráfico, pero lo cierto es que la

seguridad de pasajeros y peatones depende del funcionamiento del sistema informático que controla el vehículo. De hecho, ya se han producido fallos en estos vehículos que han ocasionado la muerte a varias personas[440]. Y, es más, el diseño de sus sistemas integra algoritmos capaces de establecer prioridades entre distintos escenarios, incluso de extrema gravedad o siniestro irremediable, en los que el algoritmo tiene que escoger qué vida posee más valor y debe priorizarse[441]. Resultan obvios los conflictos éticos que se plantean.

La conducción en situaciones reales no controladas puede estar determinada por factores imprevistos y en función del contexto muy difíciles de prever y, en consecuencia, difícil de incorporar en los datos de entrenamiento de un algoritmo.

Algunos de los accidentes acaecidos con este tipo de vehículos se han producido por situaciones no previstas en el entrenamiento del algoritmo como, por ejemplo, el atropello de peatones que cruzaban la vía por zonas indebidas, defectuosa interpretación de los datos de entrada o fallos en sensores por condiciones climatológicas.

Los defectos u omisiones en un algoritmo o sistema pueden determinar consecuencias desastrosas para los derechos e intereses de la persona, incluso poner en peligro su vida, y no sólo en sistemas dotados de inteligencia artificial más avanzada o fuerte como se expone a lo largo de esta investigación.

En relación con los vehículos autónomos -no meramente automáticos- que sean dotados de inteligencia artificial y relativa autonomía, a la vista de los dilemas analizados precedentemente, las preguntas que me surgen inmediatamente son: ¿Se pueden valorar dos vidas humanas de manera distinta?, ¿en base a qué criterios se efectúa dicha valoración y de quién?, ¿cómo debería ser programado el algoritmo?, ¿es ético y moral poner en circulación vehículos con inteligencia avanzada o fuerte?, incluso ¿se debería permitir coches autónomos sin supervisión o intervención humana y con una inteligencia artificial más débil? ¿se va a permitir en el futuro cualquier innovación disruptiva

[440] WAKABAYASHI, D. (2018). "Self-driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam". *The New York Times*. Recuperado de: www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html. Consultado el 4.02.2021.
[441] BONNEFON, J.F., SHARIF, A. Y RAHWAN, I. (2016). "The social dilemma of autonomous Vehicles". *Science*. 24.06.2016. Vol. 352, N. 6293. Pp. 1573-1576. Doi: 10.1126/science.aaf2654

cualquiera que sea el grado de valor que aporte al ser humano? Reflexiones apasionantes que me hacen cuestionar muchos aspectos, pero que no puedo abordar en el marco específico de esta investigación, si bien, concluyo con mi opinión:

La ética es más importante que nunca en nuestra sociedad digital y la innovación tecnológica debe basarse en la ética en congruencia con la finalidad última de la tecnología, que es satisfacer necesidades humanas, resolver problemas, mejorar nuestra vida y nuestro mundo.

Por más que los expertos prosigan en sus intentos de emular completamente la mente humana, no es posible todavía con el estado de la tecnología actual, por lo que un sistema de inteligencia artificial avanzado, dotado de un cierto grado de autonomía, con capacidad de autoaprendizaje y con las mayores capacidades razonamiento lógico, no tiene consciencia, carece de emociones, de estados de ánimo, no tiene sentimientos aunque los simulemos -no siente amor, odio, tristeza, simpatía, orgullo, envidia o apatía-, no puede sentir ni discernir entre lo correcto o incorrecto, entre lo bueno y lo malo, más allá de las instrucciones incorporadas en su diseño o adquiridas durante su entrenamiento y funcionamiento mediante capacidades de autoaprendizaje.

En consecuencia, conforme al desarrollo tecnológico actual y marcos jurídicos vigentes en materia de responsabilidad, un sistema de inteligencia artificial no puede ser imputable ni puede ser responsable, ni ética ni jurídicamente, ni en el ámbito civil ni en el penal. Tampoco tiene personalidad jurídica o moral.

Asimismo, durante los próximos años, salvo que se encuentren y articulen las soluciones normativas adecuadas, nos encontraremos en muchas situaciones donde resultará extremadamente complicado imputar la responsabilidad a los sujetos que realmente deberían tener o tengan un control sobre sus riesgos durante su ciclo de vida, y atribuir los daños y perjuicios causados, y ello debido también a la falta de explicación de la inteligencia artificial y de sus decisiones, muchas de las cuales pueden resultar impredecibles o inexplicables para los usuarios afectados por la misma y para cualquier ser humano en general, especialmente en caso de ausencia de límites o restricciones en su concepción que puedan limitar cualquier nivel de autonomía, impredecibilidad y demás capacidades iniciales o adquiridas por el sistema.

De ahí la necesidad de los marcos regulativos propuestos por UE, de que la transparencia y la explicabilidad integre y constituya un requerimiento ético, como ya lo es, en la mayoría de los distintos marcos éticos elaborados hasta la fecha a nivel internacional, y de que incluso algunas normas jurídicas ya hoy exijan esa explicabilidad como el Reglamento General de Protección de Datos europeo. Este es uno de los principales retos que plantea la inteligencia artificial y su denominada tecnología de "caja negra", especialmente significativo ante el denominado *Deep learning* o aprendizaje profundo.

El ser humano es a quién corresponde decidir en manos de quién deja la toma de decisiones y sobre qué desde la razón.

La cuestión es si se dejará llevar por una inercia irracional para dejar en manos de sistemas la toma de decisiones con tan alto impacto. En la cultura occidental son las empresas y las organizaciones las que están decidiendo qué decisiones quiere "automatizar" o no. En la cultura oriental, muchos gobiernos son los que deciden qué "automatizar".

En China el gobierno favorece la seguridad nacional, la armonía y la estabilidad social, el control del ciberespacio y el impacto en el crecimiento económico muy por encima de los derechos y libertades individuales.

En la UE, se pretende encontrar el difícil equilibrio entre la irrenunciable y prioritaria seguridad y confiabilidad de los sistemas y el respeto de los derechos fundamentales, con el fomento de la innovación, la inversión, el desarrollo tecnológico y la competitividad internacional.

7. Ética en el diseño

La ética en el diseño o *Ethics by design* es un requerimiento ineludible conforme a los principios y normas éticas esenciales que deben integrar los sistemas de inteligencia artificial, que integran los distintos marcos éticos objeto de análisis en este capítulo.

La existencia de unos principios y normas éticas contribuyen a garantizar y potenciar los beneficios de la inteligencia artificial, a minimizar sus riesgos y a promover un entorno de confianza y seguridad, positivo para todas las partes implicadas, en el que se debe garantizar un equilibrio, en constante tensión, entre la innovación y la competitividad empresarial, y la seguridad y confianza de los ciudadanos en el uso de la tecnología, lo que, a su vez, es un motor para la inversión y desarrollo de la tecnología.

Además, la ética es la base sobre la que se deben construir los marcos reguladores de la inteligencia artificial ante los retos y riesgos que plantea una realidad tan compleja para el ser humano y las nuevas relaciones que plantea entre hombre-máquina, máxime ante su enorme impacto en los principales bienes y derechos de la humanidad.

La reflexión consecuente a realizar sería cómo incorporar dichos principios y normas éticas asociadas a la inteligencia artificial, en el diseño y programación de los sistemas dotados de la misma, de modo que se garantice que su uso y comportamiento será respetuoso con dichos requerimientos éticos y seguro en la totalidad de contextos en los que pueda operar.

El enfoque tiene que ser global con el objetivo de construir una solución igualmente global que integre los requerimientos y aspectos organizativos, técnicos y no técnicos, a considerar en el diseño y programación de su arquitectura y funcionamiento, con el objetivo de que los sistemas se comporten de manera adecuada.

Desde el punto de vista organizativo y técnico, existen distintos métodos para alcanzar estos objetivos de manera eficaz y para asegurar que su diseño y programación esté alineado con los mismos, si bien, a mi juicio, la mejor manera de garantizarlo es basar el diseño y construcción de los sistemas inteligentes en los principios y normas éticas esenciales que he analizado, esto es, en la *Ethics by design*, *Security by design*, *Privacy by design* y *Compliance by design*, junto con la regulación jurídica y el precitado *soft law*.

Y aunque ya forma parte de estos principios precitados, es esencial incorporar la transparencia y la explicabilidad -*Explainable AI o XAI*[442]- en el diseño y programación

[442] MURDOCH, W. J., SINGH, CH., KUMBLER, K., ABBASI-ASI, R Y YU, B. (2019). "Interpretable machine learning: definitions, methods, and applications". Publicado en *PNAS*. [arXiv:1901.04592]. 2019

de sistemas, mostrando de forma clara su funcionamiento y razonamiento interno evitando la opacidad, de especial importancia en el denominado aprendizaje profundo o *Deep Learning*.

Asimismo, los sistemas deberían estar sometidos a procesos de aseguramiento de la calidad, verificación, prueba y validación, sin perjuicio del análisis de riesgos previo, coetáneo y posterior a su desarrollo, despliegue y uso, y su evaluación de conformidad. Análisis y evaluación de riesgos que deberían ser previas y preceptivas respecto de cualquier sistema inteligente a poner en el mercado, especialmente, en el caso de que por sus capacidades y características pueda impactar en aspectos tan esenciales y valiosos como la vida, la salud o los derechos fundamentales.

La explicabilidad en la inteligencia artificial constituye uno de sus principales retos directamente relacionado con la responsabilidad ética y jurídica.

Las decisiones y conductas que puedan impactar en los bienes y derechos de las personas deben ser explicables y conocer de qué manera se han adoptado, de un lado, para garantizar la transparencia, trazabilidad, rendición de cuentas, responsabilidad y seguridad -física, lógica y, sobre todo, jurídica- y, de otro, proporcionar la confianza en los sistemas inteligentes.

En relación con la explicabilidad de las conclusiones y decisiones de los sistemas inteligentes, el primer responsable debería ser su diseñador.

Los resultados de los sistemas que integran aprendizaje automático suelen ser perfectamente inteligibles constituyendo “cajas blancas”, sin embargo, los sistemas más potentes basados en *Deep Learning* son difíciles de entender en virtud de la opacidad de su proceso de desarrollo y toma de decisiones, también llamados *Black Box*.

De hecho, aunque los diseñadores prefieren usar el *Deep Learning* por sus prestaciones, son cada vez más las empresas que apuestan por usar algoritmos interpretables de “caja blanca” si sus prestaciones son suficientemente adecuadas para resolver problemas de negocio, facilitando así su explicabilidad.

No obstante, actualmente son incesantes los esfuerzos e investigaciones por mejorar la explicabilidad del *Deep Learning*[443].

Por el contrario, en los sistemas de inteligencia artificial basados en conocimiento no se producen estos problemas, en la medida que usan conocimiento explícito y reglas para llegar a sus conclusiones.

La Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica[444], incorporaba en su anexo una *Carta sobre Robótica* como código de conducta ética en el campo de la robótica, que integra un modelo licencia para los diseñadores de robots o sistemas dotados de inteligencia artificial avanzada, consistentes en un conjunto de principios y normas que los diseñadores deberían considerar y que contemplaba la denominada "*Ethics by design*", la "*Security by design*" y la "*Privacy by design*". Estos aspectos han sido analizados en el apartado 3.2. de este capítulo.

No obstante, el diseño ético de algoritmos no es un método infalible ni la adopción de patrones éticos mediante la observación del comportamiento humano no garantiza que estos patrones sean los más adecuados, sino probablemente los más comunes y habituales, por lo que pueden plantease conflictos en situaciones concretas.

Por todo ello, es necesario recurrir a mecanismos no técnicos para conseguir los objetivos pretendidos de garantizar sistemas de inteligencia artificial seguros y fiables, y a ello contribuye especialmente, entre otros, la aprobación de marcos reguladores como el propuesto por el Parlamento Europeo analizado en los apartados precedentes, a través de su Resolución de 20 de octubre de 2020 sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas[445] o en la nueva Propuesta de Reglamento de 21 de abril de 2021, con matices, de modo que se establezca el carácter

[443] BARREDO, A., DÍAZ-RODRÍGUEZ. N. ET AL. (2017). "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsable AI". *Information Fusion*. ScienceDirect. Vol. 58. Elsevier. 2020. Pp. 82-115.

[444] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))

[445] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))

vinculante de estos principios y normas éticas, que se conviertan en obligaciones, principalmente para diseñadores, desarrolladores y fabricantes, que regulen las obligaciones jurídicas asociadas a las mismas, que definan los marcos de seguridad y funcionamiento, de su evaluación, su certificación, así como sus marcos de gobernanza y control, promoción de la autorregulación, fomento de la investigación, información, concienciación social y formación.

En este sentido la reciente Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, establece un marco de requisitos y obligaciones para los sistemas inteligentes, si bien, exclusivamente para los considerados de alto riesgo, a excepción de distintas obligaciones de transparencia para determinados sistemas, cuya proyección va más allá de su puesta en funcionamiento y uso, en la medida que son exigibles ya desde su diseño y durante su ciclo de vida, conforme abordaré en el próximo capítulo.

8. Consideraciones finales

La inteligencia artificial aporta numerosos beneficios a la sociedad en general y a gobiernos, Administraciones públicas, empresas y ciudadanos en particular, y en sectores como la salud, la educación, la industria, las finanzas o el transporte.

No obstante, su potencial interactivo a nivel físico y digital, sus riesgos éticos, jurídicos y de seguridad física, lógica, jurídica y política -y tanto para la integridad física, psicológica e intelectual de las personas, sus bienes y derechos, como frente a las empresas, infraestructuras críticas, Administraciones públicas y gobiernos-, su creciente autonomía/automatización, el incremento de sus capacidades de autoaprendizaje y procesamiento, su interacción con otras tecnologías, la mayor disponibilidad de datos, su relativa impredecibilidad y su ritmo vertiginoso de desarrollo, evolución, despliegue y aplicación actual -en constante aumento-, exigen adoptar una actitud extremadamente prudente y realizar una análisis reflexivo sobre la misma, su complejidad y sus objetivos

mediatos e inmediatos, al objeto de establecer unas bases sólidas para la construcción de su regulación.

Y esa reflexión debe ser estratégica y multidisciplinar, científica, empresarial, gubernamental, ética y jurídica para definir cuáles son nuestros objetivos como seres humanos y qué mundo queremos, así como qué inteligencia artificial queremos como medio para la consecución de todo ello y no acometerla como un mero fin.

Conseguirlo precisa de la compresión de una realidad tan compleja y desde un enfoque multidisciplinar, empezando por la ética, para definir la inteligencia artificial que queremos como seres humanos.

La construcción de los futuros marcos jurídicos sobre los éticos contribuirá a que el uso de la inteligencia artificial acompañe y apoye los valores de la sociedad y opere en beneficio de la inclusión y el bien común.

Desde un punto de vista ético, debemos valorar si estamos actuando estratégicamente en la línea correcta desde la moral y la ética o, simplemente, nos hemos dejado llevar, actuando sin objetivos mediatos claros por las estrategias descarnadas de las grandes potencias mundiales y gigantes tecnológicos, para sumarnos a una carrera por liderar y dominar la inteligencia artificial, los datos y cualquier tecnología asociada a su gobierno y, en definitiva, el nuevo orden mundial en el ciberespacio.

A mi juicio, la humanidad y su inteligencia natural debería comprender que los grandes desafíos son comunes y que las estrategias deberían igualmente serlo, adoptando las decisiones y llevando a cabo las acciones necesarias para su aplicación con la finalidad de satisfacer necesidades, resolver problemas y mejorar nuestra vida y el mundo que vivimos gracias a los avances tecnológicos. ¿Cuál es sino es el objetivo de estos? ¿El poder y el dinero?

El hombre no puede ser un espectador pasivo de la historia de la inteligencia artificial, sino que es quién debe escribirla como mero medio o instrumento para conseguir sus objetivos finales.

La ética debe formar parte de los sistemas de inteligencia artificial desde su diseño y concepción y durante todo su ciclo de vida, mediante la elaboración de principios y normas éticas que rijan la misma, así como incluso la acotación de los ámbitos en los que pueda utilizarse de forma segura y fiable y sus límites.

La responsabilidad, la seguridad y la prudencia en su diseño, desarrollo, despliegue y uso debería ser algo inherente a la inteligencia artificial, especialmente por el potencial lesivo de los riesgos asociados a la misma, no tanto por la probabilidad de que se materialicen en caso de un diseño adecuado y una aplicación efectiva de la supervisión y control humanos, sino por la gravedad de su impacto en los derechos y bienes jurídicamente protegidos.

El despliegue rápido de la inteligencia artificial sin una evaluación previa de sus riesgos y sin la responsabilidad y supervisión adecuada, pueden crear graves peligros para los bienes y derechos más importantes del ser humano, por lo que es necesario una regulación adecuada y efectiva de acompañamiento que, además, debe integrar la eliminación de las barreras legales y tecnológicas que impidan o dificulten su auditoría y su comprensión.

En relación con todo ello, desde hace algún tiempo, distintas voces expertas están planteando que las empresas renuncien a sus secretos comerciales[446] en relación con su transparencia y explicabilidad anteriormente analizadas.

Los gobiernos, autoridades, Administraciones públicas, empresas y personas deben ser capaces de entender y explicar cómo y por qué se toman las decisiones, especialmente, cuando está en juego el acceso de las personas a financiación, la atención sanitaria, las prestaciones sociales, la vivienda o el empleo.

Además, el uso y aplicación de los sistemas dotados de inteligencia artificial pueden tener un potencial lesivo a gran escala, causar daños y perjuicios de distinta naturaleza a personas y entidades, lo que exige tener un marco claramente definido de responsabilidad, especialmente para su determinación dentro de la cadena y en relación con los distintos

[446] *AI Now Report 2018*. AI Now Institute. 2018. Recuperado de: https://ainowinstitute.org/AI_Now_2018_Report.pdf. Consultado el 14.01.2021.

sujetos intervinientes, desde el diseño hasta su uso y aplicación, considerando todo el ciclo de vida del producto/sistema.

Es por ello que los nuevos marcos reguladores orientados a establecer el carácter vinculante de los principios y requisitos éticos de la inteligencia artificial sobre los que construir la misma y a regular aspectos como la responsabilidad, permitirán establecer un marco de seguridad jurídica para todas las partes implicadas que, de un lado, contribuirán a promover la innovación y la calidad de los sistemas de inteligencia artificial y, de otro, su confianza y seguridad para todos, enlazando así con el modelo por el que ha optado la UE en su intervención legislativa propuesta.

En este sentido, el marco europeo coordinado a nivel ético y jurídico, instrumentado en las propuestas del Parlamento Europeo en materia de normas éticas y de responsabilidad civil, constituye un ejemplo de referencia a nivel mundial para otros países que han empezado a implementar sus estrategias de IA y su regulación.

No obstante, en mi opinión, la exigencia de algunos de estos principios y normas básicas de origen ético como, por ejemplo, el control y la supervisión humana, debería ir más allá y constituir requisitos esenciales de cualquier sistema de inteligencia artificial, sean calificados o no inicialmente de alto riesgo o no, en atención, especialmente, a su propia naturaleza, características y capacidades de los que puede estar dotada.

En consecuencia, considero que la Propuesta de Reglamento regulador de la inteligencia artificial del Parlamento Europeo y del Consejo, de 21 de abril, y su futura tramitación constituye una buena oportunidad para revisar su enfoque de riesgos y ético desde una óptica más horizontal, para valorar la introducción de cambios en el modelo optado e incorporar la exigencia de determinados principios y normas de carácter ético a cualquier sistema inteligente, cuanto menos aquellos cuyos riesgos puedan impactar en algunos de los derechos y bienes más valiosos para todos los sujetos relacionados con la misma, en los términos expuestos durante mi análisis y reflexiones.

Capítulo IV

Marco jurídico

1. Introducción

La inteligencia artificial, como hemos visto en los capítulos precedentes, afecta a todo y a todos, personas físicas y jurídicas, ciudadanos, empresas, Administraciones públicas o gobiernos, en la dimensión física como la virtual.

La sociedad y la economía digital constituyen el ecosistema necesario de relaciones e interacciones con consecuencias tanto en el mundo físico, como digital, donde la inteligencia artificial tendrá un importante despliegue y aplicación en los próximos años.

La inteligencia artificial es esencial para la transformación digital de la sociedad y de los negocios, y está presente en todo tipo de ámbitos de nuestra vida cotidiana. Su rápido avance y despliegue en todo tipo de áreas y sectores comportará enormes cambios -y cada vez más acelerados- en los modelos de relación, interacción y prestación de servicios, tanto en empresas como Administraciones públicas, sector financiero, sanidad, seguridad, relaciones laborales, agricultura, etc.

El valor que puede aportar al ser humano en modo alguno puede ser cuestionado a pesar de sus retos y riesgos de distinta naturaleza, conforme he abordado, si bien, deberá acompañarse de la creación de los marcos necesarios para permitir su desarrollo, despliegue y aplicación seguro y confiable, evitando o minimizando al máximo la probabilidad e impacto de dichos riesgos.

La necesaria seguridad y confianza en las relaciones, interacciones, prestaciones, decisiones y acciones mediante el uso de inteligencia artificial requerirán la construcción de sólidos marcos normativos de acompañamiento que garanticen las mismas a todas las partes implicadas.

La actuación legislativa deberá ser la mínima que sea necesaria para proteger los intereses, bienes y derechos protegidos por nuestro ordenamiento jurídico que en mayor medida puedan verse afectados por inteligencia artificial, ante los riesgos potenciales asociados a la misma, sus características, capacidades, sector y contexto donde pueda ser usada.

Sin embargo, el Derecho no evoluciona al mismo ritmo que evoluciona la tecnología, ni de manera uniforme y armonizada, hallándose limitado por la naturaleza territorial de los ordenamientos jurídicos, en contraposición al *ciberespacio,* donde se despliegan todo tipo de relaciones, interacciones y conductas entre personas, entes y sistemas.

Los ejemplos serían innumerables: El marco jurídico actual del contrato de obra o de arrendamiento de servicios en materia tecnológica se halla regulado en el Código Civil español de 1889, las páginas web, *apps* o videojuegos no se hayan contemplados en la vigente Ley de Propiedad Intelectual española o algunos delitos tecnológicos como el *hacking* han sido recientemente incorporados a ordenamientos jurídicos como el español en su Código Penal. Y ocurre en España como en otros países.

En materia de ciberdelincuencia, los programadores que crearon el gusano "ILOVEYOU" provocaron daños por valor de miles de millones de dólares, pero fueron absueltos porque en Filipinas no había un marco legal contra la creación de *malware*.

La regulación es necesaria, si bien, comporta riesgos asociados, especialmente, cuando se incurre en un exceso, como expuse en el capítulo II de esta investigación.

Uno de los debates que se ha suscitado en relación con la inteligencia artificial es cuál debería ser el objeto y alcance de su regulación.

Son muchos los investigadores y emprendedores exitosos relacionados con la tecnología que incluso han mostrado su oposición en algún momento de su vida a toda forma de regulación del desarrollo de la inteligencia artificial para no dificultarla o hacerlo de manera superficial, entre otros, Elon Musk, que hace algunos años hablaba de lo que se necesita ahora de los gobiernos no es supervisión sino perspicacia y acompañamiento.

Sin embargo, en los últimos 4 años, Musk ha mostrado un discurso cambiante y un posicionamiento muy distinto, en el que incluso ha llegado a afirmar la necesidad de regular la inteligencia artificial y además de manera proactiva[447].

No comparto completamente algunas de sus opiniones, en especial respecto del mero "acompañamiento", en la medida que, de un lado, se precisa ahora más que nunca una intervención legislativa para garantizar la ética, la seguridad, la confianza, el respecto de los valores y derechos fundamentales y la responsabilidad en el desarrollo y aplicación de la inteligencia artificial, especialmente ante su complejidad, características, el aumento de sus capacidades, datos tratados y su interacción con otras tecnologías.

Es ahora cuando el ser humano debe definir como la inteligencia artificial debe construirse, desarrollarse y aplicarse de manera alineada con su futuro y objetivos para ayudarle a su consecución.

Y dicha intervención debe basarse en estrategia y planificación, y en nuevas técnicas y enfoques legislativos que garanticen instrumentos eficaces, adaptativos, flexibles y evolutivos, que sean globales pero que combinen una perspectiva horizontal con las perspectivas verticales que requieran determinados sectores, aplicaciones, tipologías de sistemas inteligentes y/o niveles de riesgo. Y aquel acompañamiento al que en ocasiones ha hecho referencia el precitado Musk, no puede tratarse de un mero acompañamiento ante sus riesgos y retos que el mismo reconoce y al que se opone, como por ejemplo su uso armamentístico, sino que debe tratarse de una necesaria supervisión e intervención legislativa, adecuada, proporcionada y sin excesos.

La inteligencia artificial precisa límites, por supuesto éticos, pero indudablemente también jurídicos, para garantizar, entre otros aspectos esenciales, los derechos humanos

[447] PALAZUELOS, F. (2017). "Elon Musk: 'La inteligencia artificial amenaza la existencia de nuestra civilización'". Publicado en *El País*. 18.07.2017. Recuperado de https://elpais.com/tecnologia/2017/07/17/actualidad/1500289809_008679.html. Consultado el 12.01.2021. En el mismo sentido Condliffe, J. "Elon Musk quiere regular la inteligencia artificial antes de que 'sea demasiado tarde'". Publicado en *MIT Technology Review* el 19.07.2017. Recuperado de https://www.technologyreview.es/s/8420/elon-musk-quiere-regular-la-inteligencia-artificial-antes-de-que-sea-demasiado-tarde. Consultado el 12.01.2021.

por los que tanto se ha luchado durante toda la historia de la humanidad, conforme al modelo europeo por el que ha optado la UE.

El Derecho no puede ser considerado una amenaza sino una oportunidad, y la ética, la seguridad, el cumplimiento normativo y la responsabilidad es una fortaleza y no una debilidad para todas las partes interesadas, tanto para la industria, gracias a la seguridad y confiabilidad en la misma por parte de sus usuarios y la ventaja competitiva que pueden suponer sistemas más seguros y confiables, como para los ciudadanos.

De hecho, el propio Musk reconoce en alguno de sus reflexiones públicas que la regulación puede en ocasiones fomentar el progreso en lugar de obstaculizarlo, poniendo como ejemplo que, si los estándares públicos de seguridad para los coches autónomos pueden contribuir a la reducción del número de accidentes en los que éstos se vean involucrados, esto haría menos probable el rechazo público y podría acelerar la adopción de estas nuevas tecnologías. Todo ello comportaría confiabilidad y seguridad por parte de la sociedad en general necesaria para su despliegue.

La UE está pretendiendo liderar su regulación como base para consolidar su posición actual y pretendido liderazgo de esta tecnología a nivel mundial, en una carrera desenfrenada por la supremacía que está liderando China y EE.UU., así como algunas de las grandes tecnológicas, tal y como analice anteriormente.

Las recientes propuestas normativas realizadas por el Parlamento Europeo en materia de ética, inteligencia artificial y responsabilidad civil de la inteligencia artificial, fruto de los análisis y estudios llevados a cabo durante los últimos años, podrían convertirse o, más bien, son ya en una referencia internacional para su adecuada regulación, cuanto menos desde un enfoque europeo.

2. La inteligencia artificial en el Derecho europeo

2.1. Aspectos generales

La tecnología ha cambiado el mundo y lo seguirá cambiando en los próximos años a una velocidad de crecimiento exponencial.

Los países más desarrollados compiten en una carrera mundial por obtener su supremacía y liderazgo en las tecnologías y servicios que han identificado como de mayor impacto, así como en el gobierno de los datos, si bien, bajo distintas estrategias y enfoques.

La inteligencia artificial no es ajena a esta pugna que pretenden liderar países como China, EE.UU. o Emiratos Árabes y a la que se ha unido la UE, para la cual, la inteligencia artificial constituye un objetivo estratégico y, en congruencia con ello, acaba de lanzar nuevas inversiones para la recuperación económica focalizándose en las líneas estratégicas definidas por la misma en materia digital, especialmente en gobierno de datos, tecnología *cloud*, infraestructuras de conectividad e inteligencia artificial.

España se unió a las estrategias europeas con su *Estrategia Nacional de inteligencia artificial*[448]*-ENIA*- publicada en diciembre de 2020, anunciada en su documento de *Estrategia Española de I+D+I en inteligencia artificial de 2019*[449].

La misma incorpora un plan de acción sustentado en seis ejes estratégicos: 1) Impulsar la investigación científica, el desarrollo tecnológico y la innovación en IA; 2) Promover el desarrollo de capacidades digitales, potenciar el talento nacional y atraer talento global en inteligencia artificial; 3) Desarrollar plataformas de datos e infraestructuras tecnológicas que den soporte a la IA; 4) Integrar la IA en las cadenas de valor para transformar el tejido económico; 5) Potenciar el uso de la IA en la administración pública

[448] *Estrategia Nacional de inteligencia artificial*. Gobierno de España. Versión 1.0. Disponible en: https://www.lamoncloa.gob.es/presidente/actividades/Documents/2020/ENIA2B.pdf Consultado el 30.03.2021.

[449] Recuperado de: https://www.ciencia.gob.es/stfls/MICINN/Ciencia/Ficheros/Estrategia_Inteligencia_Artificial_IDI.pdf. Consultado el 30.03.2021.

y en las misiones estratégicas nacionales; 6) Establecer un marco ético y normativo que refuerce la protección de los derechos individuales y colectivos, a efectos de garantizar la inclusión y el bienestar social.

El último de los mismos tiene como principal objetivo asegurar si es suficiente el marco ético y jurídico actual o, en su caso, valorar su revisión y adecuación, para definir un marco ético y normativo adecuado ante las implicaciones éticas, legales, laborales, sociales y económicas de la inteligencia artificial.

Conforme recoge, el objetivo es proteger los derechos fundamentales ya reconocidos, identificando las reformas legales necesarias, así como lagunas jurídicas que requieran regulación adicional.

Según dicha estrategia, mediante los marcos éticos se pretende que el uso de la inteligencia artificial "acompañe los valores de la sociedad y opere en beneficio de la inclusión y el bienestar". En definitiva y a mi juicio, que la inteligencia artificial responda a unos objetivos definidos por el ser humano, para cuya consecución, la inteligencia artificial se constituya en un medio adecuadamente alineado para ello, construido sobre una sólida base a nivel ético, que integre los principios y valores sobre los que pretendemos construir nuestra sociedad y el mundo en el que vivimos.

2.2. Ausencia de regulación

La inteligencia artificial no se halla actualmente regulada ni en España ni en Europa, tampoco distintas dimensiones de la robótica, lo que ha obligado a los juristas a realizar una labor integradora e interpretativa a la hora de resolver los problemas y conflictos que surgen en relación con la misma, tomando como referencia la legislación de Propiedad Intelectual, Ciberseguridad, Protección de Datos, Secretos Empresariales, seguridad industrial, seguridad de los productos y consumo, así como el resto de marcos normativos generales o sectoriales en materia civil, mercantil, laboral, administrativa o penal.

La UE sentó las bases de su futuro marco regulador en el *Libro blanco sobre la inteligencia artificial*[450] que he precitado en anteriores capítulos de esta investigación.

Siguiendo sus estrategias globales, la UE pretende liderar su regulación y dio un primer paso en este sentido con las Resoluciones del Parlamento Europeo de 20 de octubre de 2020, que ya integraban sendas propuestas de Reglamento en materia ética y de responsabilidad civil de la inteligencia artificial, la primera de las cuales ha sido analizada en el capítulo anterior y la segunda lo será en el próximo.

Posteriormente, la reciente Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[451], ha apostado por una regulación más profunda y detallada, con claro enfoque a la seguridad y armonización normativa en el seno de la UE, conforme abordaré en los próximos apartados.

Hasta fechas recientes, los marcos de referencia eran principalmente éticos, conforme expuse en el capítulo anterior, y algunos estándares de la industria y de entidades como la *International Organization for Standardization* (ISO), la cual ha aprobado, entre otras, la *ISO 10218-1:2011, Robots and robotic devices - Safety requirements for industrial robots*, la *ISO 13855, Safety of machinery - Positioning of safeguards with respect to the approach speeds of parts of the human body* y la *ISO/TS 15066:2016 (en) Robots and robotic devices - Collaborative robots*.

La *International Organization for Standardization* (ISO) -como organización para la creación de estándares internacionales compuesta por diversas organizaciones nacionales de normalización-, junto con la *Comisión Electrotécnica Internacional* (IEC) -organización de normalización en el ámbito eléctrico, electrónico y tecnologías

[450] *Libro Blanco sobre la inteligencia artificial - Un enfoque europeo orientado a la excelencia y la Confianza*. Comisión Europea. Bruselas, 19.2.2020. COM (2020) 65 final.
[451] COM (2021) 206 final 2021/0106 (COD)

relacionadas-, han identificado la necesidad de desarrollar normas relativas a la inteligencia artificial capaces de beneficiar a todas las sociedades.

Como analicé en anteriores capítulos, en 2018, el comité técnico conjunto *ISO/IEC JTC 1, Tecnologías de la Información* de ISO y la *Comisión Electrotécnica Internacional* (IEC) fundaron el *Subcomité SC 42 sobre inteligencia artificial*[452], constituyendo el primer ecosistema de la inteligencia artificial para la creación de estándares y normalización.

El denominado *Subcomité ISO/IEC JTC 1/SC 42*, dedicado a la inteligencia artificial, ha publicado ya varias normas relacionadas con el Big data y con otros 13 proyectos en desarrollo en la fecha de finalización de esta investigación.

Su objetivo es desarrollar e implementar un programa de normalización de la inteligencia artificial para guiar a otros comités de ISO en el desarrollo de aplicaciones de la inteligencia artificial.

En la actualidad, existen múltiples normas de referencia, pero constituyen, por el momento, estándares y buenas prácticas no vinculantes, salvo su conversión como tales por decisiones corporativas o contractuales en el ámbito público o privado, por ejemplo, exigiendo su adhesión o futura certificación a los mismos.

Entre otras, destacar las siguientes normas ISO/IEC en distintas fases desarrollo en la fecha de cierre de esta investigación[453]:

- Inteligencia artificial - conceptos y terminología [ISO/IEC WD 22989]

- Marco para sistemas de inteligencia artificial (IA) que utilizan el aprendizaje automático (ML) [ISO/IEC WD 23053]

---

[452] ISOfocus Noviembre-diciembre 2019. P. 2. Recuperado de: https://www.iso.org/files/live/sites/isoorg/files/news/magazine/ISOfocus%20(2013-NOW)/sp/ISOfocus_137_sp.pdf. Consultado el 05.02.2021.

[453] LLORET, J.A. (2019). *Estándares y seguridad en el uso humano de la IA*. 21.10.2019. Recuperado de: https://editorialia.com/2019/10/21/estandares-y-seguridad-en-el-uso-humano-de-la-ia/. Consultado el 20.12.2020.

- Tecnología de la información - Inteligencia artificial - Gestión de riesgos [ISO/IEC AWI 23894]

- Tecnología de la información - Inteligencia artificial (IA) - Sesgo en los sistemas de IA y la toma de decisiones asistida por AI [ISO/IEC NP TR 24027]

- Tecnología de la información - Inteligencia artificial (IA - Descripción general de la confiabilidad en inteligencia artificial [ISO/IEC PDTR 24028]

- Inteligencia artificial (IA) - Evaluación de la robustez de las redes neuronales - Parte 1: Descripción general [ISO/IEC NP TR 24029-1]

- Tecnología de la información - Inteligencia artificial (IA) - Casos de uso [ISO/IEC NP TR 24030]

- Tecnología de la información - Inteligencia artificial - Visión general de las preocupaciones éticas y sociales [ISO/IEC AWI TR 24368]

- Tecnología de la información - Inteligencia artificial (IA) - Descripción general de los enfoques computacionales para sistemas de IA [ISO/IEC AWI TR 24372]

- Tecnología de la información - Inteligencia artificial - Marco de gestión de procesos para análisis de Big data [ISO/IEC AWI 24668]

- Tecnología de la información - Gobernanza de TI - Implicaciones de gobernanza del uso de inteligencia artificial por parte de las organizaciones [ISO/IEC AWI 38507].

Estas normas forman parte del denominado *soft law*, asociado a códigos de conducta, recomendaciones, códigos de buenas prácticas, etc. que podría tener carácter vinculante en caso de decisión corporativa de una entidad o en el caso de una cláusula contractual que así lo exija en el ámbito público o privado.

Estas normas son precisamente eso, estándares, a los que una organización puede decidir adherirse o no, salvo que sea un requerimiento de una norma jurídica cumplirlo, un

requerimiento administrativo, un requerimiento contractual por parte de proveedores o clientes, o un requerimiento corporativo de una organización, es decir, resultado de una decisión empresarial de adherirse, en su caso, certificarse, y convertirlos en vinculantes para toda su estructura y organización.

No obstante, pueden producir ciertos efectos jurídicos y, de hecho, la jurisprudencia europea se ha pronunciado en diversas ocasiones sobre ello, como por ejemplo la Sentencia 322/88, de 13 de diciembre de 1989 del TJCE, dictada en el asunto 322/88 -conocida como "Sentencia Grimaldi"-, en la que se asienta la doctrina de que el *soft law* debe producir efectos jurídicos porque el juez puede, y en ocasiones debe, tener estas normas en cuenta para resolver un litigio.

De hecho, tradicionalmente, para los que definen el *soft law* en relación o en contraposición con el *hard law*, destacan su dimensión de derecho sin sanción o derecho no vinculante, destacando que no son referentes obligatorios para el poder judicial, ni parámetros de enjuiciamiento de disposiciones generales o actos administrativos.

Sin embargo, a partir de la sentencia precitada, esta argumentación no es plenamente sostenible.

El TJCE consideró que las recomendaciones ilustran la interpretación de otras disposiciones nacionales y completan las disposiciones comunitarias. De este modo, si estos efectos jurídicos son predicables de las recomendaciones europeas respecto al derecho interno de los Estados, más aún lo serán las recomendaciones, normas técnicas, códigos deontológicos, etc., que se desarrollan a nivel regulativo en el interior de un mismo sistema jurídico.

Estas técnicas legislativas se han presentado doctrinalmente de distintas maneras como han significado autores como Rubio[454], en particular, como dos paradigmas regulativos distintos y opuestos, como complementarios o como una forma híbrida de regulación.

[454] RUBIO, A. (2014). "Los efectos jurídicos del soft law en materia de igualdad efectiva. La experiencia Española*"*. *Anuario de filosofía del derecho AFD*. Nº 30, 2014. ISSN 0518-0872. Universidad de Granada. Pp. 48 -53.

Según esta autora, de un lado, los que defienden el *hard law* y el *soft law* como dos modelos regulativos distintos, argumentan que el *soft law* representa un modelo regulativo, deliberativo y participativo, que fomenta e incentiva compartir información y aprendizaje mutuo entre los diferentes actores implicados, en la regulación y experimentación de nuevas estrategias de cooperación y negociación.

De otro, están los que consideran que no son dos técnicas legislativas contrapuestas, sino dos técnicas complementarias entre sí, en la medida que el *soft law* es valorado como normas de *lege ferenda* o como criterios interpretativos que han de guiar la aplicación e interpretación del *hard law*.

Y, por último, otros matizan las dos posturas anteriores y defienden que, a pesar de ser dos técnicas legislativas distintas, ambas concluyen muy eficazmente en la resolución de determinados conflictos de alta complejidad social, los cuales demandan una fuerte especialización.

Mi posicionamiento al respecto se sitúa en éstas dos últimas posturas. Cuando hablamos del desarrollo y aplicación de conjuntos de tecnologías complejas que pueden asociarse potencialmente con otras y aumentar exponencialmente tanto sus bondades como sus retos y sus riesgos me decanto por el segundo posicionamiento de base y por el tercero para dar soluciones a algunos de problemas jurídicos que ya está planteando la inteligencia artificial.

Asimismo, la doctrina se divide entre quienes niegan el valor interpretativo del *soft law* y otros que incluso le reconocen valor de fuente del derecho, al considerar que tiene una estructura normativa similar a la de los principios generales del derecho.

De este modo, el *soft law*[455] no tiene carácter vinculante y presenta dificultades para su integración en la teoría tradicional de las normas jurídicas, sin perjuicio de su acomodo en lo que Bobbio[456] denominó *la función promocional del derecho*, esto es, un derecho

[455] ALARCÓN, G. (2014). "El soft law y el sistema de fuentes*". Tratado sobre la Ley General Tributaria: Homenaje a Álvaro Rodríguez*. Vol 1. Tomo I. Editorial Aranzadi Thomson Reuters 2010. Pp. 37-68.
[456] MONTOYA, M (2005). "Derecho y política en el pensamiento de Bobbio: una aproximación". *Estudios Políticos*. Nº 26. Medellín. Enero-junio 2005. Pp. 108-114. Y ver también, BOBBIO, N. Y RUIZ DE MIGUEL, A. (1990). *Contribución a la teoría del derecho*. Editorial Debate. Pp. 371-386.

cuyo fin es proponer e incentivar su cumplimiento, no imponer determinadas conductas o prácticas, es decir, un derecho que trata de buscar, dada la naturaleza de su fin, la adhesión voluntaria de los diferentes sujetos normativos a quienes se dirige.

Según la autora precitada, la búsqueda de esta adhesión no debe interpretarse como debilidad o defecto, sino como una estrategia para favorecer un relevante cambio social y económico.

El *soft law* aporta a estas soluciones globales su capacidad de generar nuevas formas de cooperación, participación e integración en contextos complejos y en relaciones en las que confluyen diversos sujetos, intereses y niveles de regulación, que además están abiertos a fuertes incertidumbres y cambios, y ante los que la los gobiernos encuentran dificultades para regularlos como, a mi juicio podríamos considerar el escenario que supone la inteligencia artificial.

En mi opinión, tal y como he manifestado al abordar otros aspectos objeto de esta investigación, la tecnología y, en particular, la inteligencia artificial como realidad compleja, precisa y debe ser abordada igualmente de manera compleja y global desde el punto de vista ético, jurídico y de seguridad, y no exclusivamente regularla mediante normas jurídicas sino de manera global mediante *hard* y *soft law*, de modo que se articulen marcos jurídicos directamente vinculantes de manera coordinada con la promoción de la autorregulación y de normas, códigos de conducta y estándares de buenas prácticas no vinculantes inicialmente, pero que puedan conformar el marco exigido por la industria y los distintos sectores, incluso convertible en vinculante vía contractual o corporativa por las propias empresas y Administraciones públicas, adhiriéndose a los mismos, exigiendo su adhesión a proveedores o certificándose, en su caso, en los mismos. Y todo ello desde dicha perspectiva global e internacional, de modo que permiten superar las barreras que supone el concepto territorial del ordenamiento jurídico en un mercado global y mundial.

La utilización de esta técnica legislativa global, en mi opinión, contribuirá a la necesaria sensibilidad, flexibilidad y ágil adaptabilidad de los nuevos marcos a la realidad

tecnológica y social en cada momento, al objeto de disponer de marcos dúctiles y "responsive", un nuevo derecho capaz de absorber con rapidez los cambios, los retos y la diversidad, y de adaptarse al cambio constante asociado a la disrupción exponencial que la tecnología está promoviendo.

No obstante, deben identificarse aquellos principios y normas principalmente éticas esenciales que deberían contemplarse en dicho *hard law* como vinculantes y exigibles a los sistemas de inteligencia artificial en función de su naturaleza, características, contexto y riesgos, algunos de los cuales deberían ser exigibles a todos, en función el concepto de inteligencia artificial desde el que se parta y que se desea para el futuro, en relación con los bienes y derechos más preciados de todos los sujetos involucrados.

La doctrina ya ha abordado la disyuntiva entre optar por *hard law* o *soft law*, cuando considero que realmente no es tal, sino instrumentos que deben utilizarse de manera coordinada y armonizada para dar una solución flexible, adaptativa, ágil y global a una realidad compleja que requiere estas propiedades, en lo que podríamos denominar "una combinación inteligente de ambas técnicas", que a mi juicio son absolutamente complementarias.

Como he anticipado, la reciente propuesta armonizadora y reguladora de la inteligencia artificial de la UE de 21 de abril de 2021, apuesta por un *hard law* para la prohibición de determinados sistemas y para regular los requerimientos y obligaciones de los sistemas de inteligencia artificial considerados de alto riesgo conforme a la misma, y promueve el *soft law* para el resto mediante la adopción de códigos de conducta de adscripción voluntaria.

2.3. Referencias europeas e internacionales

Los distintos países comparten en mayor o menor grado distintas preocupaciones relacionadas con la inteligencia artificial y sus retos, y se ha ido conformando un consenso cada vez mayor en relación con los principios y normas éticas que deben regir el desarrollo, despliegue y aplicación de la inteligencia artificial, que pretende liderar la UE.

Una de las principales preocupaciones compartidas la constituye cómo abordar la responsabilidad en los incidentes provocados por la misma y el resarcimiento por los daños que pueda causar, si bien, en estos aspectos, los enfoques y criterios no son uniformes, especialmente en atención a su propia tradición y cultura.

No existe uniformidad en el tratamiento y regulación de la tecnología y, en particular, de la inteligencia artificial, especialmente en aspectos como los indicados, entre otros.

Los diferentes enfoques, tratamientos e interpretaciones para responder a los conflictos derivados de la inteligencia artificial con los instrumentos jurídicos vigentes son distintos y vinculados al desarrollo tecnológico de cada país.

No obstante, la preocupación por los desafíos éticos y jurídicos de la inteligencia artificial forma parte de la agencia política institucional de gobiernos, entidades y organismos internacionales.

Distintos países han abordado la inteligencia artificial de un modo más estratégico y político, antes que jurídico, para construir con posterioridad los marcos reguladores congruentes con dicha estrategia.

Emiratos Árabes fue uno de los primeros países que abordaron de modo estratégico y político la inteligencia artificial, nombrando en 2017 un "Ministro de inteligencia artificial".

Suecia creó una especie de "Ministerio del futuro", el denominado *Ministerio para el Desarrollo Estratégico y la Cooperación Nórdica*.

EE.UU. utilizó su *Office of Technology Assesment* para investigar, prever y asesorar al Congreso en el futuro. Otros países se inspiraron en la misma para crear sus propias oficinas, como la *Oficina Parlamentaria de Ciencia y Tecnología en el Reino Unido*, la *Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag* en Alemania o el *Centro Suizo de Evaluación de Tecnología*.

El Foro Económico Mundial[457] o "Foro de Davos" no ha permanecido al margen de esta escalada y elaboró el pasado año un documento sobre el uso responsable de la inteligencia artificial, que contiene una propuesta de marco para la integración de la ética y los derechos humanos en el desarrollo tecnológico por parte de gobiernos y empresas, y ello en todo el ciclo de vida del producto, esto eso, en su diseño, comercialización y uso.

La Organización Mundial de la Propiedad Intelectual -OMPI o WIPO por sus siglas en inglés- abrió un período de consultas sobre la aplicación de la inteligencia artificial para la generación la propiedad intelectual y abordar el futuro reconocimiento y registro de creaciones generadas por sistemas de IA. Sobre esta materia, el Parlamento Europeo publicó su reciente Resolución de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial[458], que será objeto de análisis en el capítulo VIII de esta investigación.

Hong Kong, Japón, Singapur y algunos países de Oriente Medio están diseñando marcos reguladores específicos sobre IA. Singapur ha aprobado su estrategia nacional sobre inteligencia artificial con la finalidad de posicionar a su país como potencia mundial en inteligencia artificial, en la que destaca su compromiso con los trabajadores de dotarlos de competencia para mejorar su competitividad.

Corea del Sur ya creó su marco regulador específico para la robótica mediante su *Korean Law on the Development and Distribution of Intelligent Robots* de 2005 y la *Legal Regulation of Autonomous Systems in South Korea* de 2012[459].

Por su parte, Japón dispone de sus *Guidelines to Secure the Safe Performance of Next Generation Robots,* así como sus estrategias en robótica mediante la *New Robot Strategy - Japan's Robot Strategy - Vision, Strategy, Action Plan* y la *Headquarters for Japan's Economic Revitalization* de febrero de 2015. Taiwan también ha creado su propio marco específico.

---

[457] Recuperado de: https://es.weforum.org/platforms/shaping-the-future-of-technology-governance-artificial-intelligence-and-machine-learning. Consultado el 2.02.2021.

[458] Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial (2020/2015(INI))

Hong Kong ha creado un marco específico para el sector bancario, exigiendo la obligación de garantizar la transparencia, resultados éticos y justos, y la protección de la privacidad, responsabilizando directamente al órgano de gobierno de las entidades bancarias del uso que hagan de la misma.

Arabia Saudí aprobó la creación de un centro de desarrollo e investigación de inteligencia artificial para acometer la transformación digital del país, con el objetivo de convertirse a una referencia en Oriente Medio.

China tiene como objetivo liderar la inteligencia artificial a nivel mundial, al igual que otras tecnologías como *Blockchain*. De hecho, ya ha ido más allá de la mera ética y ha publicado un marco de gobernanza sobre la inteligencia artificial para sus empresas, en la que se advierte de la responsabilidad penal si la tecnología no está supeditada a los humanos.

Actualmente, el gigante asiático está elaborando un borrador de Reglamento sobre algoritmos de recomendación en Internet que será presentado a consulta pública en breve, en el que se pretende regular un conjunto de obligaciones para los proveedores de recomendación algorítmica. Esta futura norma se unirá a los nuevos marcos jurídicos aprobados en materia de protección de la seguridad de las infraestructuras críticas de información que desarrolla su normativa de ciberseguridad, así como los relativos a la gestión de servicios de información en Internet, la seguridad de datos y la protección de la información personal, materia esta última en la que pretende seguir un enfoque, salvando las distancias políticas, jurídicas y culturales, similar al RGPD.

EE.UU. fue pionera en sus propuestas de regulación de la denominada "responsabilidad algorítmica". Así, en abril de 2019 se presentó al Senado la *Algorithmic Accountability Act of 2019*[460], que constituye la primera propuesta reguladora de carácter federal en esta materia.

[460] Recuperado de: https://www.congress.gov/bill/116th-congress/house-bill/2231/all-info. Consultado el 10 de diciembre de 2020.

Asimismo, al cierre de esta investigación se ha presentado un proyecto de ley denominado *Artificial Intelligence Capabilities and Transparency (AICT) Act*[461] en materia ética, seguridad y transparencia de la inteligencia artificial, con el objetivo también de aumentar la transparencia gubernamental en el uso de la misma. Del mismo, coetáneamente a dicho cierre se anunció durante la Cumbre entre la UE y EE.UU., celebrada el 5 de junio de 2021 en Bruselas, la puesta en marcha de Consejo de Comercio y Tecnología UE-EE.UU.[462] al objeto de liderar una transformación digital basada en valores compartidos.

La propuesta estadounidense constituye un avance en el reconocimiento jurídico del principio de responsabilidad algorítmica, al objeto de que el sistema utilice controles para garantizar que el operador pueda verificar que actúa de acuerdo con sus intenciones, así como identificar y rectificar resultados perjudiciales.

La propuesta exige que entidades comerciales específicas realicen evaluaciones de sistemas de alto riesgo que involucren información personal o tomen decisiones automatizadas, como los sistemas que usan inteligencia artificial o aprendizaje automático.

Asimismo, la propuesta considera que los sistemas de decisión automatizados de alto riesgo conforme a la misma, incluyen aquellos que pueden contribuir a la inexactitud, el sesgo o la discriminación, o facilitar la toma de decisiones sobre aspectos sensibles de la vida de los consumidores mediante la evaluación de su comportamiento. Además, un sistema de decisión automatizada, o un sistema de información que involucre datos personales, se considera de alto riesgo si plantea problemas de seguridad o privacidad, involucra la información personal de un número significativo de personas, o sistemáticamente supervisa una gran ubicación física de acceso público.

Los criterios de calificación de los sistemas de alto riesgo difieren sensiblemente de los previstos en las distintas propuestas europeas, lo que evidencia la diferencia de posicionamientos internacionales en relación con la inteligencia artificial y la deseable y

[461] Recuperado de: https://www.congress.gov/bill/117th-congress/senate-bill/1705/text?r=13&s=1. Consultado el 01.06.2021.
[462] "La UE y los Estados Unidos ponen en marcha el Consejo de Comercio y Tecnología para liderar la transformación digital mundial basada en valores". Publicado en *Diario La Ley – Ciberderecho*. Nº 52, 12 de julio de 2021.

actualmente utópica armonización y consenso internacional, cuanto menos, respecto de los sistemas más críticos o de mayor riesgo como, por ejemplo, el *scoring* social o determinados tratamientos de datos biométricos, identificación remota y videovigilancia.

Por lo que se refiere a las obligaciones relacionadas con estos sistemas de alto riesgo, las evaluaciones de los sistemas de decisión automatizada que integren deben describir el sistema en detalle, evaluar los costos y beneficios relativos del sistema, determinar los riesgos para la privacidad y seguridad de la información personal, y explicar las medidas tomadas para minimizar esos riesgos, si se descubren. Asimismo, las evaluaciones de los sistemas de información de alto riesgo que involucren información personal deben evaluar hasta qué punto el sistema protege la privacidad y seguridad de dicha información.

EE.UU. ha regulado otros aspectos particulares de la inteligencia artificial en los últimos años desde un enfoque vertical y en relación con determinadas aplicaciones. Illinois[463] fue uno de los primeros estados en regular el uso de la inteligencia artificial, en particular, en los procesos de selección, mediante un nuevo marco normativo vigente desde el 1 de enero de 2020, en cual se basa en el consentimiento del candidato antes de usar la inteligencia artificial para evaluar los videos de las entrevistas de trabajo. Actualmente, el país se halla inmerso en la lucha contra la pandemia y en un proceso de asentamiento del nuevo gobierno, pero tiene en su agenda una regulación más ambiciosa de la inteligencia artificial, para lo que designó un grupo de expertos que partió de la definición de sus principios éticos.

Los 37 países miembros de la *Organización para la Cooperación y el Desarrollo Tecnológico* (OCDE)[464] , incluyendo países como Japón, Corea, México y Colombia, como referí en el capítulo anterior de esta investigación, suscribieron los principios sobre inteligencia artificial alineados con valores humanos propuestos por EE.UU., supeditando

[463] Informe *Riesgos de la inteligencia artificial: ¿qué oculta 2020?.* Clifford Chance 2019.
[464] Australia, Austria, Bélgica, Canadá, Chile, Colombia, República Checa, Dinamarca, Estonia, Finlandia, Francia, Alemania, Grecia, Hungría, Islandia, Irlanda, Israel, Italia, Japón, Corea, Letonia, Luxemburgo, México, Holanda, Nueva Zelanda, Noruega, Polonia, Portugal, República de Eslovaquia, Eslovenia, España, Suecia, Suiza, Turquía, Reino Unido y Estados Unidos. Ver https://www.oecd.org/acerca/miembros-y-socios/, consultado el 02.12.2020.

el uso de esta tecnología al respeto del Estado de derecho, los derechos humanos, la democracia, la transparencia y la rendición de cuentas.

Reino Unido ha sido de los más avanzados a nivel europeo a la hora de abordar la inteligencia artificial por parte de autoridades y tribunales. La Autoridad de Competencia y Mercados publicó su "Estrategia de mercados digitales" sustentada en el seguimiento del desarrollo del aprendizaje automático y garantizar que no se perjudique la competitividad ni al consumidor. Asimismo, su oficina de patentes se ha posicionado de manera taxativa respecto de la posibilidad de protección de cualquier invención no humana, acordando el rechazo automático de cualquier solicitud de patente donde no figure su creador humano, conforme analizaré con detalle en el capítulo VIII de esta investigación.

Alemania presentó en 2019 su hoja de ruta sobre inteligencia artificial que pretende ser la base y preámbulo de un marco normativo que verá la luz en breve. Asimismo, se han publicado distintos marcos éticos, entre otros, el de la *Asociación Alemana para la Economía Digital*, en el que aborda sus desafíos e implicaciones éticas.

Polonia ha lanzado su estrategia 2019-2027 para el desarrollo de la inteligencia artificial.

España, por el momento, no ha llevado adelante todos los propósitos reiteradamente anunciados, especialmente ante su inestabilidad política y parlamentaria, si bien, ha aprobado, como se ha indicado anteriormente, su *Estrategia Nacional de inteligencia artificial* alineada con la europea y recientemente ha aprobado su *Carta de Derechos Digitales*, que comentaré en el siguiente apartado, donde la inteligencia artificial tiene especial protagonismo.

La UE, como se ha referido anteriormente, ha reconocido la importancia estratégica de la inteligencia artificial en su *Plan Coordinado sobre inteligencia artificial*[465] cuyo objetivo es armonizar y coordinar las iniciativas de inteligencia artificial en toda la UE, y definió

[465] Recuperado de: https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence. https://ec.europa.eu/info/sites/info/files/communication-eu-security-union-strategy.pdf. Consultado el 30.03.2021.

las bases de la futura regulación de la inteligencia artificial en su *Libro blanco sobre la inteligencia artificial*[466] en 2020.

Conforme a estas bases, la UE fue más allá dentro de sus objetivos de liderazgo mundial en inteligencia artificial, proponiendo un marco regulador de los principios y normas éticas y obligaciones jurídicas relacionadas, analizado en el anterior capítulo, así como un marco regulador de la responsabilidad civil derivadas de la inteligencia artificial uniforme para todos los Estados miembros, si bien, complementario a los regímenes nacionales de responsabilidad civil y penal, que será analizado con detalle en el siguiente capítulo.

Recientemente ha publicado la nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, que será objeto de análisis en los siguientes apartados.

Los principales avances internacionales se han producido principalmente en el ámbito de la ciberseguridad de los sistemas, la conducción automatizada -más que autónoma- y la ética, especialmente dentro del marco de recomendaciones, y buenas prácticas.

Por lo que se refiere a los sistemas de inteligencia artificial para la conducción autónoma, se están desarrollando varias normas y recomendaciones internacionales sobre ciberseguridad[467].

En particular, la *Comisión Económica de las Naciones Unidas para Europa (UNECE)* ha publicado una normativa sobre ciberseguridad[468] que define un conjunto de requisitos que deben cumplir los fabricantes de vehículos, los proveedores y los prestadores de servicios,

[466] Libro Blanco de la Comisión, de 19 de febrero de 2020, titulado "Inteligencia artificial - Un enfoque europeo orientado a la excelencia y la confianza" (COM(2020)0065)
[467] *Cybersecurity challenges in the uptake of artificial intelligence in autonomous driving*. UE 2021. Recuperado de: https://www.enisa.europa.eu/publications/enisa-jrc-cybersecurity-challenges-in-the-uptake-of-artificial-intelligence-in-autonomous-driving. Consultado el 22.02.2021.
[468] *Regulation on uniform provisions concerning the approval of vehicles with regards to cyber security and cyber security management system.* United Nations - Economic Commission for Europe.

y que abarca todo el ciclo de vida del vehículo (es decir, desde su desarrollo hasta su retirada).

La agencia canadiense *Transport Canada* publicó en 2020 *Canada's Vehicle Cyber Security Guidance*[469], que proporciona principios rectores para ayudar a garantizar la ciberseguridad de los vehículos. Esta guía tiene como objetivo apoyar a las partes interesadas del sector proporcionando principios rectores neutrales desde el punto de vista tecnológico y no prescriptivos para reforzar la ciberseguridad a lo largo del ciclo de vida del vehículo.

Los OEM[470] europeos publicaron un conjunto de principios de ciberseguridad, a través de los *Principios de Ciberseguridad del Automóvil* de la ACEA[471], que ya son aplicados por las empresas OEM.

La *Administración Nacional de Seguridad del Tráfico en Carreteras* (NHTSA) del gobierno de los EEUU publicó a finales de 2016 un documento en el que se introducen buenas prácticas de ciberseguridad para los coches inteligentes[472].

El *Consejo de Estándares de Singapur* publicó en 2019 una serie de directrices para el despliegue de los coches autónomos denominado *Referencia Técnica 68*[473], cuya Parte 3 está relacionada con la definición de los principios de ciberseguridad y un marco de evaluación.

La *Corporación Nacional de Desarrollo y Reforma* (NDRC) de China actualizó en febrero de 2020 su *Estrategia para el Desarrollo de los Vehículos Inteligentes*[474], que establece cinco misiones clave, entre ellas el establecimiento de un "sistema integral de ciberseguridad".

---

[469] *Canada's vehicle cyber security guidance*. Transport Canada. 2020
[470] *Original Equipment Manufacturer* o "Fabricante de Equipo Original".
[471] *Principles of Automobile Cybersecurity*. ACEA. 2017
[472] *Cybersecurity Best Practices for Modern Vehicles*. National Highway Traffic Safety Administration (NHTSA). U.S. Department of Transportation, 2016.
[473] *Technical Reference 68 - Autonomous vehicles*. Singapore Standards Council. 2019.
[474] M. SCHAUB Y A. ZHAO. (2020). *China Releases Big Plan for Autonomous Vehicles*. *China Law Insight*. 2020

El *Centro de Análisis e Intercambio de Información de Automoción* de Estados Unidos (Auto-ISAC) mantiene desde 2016 una serie de buenas prácticas de ciberseguridad en el sector de la automoción[475], que proporcionan orientación sobre la aplicación de los principios de ciberseguridad en el sector de la automoción.

En EE.UU., algunos estados han venido desarrollando durante estos últimos años normas para la prueba y circulación de vehículos autónomos como Nevada -primer Estado en permitir la circulación en 2011-, Florida, California o Louisiana. Por ejemplo, la ciudad de Boston dispone de un servicio de taxi autónomo en funcionamiento. En el capítulo V de esta investigación relativo a responsabilidad civil, abordaré algunas de estas normas en relación con la responsabilidad en sectores específicos.

En Japón, la ciudad de Fujisawa dispone de un servicio de taxis sin conductor ofrecidos por la compañía *Robot Taxi*.

En el ámbito europeo, Francia ha sido uno de los primeros países europeos en aprobar una norma que permite a los vehículos autónomos hacer pruebas por sus calles. Algunas poblaciones de Suiza tienen ya autobuses que hacen circuitos sin conductor.

Reino Unido publicó su *Automatic and Electric Vehicles Act 2018*, aprobada el 19 de julio de 2018, la cual no es muy distinta a la que tenemos en España para vehículos convencionales, si bien, se regulan aspectos muy interesantes en el ámbito de la responsabilidad, que trataré con algo más de detalle en el capítulo siguiente al abordar el marco regulador de la responsabilidad en España y la UE en la actualidad.

En concreto, esta norma establece la responsabilidad del propietario del vehículo autónomo cuando éste no disponga de seguro en vigor y cause daños a terceros, la responsabilidad asegurada por daños incluiría a los sufridos por el propietario y usuario con excepciones, se cubren daños por muerte, personales y patrimoniales con excepciones, la responsabilidad del propietario o aseguradora es compatible con otras, se contemplan las alteraciones de *software* o falta de actualización del mismo como causas de limitación o exclusión de la responsabilidad de la aseguradora y, por último, se regula

[475] *Best Practices - Executive Summary*. Auto-ISAC. 2016. Disponible en: https://automotiveisac.com/best-practices/. Consultado el 02.12.2020.

la acción de repetición de la aseguradora o del propietario que ha abonado la indemnización contra cualquier tercero responsable del accidente como, por ejemplo, el fabricante que incurrió en defectos de programación del sistema.

Londres ya ha realizado pruebas autorizadas de vehículos autónomos en zonas de conducción abierta.

En la actualidad, las principales marcas con mayor inversión en coches autónomos y en plataformas aseguran estar preparándose para el denominado "Transporte como Servicio" -*Transportation as a Service* o TaaS por sus siglas en inglés-, y en su gran mayoría están domiciliadas en Alemania, Reino Unido, EE.UU. y Corea del Sur.

Por lo que se refiere a España, esto no es una realidad por el momento. La Convención de Ginebra sobre la Circulación por Carretera de 1952[476], ratificada por España, establece en su artículo 8 que todo vehículo deberá llevar un conductor. Sin embargo, la Convención de Viena sobre Circulación por Carretera de 1968 no fue ratificada por la misma, la cual recogía idéntica prescripción. En consecuencia, esta exigencia conforme a la Convención de Viena precitada no resultaría de aplicación en España, por lo que podría considerarse un aspecto de referencia inicial para el despliegue de los vehículos sin conductor.

No obstante, la normativa española vigente contempla algunos aspectos que impediría la circulación en España de vehículos no controlados por un conductor (otra cosa es que interpretemos el concepto conductor en sentido más amplio y pueda incluirse un operador virtual).

Conforme a lo dispuesto en el artículo 13 del Real Decreto Legislativo 6/2015, de 30 de octubre, por el que se aprueba el Texto Refundido de la Ley sobre Tráfico, Circulación de Vehículos a Motor y Seguridad Vial[477], el conductor debe estar en todo momento en condiciones de controlar el vehículo.

[476] Ratificada por España mediante instrumento de adhesión de 13 de febrero de 1.958
[477] BOE 31.10.2015

Abordaré el marco legal en España sobre esta materia en el capítulo siguiente, en relación con los aspectos relacionados con la responsabilidad por daños derivados del uso de este tipo de vehículos.

Algunas compañías como General Motors, han solicitado ya los permisos necesarios para autorizar la circulación de coches sin volante ni pedales en las carreteras españolas.

Adicionalmente a todo ello, se han aprobado distintas normas y estándares a nivel internacional.

El grupo *British Standards Institution* (BSI) publicó en diciembre de 2018 dos especificaciones públicamente disponibles (PAS), a saber, PAS 1885 y PAS 11281[478].

La primera, titulada *Los principios fundamentales de la ciberseguridad en la automoción*, proporciona una orientación de alto nivel para proporcionar y mantener la ciberseguridad. En cuanto a la PAS 11281, titulada *Connected automotive ecosystems - Im- pact of security on safety - Code of practice*, ofrece recomendaciones para gestionar los riesgos de seguridad en un ecosistema automovilístico conectado.

El Instituto Europeo de Normas de Telecomunicación (ETSI) ha desarrollado un conjunto de especificaciones técnicas[479] para definir una arquitectura de seguridad del Sistema de Transporte Inteligente (ITS) junto con la especificación de los servicios para garantizar la confidencialidad de la información y evitar el acceso no autorizado a los servicios ITS. También abordan la gestión de la confianza y la privacidad en las comunicaciones de los STI. Estas normas son la base integral de las políticas europeas de certificación y seguridad de los C-ITS, que son los documentos políticos que rigen la aplicación de

[478] *PAS 1885:2018 The fundamental principles of automotive cyber security. Specification*. BS. 2018; BSI, 'PAS 11281:2018 Connected automotive ecosystems. Impact of security on safety. Code of practice'. 2018

[479] *ETSI TS 102 940 - V1.3.1 - Intelligent Tranport Systems (ITS); Security; ITS communications security architecture and security management*. ETSI. 2018; "*TS 102 941 - V1.2.1 - Intelligent Transport Systems (ITS); Security; Trust and Privacy Management*". ETSI. 2018; "*TS 102 942 - V1.1.1 - Intelligent Transport Systems (ITS); Security; Access Control*". ETSI. 2012; "*TS 102 943 - V1.1.1 - Intelligent Transport Systems (ITS); Security; Confidentiality services*". ETSI. 2012.

algunas de las normas del ETSI como base para el despliegue interoperable y seguro de los C-ITS en la UE.

La norma de la Sociedad de Ingenieros de Automoción "SAE J3061"[480], publicada oficialmente en enero de 2016, se considera la primera norma que aborda la ciberseguridad en la automoción. Proporciona un conjunto de principios de ciberseguridad de alto nivel y orientación para los sistemas ciberfísicos de los vehículos.

La Organización Internacional de Normalización (ISO) y la SAE precitada colaboraron para sustituir la práctica recomendada *SAE J3061* y propusieron la norma I*SO/SAE 21434*[481].

Esta norma se centra en la ingeniería de ciberseguridad del automóvil, especificando los requisitos y proporcionando recomendaciones para la gestión de los riesgos de ciberseguridad de los automóviles (incluidos sus componentes, software e interfaces) a lo largo de todo su ciclo de vida.

Por su parte, la *norma SAE J3101*[482] define los requisitos comunes de seguridad que deben implementarse en el hardware de los vehículos autónomos.

Asimismo, en el ámbito de la UE, sus órganos e instituciones[483] han sido numerosas las iniciativas hasta la fecha en esta materia:

A principios de 2014, la Dirección General de Movilidad y Transportes (DG MOVE) de la Comisión Europea creó una plataforma de despliegue de sistemas inteligentes de transporte cooperativos (C-ITS), concebida como un marco de cooperación que incluía a las autoridades nacionales, a las partes interesadas y a la Comisión Europea, con el objetivo de identificar y acordar cómo garantizar la interoperabilidad de los C-ITS a

[480] *SAE J3061: Cybersecurity Guidebook for Cyber-Physical Vehicle Systems*. SAE. 2016
[481] *ISO/SAE DIS 21434 - Road vehicles - Cybersecurity engineering*. ISO/SAE. 2020
[482] *SAE J3101: Hardware Protected Security for Ground Vehicles*.
[483] *Cybersecurity challenges in the uptake of artificial intelligence in autonomous driving*. UE 2021. Disponible en: https://www.enisa.europa.eu/publications/enisa-jrc-cybersecurity-challenges-in-the-uptake-of-artificial-intelligence-in-autonomous-driving. Consultado el 22.02.2021.

través de las fronteras y a lo largo de toda la cadena de valor, así como para identificar los escenarios de despliegue más probables y adecuados.

En 2016, la Comisión Europea adoptó una *Estrategia Europea sobre Sistemas Inteligentes de Transporte Cooperativos*, una iniciativa que marca un hito hacia la movilidad cooperativa, conectada y automatizada[484]. El objetivo de la Estrategia C-ITS es facilitar la convergencia de las inversiones y los marcos normativos en toda la UE, con el fin de ver el despliegue de servicios C-ITS maduros en 2019 y más allá. También en 2016, los Estados miembros y la Comisión Europea pusieron en marcha la plataforma C-Roads para vincular las actividades de despliegue de C-ITS, desarrollar y compartir conjuntamente las especificaciones técnicas y verificar la interoperabilidad mediante pruebas cruzadas. C- Roads está abierta a todas las actividades de despliegue para las pruebas de interoperabilidad.

En 2017, la Dirección General de Mercado Interior, Industria, Emprendimiento y Pequeñas y Medianas Empresas (DG GROW) de la Comisión Europea puso en marcha una iniciativa sobre la normativa de seguridad con el objetivo de contribuir a una mayor disminución del número de víctimas mortales y heridos en la carretera considerando las modificaciones del Reglamento General de Seguridad y del Reglamento de Seguridad de los Peatones.

En 2018, la Comisión Europea publicó la *Estrategia de la UE para la movilidad del futuro*[485], estableciendo una acción específica para implementar un piloto sobre las infraestructuras y los procesos de ciberseguridad comunes en toda la UE que son necesarios para una comunicación segura y fiable entre los vehículos y la infraestructura para la seguridad vial y la gestión del tráfico. Desde 2018, la Comisión Europea está implementando el *Sistema de Gestión de Credenciales de Seguridad C-ITS de la UE* (EU

[484] *A European strategy on Cooperative Intelligent Transport Systems, a milestone towards cooperative, connected and automated mobility*. European Commission. COM/2016/0766 final, 2016.
[485] *On the road to automated mobility: An EU strategy for mobility of the future*. EU. COM(2018) 283 final, 2018.

CCMS) basado en la *Política de Seguridad C-ITS (SP)* y la *Política de Certificación C-ITS (CP)* publicadas en el sitio web del *Punto de Contacto C-ITS (CPOC)*[486].

En 2019, la Comisión Europea creó un grupo de expertos de la Comisión sobre movilidad cooperativa, conectada, automatizada y autónoma, denominado "CCAM", para asesorar y apoyar a la Comisión en el ámbito de las pruebas de despliegue.

En ese mismo año, la Agencia de Ciberseguridad de la Unión Europea -ENISA- realizó, con la participación del Joint Research Centre de la Comisión Europea -JRC por sus siglas en inglés-, un estudio sobre *Buenas prácticas para la seguridad de los coches inteligentes* centrado en los coches semiautónomos y autónomos[487]. Además, ya en 2016, ENISA había realizado un estudio sobre cuestiones de seguridad de los coches inteligentes, que dio lugar a un documento titulado *Cyber Security and Resilience of smart cars*[488].

En 2020, el CCI publicó un informe sobre el futuro del transporte por carretera[489]. Ese mismo año, ENISA creó el *Grupo de Expertos en Seguridad de la Movilidad Conectada y Automatizada* (CAMSec)[490], para abordar las amenazas, los retos y las soluciones en materia de ciberseguridad de los sistemas de transporte inteligentes (ITS) y del transporte CAM.

Previamente, ENISA también había creado el grupo de trabajo *Cars and Roads SECurity* (CaRSEC), para abordar las amenazas a la ciberseguridad de los coches inteligentes, los retos y las soluciones para proteger la seguridad de los usuarios de la carretera.

En septiembre de 2020, la Comisión Europea publicó un informe elaborado por un grupo independiente de expertos sobre la ética de los vehículos conectados y automatizados. El informe incluye 20 recomendaciones que abarcan situaciones conflictivas, la creación de

[486] C-ITS Point of Contact. Recuperado de: https://cpoc.jrc.ec.europa.eu/index.html. Consultado el 02.12.2020.
[487] *Good Practices for Security Of Smart Cars*. ENISA. 2019.
[488] *Cyber Security and Resilience of smart cars - Good practices and recommendations*. ENISA. 2016.
[489] RAPOSO, M.A. ET AL. (2019). *The future of road transport*. EUR 29748 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-14318-5. DOI:10.2760/668964. 2019
[490] Los miembros de CAMSec son fabricantes de vehículos centrados en la ciberseguridad, proveedores y desarrolladores de hardware/software integrado para coches inteligentes, asociaciones y organizaciones sin ánimo de lucro relacionadas con la seguridad de los vehículos, autoridades viales y universidades, así como organismos de normalización y responsables políticos.

una cultura de la responsabilidad y la promoción de la alfabetización en materia de datos, algoritmos e inteligencia artificial mediante la participación pública.

ENISA publicó posteriormente, en diciembre de 2020, un informe bajo el título *AI Cybersecurity Challenges*[491] -Retos de Ciberseguridad de la inteligencia artificial- que sitúa la ciberseguridad y la protección de datos como los pilares esenciales para la creación y sustento de un ecosistema de inteligencia artificial seguro y confiable. Este informe fue analizado por mi parte al abordar el capítulo II de esta investigación.

Recientemente, ENISA y el *Joint Research Centre* de la Comisión Europea -JRC- publicaron el pasado 11.02.2021 un informe sobre los retos de la inteligencia artificial en la conducción autónoma bajo el título: *Cybersecurity challenges in the uptake of artificial intelligence in autonomous driving*[492]. Se trata de un informe sectorial orientado a esta aplicación de la inteligencia artificial, que fue analizado en el capítulo II de esta investigación.

Por otra parte, tanto España como la UE ha tenido una prolífica actividad legislativa en los últimos tres años en materia de aeronaves no tripuladas (drones), si bien, no abordando aquéllas gobernadas o gestionadas por un sistema inteligente, conforme trataré en el capítulo V de esta investigación.

No obstante, como he referido anteriormente y se significa en otros capítulos de esta investigación, junto a todas estas iniciativas, la UE dispone de recientes marcos reguladores generales de la seguridad y de la privacidad que se han constituido en referencia internacional.

La protección de la privacidad y la información personal de los usuarios de este tipo de vehículos se aborda en el Reglamento General de Protección de Datos (RGPD) de la UE.

---

[491] *AI Cybersecurity Challenges*. Diciembre 2020. Recuperado de: https://www.enisa.europa.eu/publications/artificial-intelligence-cybersecurity-challenges. Consultado el 22.02.2021

[492] *Cybersecurity challenges in the uptake of artificial intelligence in autonomous driving*. UE 2021. Recuperado de: https://www.enisa.europa.eu/publications/enisa-jrc-cybersecurity-challenges-in-the-uptake-of-artificial-intelligence-in-autonomous-driving. Consultado el 22.02.2021.

La Directiva de Seguridad de las Redes y de la Información (NIS) también aborda los problemas de ciberseguridad de los vehículos, ya que pretende proporcionar medidas de seguridad genéricas para mejorar la ciberseguridad en toda la UE. Actualmente, como abordé en el capítulo II de esta investigación, se está tramitando la revisión de aquella mediante una nueva Propuesta de Directiva denominada NIS 2.

Por último, la UE ha dado un paso firme para la regulación de la inteligencia artificial.

En primer lugar, mediante las precitadas Resoluciones de su Parlamento de 20 de octubre de 2020 para la regulación de la inteligencia artificial en el ámbito ético, jurídico, de responsabilidad civil por daños y propiedad intelectual, que incluyen su Propuesta de Reglamento del Parlamento Europeo y del Consejo, sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, así como su Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial. La primera de las propuestas fue analizada en el capítulo III y la segunda será objeto de análisis en el capítulo siguiente.

En segundo lugar, mediante su Propuesta de Reglamento de 21 de abril de 2021 objeto de posterior análisis en este capítulo.

Sin perjuicio de estas propuestas, debe considerarse otros marcos jurídicos generales o sectoriales de la UE que afectan igualmente a la aplicación y uso de la inteligencia artificial, por ejemplo, el Reglamento (UE) 2019/1150 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, sobre el fomento de la equidad y la transparencia para los usuarios profesionales de servicios de intermediación en línea[493], que establece un conjunto de normas en materia de transparencia, equidad y de reclamación a los usuarios profesionales de servicios de intermediación en línea y a los usuarios de sitios web corporativos en relación con los motores de búsqueda en línea.

[493] OJ L 186, 11.7.2019. Pp. 57-79

En conclusión, la gran mayoría de países más avanzados han definido sus estrategias en materia de inteligencia artificial y han iniciado sus primeros pasos en su regulación.

La UE pretende sumarse al liderazgo internacional de la inteligencia artificial y, en especial, liderar su regulación para garantizar su seguridad y confiabilidad, significando sus recientes propuestas reguladoras relacionadas con sus aspectos éticos, responsabilidad jurídica por daños y seguridad, que pretenden constituirse en referencia internacional.

3. Iniciativas y propuestas reguladoras europeas

El desarrollo de la sociedad y la economía digital crece a una velocidad en constante aumento, en paralelo al desarrollo de la tecnología, la facilidad de acceso y disponibilidad de la misma, aumento de las capacidades de computación y automatización, y a mayor disponibilidad cualitativa y cuantitativa de datos, lo que está igualmente permitiendo un crecimiento exponencial de su uso, aplicación y explotación, y de las economías y negocios sustentados en todo ello.

No obstante, la ausencia de marcos normativos que contemplen todas estas nuevas realidades y su complejidad, nos obliga como juristas a realizar importantes esfuerzos de interpretación de los existentes para su aplicación y de creación de marcos jurídicos adecuados vía contractual para regular adecuadamente los nuevos marcos de relación e interrelación entre personas y con/entre sistemas.

Las situaciones y conflictos que se plantean hoy relacionadas con la tecnología deben ser resueltas, en ocasiones, mediante la aplicación de marcos normativos nacionales aprobados hace dos siglos, como el Código Civil español y, en especial, instituciones contempladas en este que vienen del Derecho Romano.

En mi opinión, todo esto nos debe llevar a una reflexión optimista y positiva, y no es otra que la técnica legislativa del legislador de entonces tiene características que deberíamos tener en consideración hoy, dado que todavía hoy nos permite resolver conflictos

relacionados con las tecnologías más vanguardistas y con las corporaciones privadas que gobiernan en el *ciberespacio*.

La legislación civil y mercantil vigente en España no contempla realidades como *app*, *web* o *videojuego*, y mucho menos conceptos como *Blockchain*, inteligencia artificial o *cloud computing*, a pesar de que, por ejemplo, algunas de estas tecnologías, servicios o conjunto de tecnologías existan bajo dicha denominación desde hace más de 65 años.

En tecnología "el hoy, es ayer" y "el ahora, es antes". En apenas 15 años hemos vivido más innovación que en los 150 anteriores y en ese siglo y medio, más que en los últimos 150.000 años[494]. Esto no ocurre con el Derecho.

El legislador actual debe ir más allá y legislar sobre tendencias y no sobre novedades. Lo que hoy es una novedad, significa que ya está en el mercado, sin embargo, debemos pensar en lo que será novedad mañana. Y debe aplicar estrategia para identificar los objetivos, revisar los marcos existentes y crear otros nuevos que adecuen, actualicen, complementen y se integren con aquéllos, evolutivos, adaptativos, flexibles e inteligentes. Aplicar estrategia para la definición de la mejor técnica legislativa para alcanzar los objetivos propuestos.

El Derecho no puede convertirse en una fotografía de una realidad pasada, si no que el contexto es cambiante, interactivo y en constante desarrollo, y cada vez más rápido, por lo que cualquier marco normativo deberá ser lo suficientemente general, dinámico, dinamizador, flexible y adaptativo para prever el tratamiento en su marco de las realidades que vendrán. En definitiva, lo que en mi opinión considero un marco regulativo "*responsive*".

A mi juicio, las claves de una buena regulación de nuestro futuro más inmediato pasan por legislar sobre tendencias más que sobre novedades y desde una perspectiva global, con sustento en la ética y ahora más que nunca, ante la irrupción de la tecnología en todos los ámbitos de nuestra vida como seres humanos, especialmente la de alto impacto.

[494] VIDAL, M. (2019). *La era de la humanidad*. Versión Kindle. Op.cit. Pos. 2736.

En definitiva, analizar, reflexionar y legislar con una visión estratégica, pensar de forma global, buscar el máximo consenso y actuar de forma local mientras no se dispongan de otros instrumentos, en base a las características propias del *ciberespacio*, de las realidades complejas que plantean determinadas tecnologías y su interacción, de la sociedad y economía global y mundial de la que formamos parte y en la que operamos, y partiendo de los conceptos vigentes y limitaciones inherentes de "ordenamiento jurídico" y "organizaciones supranacionales".

Y a todo ello adicionar, otros instrumentos asociados que pueden contribuir a la agilidad y efectividad del Derecho, como la promoción de la autorregulación de la industria.

Si recupero de nuevo una de las frases más célebres que Stephen Hawking dejó para la posteridad, entre otras muchas que me he permitido citar a lo largo de esta investigación, "la inteligencia es la habilidad de adaptarse a los cambios*"*, es obvio que ahora más que nunca precisamos una legislación "inteligente" para una realidad "inteligente", ante la revolución que la inteligencia artificial ya ha supuesto y el cambio constante que va a suponer en todos los ámbitos, especialmente en la economía, la sociedad y el mercado laboral.

Como he expuesto, la UE está liderando en este momento, no tanto el desarrollo y despliegue de la IA a nivel mundial, a lo que sin duda está contribuyendo enormemente, sino a su regulación y, a mi juicio, tiene la oportunidad ante sí, como he puesto de manifiesto al abordar anteriores aspectos, de establecer un nuevo marco que puede constituirse en referencia mundial para su adopción o "iluminación" para los países que igualmente están trabajando ya en distintas iniciativas legislativas en materia de inteligencia artificial, como lo fue y está siendo en materia de protección datos, el Reglamento General de Protección de Datos (RGPD o GDPR por siglas en inglés).

El *Libro blanco sobre la inteligencia artificial* de la Comisión Europea, analizado en el capítulo III de esta investigación, abordó y sentó las bases de los futuros marcos reguladores de la inteligencia artificial en la UE, desde un enfoque de excelencia y confianza.

La UE dispone un amplio volumen de legislación en materia de seguridad de los productos y responsabilidad civil, tanto general -Directiva 2001/95/CE relativa a la seguridad general de los productos[495]- como sectorial, completada por la legislación nacional de sus distintos Estados miembros.

A la misma debemos adicionar otros marcos en materia de protección de derechos fundamentales y derechos de los consumidores, como el Reglamento General de Protección de Datos (RGPD o GDRP por sus siglas en inglés) y otras normas sectoriales en materia de privacidad, la Directiva sobre igualdad racial[496], la Directiva sobre igualdad de trato en el empleo y la ocupación[497], las Directivas relativas a la igualdad de trato entre mujeres y hombre con relación al empleo y el acceso a los bienes y servicios[498], Directiva sobre las prácticas comerciales desleales[499] y la Directiva sobre los derechos de los consumidores[500], así marcos de ciberseguridad, accesibilidad, propiedad intelectual e industrial, o de protección del secreto de empresa.

Estos marcos resultan potencialmente aplicables a los sistemas dotados de inteligencia artificial, si bien, el *Libro blanco* precitado plantea la necesidad de evaluar si son adecuados para abordar los retos que plantea la inteligencia artificial que he comentado anteriormente o si, por el contrario, deberían crearse instrumentos específicos.

La Comisión Europea ha considerado que debe mejorarse el marco normativo para acometer los riesgos, y así lo hace constar en el precitado *Libro blanco sobre la inteligencia artificial*.

De un lado, considera que las características de la inteligencia artificial comportan dificultades para garantizar la correcta aplicación y ejecución de la legislación nacional y

[495] Directiva 2001/95/CE del Parlamento Europeo y del Consejo, de 3 de diciembre de 2001, relativa a la seguridad general de los productos. Diario Oficial n° L 011 de 15/01/2002 P. 0004 – 0017.
[496] Directiva 2000/43/CE.
[497] Directiva 2000/78/CE.
[498] Directiva 2004/113/CE y Directiva 2006/54/CE
[499] Directiva 2005/29/CE
[500] Directiva 2011/83/UE del Parlamento Europeo y del Consejo, de 25 de octubre de 2011, sobre los derechos de los consumidores, por la que se modifican la Directiva 93/13/CEE del Consejo y la Directiva 1999/44/CE del Parlamento Europeo y del Consejo y se derogan la Directiva 85/577/CEE del Consejo y la Directiva 97/7/CE del Parlamento Europeo y del Consejo. OJ L 304, 22.11.2011, Pp. 64-88.

de la UE, especialmente ante su falta de transparencia y opacidad que dificulta detectar y demostrar posibles incumplimientos regulativos, especialmente en el ámbito de la responsabilidad y para el resarcimiento efectivo.

Por ello, considera que puede resultar necesario "adaptar o clarificar" la legislación vigente en algunos sectores, en especial, en materia de responsabilidad civil, para garantizar la aplicación y ejecución efectiva de la legislación precitada.

La respuesta a estas cuestiones, por el momento, fue la Resolución del Parlamento Europeo de 20 de octubre de 2020 que integra una Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial.

De otro, considera que la legislación general de la UE en vigor en materia de seguridad de los productos, tiene un ámbito de aplicación restringido, esto es, a los productos, no a los servicios, por lo que inicialmente no se aplicaría a los servicios basados en inteligencia artificial.

Los programas informáticos que integran un producto deben respetar las normas de seguridad del mismos, si bien, cuando nos encontramos ante un programa autónomo, se plantea la duda jurídica de si se le aplicaría el marco regulativo de seguridad de los productos de la UE, salvo en aquellos sectores que disponen de normas explícitas, como el Reglamento sobre los productos sanitarios[501], en la que los programas informáticos destinados a fines médicos por el fabricante se consideran productos sanitarios. Asimismo, pensemos por ejemplo en las soluciones basadas en sistemas inteligentes como servicio, prestados desde infraestructuras o plataformas *cloud.*

La nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, debería disipar esas dudas y resolver todas las cuestiones precitadas, incluyendo su armonización, si bien, no incluye sorprendentemente entre los

[501] Reglamento (UE) 2017/745 del Parlamento Europeo y del Consejo de 5 de abril de 2017, sobre los productos sanitarios, por el que se modifican la Directiva 2001/83/CE, el Reglamento (CE) n.o 178/2002 y el Reglamento (CE) n.o 1223/2009 y por el que se derogan las Directivas 90/385/CEE y 93/42/CEE del Consejo.

sistemas de riesgo inaceptable o de alto riesgo, algunos sistemas inteligentes en el ámbito médico o de la investigación biomédica, entre otros.

De otro, el *Libro blanco sobre inteligencia artificial* considera igualmente que la legislación vigente se centra sobre todo en los riesgos de seguridad en el momento de la comercialización, si bien, la integración de programas informáticos en los productos, puede modificar el funcionamiento de los mismos y sistemas asociados a lo largo de todo su ciclo de vida, especialmente evidente en el caso de sistemas con actualizaciones frecuentes o basados en el aprendizaje automático y cierto grado de autonomía, lo que puede comportar nuevos riesgos que no existían en el momento en que se introdujo el producto o sistema en el mercado.

Estos contextos considero que requieren la revisión y complemento de los marcos vigentes y especialmente en materia de responsabilidad, dado que podría conllevar su inaplicación o exoneración de responsabilidad, conforme analizaré en el próximo capítulo. De nuevo, la nueva Propuesta de Reglamento sobre inteligencia artificial de 21 de abril de 2021 podría dar respuesta a estas cuestiones, al menos parcialmente en la medida que no aborda cuestiones como la responsabilidad civil, al exigir un conjunto de requisitos y obligaciones durante todo el ciclo de vida del sistema, conforme analizaré posteriormente.

De otro, los marcos vigentes sobre seguridad de los productos suscitan cierta inseguridad e incertidumbre en relación con la imputación de responsabilidades entre los distintos agentes económicos que intervienen en la cadena de suministro, en la medida que imputan la responsabilidad al productor del "producto" comercializado, incluyendo todos sus componentes, como los sistemas de inteligencia artificial.

La dificultad de aplicación se plantea cuando el producto que integra la inteligencia artificial es comercializado por alguien que no es el productor. Y además de todo ello, la legislación de la UE sobre responsabilidad civil por los productos contempla la responsabilidad de los productores dejando a las normas nacionales en materia de responsabilidad la determinación de la responsabilidad de los demás participantes en la cadena de suministro. En este sentido, la solución inicial a estas cuestiones debería construirse en base a la precitada Resolución del Parlamento Europeo de 20 de octubre

de 2020 que integra su Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial.

Y, por último, la inteligencia artificial puede plantear riesgos en materia de seguridad que la legislación de la UE no aborda de manera específica -tanto en el momento de su comercialización como posteriormente ante actualizaciones o derivados del aprendizaje automáticos-, especialmente en materia de *ciberamenazas*, pérdida de conectividad o contra la seguridad personal, especialmente en relación con nuevos usos de los sistemas inteligentes, por ejemplo, en el caso de aparatos domésticos.

De nuevo, la nueva Propuesta de Reglamento sobre inteligencia artificial de 21 de abril de 2021 debería ser la base para la resolución de estas cuestiones y, conforme analicé en el capítulo II, contempla expresamente las obligaciones de ciberseguridad de los sistemas inteligentes, si bien, de manera general y exclusivamente respecto de sistemas considerados conforme al mismo de alto riesgo.

El *Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica* adjunto al *Libro Blanco sobre la inteligencia artificial*, sin embargo, concluye que la legislación vigente sobre la seguridad de los productos ya recoge una protección de la seguridad amplia frente a todo tipo de riesgos derivados de un producto en función de uso, si bien, considera viable la introducción de disposiciones que aborden de manera explícita los nuevos riesgos derivados de las tecnologías digitales emergentes, a fin de ofrecer mayor seguridad jurídica.

El mismo significa aspectos como el requerimiento de la supervisión humana como garantía desde la fase de diseño y en todo el ciclo de vida de los productos y sistemas de inteligencia artificial, evaluaciones de riesgos durante dicho ciclo de vida, especialmente ante los cambios con origen en comportamientos autónomos, la posibilidad de establecer obligaciones explícitas para los productores en relación con los riesgos para la salud mental de los usuarios, la exigencia de transparencia contra la opacidad, la posibilidad de incorporar requisitos específicos para abordar los riesgos de datos incorrectos en la fase de diseño, así como mecanismos para garantizar la calidad de los datos durante su ciclo

de vida, adaptar y clarificar las normas vigentes para dar cabida a los supuestos en los que los programas o sistemas inteligentes son comercializados separadamente a un producto o descargados en un producto tras su comercialización cuando tengan repercusiones de seguridad y, por último, la posibilidad de modificar la Directiva sobre responsabilidad por los daños causados por productos defectuosos o de adaptar la medidas nacionales en materia de responsabilidad civil al objeto de evitar la eficacia de los marcos reguladores de la misma.

Si analizamos la mayoría de las actuaciones regulatorias propuestas hasta la fecha en el ámbito de la UE, se construyen sobre la conversión de principios y normas éticas esenciales en requerimientos y obligaciones legales para abordar los riesgos adecuadamente, dotando de eficacia los marcos éticos en cuya consolidación y consenso ha trabajado duramente la UE los últimos años.

La Comisión Europea concluye en su *Libro Blanco* que, "además de las posibles adaptaciones de la legislación vigente, puede que se requiera nueva legislación específica sobre la inteligencia artificial, a fin de adaptar el marco jurídico de la UE a la evolución tecnológica y comercial actual y futura". Y con este propósito establece las bases de ese futuro marco regulador de la UE, conforme ya expuse en el capítulo III:

a) Ámbito de aplicación concreto y una definición flexible de inteligencia artificial, conforme analicé al abordar este concepto, para adaptarse al progreso técnico, pero a su vez precisa, para ofrecer la seguridad jurídica necesaria.

b) Equilibrio entre eficacia para alcanzar sus objetivos y ductibilidad para no ser excesivamente prescriptivo, para lo que considera debe seguirse un enfoque basado en el riesgo que garantice una intervención regulatoria proporcionada. En este sentido, propone una definición clara y fácil de entender y aplicar de los sistemas de riesgo elevado o no, considerando que este riesgo debe medirse en función de los bienes y derechos que puedan verse afectados y considerando si el sector y usos previstos comportan riesgos significativos, en especial desde la perspectiva de la seguridad, los derechos fundamentales y de los consumidores.

La Comisión define ya en el *Libro Blanco* cuando una aplicación debe considerarse de riesgo elevado, en particular, cuando se den los siguientes requisitos: a) La aplicación de inteligencia artificial "se emplee en un sector en el que, por las características o actividades que se llevan a cabo normalmente, es previsible que existan riesgos significativos", por ejemplo sanidad, transporte, energía y determinados ámbitos del sector público; b) La aplicación de inteligencia artificial empleada en dicho sector se use, además de manera que puedan surgir riesgos significativos, en la medida que no toda utilización de la inteligencia artificial en los sectores en cuestión implica necesariamente riesgos significativos.

A mi juicio constituye una visión parcial y sectorial de la inteligencia artificial para la valoración de riesgo, basada en el sector donde opere.

No obstante, como también se recoge en el documento, puede haber casos excepcionales en los que debido a lo que esté en peligro, el uso de aplicaciones de inteligencia artificial para determinados fines se considera de riesgo elevado con independencia del sector donde opere, por ejemplo, sistemas utilizados en los procesos de selección y contratación de personal, identificación biométrica remota u otras tecnologías de vigilancia intrusiva.

c) Determinación de los requisitos legales obligatorios, entre los que sugiere para aplicaciones de riesgo elevado, siguiendo las directrices del grupo de expertos de alto nivel, los siguientes: Garantizar el respeto de los valores y normas de la UE, especialmente de seguridad y privacidad en los datos utilizados para el entrenamiento, adecuada conservación de registros y protección de los datos, transparencia también en el suministro de información, solidez y exactitud (que incluyan su reproductibilidad, mecanismos ante errores o incoherencias durante su ciclo de vida y resiliencia), supervisión humana para garantizar una inteligencia artificial fiable, ética y antropocéntrica (que puede incluir que el resultado del sistema inteligente no sea efectivo hasta que un humano no lo haya revisado y validado o que sea efectivo, pero pueda ser revisado, así como, también, realización de seguimiento del sistema durante su funcionamiento y, si es posible, intervenir en tiempo real y desactivarlo e imponer restricciones operativas en el sistema en su fase de diseño).

d) Determinación de los requisitos específicos en el caso de identificación biométrica.

e) Identificación de los sujetos obligados y destinatarios de los requisitos, en la medida que hay diversas partes involucradas en el ciclo de vida de un sistema de inteligencia artificial, entre otras, el desarrollador, el implementador[502], el productor, distribuidor, importador, proveedor de servicios, usuario particular o profesional.

   No obstante, respecto de los sujetos que enumera, el concepto de "implementador" lo define como "la persona que utiliza un producto o servicio provisto de inteligencia artificial". No obstante, este concepto me parece excesivamente genérico y confuso respecto de la condición de usuario de un producto o servicio, por lo que considero que podría ser más adecuado asociarlo como tal a la persona que habilita o pone en funcionamiento un producto o servicio dotado de inteligencia artificial para uso propio (en cuyo caso tendrá la condición de usuario) o para su ofrecimiento o uso por tercero.

   Asimismo, no se incluye la figura del "operador" del sistema inteligente, cuya relevancia para el cumplimiento de los requisitos de los futuros marcos puede ser esencial, y que podríamos definir como la persona que mantiene en funcionamiento operativo un sistema inteligente y que podría asumir su mantenimiento preventivo, detectivo, correctivo, evolutivo o de seguridad ante comportamientos anómalos o no previstos.

   La Comisión Europea consideró, conforme a su enfoque de riesgos, que cada obligación debe dirigirse a la/s persona/a que esté/n en mejor posición para gestionar cada riesgo potencial.

   Estos aspectos son especialmente relevantes para la imputación de cualquier incumplimiento y exigencia de cualquier responsabilidad, sin perjuicio de la parte

que deba asumir la responsabilidad directa de los daños causados conforme a los marcos de responsabilidad civil vigentes.

f) Control objetivo previo de cumplimiento eficaz y evaluaciones de conformidad, por ejemplo, a través de procedimientos de ensayo, inspección y certificación, sin perjuicio de la supervisión del cumplimiento y de la posterior ejecución por parte de las autoridades nacionales competentes.

g) Garantías para la tutela judicial efectiva para las partes afectadas negativamente por los sistemas inteligentes.

h) Sistemas de etiquetado voluntario para las aplicaciones que no se consideran de riesgo elevado.

i) Definición de una estructura de gobernanza europea sobre la inteligencia artificial en forma de marco de cooperación de las autoridades nacionales competentes, especialmente para evitar la fragmentación de responsabilidades.

El enfoque europeo sobre inteligencia artificial reflejado en el *Libro Blanco* pretende promover de un lado, seguridad y confiabilidad, y de otro, la capacidad de innovación de la UE y el desarrollo de una inteligencia artificial ética y fiable en toda su economía, lo que significa que los futuros marcos normativos deben regular una inteligencia artificial ética y segura, elevando los actuales principios y normas éticas, altamente consensuados a nivel internacional, a requerimiento jurídico.

En congruencia con estas reflexiones, la UE abordó la normativización e imperatividad de determinados principios y normas éticas consideradas esenciales, así como los aspectos de seguridad, responsabilidad civil y propiedad intelectual asociados a la inteligencia artificial.

No obstante, por el camino, varios Estados miembros comenzaron a valorar distintas alternativas en su ordenamiento jurídico interno para hacer frente a los retos de la inteligencia artificial, lo que a su vez pueden generar un riesgo de fragmentación del mercado único, lo que la UE pretende evitar mediante su armonización.

Por lo que se refiere al marco ético, conforme he analizado en el capítulo III de esta investigación, el Parlamento Europeo instó a la Comisión a adoptar un nuevo marco jurídico que contemple y desarrolle los principios éticos y las obligaciones jurídicas asociadas al desarrollo, despliegue, implantación y uso de la inteligencia artificial, robótica y otras tecnologías relacionadas en la UE.

El futuro marco ético europeo vinculante, conforme recogía la Resolución del Parlamento Europeo, de 20 de octubre de 2020 sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas[503] -objeto de análisis en el anterior capítulo-, debería construirse sobre la base de principios y normas éticas de obligado cumplimiento, entre otros: Una inteligencia artificial sujeta al control y supervisión humana, seguridad, fiabilidad, transparencia, rendición de cuentas, salvaguardias contra el sesgo y la discriminación, derecho de reparación, responsabilidad social y medioambiental y respeto de la intimidad y la protección de los datos. La supervisión humana se concibe como una exigencia en este nuevo marco en determinados contextos, especialmente cuando se trate de sistemas de inteligencia artificial de riesgo alto, si bien, su posterior regulación y exigencia difiere en la última propuesta reguladora, circunscrita a los sistemas inteligentes considerados de alto riesgo.

Por lo que se refiere a la responsabilidad civil en materia de inteligencia artificial, el Parlamento Europeo ha establecido las bases del marco regulador en la Propuesta de Reglamento integrada en su Resolución de 20 de octubre de 2020 sobre un régimen de responsabilidad civil en materia de inteligencia artificial [504] con las correspondientes recomendaciones a la Comisión Europea.

Esta propuesta regulatoria se elaboró pensando en el futuro y su adaptación a las novedades de la tecnología que vengan en lo sucesivo, desde un enfoque general y adaptativo. La misma tiene un doble objetivo, de un lado proporcionar seguridad y confianza a la ciudadanía en la tecnología, en particular, en el desarrollo, despliegue y aplicación de la inteligencia artificial asegurando un resarcimiento efectivo para quien

[503] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))
[504] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))

sufra un daño como consecuencia de la misma, y de otro lado, promover la innovación y dotar de seguridad jurídica a las empresas.

El marco regulador propuesto en las Resoluciones precitadas se aplicaría, en caso de tramitación y aprobación, a cualquier actividad de inteligencia artificial, física o virtual, que provoque daños o perjuicios a la vida, la salud, la integridad física o la propiedad de personas físicas o jurídicas, o que provoque daños inmateriales significativos que den lugar a una pérdida económica verificable.

Asimismo, por lo que se refiere a los derechos de propiedad intelectual e industrial relacionados con los sistemas de inteligencia artificial, la UE está apostando por construir un sistema de derechos de propiedad intelectual efectivo y un sistema de patentes que protejan la innovación dentro de la Unión, si realmente pretende tener el liderazgo al que aspira.

En este sentido, el Parlamento Europeo aprobó su Resolución de 20 de octubre de 2020 sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial[505].

En relación con todas estas cuestiones, el Parlamento Europeo también volvió a analizar y se pronunció respecto de la posibilidad de que los sistemas de inteligencia artificial pudieran llegar a tener personalidad jurídica, negando esta posibilidad de manera taxativa, al considerar que solo los seres humanos pueden ser titulares de los derechos de protección intelectual, diferenciando entre creaciones humanas con ayuda de la inteligencia artificial y creaciones generadas directamente por la inteligencia artificial.

Adicionalmente, el Parlamento Europeo también está abogando para que los futuros marcos contemplen también aspectos relacionados con los secretos comerciales, el uso de algoritmos, los contenidos falsos (*deep fakes*) o la distorsión informativa.

Sobre este último aspecto, la Propuesta de Reglamento europeo que incorpora la Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre un marco de los

[505] Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial (2020/2015(INI))

aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas precitada, incluye la "no interferencia" o "manipulación informativa" como norma ética exigible a los sistemas de inteligencia artificial de alto riesgo.

Además de las Resoluciones precitadas, debo adicionar la *Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre inteligencia artificial: Cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal*[506], en la que el Parlamento Europeo abogó por sistemas que permitan un alto nivel de control humano sobre los sistemas de inteligencia artificial, de modo que se tenga en todo momento los medios para corregir su curso, detenerla o desactivarla en caso de comportamiento imprevisto, intervención accidental, ciberataque o interferencia de terceros con tecnología pasada en inteligencia artificial.

El Parlamento Europeo hace un llamamiento en la misma para promover un marco global sobre el uso militar de la inteligencia artificial a nivel internacional y defiende que los "Sistemas autónomos armamentísticos letales" -SAAL- únicamente sean empleados como último recurso y sólo sean lícitos si están sujetos a un control humano estricto y, consiguientemente un control "significativo", de modo que los sistemas sin control ni supervisión humana alguna, deben ser prohibidos sin excepciones.

Del mismo modo, en relación con el uso de la inteligencia artificial en el sector público, el Parlamento Europeo aboga para que el incremento del uso de sistemas de inteligencia artificial en salud y justicia nunca debería substituir al contacto humano, y que todo individuo debería ser informado cuando una decisión está siendo tomada por una inteligencia artificial y tener la opción de una segunda opinión.

Del mismo modo, significa la necesidad de proteger los datos de pacientes en el uso de sistemas inteligentes en el ámbito de la salud pública y la igualdad de trato, así como

[506] Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre inteligencia artificial: cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal (2020/2013(INI))

garantizar que las decisiones judiciales finales en el ámbito de la justicia sean tomadas por humanos, ser verificadas por el mismo y sujetas a un proceso con todas las garantías.

En relación con la vigilancia masiva y *deep fakes*, el Parlamento Europeo reitera en esta Resolución su preocupación por la amenaza que ello supone para los derechos humanos y la soberanía estatal, instando a que se prohíba a las autoridades públicas el uso de "aplicaciones de calificación social masiva altamente intrusivas" que permitan controlar y calificar a los ciudadanos, lo que ha tenido su reflejo en la última Propuesta de Reglamento sobre inteligencia artificial de 21 de abril de 2021. Conforme analizaré posteriormente, esta nueva Propuesta de Reglamento prohíbe directamente estos sistemas.

Adicionalmente, se tramitó la iniciativa sobre *La inteligencia artificial en derecho penal y su uso por parte de las autoridades policiales y judiciales en materia penal*, pendiente de decisión de la Comisión de Libertades Civiles, Justicia y Asuntos de Interior, y, durante la revisión de esta investigación previa a su cierre, el Parlamento Europeo adoptó la Resolución del Parlamento Europeo, de 19 de mayo de 2021, sobre inteligencia artificial en la educación, la cultura y el sector audiovisual[507].

Asimismo, el Parlamento Europeo ha venido trabajando en paralelo en muchos otros aspectos relacionados con la inteligencia artificial, por ejemplo, en su uso civil y militar a través de su Comisión de Asuntos Jurídicos, su uso en la educación, cultura y sector audiovisual, a través de la Comisión de Cultura, o su uso en el Derecho Penal, a través de la Comisión de Libertades Civiles.

Por lo que se refiere a nuevas propuestas reguladoras, por último, se publicó la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, que será objeto de análisis en los posteriores apartados.

[507] Resolución del Parlamento Europeo, de 19 de mayo de 2021, sobre inteligencia artificial en la educación, la cultura y el sector audiovisual (2020/2017 (INI)).

A nivel nacional, en España, merece destacarse la reciente y precitada *Carta de Derechos Digitales*[508] elaborada por la Secretaría de Estado de Digitalización e inteligencia artificial, del Ministerio de Asuntos Económicos y Transformación Digital, que recoge en su apartado 5. XXV los Derechos ante la inteligencia artificial.

Los derechos que regulan se basan en los principios y normas éticas esenciales sobre las que se han basado los trabajos, resoluciones y propuestas de la UE hasta la fecha.

En particular, establece que la inteligencia artificial deberá asegurar un enfoque centrado en la persona y su inalienable dignidad, perseguirá el bien común y asegurará cumplir con el principio de no maleficencia. Lo que constituye ya un derecho y una correlativa obligación para diseñadores, desarrolladores y productores de sistemas inteligentes, sin diferenciar niveles de riesgo de los sistemas que conciban.

Considero que, conforme a mis opiniones manifestadas a lo largo de esta investigación, no se podría empezar mejor esta declaración, si bien, veremos como todo ello se hace efectivo en la práctica, en función de las técnicas que se utilicen para abordar los retos y riesgos de la inteligencia artificial, especialmente su regulación y su alcance.

Asimismo, merece destacarse que la *Carta de Derechos Digitales* exige también que, en el desarrollo y ciclo de vida de los sistemas de inteligencia artificial, se garantice el derecho a la no discriminación -cualquiera que fuera su origen, causa o naturaleza, en relación con las decisiones, uso de datos y procesos basados en inteligencia artificial-, que se establezcan condiciones de transparencia, auditabilidad, explicabilidad, trazabilidad, supervisión humana y gobernanza -y en cualquier caso, la información facilitada deberá ser accesible y comprensible- y, por último, que se garantice la accesibilidad, usabilidad y fiabilidad.

Y, por último, también contempla el derecho de las personas a solicitar una supervisión e intervención humana y a impugnar las decisiones automatizadas tomadas por sistemas de

[508] Recuperado de: https://www.lamoncloa.gob.es/presidente/actividades/Documents/2021/140721-Carta_Derechos_Digitales_RedEs.pdf.

inteligencia artificial que produzcan efectos en su esfera personal y patrimonial, alienado con el derecho ya contemplado en el ámbito de la privacidad, en el RGPD.

La definición y reconocimiento formal de estos derechos es un importante paso adelante, si bien, el mayor desafío se plantea ahora para garantizar su eficacia y respeto en el mercado, manteniendo el equilibrio entre todos los intereses en juego y conforme a los marcos jurídicos vigentes y los nuevos que deberán construirse considerando los derechos e intereses de todos los agentes participantes en el mercado, la innovación, el desarrollo y la competitividad, para alcanzar el deseable y complejo equilibrio y armonización.

4. Innovación y competitividad frente al marco regulador

Como he expuesto anteriormente, tanto en el capítulo II de esta investigación en relación con los retos y riesgos de la inteligencia artificial, así como en otros apartados de la misma en los que he abordado su regulación, uno de los principales riesgos que plantea es el de su regulación excesiva.

El riesgo de una hiperregulación puede afectar a la competitividad empresarial en el ámbito de la UE, especialmente al sector desarrollador, comercializador y aplicador de la inteligencia artificial, por lo que debe mantenerse el equilibrio al que he hecho referencia en distintos puntos de esta investigación, entre la seguridad y la confiabilidad, de un lado, y el avance, la innovación y la competitividad empresarial, de otro.

Este equilibrio se haya en constante tensión ante el desarrollo, despliegue y aplicación incesante de la inteligencia artificial en nuevos ámbitos y sectores, la potenciación de sus capacidades asociadas y relacionadas, relación e interrelación con otras tecnologías, el avance científico incesante y los potenciales riesgos asociados de alto impacto.

Salvaguardar ese equilibrio con medidas adicionales podría posicionar a la industria europea de la inteligencia artificial a la vanguardia en seguridad y confiabilidad de sus sistemas, conforme ya lo es en otros sectores, y compensar la inversión de la industria en seguridad y fiabilidad con una ventaja competitiva.

Sin embargo, no definir y sustentar adecuadamente este equilibrio podría hacer que la industria europea compita con desigualdad de armas en el mercado mundial, restringiendo su mercado al europeo y a aquellos países que apuesten por el enfoque y modelo de la UE.

Del mismo modo, se podría producir una cierta "competencia desleal" de la industria con origen en otros países de fuera de la UE, que podrían ofrecer sistemas inteligentes que pueden proporcionar soluciones a problemas actuales, satisfacer necesidades, mejorar la vida e incluso salvarla, pero que no cumplirían los marcos y estándares éticos y marcos de seguridad europeos, dejando a un lado los jurídicos, lo que colocaría a los ciudadanos, empresas, Administraciones públicas e incluso a los propios gobiernos ante la disyuntiva de disponer o renunciar a los mismos a pesar de su benevolencia.

Las propuestas reguladoras de la UE en materia de inteligencia artificial recogen la necesidad y objetivo de salvaguardar este equilibrio.

5. Necesidad de un enfoque ético, jurídico y de seguridad

La imparable carrera por liderar la inteligencia artificial a nivel internacional por parte de Estados y grandes empresas empezó hace varios años sin tener un escenario ético y regulador definido ni plenamente consensuado a fecha actual, y su ritmo creciente y vertiginoso tanto en su desarrollo como en su despliegue y aplicación requieren marcos de referencia en ambos ámbitos -ético y jurídico-, pero también de seguridad.

Respecto a la seguridad, esta constituye ya un requerimiento ético y jurídico en distintos instrumentos reguladores generales como, por ejemplo, en materia de privacidad o servicios esenciales, y lo será en mayor medida en la UE en caso de aprobación final de las propuestas reguladoras presentadas por sus órganos legislativos en el ámbito de la inteligencia artificial.

La complejidad y enorme impacto de inteligencia artificial exige su acompañamiento de marcos de seguridad. El marco actual y preocupación por sus riesgos y amenazas ha sido analizado en el capítulo II.

Respecto de la dimensión ética, como he analizado en el capítulo III, durante estos últimos años se han ido consensuando a nivel internacional los principios y normas éticos básicas que debería cumplir cualquier elemento, sistema, tecnología, producto o servicio dotado de inteligencia artificial, e incluso la UE fue más allá ante la insuficiencia de la ética para garantizar seguridad y confianza en gobiernos, operadores y ciudadanos, mediante la propuesta de regulación legal de los principios y normas éticas esenciales orientadas a su carácter vinculante y exigencia obligatoria, y el establecimiento de un conjunto de obligaciones jurídicas asociadas.

De nuevo, la complejidad y enorme impacto de inteligencia artificial exige también su acompañamiento de marcos éticos sobre los que deben construirse los jurídicos, convirtiendo en vinculantes algunos de sus principios y normas. Además, deben asegurar que los derechos fundamentales prevalezcan a cualquier tecnología que se desarrolle ahora o en el futuro, de modo que la misma sea respetuosa con los mismos en su concepción, despliegue y funcionamiento[509].

Las leyes no son el único ni frecuentemente el mejor mecanismo para ajustar conductas y actividades a unos objetivos y bases pretendidas, si bien, en ocasiones, se hacen absolutamente necesarias para garantizarlo, pero no de manera aislada sino bajo un enfoque estratégico, "inteligente" y global.

Los mecanismos complementarios para conseguir estos ajustes son diversos, desde la educación, la autorregulación, la generación de incentivos, promoción de códigos de conducta o de buenas prácticas, definición de estándares o mecanismos de evaluación y certificación.

Como expuse en el apartado 2.2. al analizar la ausencia de regulación y la conveniencia de combinar el denominado *hard law* y el *soft law*, con su reconocimiento jurisprudencial, se evidencia la necesidad de actuar de manera global y combinada ante una realidad tan compleja y con tal alto impacto en todos los ámbitos. No es lo mismo regular aspectos de

[509] En este sentido, destacar el criterio y doctrina ya recogida en la Sentencia del 13.05.20214 (Sentencia Google), Asunto C-131/12, del Tribunal de Justicia de la Unión Europea, alineada con la previa anulación por el TJUE de la Directiva 2006/24/CE de conservación de datos por sentencia de 8 de abril de 2014, sobre la necesidad de protección de los derechos fundamentales, especialmente en materia de protección de datos, en el desarrollo de la economía y sociedad digital.

alto impacto en los derechos y bienes de los sujetos involucrados, incluso sean o no de carácter fundamental los primeros o esencial los segundos, que cualquier otro que pudiera dejarse a la autorregulación o *soft law*.

Los enormes retos éticos y jurídicos que plantea la inteligencia artificial y su complejidad a nivel internacional exigen, a mi juicio, adoptar un enfoque y actuación global desde distintos enfoques, disciplinas, medios y canales, que debe incluir mecanismos de consenso y de cooperación internacional, revisión y complementación de los marcos reguladores existentes, definición de estándares, mecanismos de evaluación y certificación, promoción de la autorregulación y códigos de conducta de la industria, políticas de incentivos al I+D+I en tecnología confiable y segura, la concienciación y la educación.

La creación de un marco regulador para la inteligencia artificial es, a mi juicio, necesario para adecuar y armonizar el marco vigente, especialmente a nivel de la UE, así como para dotar de seguridad jurídica a todas las partes implicadas en el diseño, desarrollo, despliegue y uso de esta, garantizar el respeto de los derechos fundamentales y promover una adecuada autorregulación e incentivos para su cumplimiento.

No obstante, debe ser proporcionada y considerar adecuadamente sus retos y riesgos, pero sin obstaculizar la inversión, el desarrollo, la innovación y las iniciativas empresariales, como expresamente recogen las nuevas propuestas europeas que serán analizadas a lo largo de esta investigación.

No es tarea fácil la creación de marcos reguladores que garanticen este equilibrio entre todos estos objetivos, que, además, se halla en constante tensión debido a los incesantes cambios cualitativos y cuantitativos en los retos y riesgos que plantea la inteligencia artificial, especialmente por el crecimiento exponencial de su desarrollo y aplicación, capacidades asociadas e interacción con otras tecnologías.

La definición de estos nuevos marcos debe partir desde dicha reflexión global, pero también integradora y armonizadora, identificando los aspectos del ordenamiento jurídico vigente que sea necesario regular, adaptar, mejorar, complementar, derogar y armonizar, y en tanto en el ámbito civil y de las distintas responsabilidades derivadas de la

inteligencia artificial, como en el administrativo o laboral, con el objeto de definir marcos sólidos que den soluciones adecuadas a la realidad de hoy como a la de mañana.

Asimismo, estos nuevos marcos deben sustentarse y acompañar a los marcos éticos, y prever mecanismos para garantizar y exigir su cumplimiento.

El temor por regular en exceso y obstaculizar la innovación, la inversión y el desarrollo de las economías locales es inevitable, pero debemos seguir avanzando para ofrecer un ecosistema equilibrado donde se garanticen todos los derechos e intereses en juego al objeto de garantizar una tecnología segura al servicio del ser humano.

Como he expuesto, el Derecho y la tecnología deben estar alineados para conseguir los objetivos estratégicos del ser humano, de modo que corresponde a éste ahora definir qué futuro y mundo desea y el papel de la inteligencia artificial en el mismo.

Para conseguir estos objetivos debemos fijar el punto de inicio y el punto deseado y establecer la hoja de ruta para su consecución mediante estrategias, y en ello están trabajando ya los distintos países y gigantes tecnológicos, si bien, con objetivos muy diversos.

Si como he expuesto, la tecnología debe ser considerada un medio al servicio del hombre para alcanzar sus objetivos y el bien común, el Derecho debería constituir un instrumento con idéntica finalidad, es decir, desde una perspectiva keynesiana debería erigirse como el conjunto de normas orientadas a ordenar la vida del hombre en sociedad para conseguir el bien común de todos y cada uno de sus miembros.

En consecuencia, ante una realidad tan compleja, en constante evolución y con tal alto impacto en la sociedad como la inteligencia artificial, es imprescindible su regulación para garantizar la ética, la seguridad y el cumplimiento para la consecución dicho bien común.

La cuestión esencial es cómo hacerlo. A lo largo de esta investigación he expuesto mi postura ante esta cuestión.

La doctrina también ha ido posicionándose al respecto. Robles Carrillo considera que el modelo de regulación de la inteligencia artificial debe adoptar un enfoque proactivo y abierto, no formalista, para la organización de su gobernanza[510].

Su regulación puede abordarse desde dos enfoques o perspectivas, mediante una aproximación y regulación "suave" o *soft law*, con especial protagonismo de la autorregulación de la industria que se sustentaría en un enfoque de abajo hacia arriba -*down-top*-, lo que supuestamente podría facilitar la creación de normas más ajustadas a las necesidades en cada momento ante una realidad tan compleja y promovidas por quienes mejor entienden y conocen el sector donde operan, o mediante una regulación legal de la misma de arriba abajo -*top-down*- basada en el *hard law*, si bien, las posturas no son unánimes en cuanto al grado de intensidad de la misma, como incluso queda reflejado en las propuestas de regulación del Parlamento Europeo y de la Comisión objeto de análisis en esta investigación.

La apuesta por la definición de unos principios básicos sobre los que deben construirse los futuros marcos normativos, que constituiría la base del precitado *soft law* ha caracterizado las estrategias y trabajo de la UE en los últimos tres (3) años, conforme he expuesto en el capítulo III y en este capítulo.

Por su parte, la OCDE, entre otras, adoptó el 22 de mayo de 2019 una recomendación continente de un conjunto de principios bajo el título *Principles for responsible stewardship of trustworthy AI*[511] (Principios para la gestión responsable de una IA confiable). Posteriormente, China, entre otros países, publicó sus *Governance Principles for a New Generation of Artificial Intelligence: Develop Responsible Artificial Intelligence*[512] (Principios de gobernanza para una Inteligencia responsable), para garantizar su gobernanza y desarrollo sostenible, su seguridad fiabilidad y control. En paralelo, otros países han seguido distintos enfoques para acometer la inteligencia artificial, especialmente mediante el diseño de estrategias como España. Actualmente el

[510] ROBLES CARRILLO, M. (2020). "La gobernanza de la inteligencia artificial: contexto y parámetros generales". *Revista electrónica de estudios internacionales (REEI)*. Nº. 39 2020.

[511] Recuperado de: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449. Consultado el 14.01.2021.

[512] Recuperado de: https://www.loc.gov/item/global-legal-monitor/2019-09-09/china-ai-governance-principles-released/. Consultado el 14.01.2021.

Observatorio de Políticas de la OCDE ha identificado más de 600 iniciativas políticas en materia de inteligencia artificial en 60 países[513], que incluyen tanto recomendaciones, como propuestas de códigos de conducta y regulación, disponibles hasta 2020 en su página web.

Algunos países optaron inicialmente por sistemas de autorregulación en ausencia de un marco regulador, como Dinamarca, Francia, Finlandia o Estonia, junto con procesos de normalización y etiquetado voluntario, al que se han unido otros países posteriormente.

En relación con todo ello, parece inevitable considerar que la irrupción de tecnologías disruptivas en la historia más reciente del ser humano nos ha llevado a considerar la correlativa inseguridad jurídica y la consciencia de ésta parece llevar aparejado un impacto en la innovación, el desarrollo y la inversión. No siempre ha sido así ni lo será, partiendo desde una simple perspectiva de riesgos.

El Derecho constituye un instrumento esencial para generar un marco de seguridad jurídica para todas las partes implicadas, garantizar derechos y para propiciar una inteligencia artificial segura y confiable, que debe integrarse y operar en coordinación con mecanismos de autorregulación y estandarización apropiados, como así se destaca en la Comunicación de la Comisión "*Generar confianza en la inteligencia artificial centrada en el ser humano*"[514].

La estrategia europea apuesta por todo ello como se recoge tanto en las Comunicaciones precitadas, como en el precitado *Libro blanco sobre la inteligencia artificial*[515] y en su informe denominado *Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica*[516], en los que propone crear un ecosistema de confianza sustentado en la creación de un nuevo marco jurídico que revise y complemente el existente, si bien, lo centra en los

[513] Recuperado de: https://oecd.ai/. Consultado el 14.01.2021.
[514] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones *Generar confianza en la inteligencia artificial centrada en el ser humano*. 8 de abril de 2019. COM/2019/168 final.
[515] *Libro Blanco sobre la inteligencia artificial - Un enfoque europeo orientado a la excelencia y la Confianza*. Comisión Europea. Bruselas, 19.2.2020. COM(2020) 65 final.
[516] *Informe de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo: Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, Il internet de las cosas y la robótica*. 19.02.2020. COM/2020/64 final.

sistemas de alto riesgo relacionados con sectores con riesgos significativos o con afectación de derechos fundamentales. Este propósito se ha materializado en las propuestas de regulación europeas en materia ética, inteligencia artificial y de responsabilidad civil, objeto de análisis en esta investigación.

En mi opinión, como he anticipado al abordar las cuestiones de seguridad, la posibilidad de dejar en manos de la propia industria la regulación de un conjunto de tecnologías y capacidades tan complejas con riesgos potenciales y de impacto tan elevado como los que comporta la inteligencia artificial, esto es, la autorregulación, es impensable y ya ha evidenciado su fracaso en relación con otras tecnologías, especialmente cuando la seguridad, confianza y responsabilidades asociadas a estas comportan fuertes inversiones y reducción de los beneficios para la propia industria, por lo que es necesario establecer un marco legal obligatorio y vinculante, sin perjuicio que se dejen determinadas parcelas a la autorregulación mediante estándares o códigos de buenas prácticas.

De hecho, es la industria quién mejor conoce sus actividades, tecnologías, necesidades y riesgos por lo que está en la mejor posición para definir las normas por las que regirse mediante códigos de conducta sectoriales o corporativos propios de una empresa, a los que decida adherirse y comprometerse.

Sin embargo, el impacto social y económico, y los importantes aspectos técnicos y éticos que comporta la inteligencia artificial imposibilita dejar la regulación de una realidad tan compleja en la industria, sin perjuicio de que forme parte protagonista en su definición.

A mi juicio, no se puede garantizar que la industria garantice las mejores prácticas para combatir los múltiples retos y riesgos que plantea la inteligencia artificial, incluyendo la responsabilidad y resarcimiento efectivo, máxime cuando ello supone mayor inversión, incremento de costes y afectación de la competitividad, por lo que estas prácticas deben ser promovidas y exigidas desde fuera, esto es, desde el entorno regulatorio.

De hecho, la propia codificación del *software* sobre el que se sustenta el sistema inteligente ya es un en sí misma una autorregulación propia del diseñador o fabricante.

En mi opinión la autorregulación puede ser una solución eficaz para algunas de las cuestiones que plantea la inteligencia artificial, pero en absoluto para acometer todos los riesgos y retos que supone y que fueron expuestos en el capítulo II de esta investigación.

Como ha puesto de manifiesto el Comité Económico y Social Europeo en alguno de sus dictámenes que citaré más adelante, en especial, el emitido bajo el título *Fomentar la confianza en la inteligencia artificial centrada en las personas*, es necesario identificar con todas las partes interesadas qué aspectos deben abordarse a través de la regulación legal junto con mecanismos de control que, en caso de incumplimiento, incluyan sanciones, y cuáles deben abordarse mediante códigos éticos autorregulación o compromisos corporativos voluntarios. Asimismo, ha significado en diversas ocasiones, en especial, en su Dictamen sobre *Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad*, de 31 de mayo de 2017[517], la necesidad de acometer su regulación de manera global y no meramente regional, dado que sería insuficiente y generaría efectos no deseados.

El deseable consenso global es todavía una utopía, pero a nivel ético los esfuerzos han sido grandes, especialmente por la UE. De hecho, son múltiples las iniciativas internacionales y nacionales para ayudar a la definición de principios y estándares bajo el mayor consenso posible, y no solo desde la perspectiva ética como ODISEIA[518] en España, sino desde un enfoque más amplio como la *Partnership on AI*[519] -PAI por sus siglas en inglés-. A nivel jurídico, la complejidad es todavía mayor, pero al menos sería deseable disponer cuanto antes de un marco armonizado en el seno de la UE que podría ser una referencia internacional.

En relación todos estos aspectos, me permito significar igualmente el Dictamen del Comité Económico y Social Europeo sobre *Autorregulación y corregulación en el marco legislativo de la Unión Europea*, de 22.04.2015[520], que considera la autorregulación y la

[517] Dictamen del Comité Económico y Social Europeo sobre la "Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad» (Dictamen de iniciativa). OJ C 288, 31.8.2017. Pp. 1-9.
[518] https://www.odiseia.org/
[519] https://www.partnershiponai.org/
[520] Dictamen del Comité Económico y Social Europeo sobre el tema "Autorregulación y corregulación en el marco legislativo de la Unión Europea". DOUE 04.09.2015. C291. Pp. 29-38.

corregulación como instrumentos complementarios o suplementarios del Derecho positivo -*hard law*-, pero nunca alternativos al mismo, "salvo que en las normas fundamentales exista una base habilitante adecuada".

La autorregulación ha funcionado en otros sectores como en el del videojuego en aspectos concretos y también en la UE, pero no considero que la complejidad y el contexto actual de feroz competencia internacional en el ámbito de la inteligencia artificial permita consensuar y aplicar normas suficientemente protectoras y garantistas por la propia industria en base a un consenso mundial, máxime ante las fuertes inversiones que supone en materia de seguridad y su concepción como obstáculos para la innovación, comercialización y la competitividad en determinados sectores.

Los códigos de conducta o estándares de buenas prácticas se han mostrado insuficientes como he referido, por ejemplo, en relación con la seguridad de dispositivos IoT en Reino Unido, por lo que, en mi opinión, al igual que la de otros autores como Muñoz Villarreal[521] y Diaz Alabart[522], estos instrumentos deben ir acompañados de instrumentos jurídicos vinculantes que garanticen aspectos esenciales como su seguridad, explicabilidad, trazabilidad o su control y supervisión humana.

Asimismo, autores como Gijs Leenders[523] consideran que la autorregulación no ofrece una solución viable y efectiva a los retos de la regulación de la inteligencia artificial y otros, en los que la misma nunca ha funcionado bien.

En este sentido, el posicionamiento en relación con la necesidad de regulación e insuficiencia de la autorregulación ante una realidad tan compleja como la inteligencia artificial se halla igualmente avalado por el mundo científico y, en especial, por Stephen Hawking, al que he citado y citaré en numerosas ocasiones en esta investigación.

---

[521] MUÑOZ VILLARREAL, A. Y GALLEGO CORCHERO, V. (2019). "Inteligencia artificial e irrupción de una nueva personalidad en nuestro ordenamiento jurídico ante la imputación de responsabilidad a los robots". En Monterrosso Casado, E. (Dir.). "*Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*". Tirant lo Blanch 2019. P. 85.

[522] DIAZ ALABART, S. (2018). Conferencia "Robótica y Responsabilidad Civil". Real Academia de Jurisprudencia y Legislación. 31.05.2018.

[523] LEENDERS, G. (2019). "The Regulation of Artificial Intelligence - A Case Study of the Partnership on AI". *Medium*. Publicado el 13.04.2019. Recuperado de: https://becominghuman.ai/the-regulation-of-artificial-intelligence-a-case-study-of-the-partnership-on-ai-c1c22526c19f. Consultado el 14.01.2021.

Durante su discurso en la Web Summit de Lisboa de 2017, Hawking[524] afirmó que "el éxito en la creación de una inteligencia artificial eficaz podría ser el evento más importante en la historia de nuestra civilización. O lo peor. Simplemente no lo sabemos. Por lo tanto, no podemos saber si la inteligencia artificial nos ayudará infinitamente, o si nos ignorará y dejará de lado, o posiblemente nos destruirá". Y adicionó que "a menos que aprendamos a prepararnos y evitar los riesgos potenciales, la inteligencia artificial podría ser el peor evento en la historia de nuestra civilización. Trae peligros, como poderosas armas autónomas, o nuevas formas para que unos pocos opriman a la mayoría. Podría traer grandes trastornos a nuestra economía". Asimismo, puso en valor ya entonces las propuestas regulatorias europeas.

En el mismo sentido se ha pronunciado el sector de la industria con voces tan autorizadas como Elon Musk en algunas de sus declaraciones públicas precitadas anteriormente, sin perjuicio de que mantenga su postura en el futuro, dados los giros que ha mostrado en su posicionamiento en relación con la regulación de la inteligencia artificial.

Los retos para el legislador son importantes, especialmente al tener que abordar aspectos transnacionales y globales, la complejidad técnica de su objeto, su rápida evolución e implementación, su interacción con otras tecnologías o el incesante aumento de sus capacidades y nuevos riesgos asociados, si bien, el Derecho exige la aplicación de nuevos enfoques y técnicas legislativas más *responsive*, ágiles, dinámicas, adaptativas e "inteligentes" que nunca y legislar sobre tendencias más que sobre novedades y estratégicamente, de modo que las respuestas jurídicas de hoy puedan servir también para resolver problemas del mañana.

En definitiva, reiterando mi opinión manifestada en los apartados precedentes, considero necesario un enfoque global jurídico y ético que integre un marco regulativo global e integrado con principios y normas éticas globalmente aceptadas y que igualmente contemple y se combine con mecanismos de autorregulación como códigos y estándares de buenas prácticas, al objeto de abordar los retos que plantea la inteligencia artificial en

[524] KHARPAL, A. "Stephen Hawking says A.I. could be 'worst event in the history of our civilization'". Publicado en *CBNC* el 06.11.2017. Recuperado de: https://www.cnbc.com/2017/11/06/stephen-hawking-ai-could-be-worst-event-in-civilization.html. Consultado el 14.01.2021.

el ámbito jurídico, ético y de seguridad y que promueva la exigencia de marcos de gobernanza públicos y privados.

La necesidad de nuevos marcos y de la revisión y adecuación de los existentes a los nuevos retos que plantea la inteligencia artificial es incuestionable y así se refleja en los documentos de trabajo de la UE, en particular, en el Dictamen del Comité Económico y Social europeo precitado en otros apartados y emitido bajo el título *Fomentar la confianza en la inteligencia artificial centrada en las personas*[525]. Entre otras conclusiones, considera que "es imprescindible revisar la legislación para regular las nuevas situaciones que la introducción de esta tecnología traerá consigo", entre otras y me permito significar, las relativas a la responsabilidad por daños. Además, los nuevos marcos deberán integrarse con los preexistentes, en una necesaria convergencia en sus distintos objetivos y conseguir la deseable armonización.

6. La nueva propuesta europea. Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, por el que se establecen normas armonizadas sobre la inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión

6.1. Aspectos generales

La UE ha dado un paso más en materia de inteligencia artificial, en la medida que, tras las resoluciones precitadas del Parlamento Europeo de octubre de 2020 en materia ética, de responsabilidad y de propiedad intelectual, que incluían las sendas propuestas de Reglamento en materia ética -analizada en el capítulo III- y de responsabilidad civil -que será objeto de análisis en el próximo capítulo-, se ha publicado en el seno de la UE una nueva Propuesta de Reglamento que pretende regular el uso de la inteligencia artificial basado hipotéticamente en el trabajo y conclusiones previas que condujeron a las resoluciones y propuestas precitadas, si bien, como se expondrá a lo largo de su análisis

[525] 8.4.2019. COM (2019) 168 final

en relación con los aspectos objeto de esta investigación, se aparta en distintos aspectos de las mismas, empezando por la propia definición de inteligencia artificial.

La Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, por el que se establecen normas armonizadas sobre la inteligencia artificial (Ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión[526], se focaliza en la regulación de la inteligencia artificial y, en particular, en la prohibición de determinados sistemas por su riesgo inaceptable y en la amplia regulación de los denominados sistemas de alto riesgo, dejando al margen al resto de sistema de inteligentes, invitando a su sujeción voluntaria a futuros códigos de buenas prácticas.

Del mismo modo que la propuesta precedente del Parlamento Europeo de 20 de octubre de 2020, regula y exige algunos de los principios éticos más relevantes en relación con los sistemas inteligentes de alto riesgo para transformarlos en vinculantes, con su integración y/o conversión en requisitos y obligaciones para los mismos.

La nueva propuesta no sólo regula la inteligencia artificial, sino también sus usos.

La propuesta precedente más cercana a la publicada es la integrada en la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas, que incorpora una Propuesta de Reglamento del Parlamento Europeo y del Consejo, sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas.

En la misma fecha, el Parlamento Europeo publicó su Resolución sobre responsabilidad civil de la inteligencia artificial, a la que incorporó una Propuesta de Reglamento del Parlamento Europeo y del Consejo, relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial.

Las cuestiones relativas a la responsabilidad civil de la inteligencia artificial no son abordadas en la nueva propuesta de 21 de abril de 2021, salvo algunos aspectos

[526] COM (2021) 206 final.

específicos que comentaré en su análisis, como tampoco aborda los aspectos relacionados con la propiedad intelectual, que fueron objeto de la tercera Resolución del Parlamento Europeo de 20 de octubre de 2020.

Los antecedentes de esta nueva propuesta han sido analizados a lo largo de los capítulos precedentes de esta investigación o serán objeto de análisis en capítulos posteriores.

Los trabajos preparatorios se iniciaron principalmente con la creación del *Grupo de expertos de alto nivel sobre inteligencia artificial* -HLEG-, prosiguieron con el respaldo de la Comisión Europea a los 23 requisitos clave establecidos en las directrices éticas del grupo precitado para una inteligencia artificial fiable, al que prosiguió el *Libro Blanco sobre inteligencia artificial* y, con posterioridad, la *Evaluación de Impacto Inicial*[527] sobre esta propuesta que fue examinada por el Comité de Control Reglamentario de la Comisión.

A la reunión del precitado Comité el 16 de diciembre de 2020, prosiguió un primer dictamen negativo sobre la misma. Posteriormente, tras una revisión sustancial de la evaluación de impacto, se emitió un dictamen positivo el 21 de marzo de 2021.

Conforme a los documentos preparatorios, los objetivos finales propuestos, especialmente garantizar el correcto funcionamiento del mercado único y la creación de un ecosistema de inteligencia artificial fiable en el seno de la UE, podían abordarse a través de diferentes políticas con distintos grados de intervención regulatoria.

Las posibles opciones incluían la creación de un instrumento legislativo de la UE que estableciera un etiquetado único, un enfoque sectorial *ad hoc*, un instrumento legislativo horizontal siguiendo un enfoque basado en el riesgo, un instrumento legislativo horizontal siguiendo dicho enfoque pero que contemplara códigos de conducta (*soft law*) para los sistemas de inteligencia artificial no calificados de alto riesgo o un instrumento legislativo horizontal de la UE que regulara requerimientos obligatorios para todos los sistemas de la inteligencia artificial, independientemente del riesgo que supongan.

[527] Disponible en: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=PI_COM:Ares(2020)3896535&from=EN. Consultado el 30.04.2021

Las opciones políticas fueron evaluadas en función de los impactos económicos y sociales y en los derechos fundamentales, apostando finalmente por un marco regulatorio solo para sistemas de inteligencia artificial de alto riesgo, con definición de una serie de requerimientos exigibles exclusivamente a los mismos -en materia de datos, documentación, trazabilidad, información, transparencia, supervisión y control humano, solidez y precisión-, contemplando la posibilidad de que los proveedores de sistemas de inteligencia artificial que no sean de alto riesgo puedan adscribirse a códigos de conducta de manera facultativa, y ello por considerar ésta la opción "más adecuada para abordar de la manera más eficaz los objetivos de esta propuesta".

Los impactos valorados de las diferentes opciones se recogen en la evaluación de impacto que se adjunta como Anexo 3 a la Propuesta de Reglamento.

De antemano, me permito anticipar mi opinión sobre esta importante decisión, en la medida que no comparto esta elección desde un enfoque de riesgos y de seguridad, en la medida que, conforme ya recogía la Propuesta de Reglamento del Parlamento Europeo y del Consejo, sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas -incorporado a la Resolución del Parlamento Europeo de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas-, aunque respecto de concretos principios y normas éticas, considero que deben requerirse una serie de principios y normas éticas y jurídicas comunes para cualquier sistema de inteligencia artificial, precisamente por sus propias características, capacidades y riesgos potenciales asociados a la misma, con independencia del sector o contexto donde opere y nivel de riesgo inicial, sin perjuicio de los adicionales para los sistemas considerados de alto riesgo, de riesgo limitado o medio o de riesgo bajo o mínimo.

En este sentido, el Parlamento Europeo se ha posicionado reiteradamente respecto del necesario control humano de la inteligencia artificial en sus resoluciones precitadas, si bien, se introduce en la última propuesta reguladora de la Comisión como exigencia jurídica exclusivamente para sistemas de alto riesgo.

El Parlamento Europeo, sin embargo, ha venido recogiendo una visión más amplia de esta exigencia no sólo en el ámbito militar, sino civil en su reciente y precitada *Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre Inteligencia artificial: Cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal*[528], en la que considera que la inteligencia artificial utilizada tanto en un contexto militar, como civil, debe estar sujeta a un control humano apropiado, de modo que un ser humano tenga en todo momento los medios para corregir su curso, detenerla o desactivarla en caso de comportamiento imprevisto, intervención accidental o ciberataque, entre otros supuestos.

Además, es necesario insistir que los distintos retos y riesgos que plantea la inteligencia artificial no sólo afectan a personas físicas y a sus datos personales, a su seguridad o a sus derechos fundamentales, sino que afectan a otros derechos y bienes, así como también a entidades públicas y privadas o a gobiernos, donde no sólo puede estar en juego el tratamiento no autorizado de datos personales, sino la vida y la salud de una persona, las comunicaciones regionales o mundiales, el equilibrio de un mercado o el medioambiente. Pensemos simplemente en un sistema de inteligencia artificial que gestione una red de comunicaciones y lo que puede suponer su caída, bloqueo o manipulación por terceros.

De este modo, únicamente concibo un enfoque global de los retos y riesgos de la inteligencia artificial y, consiguientemente su regulación.

En consecuencia, no puedo compartir totalmente el posicionamiento y opción final sobre la que se ha construido la reciente Propuesta de Reglamento del Parlamento Europeo y del Consejo de 21 de abril de 2021, en la medida que, tal y como he expuesto, considero que existen principios y normas éticas esenciales que deberían ser exigibles a cualquier sistema de inteligencia artificial, como seguridad o control y supervisión humana, precisamente por la complejidad, naturaleza, características y capacidades de las que puede estar inicialmente dotado o que pueden ser posteriormente adquiridas, y que

[528] Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre inteligencia artificial: Cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal (2020/2013(INI))

deberían tener carácter vinculante y ser jurídicamente obligatorias para cualquier sistema, cualquier que sea su nivel de riesgo inicial, y ello precisamente también desde un enfoque de riesgos.

A mi juicio, cualquier sistema de inteligencia artificial por sus propias características puede conllevar riesgos intrínsecos para la seguridad, bienes, derechos y libertades de las personas físicas y jurídicas, debiendo considerarse cualquier probabilidad por baja que sea, así como su impacto. ¿Cuáles son los riesgos que plantea un simple *chatbot* operado por un tercero con el que interaccionan todos los miembros de una unidad familiar, incluidos menores?

Sin embargo, la opción elegida, a juicio de la Comisión, evitará las acciones unilaterales de los Estados miembros que puedan provocar una desfragmentación y proporcionará más seguridad jurídica a las empresas, pero ¿también a los ciudadanos en caso de sistemas inteligentes no calificados de alto riesgo?

A mi juicio, parece una propuesta más enfocada a la protección de la competitividad empresarial y la innovación que al ser humano, con excepción, obviamente, de los sistemas de alto riesgo regulados con minuciosidad inicial en el Reglamento propuesto. De hecho, es muy significativo que no se hace mención alguna a los derechos de los posibles afectados por el funcionamiento de sistemas inteligentes, ni garantías ni mecanismos de tutela y se habla poco de derechos fundamentales.

Se presenta como un reglamento técnico y complejo, con un enfoque industrial, ético y de seguridad, y muy orientado a la gestión del riesgo.

La nueva propuesta recoge muchas de las cuestiones abordadas en las Resoluciones y propuestas precedentes y precitadas del Parlamento Europeo, si bien, no todas, significando la prohibición directa de los sistemas de videovigilancia masiva en tiempo real en espacios públicos, consecuencia de la carta dirigida el pasado 15 de abril de 2021 a la Presidenta de la Comisión Europea por varios eurodiputados, en la que manifestaban su oposición al régimen de videovigilancia masiva en el borrador de Reglamento que se filtró días de antes de su publicación.

El texto final publicado recoge ya la prohibición de estos sistemas, permitiéndose solo en circunstancias especiales y bajo autorización judicial o administrativa, lo que evidencia una cierta precipitación final en la publicación del texto final de la propuesta ante cuestiones tan relevantes como una prohibición, aunque "descafeinada", habida cuenta de las excepciones previstas.

Conforme al marco propuesto, la clasificación del riesgo queda en manos de los legisladores en un enfoque *top-down*, lo que podría generar problemas en el futuro, especialmente en la medida que es la industria quién tiene una mayor proximidad a la investigación e innovación en proceso y la que tienen mayor proximidad, control y supuesto conocimiento del riesgo. Además, la clasificación de efectúa de manera objetiva sin responder a una clasificación efectiva del riesgo.

La Propuesta de Reglamento objeto de análisis ha sido acompañada de su Estudio de impacto, así como del Estudio de apoyo de esta. En fechas coincidentes con la finalización de la versión revisada de esta investigación se ha emitido el Dictamen conjunto 5/2021, del Comité Europeo de Protección de Datos con el Supervisor Europeo de Protección de Datos sobre el Reglamento propuesto[529].

Ambos organismos valoran positivamente la propuesta, si bien, consideran que deben revisarse determinados aspectos para garantizar su aplicabilidad y eficacia.

Entre otros aspectos, proponen prohibir cualquier tipo de sistema inteligente de *scoring* o puntuación social y no sólo el llevado a cabo por entidades públicas, así como los sistemas inteligentes de identificación biométrica remota de personas en espacios de acceso público que permitan el reconocimiento automático de rasgos humanos, ya sea el rostro, la marcha, huellas dactilares, voz o comportamiento.

Asimismo, recomiendan que se prohíban los sistemas de inteligencia artificial que permitan la inferencia de emociones o que clasifiquen a los individuos a partir de la

[529] Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). EDPB-EDPS 18.06.2021. Recuperado de: https://edpb.europa.eu/system/files/2021-06/edpb-edps_joint_opinion_ai_regulation_en.pdf. Consultado el 20.06.2021.

biometría en grupos, según la etnia, género, orientación sexual o política y otros criterios discriminatorios.

El Reglamento propuesto se orienta principalmente a regular el uso de los sistemas de inteligencia artificial de forma segura y ética, pretendiendo armonizar el marco regulador actual, especialmente cuando un sistema inteligente sea un componente de seguridad de determinados productos o constituya el producto en sí mismo, y estableciendo una serie de requerimientos regulatorios dirigidos a reducir sus riesgos, aunque exclusivamente aplicables a sistemas de alto riesgo, lo que supone de antemano una restricción no razonable desde un punto de vista de gestión de los riesgos asociados a la inteligencia artificial, que responde a la opción elegida precitada para construir esta propuesta regulatoria, conforme he comentado anteriormente.

Asimismo, como igualmente he anticipado, no contempla a las personas afectadas en todo su articulado, sin perjuicio, obviamente, de que sean también uno de los principales destinatarios de la protección que pretende dispensar para mitigar los riesgos para su seguridad y sus derechos fundamentales, conforme recoge la Evaluación de Impacto adjunta al mismo. Entre otros aspectos, no se contemplan mecanismos de garantías o tutela de derechos de las personas afectadas y se echan en falta evaluaciones de impacto en los derechos humanos.

El ámbito de aplicación subjetivo y territorial, siguiendo las propuestas anteriores, es amplio, comprendiendo a los distintos agentes dentro de la cadena de valor y ciclo de vida de la inteligencia artificial, esto es, proveedores, importadores, distribuidores, fabricantes y usuarios. De inicio, parece orientarse a su despliegue, explotación y uso, si bien, una vez analizado su contenido y requerimientos, regula e impacta directamente en su diseño y concepción, como luego expondré. Precisamente por ello, hubiera sido deseable una mayor atención a diseñadores y desarrolladores.

Del mismo modo, extiende su ámbito de aplicación territorial tanto a los sistemas utilizados en la UE como los utilizados en un tercer país pero que desplieguen efectos en la UE.

Por otra parte, la propuesta se sustenta en un enfoque de riesgos, estableciendo distintos niveles, a los que resultarán de aplicación distintos requisitos y obligaciones en atención a su clasificación conforme a los mismos, alejándose del concepto de autonomía que pueda ostentar o no el sistema, que hasta la fecha formaba parte de las distintas definiciones en los documentos de trabajo, propuestas y resoluciones precedentes.

Por lo que se refiere a dicha clasificación, en primer lugar, la propuesta prohíbe una serie de sistemas de inteligencia artificial, mediante la inclusión de un listado tasado que deberá ser periódicamente revisado, cuyo uso estaría prohibido, por considerar que implican un riesgo inadmisible para la vida, la seguridad y los derechos fundamentales.

Como analizaré con posterioridad, este listado incluye sistemas capaces de manipular el comportamiento humano, predecir información respecto a colectivos o grupos para identificar sus vulnerabilidades o circunstancias especiales o sistemas que impliquen la identificación biométrica o la videovigilancia masiva en espacios públicos y en directo por parte de las autoridades públicas.

En relación con éstos últimos, se parte de la prohibición partiendo de un principio de prudencia, si bien, se permitiría su uso conforme a dicha Propuesta cuando lo sea en cumplimiento de la Ley y bajo previa autorización, judicial o administrativa, previendo incluso que esta puede ser solicitada con posterioridad a su implementación en casos de "extrema urgencia", concepto jurídico indeterminado que, de no concretarse en su tramitación, constituirá un foco de inseguridad jurídica.

En segundo lugar, la propuesta lista un conjunto de sistemas que considera de alto riesgo para los derechos y libertades de los individuos -especialmente la salud, la seguridad y los derechos fundamentales-, permitidos, pero sujetos a un conjunto de requisitos y obligaciones con el objetivo de garantizar su uso legal, ético, robusto y seguro, con afectación también a su propio diseño, considerando pues la *Ethics by design* y la *Security by design* en el conjunto de requerimientos que deben cumplir estos sistemas.

En tercer lugar, la propuesta no regula el resto de sistemas de inteligencia artificial, que siguiendo un enfoque de riesgos, se trataría de los sistemas de riesgo medio o limitado, donde podrían incardinarse inicialmente asistentes virtuales o *chatbots*, y los de riesgo

bajo o mínimo, todos ellos no sujetos a un marco específico de requisitos y obligaciones concretas, salvo obligaciones concretas de transparencia e información para determinados sistemas de inteligencia artificial -sean o no calificables de alto riesgo-.

En consecuencia, el resto de los sistemas distintos a los anteriores quedarían fuera del ámbito de aplicación del Reglamento propuesto, no quedando sujetos a ninguna obligación en particular.

Conforme a lo previsto en el Reglamento propuesto, los sistemas inteligentes distintos a los prohibidos y de alto riesgo, a excepción de las obligaciones precitadas de transparencia para concretos sistemas tipificados, quedarían fuera del marco regulador previsto en el mismo, sin perjuicio de su cumplimiento voluntario y adscripción a futuros códigos de conducta. En este sentido existen ya países avanzados en la creación de estos marcos como Dinamarca o Malta.

Dentro de estos sistemas no contemplados, nos podríamos encontrar sistemas de riesgo medio o limitado, de riesgo bajo o mínimo (por ejemplo, sistemas de filtro Antispam), en su caso, de riesgo inexistente, en mi opinión inconcebibles por las propias características y posibles capacidades de la inteligencia artificial o simplemente sistemas inteligentes no clasificados. ¿Se permitirá el uso de sistemas inteligentes en el ámbito de la UE que no hayan sido clasificados, bien desde el punto de vista de los marcos reguladores de la inteligencia artificial o, por ejemplo, de la privacidad?

El Reglamento propuesto no prevé la exigencia de un análisis, revisión y clasificación de cualquier sistema inteligente que se desee poner en el tráfico desde un punto de vista de riesgos para la vida, la salud o los derechos fundamentales, y mucho menos para otros de los bienes e intereses en juego a los que he aludido en esta investigación, como son los daños materiales e impactos graves en infraestructuras críticas, servicios esenciales, actividades económicas, mercados o para la sociedad en su conjunto, en definitiva bienes o derechos protegidos a nivel legal y, en algunos casos, a nivel constitucional. Ni tan siquiera se exige una declaración responsable.

Desde luego, en la medida que involucre el tratamiento de datos personales, los análisis de riesgos y las evaluaciones de impacto requeridas por el RGPD obligarían a la

calificación del riesgo y el establecimiento de medidas para su gestión, si bien, en caso de que no implique inicialmente este tratamiento, no serían preceptivas conforme a este marco especial.

Desde un enfoque de riesgos y aun considerando exclusivamente un conjunto de riesgos estrictamente tasados como punto de partida al abordar la evaluación y clasificación de un sistema -ya sea para la salud, la seguridad o los derechos fundamentales-, me resulta difícil concebir sistemas inteligentes carentes de riesgo inherente o residual para los mismos y otros bienes o derechos en base a la propia naturaleza, características y capacidades de las que pueda estar dotado cada sistema inteligente, salvo contextos muy concretos. Y ello, en la medida que debe considerarse desde dicho enfoque no sólo la probabilidad por muy baja que sea, sino también el impacto los distintos bienes y derechos objeto de protección, que puede ser muy alta.

En este sentido, es complicado pensar en la existencia de sistemas inteligentes que se relacionen o interaccionen con personas y cosas -cuyos riesgos deberían ser previamente analizados para posibilitar su clasificación previa en cualquier caso-, que no comporten algún tipo de riesgo inherente o residual por su propia naturaleza, características y posibles capacidades, ya sea en función de su probabilidad o de su impacto (aunque ambos fueran poco significativos), en la medida que el riesgo, cuando menos, debería ser clasificado como bajo. Y, aun así, me resulta difícil imaginar, desde un punto de vista de análisis de riesgos, la existencia de amenazas de probabilidad muy baja que no puedan comportar impactos altos en todos estos bienes y derechos sobre los que puede materializarse el riesgo en función de las características y capacidades del sistema inteligente.

Desde mi punto de vista, hemos pasado de un reglamento más general focalizado en los principios y normas éticas esenciales y su exigibilidad jurídica al que respondía la propuesta de 20 de octubre de 2020, a un reglamento más específico y enfocado en el cumplimiento regulatorio y la seguridad, por lo que sería deseable una revisión híbrida desde un enfoque ético y jurídico para alcanzar, a mi juicio, el equilibrio necesario y desde un enfoque global de la inteligencia artificial y no sesgado a los sistemas calificados de alto riesgo.

El Reglamento propuesto parte de la consideración de la inteligencia artificial como una actividad de riesgo, que precisa regulación en los niveles más altos, por lo que abordar su regulación podría hacerse mediante distintos instrumentos o técnicas, desde la prohibición total o parcial de ciertas actividades para evitar el riesgo conforme al principio de precaución, controles preventivos suaves como declaraciones responsables, controles preventivos más duros como el sometimiento a regímenes de autorización, controles exclusivamente posteriores con eficacia depuradora, reparadora y, en su caso sancionadora, como la responsabilidad civil y, en su caso penal, y los regímenes sancionadores, así como instrumentos de gobernanza, inspección y supervisión.

El Reglamento combina algunos de estos instrumentos y no entra en otros, como en la responsabilidad.

La propuesta establece un marco obligacional distinto en función del tipo de sistema inteligente, conforme a la clasificación precitada.

Los sistemas de alto riesgo solo "se hallarán permitidos" siempre que cumplan los requisitos y obligaciones requeridas por el Reglamento propuesto, debiendo ser sometidos a una evaluación de conformidad y garantizar la gestión del riesgo durante toda su vida útil.

Estos sistemas estarían sujetos a obligaciones específicas de seguridad, control y supervisión humana -de modo que siempre tendrá que haber una persona con capacidad de control para mitigar un riesgo-, de transparencia, de gobernanza de datos, de inscripción previa y de conformidad y certificación previa, conforme a las especificaciones técnicas que habrá que cumplir.

Algunos sistemas de riesgo medio o limitado (además de los de alto riesgo igualmente sujetos a las mismas) quedarían exclusivamente sujetos a un conjunto de obligaciones concretas de transparencia e información para garantizar que su funcionamiento, características e implicaciones inherentes al uso de estos sistemas son conocidos por los usuarios.

Y el resto de los sistemas inteligentes, como he referido, tanto los de riesgo medio o limitado, bajo o mínimo, como no clasificados, de mantenerse la redacción actual, no se hallarían sujetos a obligación alguna conforme al marco propuesto, sin perjuicio de que se hallen sujetos a otros vigentes, especialmente de seguridad o privacidad, conforme al Reglamento General de Protección de Datos, ni tan siquiera a un análisis de riesgos previo preceptivo.

Considerar la inexistencia de riesgo alguno, como he referido anteriormente, lo considero complicado, dado que el riesgo, aunque sea muy bajo, es previsible que exista en atención a las características de estos sistemas y posibles capacidades de las que pueden estar dotados en muchos supuestos, especialmente ante su funcionamiento y uso posterior.

Según la propuesta, todos estos sistemas estarían sujetos a sistemas voluntarios de autorregulación como la adhesión a códigos de conducta voluntarios, dejando abierta la regulación de estos sistemas.

En definitiva, considero que estas categorías de sistemas no regulados y liberados de cualquier obligación conforme a lo previsto en esta propuesta, en las que además tendrían supuestamente su encaje los sistemas de menor complejidad -que actualmente constituyen la gran mayoría de los sistemas que se utilizan en el mercado-, lo único que genera es confusión e inseguridad jurídica que deberá resolverse en las futuras revisiones de esta propuesta, ante la oportunidad que supone.

En este sentido, considero que el papel del Parlamento Europeo será determinante en este sentido, así como en relación con la ausencia de un marco de protección, de garantías y de tutela para las personas afectadas en el mismo, salvo la remisión a otros futuros marcos que las contemplen.

En cuanto al despliegue de este nuevo marco normativo, la aplicación de este deberá llevarse a cabo a través de un marco de gobernanza a nivel de los Estados miembros conforme a estructuras ya existentes y con la creación de otras nuevas bajo mecanismos de cooperación que debe liderar el Comité Europeo de Inteligencia Artificial. El régimen propuesto se ha concebido bajo un enfoque similar al previsto en el RGPD.

La propuesta incorpora una serie de medidas adicionales para apoyar la innovación, especialmente a través de *sandboxes* regulatorios y otras medidas adicionales que, según su memoria expositiva, servirán para reducir la carga regulatoria y para apoyar a las PYMES y *startups*, si bien, como luego analizaré con mayor profundidad, el análisis cualitativo y cuantitativo de los requerimientos y obligaciones regulados en el Reglamento propuesto puede chocar frontalmente con este propósito anunciado.

Por último, la propuesta incorpora un régimen sancionador del que adolecía, como he referido en su análisis, la Propuesta de Reglamento del Parlamento Europeo y del Consejo, sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, de 20 de octubre de 2020, como herramienta conminatoria para el cumplimiento de las obligaciones previstas en el mismo. El régimen sancionador propuesto se ha concebido bajo un enfoque similar al que integra el RGPD.

El incumplimiento relativo a prácticas prohibidas y las obligaciones de gestión y gobernanza de datos de los sistemas de alto riesgo podrá ser sancionado con hasta treinta millones de euros o el 6% del volumen de negocio anual total a escala mundial del ejercicio económico anterior.

El incumplimiento de cualquier otro requisito u obligación podrá ser sancionado con hasta veinte millones de euros o el 4% del volumen de negocio anual a total a escala mundial del ejercicio económico anterior.

Y, por último, el suministro de información incorrecta, incompleta o engañosa a los organismos y/o autoridades nacionales podrá ser sancionado con hasta diez millones de euros o el 2% del volumen de negocio anual total a escala mundial del ejercicio económico anterior.

La Propuesta de Reglamento inicia ahora su recorrido y deberá ser revisada y debatida por el Parlamento y el Consejo, pudiendo sugerir enmiendas hasta su aprobación final. Una vez aprobado, el Reglamento será aplicable en todos los países de la UE, aunque se espera una importante asincronía entre su publicación y su exigencia que ya recoge la

propuesta, al igual que ocurrió con el RGPD, al objeto de permitir la adaptación por parte del mercado a sus disposiciones.

6.2. Análisis

La propuesta se presenta en su sede expositiva bajo en enfoque horizontal, equilibrado y proporcionado para la inteligencia artificial, con el propósito de ofrecer un marco legal sólido y flexible y bajo una premisa de intervención mínima limitada a los requisitos mínimos necesarios para abordar los riesgos y los problemas relacionados con la inteligencia artificial en equilibrio con el desarrollo tecnológico, siguiendo así el enfoque propuesto en el *Libro blanco sobre inteligencia artificial* y sobre el que se han construido las propuestas precedentes, y todo ello con el objetivo de garantizar el necesario equilibrio entre seguridad, confiabilidad e innovación, desarrollo tecnológico y competitividad.

Asimismo, la memoria expositiva de la propuesta califica el marco legal recogido en la misma como sólido, flexible, proporcionado, integral y evolutivo o adaptativo, esto es, preparado para el futuro, bajo un enfoque basado en el riesgo como las propuestas precedentes.

Según la memoria expositiva, la intervención legal se circunscribe y adapta a aquellas situaciones concretas en las que se considera que existe una causa justificada de preocupación actual o potencial para el futuro, y además bajo mecanismos flexibles que permiten adaptarse dinámicamente a la evolución tecnológica y los nuevos escenarios que conlleve.

De este modo, la propuesta no se presenta como una regulación global de la inteligencia artificial, sino de algunas de sus aplicaciones y usos de determinados sistemas, y bajo normas armonizadas para el desarrollo, comercialización y uso de estos sistemas en la UE.

Otra cosa distinta es que realmente el texto final propuesto permita alcanzar todos los objetivos propuestos precitados y pueda ser calificable en el modo anunciado dado que, como posteriormente se analizará y, a modo de ejemplo, el marco obligacional que

propone, en mi opinión, puede producir un resultado sensiblemente distante de la realidad de la innovación tecnológica actual en materia de inteligencia artificial, impulsada por grandes tecnológicas, pero también por *startups* y PYMEs, siendo estas últimas las que van a tener verdaderas dificultades de antemano para cumplir los múltiples requerimientos para poder ver sus productos en el mercado, de modo que es previsible que, en caso de productos y sistemas especialmente innovadores se vean abocadas a entrar en procesos de adquisición o inversión en los que las *due diligence* recojan importantes minoraciones del precio y retenciones de garantía hasta la verificación de cumplimiento, clasificación y, en su caso, certificación. Y ello sin olvidar la posible deslocalización del talento y proyectos para desarrollarse en entornos más laxos y flexibles para favorecer la innovación, especialmente para el bien común.

Por otra parte, una laxitud no equilibrada de determinados requisitos para este tipo de empresas puede conllevar un efecto perverso para el mercado, en la medida que puede potenciar la creación de nuevas empresas por las grandes corporaciones para acometer proyectos y su posterior explotación, con sujeción a marcos obligacionales más laxos.

6.2.1. Razones y objetivos

La memoria explicativa que acompaña a la propuesta y constituye su exposición de motivos, especifica las razones y objetivos generales de la propuesta que, en congruencia con su propio título, parece tener como fin inmediato la urgente armonización de las normas sobre inteligencia artificial en la UE, especialmente ante las iniciativas de algunos Estados miembros en esta materia como reitera en el Considerando 2º, en paralelo a su objetivo mediato que no es otro que el mismo definido en el *Libro blanco sobre inteligencia artificial* que fue objeto de análisis en anteriores capítulos y sobre el que se sustentan las resoluciones y propuestas de octubre de 2020 del Parlamento Europeo, esto es, promover la adopción de la inteligencia artificial y abordar sus riesgos asociados con ciertos usos de dicha tecnología, aunque no todos, para el desarrollo de un ecosistema de confianza basado en un marco legal confiable.

El Considerando 1º de la propuesta define el objetivo del futuro Reglamento desde un prisma aparentemente de mercado, esto es, mejorar el funcionamiento del mercado interior estableciendo un marco jurídico uniforme para el desarrollo, la comercialización y uso de la inteligencia artificial de conformidad con los valores de la UE, si bien, reitero, exclusivamente en el marco de los sistemas de alto riesgo.

Del mismo modo, aboga por que las normas sobre inteligencia artificial en el seno de la UE se centren en el ser humano, se sustenten en la ética -*Ethics by design*- y que apoyen el ansiado objetivo expuesto en las propuestas previas, de que la UE sea un líder mundial en el desarrollo de inteligencia artificial segura, confiable y ética, conforme a la voluntad reiteradamente declarada del Consejo Europeo[530] y conforme instó el Parlamento Europeo en su Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas, que fue objeto de análisis en anteriores capítulos.

Sin embargo, a pesar de ese propósito inicial, hay que reiterar que se focaliza en prohibir determinados sistemas y a regular exclusivamente los calificados de alto riesgo conforme al mismo, dejando a un lado todos los demás, remitiéndose a códigos de conducta de adscripción voluntaria.

Según esta exposición de motivos, la propuesta también obedece a las solicitudes del Consejo Europeo[531] de determinar claramente las aplicaciones de inteligencia artificial que deben considerarse de alto riesgo, así como de abordar adecuadamente la opacidad, la complejidad, el sesgo, el cierto grado de imprevisibilidad y autonomía parcial que pueda asociarse a ciertos sistemas inteligentes, al objeto de garantizar su compatibilidad con los derechos fundamentales y facilitar la aplicación del marco legal.

[530] Consejo de Europa. Reunión especial del Consejo de Europa el 1 y 2 de octubre de 2020. Conclusiones. EUCO 13/20, 2020. P.6.
[531] Consejo Europeo. Reunión especial del 1 y de octubre de 2020. Conclusiones. EUCO 13/20, 2020; Council of the European Union, *Presidency Conclusions - The Charter of Fundamental Rights in the context of Artificial Intelligence and Digital Change*, 11481/20, 2020.

La memoria explicativa hace referencia a la propuesta previa del Parlamento Europeo sobre aspectos éticos de la inteligencia artificial de 20 de octubre de 2020, si bien, no identifica ni motiva los cambios en la nueva propuesta respecto de la precedente, focalizándose en los objetivos específicos de esta nueva propuesta orientada a sistemas de alto riesgo:

a) Garantizar que los sistemas de inteligencia artificial comercializados y utilizados en la UE sean seguros y respeten la legislación vigente sobre valores y derechos fundamentales de la UE.

b) Garantizar la seguridad jurídica para facilitar la inversión y la innovación en inteligencia artificial.

c) Mejorar la gobernanza y la aplicación efectiva de la ley existente sobre derechos fundamentales y requisitos de seguridad aplicables a los sistemas de inteligencia artificial.

d) Facilitar el desarrollo de un mercado único para aplicaciones de inteligencia artificial legales, seguras y confiables y, prevenir la fragmentación del mercado.

Según la memoria explicativa, la propuesta mejorará y promoverá la protección de los derechos fundamentales propugnados por la *Carta de los Derechos Fundamentales de la UE*, en particular: El derecho a la dignidad humana (Artículo 1), el respeto a la vida privada y la protección de datos personales (Artículos 7 y 8), no discriminación (Artículo 21) e igualdad entre mujeres y hombres (Artículo 23), libertad de expresión (Artículo 11), la libertad de reunión (Artículo 12), la protección del derecho a un recurso efectivo y a un juicio justo, los derechos de defensa y la presunción de inocencia (Artículos 47 y 48), así como el principio general de buena administración.

Conforme sea aplicable en determinados ámbitos, la Comisión considera que la propuesta afectará positivamente a los derechos de un conjunto de grupos especiales, como los derechos de los trabajadores a unas condiciones de trabajo justas y equitativas (Artículo 31), un alto nivel de protección del consumidor (Artículo 38), los derechos del niño (Artículo 24) y la integración de las personas con discapacidad (Artículo 26).

Y también repercutirá del mismo modo en el derecho a un alto nivel de protección ambiental y la mejora de la calidad del medio ambiente (Artículo 37), incluso en relación con la salud y la seguridad de las personas.

No obstante, debo destacar que la propuesta contiene algunas limitaciones "proporcionadas y limitadas" a la libertad de empresa (Artículo 16) y a la libertad artística y científica (Artículo 13) por razones de interés público y para proteger derechos fundamentales como la salud, la seguridad, la protección del consumidor u otros, así como a la propiedad intelectual (Artículo 17.2), ante los requerimientos de transparencia e información, cuya divulgación estaría sometida en cualquier caso a la Directiva 2016/943[532].

La Propuesta de Reglamento objeto de análisis incorpora en su Exposición de Motivos el reconocimiento de la labor del Parlamento Europeo, afirmando que la misma tiene en cuenta la citada Resolución del mismo en materia ética de 20 de octubre de 2020.

6.2.2. Convergencia, interoperabilidad, integración y coherencia con el marco legal vigente y otras políticas.

El carácter horizontal de la propuesta justifica la aplicación de mecanismos de coherencia con los marcos reguladores en la UE de aplicación a los distintos sectores donde ya se utilizan o es previsible que se utilicen sistemas de inteligencia artificial, así como con la Carta de Derechos Fundamentales de la UE y la legislación en materia de protección de datos, protección del consumidor, no discriminación e igualdad de género.

En este sentido, esta propuesta se remite y complementa en distintos aspectos al RGPD, con un conjunto de normas armonizadas aplicables al diseño, desarrollo y uso de ciertos sistemas de inteligencia artificial de alto riesgo y restricciones sobre ciertos usos de los

[532] Directiva (UE) 2016/943 del Parlamento Europeo y del Consejo, de 8 de junio de 2016, relativa a la protección de los conocimientos técnicos y la información empresarial no divulgados (secretos comerciales) contra su obtención, utilización y revelación ilícitas. OJ L 157, 15.6.2016. Pp. 1-18.

sistemas de identificación biométrica remota. E igualmente lo hace respecto de la legislación de la UE en materia de no discriminación, al objeto de minimizar la denominada "discriminación algorítmica", en especial, en relación con el diseño y la calidad de los conjuntos de datos utilizados para el desarrollo de sistemas de inteligencia artificial.

Del mismo modo, la propuesta pretende integrarse con la legislación de seguridad de los productos existente con el objetivo precitado de garantizar la coherencia, evitar doble regulación y minimizar cargas adicionales, en el caso de sistemas de inteligencia artificial de alto riesgo que constituyan componentes de seguridad de los productos.

En este sentido, la UE pretende adoptar un Nuevo Marco Legislativo -*NLF* por sus siglas en inglés- que establezca las reglas para poner un producto en el mercado, basándose en la evaluación de la conformidad con los requisitos de salud y seguridad existentes. El *NLF* definirá el papel de los distintos operadores económicos involucrados -fabricantes, importadores o distribuidores-, los principios para la acreditación de organismos habilitados, su papel y sus funciones, y el marcado CE, que deberá estar presente en el producto.

La Comisión Europea publicó en fechas coetáneas a la de la publicación de la Propuesta de Reglamento sobre inteligencia artificial objeto de análisis, una Propuesta de Reglamento sobre máquinas para la revisión de la Directiva de Máquinas 2006/42/CE.

El objetivo de este nuevo marco legislativo es garantizar la seguridad general del producto final, por lo que sin perjuicio de los requisitos de seguridad exigidos en la Propuesta de Reglamento por el que se establecen normas armonizadas sobre la inteligencia artificial de 21 de abril de 2021, la nueva regulación sobre máquinas se focalizará en los requisitos específicos y adicionales con respecto a la integración segura de un sistema de inteligencia artificial en el producto final.

En este sentido, respecto de los productos regulados por el precitado Nuevo Marco Legislativo -*NLF*-, entre otros, maquinaria, dispositivos médicos o juguetes, cuando integren sistemas de inteligencia artificial de alto riesgo, los requisitos de éstos requeridos

en la Propuesta de Reglamento objeto de análisis serán comprobados en los procedimientos de evaluación de conformidad generales previstos en el precitado *NLF*.

Sin embargo, el Reglamento propuesto no será de aplicación directa en el caso de sistemas de inteligencia artificial de alto riesgo relacionados con productos regulados por la legislación del denominado "antiguo enfoque" que incluye aviación o automóviles, sin perjuicio de la necesaria consideración de los requisitos establecidos para los sistemas de inteligencia artificial de alto riesgo, para la adopción de legislación de ejecución o delegada.

La precitada coherencia también se prevé en el marco de los servicios financieros, de modo que las autoridades responsables de la supervisión de la legislación de servicios financieros de la UE deben ser designadas como autoridades competentes para supervisar los requisitos de la Propuesta de Reglamento objeto de análisis, considerando que "los sistemas de inteligencia artificial están, en cierta medida, regulados implícitamente en relación con el sistema de gobierno interno de las entidades de crédito".

También prevé su coherencia con la legislación en materia de servicios, incluido los de intermediación regulados por la Directiva 2000/31/CE sobre comercio electrónico[533] y la reciente propuesta de la Comisión Europea para la *Ley de servicios digitales* (DSA)[534].

Respecto de los sistemas de inteligencia artificial que se hallen integrados en sistemas de TI a gran escala en el denominado Espacio de Libertad, Seguridad y Justicia[535], la propuesta no se aplicará a los sistemas inteligentes que sean comercializados o puestos

[533] Directiva 2000/31/CE del Parlamento Europeo y del Consejo, de 8 de junio de 2000, relativa a determinados aspectos jurídicos de los servicios de la sociedad de la información, en particular el comercio electrónico en el mercado interior (Directiva sobre el comercio electrónico). Diario Oficial n° L 178 de 17/07/2000. Pp. 1-16

[534] Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a un mercado único de servicios digitales (Ley de servicios digitales) y por el que se modifica la Directiva 2000/31/CE.
COM/2020/825 final

[535] El espacio de libertad, seguridad y justicia (ELSJ) es el nombre con que se designa un conjunto de políticas y actuaciones que la Unión Europea despliega, esencialmente dentro pero también fuera de sus fronteras, para lograr el objetivo1 de crear un área compartida entre sus Estados miembros donde se alcance un alto grado de cooperación y coordinación política, policial y judicial a nivel comunitario que facilite la seguridad interior, una justicia eficaz y una fuerte protección de las libertades públicas para sus ciudadanos.

en servicio antes de que haya transcurrido un año desde la fecha de aplicación del futuro Reglamento, salvo excepciones.

La propuesta integra un paquete de medidas más amplio que pretende abordar los problemas que supone el desarrollo y uso de la inteligencia artificial reflejados en el *Libro blanco sobre inteligencia artificial*, y con las que pretende garantizar su coherencia y complementariedad y las políticas relacionadas.

Este paquete incluye la revisión de la legislación sectorial de productos -Directiva de maquinaria precitada o Directiva de seguridad general de productos precitadas- e iniciativas en curso sobre responsabilidad relacionadas con nuevas tecnologías, incluyendo la inteligencia artificial, en donde considero podría entenderse enmarcada la previa Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial, de 20 de octubre de 2020, la cual deberá ser congruente con esta nueva propuesta y debe complementarse con el objeto de proporcionar seguridad jurídica y crear el pretendido ecosistema de confianza en la UE.

También pretende ser coherente con la estrategia digital global de la Comisión, especialmente en su objetivo de que la tecnología funcione para las personas y los objetivos definidos en la Comunicación *Dar forma al futuro digital de Europa*[536].

También con la Ley de Gobernanza de Datos[537], la Directiva de Datos abiertos[538] y otras iniciativas bajo la *Estrategia Europea de Datos*[539]. En la fecha de cierre de esta investigación el Comité de Industria, Investigación y Energía del Parlamento Europeo ha adoptado una posición favorable a la nueva Ley de Gobernanza de Datos, por lo que prosigue su tramitación.

---

[536] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, Brújula digital 2030: el camino europeo para la década digital COM / 2021/118 final.
[537] Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre gobernanza europea de datos (Ley de gobernanza de datos). COM / 2020/767 final
[538] Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, sobre datos abiertos y reutilización de la información del sector público. DO L 172 de 26.6.2019. P. 56–83
[539] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones Una estrategia europea para los datos. COM / 2020/66 final

Y, por último, la propuesta pretende contribuir a conformar las normas y estándares mundiales y promover una inteligencia artificial confiable, conforme una base para continuar la colaboración internacional sobre esta materia.

6.2.3. Fundamento legal

La propuesta pretende garantizar el correcto funcionamiento del mercado interior mediante una armonización legislativa en esta materia al amparo del artículo 114 y 116 del Tratado de Funcionamiento de la Unión Europea (TFUE) que evite la fragmentación, especialmente ante las iniciativas de algunos Estados miembros para regular la inteligencia artificial que ya se están produciendo.

Conforme recoge su memoria explicativa los enfoques nacionales sólo crearán inseguridad jurídica, barreras adicionales y, consecuentemente, ralentizarán la adopción de la inteligencia artificial por el mercado.

En consecuencia, considera que un marco normativo europeo armonizado garantizará la igualdad de condiciones, protegerá a las personas y fortalecerá la competitividad, lo que de antemano podría ser cuestionado desde un enfoque internacional, especialmente a la vista de los requerimientos que regula para *startups,* que pueden considerar este marco un posible obstáculo inicial y valorar constituirse en otros países terceros, llevándose consigo el talento.

La propuesta define requisitos obligatorios comunes *ex ante,* aplicables al diseño y desarrollo de determinados sistemas de inteligencia artificial antes de su comercialización, así como controles *ex post,* una vez comercializados o puestos en funcionamiento en el mercado, por lo que su eficacia real, como he expuesto anteriormente, alcanza no solo la comercialización, puesta a disposición o uso de los sistemas regulados, sino también a su diseño.

Asimismo, hace referencia en su memoria expositiva a futuras guías de apoyo y herramientas de cumplimiento que ayudarán a proveedores y usuarios a cumplir con sus requerimientos y minimizar sus costes, los cuales considera firmemente proporcionados

a los objetivos pretendidos y supuestos beneficios económicos y reputacionales de los operadores.

Por último, la elección de un reglamento como instrumento legal se fundamenta en la necesidad de una aplicación uniforme del nuevo marco de directa aplicación en todos los Estados miembros, argumentando que sus disposiciones no son demasiado prescriptivas y permiten cierto margen de acción a los Estados miembros en determinados aspectos que no afecten a sus objetivos principales, en particular, la organización interna del sistema de vigilancia del mercado y la adopción de medidas para fomentar la innovación.

6.2.4. Proporcionalidad

La propuesta se presenta formalmente como "proporcionada", basada en los marcos legales existentes y necesaria, construida sobre un enfoque basado en el riesgo y que impone cargas regulatorias cuando es probable que un sistema inteligente represente un alto riesgo para los derechos fundamentales y la seguridad.

El enfoque inicial es positivo, en la medida que la intervención legislativa no se plantea exclusivamente necesaria cuando dicho alto riesgo pueda afectar a la vida, la salud o los derechos fundamentales de las personas, sino también a la seguridad.

No obstante, en mi opinión, dicha proporcionalidad resulta discutible en determinados aspectos y, desde luego, considero la propuesta insuficiente, al no contemplar principios y normas éticas básicas y obligaciones consecuentes para el resto de sistemas de inteligencia artificial distintos a los considerados de alto riesgo.

6.2.5. Consultas previas

La propuesta hace referencia en su memoria expositiva a los resultados de la consulta pública efectuada, en la que se recogen las advertencias de varias partes interesadas de evitar la duplicidad legislativa, el conflicto de obligaciones y la regulación excesiva, así

como la proporcionalidad y neutralidad tecnológica del marco regulatorio, lo que justifica la misma y los cambios frente a las Resoluciones previas de 20 de octubre de 2020 del Parlamento Europeo precitadas.

Como no podría ser de otra forma, en la consulta pública indicada se puso de relieve la necesidad de una definición estrecha, clara y precisa de la inteligencia artificial, así como la necesidad de definir conceptos como "riesgo", "alto riesgo", "bajo riesgo", "identificación biométrica remota" o "daño", lo que es una necesidad por razones de seguridad jurídica.

La mayoría de las aportaciones se mostraron favorables a un enfoque basado en el riesgo, si bien, los tipos de riesgos y amenazas se consideró que deben basarse en un enfoque sector a sector y caso a caso, debiendo tener en cuenta para su evaluación, no sólo estos factores sino el impacto en los derechos y la seguridad. De nuevo, no se habla de la probabilidad de esos riesgos para su calificación, focalizándose en su impacto y activos a proteger.

Por último, más del 50% de las opiniones recabadas fueron partidarias de un sistema que combinara una autoevaluación de riesgos inicial con una posterior para los sistemas de alto riesgo, conforme recoge su memoria expositiva precitada.

6.2.6. Estructura del Reglamento propuesto.

La Propuesta del Reglamento consta de ochenta y cinco (85) artículos agrupados en doce (12) Títulos, precedidos de una extensa sede expositiva con ochenta y nueve (89) Considerandos, siguiendo la estela de las propuestas previas de 20 de octubre de 2020 y Reglamentos previos, como el Reglamento General de Protección de Datos (RGPD).

6.3. Objeto.

El artículo 1 del Reglamento propuesto regula su objeto.

Su tenor literal invita a pensar que el mismo parece alejarse de la pretensión de regular la tecnología, en este caso, la inteligencia artificial, y aproximarse a una pretendida neutralidad tecnológica, cuando sus prescripciones y requerimientos van más allá de la mera comercialización, puesta en servicio o uso de sistemas de inteligencia artificial, en la medida que establece requerimientos en su propio diseño y concepción.

El Reglamento propuesto regula la inteligencia artificial, pero se focaliza en la denominada inteligencia artificial "débil", conforme a la definición "sesgada" de la misma, siendo ésta a partir de la que se construyen sus disposiciones.

Según establece el precepto indicado, el Reglamento propuesto regula, de un lado, un conjunto de normas armonizadas para la comercialización, la puesta en servicio y el uso de sistemas de inteligencia artificial en la UE, de otro un conjunto de prohibiciones de determinados usos de inteligencia artificial, de otro un conjunto de requisitos específicos para los sistemas de inteligencia artificial de alto riesgo y de obligaciones para los operadores de dichos sistemas, de otro reglas de transparencia armonizadas para los sistemas de inteligencia artificial específicos -sistemas destinados a interactuar con personas físicas, reconocimiento de emociones, categorización biométrica y para generar o manipular contenido de imagen, audio o video- y, por último, un conjunto de normas sobre seguimiento y vigilancia del mercado.

6.4. Una nueva definición de inteligencia artificial.

La necesidad de definir previamente la realidad que pretende regular la propuesta es incuestionable, especialmente ante la falta de consenso a nivel científico, ético y jurídico, conforme ha sido analizado con profundidad en el capítulo I de esta investigación.

Conforme recoge el Considerando 6 de la Propuesta de Reglamento, la noción de sistema de inteligencia artificial debe definirse claramente para garantizar la seguridad jurídica y, al mismo tiempo, proporcionar la flexibilidad necesaria para adaptarse a futuros desarrollos tecnológicos.

Según la misma, la definición debe basarse en las características funcionales clave del *software*, en particular la capacidad, para un conjunto dado de objetivos definidos por humanos, de generar resultados como contenido, predicciones, recomendaciones o decisiones que influyen en el entorno con el que el sistema interactúa, ya sea en una dimensión física o digital.

Los sistemas de inteligencia artificial pueden diseñarse para funcionar con distintos niveles de autonomía y utilizarse de forma independiente o como componente de un producto, independientemente de si el sistema está integrado físicamente en el producto (integrado) o sirve a la funcionalidad del producto. sin estar integrado en el mismo (no incrustado).

El artículo 3 de la Propuesta de Reglamento define "Sistema de inteligencia artificial" (sistema de IA) como el *software* desarrollado con una o más de las técnicas y enfoques enumerado en el Anexo I y que puede, para un conjunto dado de objetivos definidos por el ser humano, generar resultados como contenido, predicciones, recomendaciones o decisiones que influyen en los entornos con los que interactúan. Como he anticipado, desaparece de dicha definición el atributo de "autonomía".

De inicio, la definición dista mucho de lo que considero debe ser la deseable claridad, concreción y sencillez a la hora de definir un concepto jurídico esencial, técnico y tan complejo, que además constituye el objeto de regulación de esta propuesta y de todo un conjunto de normas derivadas tanto de *hard law* como *soft law*.

Adicionalmente, la definición se integra y complementa necesariamente por las complejas técnicas y enfoques listados en el precitado anexo que luego analizaré, que se presentan como un *numerus clausus*, si bien, la Comisión Europea quedaría facultada por el futuro Reglamento para adoptar actos delegados para actualizar dichos listados conforme a la evolución de la tecnología y del mercado, conforme regula su artículo 4, si bien, "sobre la base de característica similares a las técnicas y enfoques enumerados en el mismo", lo que a mi juicio constituye un reduccionismo innecesario.

Este último inciso no me parece afortunado en la medida que desde un enfoque de seguridad jurídica y de riesgos, no debería circunscribirse esas actualizaciones a que se

traten de técnicas o enfoques con características similares dado que, desde un punto de vista técnico, supone una restricción inicial innecesaria e inadecuada, en la medida que podrán aparecer técnicas y enfoques con características distintas con la evolución de inteligencia artificial y su interacción o integración con otras tecnologías que exigirán su inclusión en dicho listado.

Las técnicas y enfoques de inteligencia artificial que integra el Anexo I precitado son los siguientes:

a) Enfoques de aprendizaje automático, incluido el aprendizaje supervisado, no supervisado y por refuerzo, utilizando una amplia variedad de métodos, incluido el aprendizaje profundo, es decir, los enfoque sobre los que se sustenta la inteligencia artificial actual.

b) Enfoques basados en la lógica y el conocimiento, incluida la representación del conocimiento, la programación inductiva (lógica), las bases de conocimiento, los motores de inferencia y deductivos, el razonamiento (simbólico) y los sistemas expertos.

c) Enfoques estadísticos, estimación bayesiana, métodos de búsqueda y optimización, cuya inclusión comportará la aplicación de este Reglamento propuesto a los sistemas que sustenten en los mismos.

En definitiva, sólo el *software* que sea desarrollado bajo cualquiera de estas técnicas o enfoques o con una combinación de todas o algunas de ellas, será considerado un sistema de inteligencia artificial a los efectos del Reglamento propuesto, siempre que tenga la capacidad de generar resultados como contenido, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúan, en relación con un conjunto dado de objetivos definidos por el ser humano.

Sin duda, una definición compleja a nivel técnico y a mi juicio excesivamente genérica y amplia respecto de los sistemas contemplados y no exhaustiva, donde el grado de autonomía del sistema desaparece de la definición y se muestra como irrelevante para su

calificación como tal, por lo que casi cualquier *software* común con cualquiera de esas capacidades y características sería considerado inteligencia artificial.

Conforme analicé en el capítulo I, la autonomía ha integrado la definición de la inteligencia artificial en todos los documentos de trabajo, resoluciones y propuestas del Parlamento Europeo y de la Comisión hasta la irrupción de esta nueva definición incorporada en el Reglamento propuesto.

De hecho, la precedente *Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre inteligencia artificial: Cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal*[540], reiteraba la necesidad de adoptar un marco jurídico europeo común con definiciones armonizadas y principios éticos comunes, pidiendo formalmente a la Comisión que adoptara las definiciones de "Inteligencia artificial" y de "autonomía" del siguiente modo, en la línea de las Resoluciones y Propuestas precedentes de 20 de octubre de 2020:

- "Sistema de IA": Todo sistema basado en programas informáticos o incorporado en dispositivos físicos que muestra un comportamiento que simula la inteligencia, entre otras cosas, mediante la recopilación y el tratamiento de datos, el análisis y la interpretación de su entorno y la adopción de medidas, con cierto grado de autonomía, para lograr objetivos específicos.

- "Autónomo": Todo sistema de inteligencia artificial que funciona interpretando determinados datos de entrada y utilizando un conjunto de instrucciones predeterminadas, sin limitarse a ellas, a pesar de que el comportamiento del sistema esté limitado y orientado a cumplir el objetivo que se le haya asignado y otras decisiones pertinentes de diseño tomadas por su desarrollador.

[540] Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre inteligencia artificial: Cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal (2020/2013(INI)).

De nuevo, esta definición incorpora la alusión a que dicha autonomía comporta el funcionamiento del sistema conforme a unas instrucciones predeterminadas. pero no hallándose limitado a las mismas, lo que parece referirse a sistemas dotados de una inteligencia artificial más avanzada que la considerada "débil", como he comentado en capítulos anteriores, al abordar estas definiciones.

La definición de la reciente propuesta objeto de análisis difiere absolutamente de las definiciones propuestas por el Parlamento Europeo a la Comisión.

De manera asociada a este concepto, la Propuesta de Reglamento también define jurídicamente algunos sistemas inteligentes objeto de regulación específica en función de sus características.

En primer lugar, define los "sistemas de reconocimiento de emociones" como aquellos sistemas de inteligencia artificial destinados a identificar o inferir emociones o intenciones de personas físicas sobre la base de sus datos biométricos.

En segundo lugar, define los "sistemas de categorización biométrica" como aquellos sistemas de inteligencia artificial destinados a asignar a las personas físicas categorías específicas, como sexo, edad, color de cabello, color de ojos, tatuajes, origen étnico u orientación sexual o política, sobre la base de sus datos biométricos.

Y, en tercer lugar, define con detalle el concepto de "sistema de identificación biométrica remota", diferenciando entre aquellos que dispongan de la posibilidad de hacerlo en tiempo real y los que no a los efectos de cualificar el riesgo como inadmisible y consecuente prohibición de estos.

Estos sistemas inteligentes se definen como aquellos destinados a identificar a las personas físicas a distancia mediante la comparación de los datos biométricos de una persona con los datos biométricos contenidos en una base de datos de referencia, y sin el conocimiento previo del usuario del sistema de inteligencia artificial si la persona estará presente y podrá ser identificada.

Entre los mismos, define los que permiten la identificación en tiempo real, que son aquellos que captan los datos biométricos, comparan e identifican sin un retraso

significativo, lo que comprende tanto la identificación instantánea como aquella que se produzca "con breves retrasos limitados para evitar la elusión".

6.5. Ámbito objetivo

El Reglamento propuesto se aplicará a los sistemas de inteligencia artificial expresamente regulados en el mismo, de modo que no será exigible al resto, tal y como se infiere del artículo 2 del mismo. Además, el mismo recoge expresamente la exclusión de los sistemas de inteligencia artificial desarrollados o utilizados exclusivamente con fines militares, conforme se recogía en el *Libro blanco de inteligencia artificial* y en las propuestas precedentes de 20 de octubre de 2020.

No obstante, para los sistemas de inteligencia artificial que sean considerados de alto riesgo que sean componentes de seguridad de productos o sistemas, o que sean en sí mismos productos o sistemas, y que entren en el ámbito de aplicación de las normas que relaciona su artículo 2.2 y que expongo a continuación, únicamente se aplicará el artículo 84 del Reglamento propuesto (evaluación y revisión del Reglamento), quedando sujetos los mismos a su específico marco regulador.

En relación con esta exclusión parcial expresa considero necesario abordar, en primer lugar, cuando un sistema de este tipo se considera componente de seguridad de productos o sistemas a los efectos del Reglamento propuesto.

En este sentido, la Propuesta de Reglamento aclara esta cuestión en las definiciones contenidas en su artículo 3, considerando que se tratará de aquellos sistemas inteligentes de alto riesgo que constituyan un componente de un producto o de un sistema cumpliendo una función de seguridad de ese producto o sistema, o cuyo fallo o mal funcionamiento pone en peligro la salud y la seguridad de las personas o de la propiedad. Estos sistemas quedarán sujetos a los marcos específicos que regulan dichos productos o sistemas.

El artículo 2.2. del Reglamento propuesto relaciona los marcos normativos a los que quedarían sujetos los sistemas precitados en el párrafo anterior, que he agrupado en los distintos sectores:

- Aviación civil: Reglamento (CE) n° 300/2008 del Parlamento Europeo y del Consejo, de 11 de marzo de 2008, sobre normas comunes para la seguridad de la aviación civil y por el que se deroga el Reglamento (CE) n° 2320/2002[541]. Y Reglamento (UE) 2018/1139 del Parlamento Europeo y del Consejo, de 4 de julio de 2018, sobre normas comunes en el ámbito de la aviación civil y por el que se crea una Agencia de la Unión Europea para la Seguridad Aérea y por el que se modifican los Reglamentos (CE) n.° 2111/2005, (CE) n.° 1008/2008, (UE) n.° 996/2010, (CE) n.° 376/2014 y las Directivas 2014/30/UE y 2014/53/UE del Parlamento Europeo y del Consejo y se derogan los Reglamentos (CE) n.° 552/2004 y (CE) n.° 216/2008 del Parlamento Europeo y del Consejo y el Reglamento (CEE) n.° 3922/91 del Consejo[542].

- Vehículos agrícolas o forestales: Reglamento (UE) n ° 167/2013 del Parlamento Europeo y del Consejo, de 5 de febrero de 2013, relativo a la homologación de los vehículos agrícolas o forestales, y a la vigilancia del mercado de dichos vehículos[543].

- Vehículos de dos o tres ruedas o cuatriciclos: Reglamento (UE) n.° 168/2013 del Parlamento Europeo y del Consejo, de 15 de enero de 2013, relativo a la homologación de los vehículos de dos o tres ruedas y los cuatriciclos, y a la vigilancia del mercado de dichos vehículos[544].

- Equipos marinos: Directiva 2014/90/UE del Parlamento Europeo y del Consejo, de 23 de julio de 2014, sobre equipos marinos, y por la que se deroga la Directiva 96/98/CE del Consejo[545].

[541] OJ L 97, 9.4.2008. Pp. 72-84
[542] OJ L 212, 22.8.2018. Pp. 1-122
[543] OJ L 60, 2.3.2013. Pp. 1-51
[544] OJ L 60, 2.3.2013. Pp. 52-128
[545] OJ L 257, 28.8.2014. Pp. 146-185

- Interoperabilidad sistema ferroviario: Directiva (UE) 2016/797 del Parlamento Europeo y del Consejo, de 11 de mayo de 2016, sobre la interoperabilidad del sistema ferroviario dentro de la Unión Europea[546].

- Vehículos a motor y remolques: Reglamento (UE) 2018/858 del Parlamento Europeo y del Consejo, de 30 de mayo de 2018, sobre la homologación y la vigilancia del mercado de los vehículos de motor y sus remolques y de los sistemas, los componentes y las unidades técnicas independientes destinados a dichos vehículos, por el que se modifican los Reglamentos (CE) n.° 715/2007 y (CE) n° 595/2009 y por el que se deroga la Directiva 2007/46/CE[547]. Reglamento (UE) 2019/2144 del Parlamento Europeo y del Consejo de 27 de noviembre de 2019 relativo a los requisitos de homologación de tipo de los vehículos de motor y de sus remolques, así como de los sistemas, componentes y unidades técnicas independientes destinados a esos vehículos, en lo que respecta a su seguridad general y a la protección de los ocupantes de los vehículos y de los usuarios vulnerables de la vía pública, por el que se modifica el Reglamento (UE) 2018/858 del Parlamento Europeo y del Consejo y se derogan los Reglamentos (CE) n.º 78/2009, (CE) n.º 79/2009 y (CE) nº 661/2009 del Parlamento Europeo y del Consejo y los Reglamentos (CE) n.º 631/2009, (UE) n.º 406/2010, (UE) n.º 672/2010, (UE) n.o 1003/2010, (UE) n.º 1005/2010, (UE) n.º 1008/2010, (UE) n.º 1009/2010, (UE) n.o 19/2011, (UE) n.º 109/2011, (UE) n.º 458/2011, (UE) n.º 65/2012, (UE) n.o 130/2012, (UE) n.º 347/2012, (UE) n.º 351/2012, (UE) n.º 1230/2012 y (UE) 2015/166 de la Comisión[548].

[546] OJ L 138, 26.5.2016. Pp. 44-101
[547] OJ L 151, 14.6.2018. Pp. 1-218
[548] OJ L 325, 16.12.2019. Pp. 1-40

6.6. Ámbito subjetivo

El marco jurídico establecido por el texto propuesto se aplicará tanto a agentes públicos como privados ubicados en la UE o fuera de la misma cuando el sistema inteligente se introduzca en la UE o su uso afecte a personas establecidas en la misma.

El Reglamento propuesto, conforme regula en su artículo 2.1., se aplicará a los proveedores que comercialicen o pongan en servicio sistemas de inteligencia artificial en la UE, con independencia de que dichos proveedores estén establecidos en la UE o en un tercer país, cuando la comercialización o puesta en servicio se produzca en la misma, lo que comporta la eficacia transnacional del instrumento normativo propuesto y la protección del mercado unionista, de modo que los requerimientos serán exigibles cualquiera que sea la ubicación geográfica del proveedor del sistema.

A los efectos de esta propuesta, se considera "proveedor", conforme a la definición contenida en su artículo 3, como toda persona física o jurídica, autoridad pública, agencia u otro organismo que desarrolle un sistema de inteligencia artificial o que disponga de un sistema de inteligencia artificial desarrollado con el fin de comercializarlo o ponerlo en servicio con su propio nombre o marca comercial, ya sea de pago o de forma gratuita[549].

En consecuencia, el proveedor podrá tener o no la condición de diseñador, desarrollador, fabricante o ser el mero comercializador o distribuidor.

En este sentido, el texto propuesto define igualmente el concepto de "Proveedor a pequeña escala", como aquellos proveedores que sean microempresas o una pequeña

[549] Al cierre esta investigación se ha generado una primera versión en castellano del Reglamento propuesto para su publicación. En relación con la misma, destacar que esta versión en castellano evidencia una traducción de la definición de "proveedor" que se aleja parcialmente de su redacción original en inglés, en particular, lo define como "*toda persona física o jurídica, autoridad pública, agencia u organismo de otra índole que desarrolle un sistema de IA o para el que se haya desarrollado un sistema de IA con vistas a introducirlo en el mercado o ponerlo en servicio con su propio nombre o marca comercial, ya sea de manera remunerada o gratuita*", en la medida que considera "proveedor" a esa persona, entidad o autoridad "*para la que se haya desarrollado un sistema de IA*" con estos objetivos, cuando la versión originariamente redactada y publicada en inglés no recoge esto, sino "que disponga de un sistema de IA" con el fin de comercializarlo o ponerlo en servicio bajo su propio nombre o marca, ya sea a título oneroso o gratuito.

empresa en el sentido previsto en la Recomendación 2003/261/CE[550] de la Comisión Europea. Posteriormente analizaré algunas particularidades previstas en distintos aspectos por el Reglamento para este tipo de proveedores.

Del mismo modo, el Reglamento propuesto se aplicará a los usuarios de sistemas de inteligencia artificial ubicados dentro de la UE, así como a los proveedores y usuarios de sistemas de inteligencia artificial ubicados en un tercer país, cuando el resultado producido por el sistema se utilice en la Unión.

El concepto usuario es esencial acotarlo para delimitar el ámbito subjetivo de aplicación.

El artículo 3 de la Propuesta define el concepto "usuario", y lo hace considerando como tal cualquier persona física o jurídica, autoridad pública, agencia u otro organismo que utilice un sistema de inteligencia artificial bajo su autoridad, excepto cuando el sistema de inteligencia artificial se utilice en el curso de una actividad personal no profesional, por lo tanto, cuando una persona utilice un sistema de inteligencia artificial en el ámbito doméstico o personal no se considerará "usuario" a los efectos de este Reglamento y, consecuentemente, no se hallará sujeto a sus obligaciones.

En relación con ambos conceptos y a los efectos del análisis y comentarios que efectuaré a continuación, el artículo 3 del Reglamento propuesto define igualmente el concepto de "operador", considerando como tal al proveedor, el usuario, el representante autorizado, el importador y el distribuidor, en atención al contexto.

El Reglamento propuesto no se aplicará a las autoridades de un tercer país ni a las organizaciones internacionales que entren en el ámbito de aplicación del presente Reglamento con arreglo al apartado 1, cuando dichas autoridades u organizaciones utilicen sistemas de inteligencia artificial en el marco de acuerdos internacionales de cooperación policial y judicial con la Unión o con uno o varios Estados miembros.

A sensu contrario, sí resultaría de aplicación a las mismas, cuando lo hagan al margen de estos acuerdos, lo que a mi juicio supone un avance respecto de las actividades que vienen

[550] Recomendación de la Comisión, de 6 de mayo de 2003, sobre la definición de microempresas, pequeñas y medianas empresas. OJ L 124, 20.5.2003. P.p. 36-41.

llevando a cabo determinadas autoridades y servicios de determinados gobiernos mediante la inteligencia artificial, supuestamente en interés público y de la ciudadanía, lo que sería un tema apasionante sobre el que profundizar, pero que excede del objeto y alcance de esta investigación.

Sobre esta cuestión, el precitado Dictamen conjunto 5/2021, del Comité Europeo de Protección de Datos con el Supervisor Europeo de Protección de Datos sobre el Reglamento propuesto, refleja una gran preocupación por parte de ambos organismos ante el riesgo de elusión de la normativa.

Asimismo, el Reglamento propuesto también afectará a los importadores de sistemas de inteligencia artificial, en la medida que estos deberán asegurarse que el proveedor extranjero haya efectuado el procedimiento de evaluación de conformidad exigido en el mismo y de que aquél disponga de la documentación técnica requerida por el Reglamento propuesto. Además, los importadores deberán asegurarse de que su sistema lleve el marcado europeo de conformidad (CE) y vaya acompañado de la documentación e instrucciones precisas.

Por último, el artículo 2.5 del Reglamento propuesto introduce un aspecto directamente relacionado con el eje sobre el que se ha construido esta investigación, esto es, la responsabilidad, y en particular, sobre la responsabilidad de los prestadores de servicios intermediarios en la Directiva 2000/31/CE del Parlamento Europeo y del Consejo -Capítulo II, Sección IV-, que será sustituida por la futura Ley de Servicios Digitales[551], de modo que las disposiciones sobre responsabilidad recogidas en ésta serán plenamente aplicables a los sistemas de inteligencia artificial.

[551] Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a un mercado único de servicios digitales (Ley de servicios digitales) y por el que se modifica la Directiva 2000/31/CE. COM/2020/825 final

6.7. Clasificación de los sistemas de inteligencia artificial

Durante esta investigación se han analizado las distintas clasificaciones de los sistemas de inteligencia artificial desde distintos enfoques y disciplinas, tanto científicas, filosóficas, éticas como jurídicas, si bien, el Reglamento propuesto se aleja del concepto de autonomía y otras capacidades de los sistemas inteligentes, a diferencia de las definiciones contenidas en los informes, documentos de trabajo, resoluciones y propuestas en el seno de la UE previamente, para clasificar los sistemas desde un enfoque puramente de riesgos y de manera *seudo objetiva*, en función de sus funcionalidades principales y usos potenciales.

De este modo, la Comisión Europea, a través del Reglamento propuesto, para distinguir cinco (5) tipos de sistemas inteligentes en atención a la valoración del riesgo de los mismos: Sistemas prohibidos debido a su riesgo inaceptable y permitidos exclusivamente en determinados contextos, sistemas de alto riesgo, sistema de riesgo medio o limitado, sistemas de riesgo bajo o mínimo y luego el resto de sistemas, de riesgo inexistente o no clasificados, en relación con los cuales, como he referido anteriormente, se me hace difícil incardinar un sistema inteligente de riesgo "0" por las propias características intrínsecas a nivel técnico de estos sistemas y las posibles capacidades de las que puedan estar dotados.

6.7.1. Sistemas de inteligencia artificial de riesgo inadmisible: Prohibidos

El Reglamento propuesto recoge un conjunto de sistemas de riesgo inaceptable que se hayan sujetos a una prohibición absoluta o relativa por contravenir los valores de la UE, en caso de concurran determinadas circunstancias.

a) Sistemas con capacidad para manipular el comportamiento humano.

El Reglamento propuesto, como he comentado anteriormente, los define en su artículo 5.1.a) como los sistemas de inteligencia artificial que despliegan técnicas subliminales más allá de la conciencia de una persona para distorsionar materialmente el comportamiento de una persona de una manera que causa o pueda causarle a esa persona u otra daños físico o psicológico.

Es decir, sistemas inteligentes que manipulen el comportamiento de la persona sin consciencia por parte de la misma y que comporten un daño real o potencial a la misma o a terceros, ya sea físico o psicológico.

En mi opinión sería deseable una mayor concreción del concepto "distorsión", que debería incluir aspectos como la modificación, manipulación o control de la conducta. ¿Quedarían prohibidos los algoritmos que hoy manipulan y dirigen nuestra navegación y comportamiento *online* en redes y plataformas? No parece que se pretenda llegar hasta este punto, pero exigiría una revisión de su tenor literal, sin perjuicio de que debieran estar prohibidos o exigir información previa adecuada, cuando menos, respecto de colectivos vulnerables.

Asimismo, en su artículo 5.1.b) se prohíben igualmente aquellos sistemas de inteligencia artificial que exploten "cualquiera de las vulnerabilidades de un grupo específico de personas debido a su edad, discapacidad física o mental", con el fin de distorsionar materialmente el comportamiento de una persona perteneciente a ese grupo de manera que cause o pueda causar a esa persona o a otra un daño físico o psicológico.

De este modo, respecto de estos colectivos -y considero incluidos tanto menores de edad y personas de edad avanzada en función del contexto-, se prohíben expresamente los sistemas que exploten su vulnerabilidad con intención de manipular su comportamiento, con independencia de que lleguen a hacerlo o no, es decir, basta la mera existencia del riesgo, lo que supone un matiz diferenciador respecto de apartado anterior, en el que se requiere la manipulación efectiva.

Conforme establece el precepto indicado, se prohíbe en ambos supuestos tanto su puesta en el mercado, como su puesta en servicio o uso.

b) Sistemas de puntuación social o "social scoring" público.

El Reglamento propuesto prohíbe igualmente en su artículo 5.1.c) -cuando su puesta en el mercado, en servicio o su utilización sea por parte de las autoridades públicas o en su nombre-, los sistemas de inteligencia artificial para la evaluación o la clasificación de la fiabilidad de las personas físicas durante un determinado período de tiempo, sobre la base de su comportamiento social o de sus características personales o de personalidad conocidas o previstas, con otorgamiento de una puntuación social que conduzca a los siguientes resultados, todos o cualquiera de ellos:

- El tratamiento perjudicial o desfavorable de determinadas personas físicas o grupos enteros de ellas en contextos sociales que no guarden relación con los contextos en los que se generaron o recopilaron originalmente los datos.

- El tratamiento perjudicial o desfavorable de determinadas personas físicas o grupos enteros de ellas que no esté justificado o sea desproporcionado con respecto a su comportamiento social o a su gravedad.

Es decir, para su prohibición, estos sistemas tienen que tener estos efectos, todos o uno de ellos, de modo que la evaluación y perfilado de la población por los poderes públicos mediante sistemas inteligentes podría no estar inicialmente prohibido, salvo que produzcan cualquiera de los resultados precitados. Como he referido en los capítulos precedentes, hemos tenido un caso reciente en Países Bajos.

A sensu contrario, ello conlleva a considerar que dichos sistemas se hallarían permitidos cuando la evaluación y clasificación se refiera a personas jurídicas así como, aun refiriéndose a personas físicas, sean puestos en el tráfico, en

funcionamiento o uso por entidades privadas -salvo que actúen en nombre de autoridades públicas-, por ejemplo redes sociales, que podrían puntuar a los usuarios conforme a la evaluación o clasificación de su fiabilidad en base de su comportamiento social o de sus características personales o de personalidad, con las consecuencias que las mismas definan.

No obstante, recordar que el tratamiento no discriminatorio, injustificado o desproporcionado por los poderes públicos se haya ya contemplado en otras normas, así como en la propia Constitución Española a nivel nacional y en el Convenio Europeo de Derechos Humanos.

c) Sistemas de identificación biométrica remota "en tiempo real" en espacios de acceso público con el propósito de hacer cumplir la ley.

El artículo 5.1.d) prohíbe estos sistemas, sean públicos o privados, salvo que su uso sea estrictamente necesario para cumplir cualquier de los objetivos relacionados en dicho precepto:

- La búsqueda selectiva de posibles víctimas específicas de delitos, incluidos los niños desaparecidos;

- La prevención de una amenaza específica, sustancial e inminente para la vida o la seguridad física de personas físicas o de un atentado terrorista;

- La detección, localización, identificación o enjuiciamiento de un autor o sospechoso de un delito contemplado en el artículo 2, apartado 2, de la Decisión Marco 2002/584/JAI[552] del Consejo y sancionable en el Estado miembro de que se trate con una pena o una medida de seguridad

[552] 2002/584/JAI: Decisión Marco del Consejo, 13 de junio de 2002, relativa a la orden de detención europea ya los procedimientos de entrega entre Estados miembros - Declaraciones realizadas por algunos Estados miembros con ocasión de la adopción de la Decisión marco. DO L 190 de 18.7.2002. Pp. 1-20.

> privativa de libertad de un máximo de, al menos, tres años, según determine la legislación de dicho Estado miembro.

De inicio, en mi opinión, quizás debería realizarse una reflexión más profunda y matización sobre la tipología y gravedad de los delitos en cuestión para valorar la concurrencia de la excepción.

El uso de los sistemas de identificación biométrica remota "en tiempo real" en espacios de acceso público con fines policiales para cualquiera de los objetivos precitados deberá, en todo caso, considerar la naturaleza de la situación que da lugar al posible uso, es decir el contexto, en términos de análisis de riesgos, debiendo considerar la gravedad, probabilidad y escala del daño causado en ausencia de uso de dicho sistema, así como las consecuencias del uso del sistema para los derechos y libertades de las personas interesadas, esto es, la gravedad, probabilidad y escala de estas consecuencias.

En definitiva, conforme establece el artículo 5.2. del Reglamento propuesto, se exigiría hacer un análisis de riesgos y evaluación de impacto previa donde se analice el contexto, así como la probabilidad, el impacto y la magnitud del daño consecuente y se concluya su necesidad y proporcionalidad. Y, además, de todo ello deberá cumplir las salvaguardias y condiciones necesarias y proporcionadas en relación con su uso, en especial respecto a limitaciones temporales, geográficas y personales conforme a dicho precepto.

Adicionalmente a todo ello, de conformidad con lo dispuesto en el artículo 5.3, cada uso individual de estos sistemas deberá estar sujeto a autorización previa otorgada por una autoridad judicial o administrativa independiente del Estado miembro en el que vaya a producirse el uso, previa solicitud motivada -que entiendo debería acompañar la evaluación de impacto precitada aunque no se indica nada al respecto en el texto propuesto ni su mera referenciación-, salvo en casos de urgencia debidamente justificada, en cuyo caso, podría iniciarse su uso sin autorización, aunque no se exime de su solicitud posterior durante o después de su uso.

Como he referido anteriormente, este concepto de "urgencia" debería ser concretado para evitar la inseguridad jurídica que conlleva.

En este sentido, el texto podría haber adicionado alguna concreción más específica respecto del momento en que deba producirse esta solicitud en casos de urgencia, dado que considero que se debería producir a la mayor posible desde su puesta en uso.

La autorización se otorgará cuando se determine que su uso es necesario y proporcionado en términos de riesgos para lograr los objetivos precitados, conforme regula dicho artículo.

No obstante, los Estados miembros tendrán la facultad de decidir la posibilidad de autorizar total o parcialmente el uso de este tipo de sistemas, dentro de los límites y en las condiciones precitadas.

El Reglamento propuesto establece en su artículo 5.4 que el Estado que así lo determine, deberá regular en su legislación nacional las normas necesarias para la solicitud, la expedición y el ejercicio de las autorizaciones indicadas, así como para su supervisión, especificando respecto de que objetivos y delitos recogidos en dicho precepto, las autoridades competentes podrán estar autorizadas a utilizar dichos sistemas con fines policiales.

Sin perjuicio de las consideraciones precedentes, quizás se debería contemplar expresamente que, en el caso de que la prohibición inicial de estos sistemas quede sin efecto en virtud de las excepciones contempladas, dichos sistemas serán en cualquier caso considerados de alto riesgo con sujeción a lo dispuesto en el Reglamento propuesto. Del mismo sería interesante determinar el alcance de "remoto" y "tiempo real", ante el uso de drones en la captación o cámaras de policías.

6.7.2. Sistemas de inteligencia artificial de alto riesgo

El Reglamento propuesto contempla un conjunto limitado de sistemas inteligentes que considera de alto riesgo para los derechos y libertades de los individuos -especialmente la salud, la seguridad y los derechos fundamentales-, los cuales se hayan permitidos pero sujetos un marco de requisitos y obligaciones reforzadas con el objetivo de garantizar su uso legal, ético, robusto y seguro.

Los sistemas de inteligencia artificial calificados como tales conforme al mismo constituye el núcleo del instrumento normativo propuesto.

El artículo 6 del Reglamento propuesto regula las reglas para clasificar un sistema de inteligencia artificial de alto riesgo, basándose en la finalidad prevista del sistema de inteligencia artificial, en congruencia con la legislación vigente de la UE en materia de seguridad de los productos, de modo que la clasificación del riesgo dependerá esencialmente de la función desempeñada por el sistema, de la finalidad y de los usos específicos.

La Comisión incluye entre los criterios de referencia para la clasificación propuesta, el alcance de uso de la aplicación y su finalidad, el número de personas potencialmente afectadas, la dependencia respecto del resultado, la irreversibilidad de los daños, así como el grado en el que la legislación vigente de la UE prevé medidas eficaces para prevenir o minimizar sustancialmente sus riesgos.

El Reglamento propuesto aporta pues la seguridad jurídica de su inclusión expresa y se centra en los sectores más proclives a generar daños, si bien, de antemano, echo en falta algunos, como luego abordaré.

En primer lugar, se consideran sistemas de inteligencia artificial de alto riesgo los especificados “objetivamente” en el Anexo III del Reglamento propuesto en atención a su objeto o sector donde operan:

La clasificación del riesgo se basa principalmente en la finalidad prevista del sistema, en consonancia con la legislación vigente de la UE en materia de seguridad

de los productos, a la que hice referencia en el capítulo II de esta investigación. Realmente en su finalidad prevista y modalidad de uso:

1) Identificación biométrica y categorización de personas físicas.

Esta categoría incluye los sistemas de inteligencia artificial destinados a ser utilizados para la identificación biométrica remota "en tiempo real" y "posterior" de personas físicas, a excepción de los prohibidos en el artículo 5.

2) Gestión y operación de infraestructuras críticas.

Esta categoría hace referencia a los sistemas de inteligencia artificial destinados a ser utilizados como componentes de seguridad en la gestión y operación del tráfico rodado y el suministro de agua, gas, calefacción y electricidad.

3) Educación y formación profesional.

Esta categoría incluye:

a) Sistemas de inteligencia artificial destinados a ser utilizados con el fin de determinar el acceso o asignar personas físicas a instituciones educativas y de formación profesional.

b) Sistemas de inteligencia artificial destinados a ser utilizados con el fin de evaluar a los estudiantes en instituciones educativas y de formación profesional y para evaluar a los participantes en las pruebas comúnmente requeridas para la admisión a las instituciones educativas.

4) Empleo, gestión de trabajadores y acceso al autoempleo.

Esta categoría incluye:

a) Sistemas de inteligencia artificial destinados a ser utilizados para la contratación o selección de personas físicas, en particular para publicitar vacantes, seleccionar o filtrar solicitudes, evaluar candidatos en el curso de entrevistas o pruebas.

b) Sistema de inteligencia artificial destinados a ser utilizados para tomar decisiones sobre la promoción y terminación de relaciones contractuales relacionadas con el trabajo, para la asignación de tareas y para monitorear y evaluar el desempeño y el comportamiento de las personas en tales relaciones.

5) Acceso y disfrute de servicios privados esenciales y servicios y beneficios públicos.

Esta categoría incluye:

a) Los sistemas de inteligencia artificial destinados a ser utilizados por las autoridades públicas o en nombre de las autoridades públicas para evaluar la elegibilidad de las personas físicas para los beneficios y servicios de asistencia pública, así como para otorgar, reducir, revocar o reclamar dichos beneficios y servicios;

b) Los sistemas de inteligencia artificial destinados a ser utilizados para evaluar la solvencia de las personas físicas o establecer su puntaje crediticio, con la excepción de los sistemas de inteligencia artificial puestos en servicio por proveedores de pequeña escala, conforme definí anteriormente, para su propio uso.

c) Los sistemas de inteligencia artificial destinados a ser utilizados para el envío o para establecer la prioridad en el envío de servicios de primera respuesta de emergencia, incluidos los bomberos y la asistencia médica.

6) Cumplimiento de la ley.

Esta categoría incluye:

a) Sistemas de inteligencia artificial destinados a ser utilizados por las autoridades encargadas de hacer cumplir la ley para realizar evaluaciones de riesgo individuales de personas físicas con el fin de evaluar el riesgo de una persona física de cometer un delito o reincidencia o el riesgo de posibles víctimas de delitos penales. En relación con la problemática que plantean estos sistemas en la actualidad, me remito a su análisis en los capítulos anteriores, especialmente respecto del sesgo presente;

b) Sistemas de inteligencia artificial destinados a ser utilizados por las autoridades policiales como polígrafos y herramientas similares o para detectar el estado emocional de una persona física;

c) Los sistemas de inteligencia artificial destinados a ser utilizados por las autoridades encargadas de hacer cumplir la ley para detectar falsificaciones profundas o *deep fakes* reguladas en el artículo 52.3. del Reglamento propuesto;

d) Sistemas de inteligencia artificial destinados a ser utilizados por las autoridades encargadas de hacer cumplir la ley para evaluar la fiabilidad de las pruebas en el curso de la investigación o el enjuiciamiento de delitos penales;

e) Sistemas de inteligencia artificial destinados a las autoridades encargadas de hacer cumplir la ley para predecir la aparición o repetición de un delito real o potencial basándose en la elaboración de perfiles de personas físicas a que se refiere el artículo 3, apartado 4, de la *Directiva (UE) 2016/680* o en la evaluación de rasgos de personalidad. y características o comportamiento delictivo pasado de personas físicas o grupos;

f) Sistemas de inteligencia artificial destinados a ser utilizados por las autoridades policiales para la elaboración de perfiles de personas físicas a que

se refiere el artículo 3, apartado 4, de la Directiva (UE) 2016/680 en el curso de la detección, investigación o enjuiciamiento de delitos penales;

g) Sistemas de inteligencia artificial destinados a ser utilizados para análisis de delitos relacionados con personas físicas, lo que permite a las autoridades policiales buscar grandes conjuntos de datos complejos relacionados y no relacionados disponibles en diferentes fuentes de datos o en diferentes formatos de datos para identificar patrones desconocidos o descubrir relaciones ocultas en los datos.

7) Gestión de migración, asilo y control de fronteras.

Esta categoría incluye:

a) Sistemas de inteligencia artificial destinados a ser utilizados por las autoridades públicas competentes como polígrafos y herramientas similares o para detectar el estado emocional de una persona física;

b) Sistemas de inteligencia artificial destinados a ser utilizados por las autoridades públicas competentes para evaluar un riesgo, incluido un riesgo para la seguridad, un riesgo de inmigración irregular o un riesgo para la salud, planteado por una persona física que tiene la intención de entrar o ha entrado en el territorio de un Estado miembro;

c) Sistemas de inteligencia artificial destinados a ser utilizados por las autoridades públicas competentes para la verificación de la autenticidad de los documentos de viaje y la documentación justificativa de las personas físicas y detectar documentos no auténticos comprobando sus características de seguridad;

d) Sistemas de inteligencia artificial destinados a ayudar a las autoridades públicas competentes a examinar las solicitudes de asilo, visados y permisos de residencia y las quejas asociadas con respecto a la elegibilidad de las personas físicas que solicitan un estatuto.

8) Administración de justicia y procesos democráticos.

Esta categoría hace referencia a los sistemas de inteligencia artificial destinados a ayudar a una autoridad judicial a investigar e interpretar los hechos y la ley y a aplicar la ley a un conjunto concreto de hechos.

Se trata pues de una lista cerrada, si bien, al igual que las propuestas de reglamentos precedentes, se introducen mecanismos ágiles de actualización de los mismos, con el objetivo de mantenerlos permanentes actualizados a la realidad tecnológica y social, en particular, mediante la facultad de la Comisión Europea para adoptar actos delegados que actualicen y adicionen nuevos sistemas inteligentes, conforme recogen los artículos 7, 73, siguientes y concordantes de la Propuesta de Reglamento, si bien, cuando se cumplan las dos condiciones siguientes:

- Los sistemas de inteligencia artificial están destinados a utilizarse en cualquiera de los ámbitos generales enumerados anteriormente;

- Los sistemas de inteligencia artificial planteen un riesgo de daño a la salud y la seguridad, o un riesgo de impacto adverso sobre los derechos fundamentales que sea, con respecto a su probabilidad y gravedad, equivalente o mayor al riesgo de daño o de impacto adverso que suponen los sistemas de inteligencia artificial de alto riesgo anteriormente relacionados.

  A la vista de la redacción contenida en el artículo 7.1 del Reglamento propuesto que exige las dos condiciones precitadas, es obvio que se requerirá un análisis de riesgos y una evaluación de impacto para determinar el nivel de riesgo y si el mismo sería equivalente, en términos de impacto, al que suponen los sistemas de inteligencia artificial ya clasificados.

  Además, conforme anticipé en mis consideraciones generales sobre el Reglamento propuesto, plantea un reduccionismo innecesario que, a mi juicio, debe revisarse, en la medida que exige que esos sistemas de inteligencia artificial susceptibles de clasificación de alto riesgo estén destinados a utilizarse en cualquiera de los ámbitos generales enumerados anteriormente en el precepto precitado. Ello excluiría otros

ámbitos que pueden ser igualmente críticos en el futuro ante la evolución incesante del desarrollo, despliegue y aplicación de la inteligencia artificial.

La redacción propuesta plantea, en mi opinión, un escenario confuso, en la medida que, respecto de los nuevos sistemas, se deberá calificar previamente el riesgo en función de su probabilidad y de su impacto, lo que podría comportar, en función de los parámetros y criterios de valoración del riesgo previamente adoptados para su análisis, que determinados riesgos sean considerados de probabilidad muy baja o despreciable pero que su impacto sea alto o máximo, pero que sin embargo al asociar ambos criterios, el valor del riesgo final específico sea medio y no alto o máximo, cuando los sistemas que integran el listado incorporada en el Anexo III se consideran de manera objetiva como de alto riesgo en base principalmente a su finalidad y sector, con independencia de su probabilidad, que podría ser en muchos casos baja, lo que comportaría en el futuro tener que dejar fuera de esta consideración y de su inclusión en este listado, sistemas con riesgos de impacto flagrantemente alto pero de probabilidad baja.

A título meramente ilustrativo sobre esta cuestión, incorporo esta tabla muy básica de asociación entre probabilidad e impacto, obviamente adaptable en función de los activos a proteger y su tolerancia al riesgo, para detenernos en cuál podría ser el valor del riesgo final en caso de probabilidad baja o muy despreciable:

RIESGO		**IMPACTO**		
		Bajo	**Medio**	**Alto**
PROBABILIDAD	**Baja**	Insignificante	Bajo	Medio
	Media	Bajo	Medio	Alto
	Alta	Medio	Alto	Muy alto

Es por ello, por lo que considero que debería revisarse y reflexionar sobre la consideración de la “la probabilidad” en el análisis y valoración para la adopción de un acto de delegado, en congruencia con los criterios utilizados para clasificar sistemáticamente los sistemas que integran la lista del Anexo III.

Adicionalmente, el precitado artículo 7.2 incorpora distintos criterios a considerar por la Comisión para la evaluación de un sistema para su inclusión en dicho listado, en particular, el objetivo previsto del sistema, su utilización, la causación previa de daños, el alcance potencial del daño o impacto adverso, dependencia de las personas potencialmente afectadas, vulnerabilidad de las mismas, su reversibilidad o la existencia en la legislación vigente de la UE medidas efectivas de reparación en relación con los sus riesgos -con exclusión de las reclamaciones de los daños y perjuicios-, así como medidas eficaces para prevenir o minimizar sustancialmente esos riesgos.

En este sentido, entiendo que estos son los criterios que también han llevado a la Comisión a proponer los sistemas listados en el Anexo III para su catalogación como de alto riesgo.

En segundo lugar, también se consideran sistemas de inteligencia artificial de alto riesgo aquellos que constituyan un componente de seguridad de un producto contemplado en la legislación sectorial de la UE o que constituyan el propio producto, siempre que estén sujetos a una evaluación de conformidad por terceros conforme a dicha legislación.

De este modo, en adición a los sistemas incluidos en el precitado Anexo III, el artículo 6.1 del Reglamento propuesto establece que se considerarán sistemas de inteligencia artificial de alto riesgo, con independencia de que se comercialicen o se pongan en servicio independientemente de los productos que se indican a continuación, los que cumplan las dos condiciones siguientes:

a) El sistema de inteligencia artificial vaya ser utilizado como componente de seguridad de un producto, o sea en sí mismo un producto, cubierto por la legislación de armonización de la Unión enumerada en el Anexo II del Reglamento propuesto;

b) El producto cuyo componente de seguridad es el sistema de inteligencia artificial o el propio sistema de inteligencia artificial como producto, se halle sometido a una evaluación de la conformidad por parte de un tercero con

vistas a la comercialización o puesta en servicio de dicho producto conforme a la legislación de armonización de la Unión enumerada en el precitado Anexo II.

La legislación de armonización a la que alude dicho Anexo II, que integra la que constituye el denominado *Nuevo Marco Legislativo* -NLF- al que he aludido anteriormente, es la siguiente:

A. Legislación de armonización de la UE basada en el *Nuevo Marco Legislativo*.

1. Directiva 2006/42/CE del Parlamento Europeo y del Consejo, de 17 de mayo de 2006, sobre máquinas, y por la que se modifica la Directiva 95/16/CE[553];

2. Directiva 2009/48/CE del Parlamento Europeo y del Consejo, de 18 de junio de 2009, sobre la seguridad de los juguetes[554];

3. Directiva 2013/53/UE del Parlamento Europeo y del Consejo, de 20 de noviembre de 2013, sobre embarcaciones de recreo y motos acuáticas y por la que se deroga la Directiva 94/25/CE[555].

4. Directiva 2014/33/UE del Parlamento Europeo y del Consejo, de 26 de febrero de 2014, sobre la armonización de las legislaciones de los Estados miembros relativas a ascensores y componentes de seguridad para ascensores[556];

5. Directiva 2014/34/UE del Parlamento Europeo y del Consejo, de 26 de febrero de 2014, sobre la armonización de las legislaciones de los Estados

[553] DO L 157 de 9.6.2006. P. 24. Derogada por el Reglamento de máquinas
[554] DO L 170 de 30.6.2009. P. 1
[555] DO L 354 de 28.12.2013.P. 90
[556] DO L 96 de 29.3.2014. P. 251

miembros relativas a equipos y sistemas de protección destinados a su uso en atmósferas potencialmente explosivas[557];

6. Directiva 2014/53 / UE del Parlamento Europeo y del Consejo, de 16 de abril de 2014, sobre la armonización de las leyes de los Estados miembros relativas a la comercialización de equipos radioeléctricos y por la que se deroga la Directiva 1999/5/CE[558];

7. Directiva 2014/68/UE del Parlamento Europeo y del Consejo, de 15 de mayo de 2014, sobre la armonización de las legislaciones de los Estados miembros relativas a la comercialización de equipos a presión[559];

8. Reglamento (UE) 2016/424 del Parlamento Europeo y del Consejo, de 9 de marzo de 2016, sobre instalaciones de transporte por cable y por el que se deroga la Directiva 2000/9/CE[560];

9. Reglamento (UE) 2016/425 del Parlamento Europeo y del Consejo, de 9 de marzo de 2016, sobre equipos de protección individual y por el que se deroga la Directiva 89/686/CEE del Consejo[561];

10. Reglamento (UE) 2016/426 del Parlamento Europeo y del Consejo, de 9 de marzo de 2016, sobre aparatos que queman combustibles gaseosos y por el que se deroga la Directiva 2009/142/CE[562];

11. Reglamento (UE) 2017/745 del Parlamento Europeo y del Consejo, de 5 de abril de 2017, sobre productos sanitarios, que modifica la Directiva 2001/83/CE, el Reglamento (CE) no 178/2002 y el Reglamento (CE) no 1223/2009 y deroga Directivas del Consejo 90/385/CEE y 93/42/CEE[563];

[557] DO L 96 de 29.3.2014. P. 309
[558] DO L 153, 22.5.2014. P. 62
[559] DO L 189 de 27.6.2014. P. 164
[560] DO L 81 de 31.3.2016. P. 1
[561] DO L 81 de 31.3.2016. P. 51
[562] DO L 81 de 31.3.2016. P. 99
[563] DO L 117 de 5.5.2017. P. 1

12. Reglamento (UE) 2017/746 del Parlamento Europeo y del Consejo, de 5 de abril de 2017, sobre productos sanitarios para diagnóstico in vitro y por el que se deroga la Directiva 98/79/CE y la Decisión 2010/227/UE de la Comisión[564].

B. Otra legislación de armonización de la UE

1. Reglamento (CE) 300/2008 del Parlamento Europeo y del Consejo, de 11 de marzo de 2008, sobre normas comunes en el ámbito de la seguridad de la aviación civil y por el que se deroga el Reglamento (CE) 2320/2002[565].

2. Reglamento (UE) 168/2013 del Parlamento Europeo y del Consejo, de 15 de enero de 2013, sobre la homologación y la vigilancia del mercado de vehículos y cuatriciclos de dos o tres ruedas[566];

3. Reglamento (UE) 167/2013 del Parlamento Europeo y del Consejo, de 5 de febrero de 2013, sobre la homologación y la vigilancia del mercado de vehículos agrícolas y forestales[567];

4. Directiva 2014/90/UE del Parlamento Europeo y del Consejo, de 23 de julio de 2014, sobre equipos marinos y por la que se deroga la Directiva 96/98/CE del Consejo[568];

5. Directiva (UE) 2016/797 del Parlamento Europeo y del Consejo, de 11 de mayo de 2016, sobre la interoperabilidad del sistema ferroviario dentro de la Unión Europea[569].

6. Reglamento (UE) 2018/858 del Parlamento Europeo y del Consejo, de 30 de mayo de 2018, sobre la homologación y vigilancia del mercado de los vehículos de motor y sus remolques, y de los sistemas, componentes y

564 DO L 117 de 5.5.2017. P. 176
565 DO L 97 de 9.4.2008. P. 72
566 DO L 60 de 2.3.2013. P. 52
567 DO L 60 de 2.3.2013. P. 1
568 DO L 257 de 28.8.2014. P. 146
569 DO L 138 de 26.5.2016. P. 44

unidades técnicas independientes destinados a dichos vehículos, que modifica los Reglamentos (CE) No 715/2007 y (CE) no 595/2009 y por la que se deroga la Directiva 2007/46/CE[570]; Reglamento (UE) 2019/2144 del Parlamento Europeo y del Consejo, de 27 de noviembre de 2019, sobre los requisitos de homologación de tipo de los vehículos de motor y sus remolques, y los sistemas, componentes y unidades técnicas independientes destinados a dichos vehículos, en lo que respecta a su seguridad y protección de los ocupantes de vehículos y usuarios vulnerables de la vía pública, que modifica el Reglamento (UE) 2018/858 del Parlamento Europeo y del Consejo y deroga el Reglamento (CE) 78/2009.

7. Reglamento (UE) 2018/1139 del Parlamento Europeo y del Consejo, de 4 de julio de 2018, sobre normas comunes en el ámbito de la aviación civil y por el que se crea una Agencia de seguridad aérea de la Unión Europea, y se modifica el Reglamento (CE) n.o 2111/2005, (CE) 1008/2008, (UE) 996/2010, (UE) 376/2014 y las Directivas 2014/30/UE y 2014/53/UE del Parlamento Europeo y del Consejo, y por el que se derogan los Reglamentos (CE) 552/2004 y (CE) no 216/2008 del Parlamento Europeo y del Consejo y el Reglamento (CEE) 3922/91 del Consejo[571], en lo que respecta al diseño, la producción y comercialización de las aeronaves contempladas en el artículo 2, apartado 1, letras a) y b), cuando se trate de aeronaves no tripuladas y sus motores, hélices, piezas y equipos para controlarlas a distancia.

Todos estos sistemas de alto riesgo deberán cumplir una serie de requisitos que serán analizados con posterioridad, en particular, de calidad de los conjuntos de datos utilizados, de transparencia, de supervisión humana, de solidez, de exactitud y de ciberseguridad, que se consideran necesarios para mitigar los riesgos para los derechos fundamentales y la seguridad y que no están cubiertos o, el menos adecuadamente, por los marcos jurídicos existentes. Muchos de ellos constituyen algunos de los principios y valores éticos esenciales que conforman la gran mayoría

[570] DO L 151 de 14.6.2018. P. 1
[571] DO L 212 de 22.8.2018. P. 1

de los marcos éticos propuestos a nivel internacional y también europeo, de modo que el Reglamento propuesto convertiría estos en vinculantes y su implementación en un requerimiento jurídico.

La infracción de los mismos permitirá a las autoridades nacionales acceder a la información necesaria para su investigación, sin perjuicio de su sanción conforme a los sistemas sancionadores que establezcan los Estados miembros conforme a las bases ya previstas en el Reglamento propuesto, así como a la depuración de las responsabilidades civiles de naturaleza contractual y extracontractual para el resarcimiento de los daños sufridos.

En mi opinión, considero que dentro de los sistemas de riesgo inaceptable como de alto riesgo, deberían aparecer otros sectores y usos específicos dentro de los mismos, que tampoco aparecen en la categoría de alto riesgo en las Resoluciones y Propuestas precedentes del Parlamento Europeo, de 20 de octubre de 2020, en particular, el sector social-sanitario y atención a la discapacidad, el sector salud en general o la biomedicina, sin perjuicio de que existan marcos reguladores sectoriales que contemplan y den solución a muchos de los aspectos relacionados con sus riesgos, incluso que permiten su consideración como componentes de seguridad o su sometimiento a una evaluación de conformidad, pero no a todos.

Asimismo, remitiéndome a mis consideraciones realizadas al analizar la Resolución y Propuesta de Reglamento en el ámbito ético de 20 de octubre de 2020, considero que, sin perjuicio de que se establezca un *numerus clausus* de sistemas de alto riesgo, el futuro marco regulador de la inteligencia artificial debe prever la consideración como tales de otros que, con independencia del sector y uso predeterminado en el que operen, evidencien un alto riesgo en relación con la salud, la seguridad y los derechos fundamentales, a la vista de su probabilidad, impacto y contexto. Este aspecto ya se contemplaba en la precitada propuesta de octubre de 2020.

Del mismo modo, resulta sorprendente que no se contemplen como alto riesgo los sistemas inteligentes que permitan el reconocimiento y manipulación de emociones, especialmente significativo en el caso de afectación a un colectivo de personas

especialmente vulnerables por su edad o discapacidad física o mental, así como sistemas de *profiling* con ciertas finalidades en el sector privado.

Por otra parte, respecto de algunos sistemas de alto riesgo, podría ser cuestionable su clasificación como tales ante su posible consideración como de riesgo inaceptable.

Y, adicionalmente a las limitaciones comentadas para su calificación como tales, debo significar mi percepción de la excesiva focalización en ocasiones en la perspectiva de la inteligencia artificial desde la protección de datos de datos, cuando sus retos y riesgos, como he expuesto anteriormente, no se circunscriben exclusivamente a personas físicas, sino a todo tipo de entidades públicas y privadas, gobiernos y bienes supranacionales, como es el propio medio ambiente y la conservación del mundo en el que vivimos. Existen otros bienes y derechos y otros sujetos a los que pueden afectar los riesgos potenciales de la inteligencia artificial.

En mi opinión, me parece positivo partir del concepto de seguridad del propio Reglamento General de Protección de Datos (RGPD), como la capacidad de garantizar la confidencialidad, la integridad, disponibilidad y resiliencia permanentes de los sistemas y servicios de tratamiento, si bien, los activos implicados en la seguridad en relación con sistemas inteligentes y sus riesgos, son todos, personales, empresariales o de negocio o gubernamentales, así como públicos y privados.

En conclusión, considero que estos aspectos deberán ser considerados en su tramitación, donde el papel del Parlamento Europeo será determinante para la reflexión y maduración de estas y otras cuestiones.

6.7.3. Riesgo limitado

El Reglamento propuesto contempla un conjunto de sistemas de inteligencia artificial específicos que, aun no siendo calificables de alto riesgo, conllevan determinados riesgos que motivan la imposición de obligaciones específicas,

especialmente cuando concurran riesgos de confusión para la persona sobre si está interactuando con una persona o una máquina y de manipulación.

Sin embargo, no clasifica estos sistemas desde un enfoque de riesgos, de modo que siguiendo la clasificación utilizada por el Reglamento propuesto para el resto podrían considerarse de riesgo limitado o medio, como asistentes virtuales o *chatbots*.

La realidad práctica evidencia la existencia de sistemas de inteligencia artificial que no son calificables como de riesgo inaceptable o alto riesgo, pero que conforme a cualquier metodología de análisis de riesgos debería ser clasificados, cuanto menos de riesgo medio, cuya probabilidad posiblemente sea en la mayoría de los casos baja (o no) pero el impacto alto.

Estos sistemas clasificables de este modo no estarían sujetos al Reglamento propuesto salvo en el caso de los sistemas expresamente previstos, de modo que sería absolutamente voluntario el cumplimiento de las distintas obligaciones previstas en el mismo, sin perjuicio de que, obviamente, deban desarrollarse y utilizarse de conformidad con la legislación vigente.

Como luego referiré al analizar los sistemas considerados de riesgo mínimo o no clasificados, considero que se debería contemplar y exigir a todo sistema de inteligencia artificial clasificable en este nivel distinto al alto riesgo, un análisis de riesgos previo y una evaluación de impacto por parte de proveedores y usuarios antes de su comercialización o uso, es decir, una autoevaluación, similar a los análisis de riesgos y evaluaciones que exige el actual RGPD, y ello en base a la características y contexto de la propia tecnología sobre la que se sustentan, capacidades de las que pueden estar dotados y en congruencia con el enfoque de riesgos sobre el que se ha construido el Reglamento propuesto.

Todo ello permitiría que los riesgos serían previamente identificados y analizados, el impacto en la salud, la seguridad, los derechos fundamentales y en los demás derechos y bienes implicados susceptibles de afectación evaluado, y el sistema clasificado, aunque ello no comporte obligaciones adicionales para proveedores y

usuarios distintas a las generales contempladas en otros marcos jurídicos aplicables a los que se hallen sujetos, permitiendo una adecuada y consecuente gestión de dichos riesgos.

6.7.4. Riesgo mínimo

Sistemas inteligentes en los que el riesgo resultante de evaluar, supuestamente y desde un enfoque de riesgos, el sector en el que opere, el uso previsto, el contexto, la probabilidad y el impacto de sus riesgos para la salud, la seguridad y los demás derechos y bienes implicados susceptibles de afectación, resulta bajo y mínimo. De modo que, a juicio de la Comisión Europea, todos estos sistemas pueden desarrollarse y utilizarse con arreglo a la legislación vigente sin obligaciones jurídicas adicionales.

6.7.5. Sistemas no clasificados o de riesgo inexistente

Los sistemas inteligentes de riesgo inexistente, como he manifestado en mis consideraciones previas sobre el Reglamento propuesto, los considero inconcebibles, salvo en supuestos específicos por las propias características y posibles capacidades de la inteligencia artificial, dado que considero que siempre existirá un riesgo inherente y residual, aceptable o no, en virtud de su graduación. En mi opinión, el "riesgo algorítmico" es inherente a la inteligencia artificial, tal y como está siendo concebida en las distintas propuestas regulatorias hasta la fecha.

La última categoría correspondería a los sistemas inteligentes no clasificados, cuestión que debería estar expresamente prohibida por los futuros marcos regulatorios, cuando menos de cara a su puesta en funcionamiento, comercialización o uso, cualquiera que sea el nivel de riesgo finalmente resultante de la misma, debiendo exigirse como mínimo una autoevaluación previa de todo sistema de inteligencia artificial.

En consecuencia, a excepción de los sistemas de riesgo inaceptable o alto riesgo, el resto de sistemas de inteligencia artificial quedarían fuera del ámbito de aplicación del futuro Reglamento, no quedando sujetos pues a ninguna obligación en particular, de modo que sería absolutamente voluntario el cumplimiento de los distintos requisitos y obligaciones previstas en el Reglamento propuesto, sin perjuicio de que, obviamente, deban desarrollarse y utilizarse de conformidad con la legislación vigente.

El Reglamento propuesto se remite a los códigos de conducta de adscripción voluntaria para estos sistemas.

Para la Comisión Europea la inmensa mayoría de los sistemas de inteligencia artificial utilizados en la actualidad pertenecería a estas otras categorías.

Se trataría pues de sistemas de riesgo limitado o mínimo que el Reglamento propuesto opta por no clasificar formalmente siguiendo la estrategia por la que ha optado la Comisión, así como ante la complejidad o más bien imposibilidad de su clasificación previa, dada la inmensidad de tipologías, características y usos, si bien, a mi juicio, esto no debería obstar para identificarlos y establecer algún marco mínimo para estos sistemas inteligentes en congruencia con el enfoque de riesgos sobre el que se ha construido esta propuesta, a pesar del modelo regulador por el que ha optado la Comisión Europea.

En este sentido, como he expuesto anteriormente, se me hace francamente difícil pensar en la existencia de sistemas inteligentes que se relacionen e interaccionen con personas y cosas que no comporten algún tipo de riesgo inherente o residual, ya sea en función de su probabilidad o de su impacto (aunque ambos fueran poco significativos), en la medida que el riesgo, cuando menos, debería ser clasificado como bajo. Y, aun así, me resulta difícil pensar, desde un punto de vista de análisis de riesgos, en la existencia de amenazas de probabilidad muy baja que no puedan conllevar impactos altos.

En mi opinión, como he referido, considero que, respecto de cualquier otro sistema, se debería contemplar un análisis de riesgos previo preceptivo y evaluación de impacto por parte de proveedores y usuarios antes de su comercialización o uso, es decir, una autoevaluación, al objeto de que los riesgos sean previamente identificados y analizados,

el impacto evaluado y el sistema clasificado, posibilitando una adecuada gestión de dichos riesgos.

No obstante, como luego reiteraré, esta obligación de análisis, evaluación y clasificación debería venir acompañada de otras exigencias mínimas con origen en los marcos éticos analizados en anteriores capítulos, incorporados incluso en la Propuesta de Reglamento de octubre de 2020, de modo que se promuevan aspectos como la ética, la seguridad, la privacidad y el cumplimiento regulatorio en el diseño para aquellos sistemas que pretendan ser puestos en el mercado y usados.

6.8. Requerimientos y obligaciones para los sistemas de IA de alto riesgo

El Reglamento propuesto, como he anticipado, se focaliza en los sistemas de inteligencia artificial considerados de alto riesgo estableciendo un marco de *hard law* para los mismos, y siempre en referencia a una inteligencia artificial "débil".

Por el contrario, respecto del resto de sistemas de inteligencia artificial se limita a regular algunas obligaciones generales de transparencia e información para concretos sistemas, pero no de calidad, gobernanza de datos, control y supervisión humana, precisión, robustez o seguridad, invitando a que se creen marcos de *soft law* a los que puedan sujetarse voluntariamente, en particular, futuros códigos de conducta.

El cumplimiento de estos requisitos legalmente exigidos no evita todos los riesgos que pueden plantear estos sistemas, en la medida que, si a pesar de su cumplimiento y adopción de medidas, se producen daños, podrá existir responsabilidad civil y, su caso, penal, aunque será necesario valorar en qué medida la responsabilidad queda excluida o reducida, en su caso, como consecuencia del cumplimiento y aplicación de la medidas exigidas, que evidenciaría una actuación diligente dirigida a minimizar los riesgos, en función del régimen de responsabilidad exigible que pueda considerar ésta. En este sentido, como he referido, el Reglamento propuesto no aborda las cuestiones de responsabilidad, que si fueron objeto de una propuesta previa en la Recomendación del

Parlamento Europeo de 20 de octubre de 2020 sobre esta materia, que será analizada en el próximo capítulo.

Asimismo, los requisitos y posteriores obligaciones que regula el Reglamento propuesto no suponen el único y exclusivo marco a cumplir por los sistemas inteligentes sujetos a su ámbito de aplicación, sino que los mismos se hallarán sujetos a los demás marcos generales o sectoriales de aplicación que puedan afectar a su diseño, comercialización y uso, por ejemplo, en materia de privacidad o marcos reguladores de procedimientos, y tanto en el ámbito público o privado.

Esta cuestión ha sido refrendada en el reciente y precitado Dictamen conjunto 5/2021, del Comité Europeo de Protección de Datos con el Supervisor Europeo de Protección de Datos sobre el Reglamento propuesto, el cual además considera que el cumplimiento de las obligaciones derivadas del marco jurídico de la UE, incluyendo el relativo a la protección de datos personales, debería ser una condición previa para que se permita la introducción en el mercado europeo de un sistema inteligente como producto con marcado CE. Y de nuevo, a mi juicio, esta exigencia debería ser general respecto de cualquier sistema de inteligencia, cualquiera que sea su nivel de riesgo inicial, como pueden ser un *chatbot*.

Sin embargo, ambos organismos consideran que el requisito de garantizar el cumplimiento de estos marcos debería incluirse en el capítulo relativo a los requisitos exigidos a los sistemas de inteligencia artificial de alto riesgo, y que se debería adaptar el procedimiento de evaluación de conformidad de la propuesta para que los terceros realicen siempre las evaluaciones de conformidad de los sistemas de inteligencia artificial de alto riesgo *ex ante*.

Durante los siguientes subapartados expondré los requisitos y obligaciones a los que se hallan sujetos dichos sistemas y los agentes que se relacionen con los mismos desde un inicio, esto es, desde su diseño, pasando por su comercialización o puesta a disposición, así como con posterioridad mediante su certificación, seguimiento y monitorización.

6.8.1. Requerimientos para los sistemas de inteligencia artificial de alto riesgo

La Propuesta de Reglamento del Parlamento Europeo y del Consejo, sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, de 20 de octubre de 2020, analizada en anteriores apartados, recogía un conjunto de principios éticos generales exigibles a la inteligencia artificial y a las tecnologías conexas exigibles a todos los sistemas de inteligencia artificial, sin perjuicio de regular un marco concreto de principios y obligaciones éticas y jurídicas específicas para los sistemas de alto riesgo, en particular y entre otros, los de control y supervisión humana, seguridad, transparencia, trazabilidad, información, ausencia de sesgo y discriminación, responsabilidad social, igualdad de género, sostenibilidad medioambiental, derecho de resarcimiento, evaluación de riesgos, evaluación de conformidad y el certificado europeo de conformidad ética.

La nueva Propuesta de Reglamento, sin embargo, establece un conjunto de requerimientos y obligaciones exclusivamente para los sistemas de alto riesgo que difiere sensiblemente de la propuesta precitada, omitiendo igualmente, entre otras, obligaciones recogidas en aquella propuesta directamente relacionadas con el hilo conductor de esta investigación, esto es, la responsabilidad y el derecho de resarcimiento (Artículo 13 de la precitada propuesta).

El Reglamento propuesto establece un conjunto de requerimientos para los sistemas de alto riesgo, no exigibles, por tanto, a sistemas de riesgo medio, limitado o bajo, y exigibles desde el diseño y desarrollo del sistema y, por supuesto en su comercialización, puesta en funcionamiento y uso posterior.

De este modo, el Reglamento propuesto constituye verdaderamente una regulación de la inteligencia artificial, cuanto menos de determinados sistemas y usos considerados de riesgo inaceptable o alto riesgo, no circunscrita meramente a su comercialización y uso, sino que afecta también a su diseño y concepción, estableciendo un marco de referencia para la *Ethics, security, privacy and compliance by design*, si bien, no global y limitado.

Los requisitos exigidos para los sistemas de inteligencia artificial de alto riesgo son objeto de análisis en los siguientes apartados.

1º) Sistema de gestión de riesgos

El artículo 9 del Reglamento propuesto exige, sin indicar, inicial y expresamente, en quién debe recaer dicha obligación, que se establezca, implemente, documente y mantenga un sistema de gestión de riesgos sobre los sistemas de inteligencia artificial de alto riesgo, continuo, actualizado regularmente y durante todo el ciclo de vida de dichos sistemas. De este modo, considero que inicialmente esta obligación afectaría al proveedor sin perjuicio que sea recomendable exigir expresamente la misma al usuario de este tipo de sistemas en base al propio enfoque de riesgos sobre el que se ha construido la propuesta.

El sistema de gestión de riesgos, conforme recoge expresamente el artículo 9.2 del Reglamento propuesto, requerirá distintas acciones:

a) Identificar y analizar los riesgos, tanto los conocidos como los previsibles asociados al sistema inteligente de alto riesgo.

b) Estimar y evaluar los mismos, si bien, no evaluarlos en relación con cualquier uso potencial, sino que el Reglamento propuesto establece que dicha evaluación se circunscribirá a los riesgos que puedan surgir "cuando el sistema de inteligencia artificial de alto riesgo se utilice de acuerdo con su finalidad" así como "en condiciones de uso indebido razonablemente previsible".

   De este modo, requiere que se evalúen también los riesgos respecto de uso indebidos alejados de su propósito inicial, pero siempre y cuando sean razonablemente previsibles.

   Entiendo la inclusión de esta acotación ante la variabilidad de supuestos que se pueden plantear y de difícil previsión para el proveedor o usuario, que además se haya en directa conexión con aspectos de responsabilidad, si bien, plantea un problema respecto del concepto "razonablemente previsible" que comporta una cierta inseguridad jurídica y nos lleva a los conceptos analizados anteriormente de la *Etichs by design* y el *Compliance by design*, en la medida que esta evaluación de

riesgos deberá realizarse ya en su diseño y desarrollo sobre los posibles usos indebidos que razonablemente puedan esperarse, acordes, en mi opinión, a las características y capacidades de las que esté dotado el sistema.

Por otra parte, conforme establece el artículo 9.8 del Reglamento propuesto, se deberá prestar especial atención si es probable que accedan al sistema niños o que éste pueda tener un impacto en los mismos.

c) Evaluar otros riesgos que puedan surgir sobre la base del análisis de los datos recopilados en virtud del sistema de seguimiento posterior a la comercialización preceptivo, previsto en el artículo 61 del Reglamento propuesto.

d) Y por último, en congruencia con todo lo anterior, adoptar las medidas adecuadas para la gestión del riesgo identificado y evaluado conforme a lo previsto en el artículo 9 del Reglamento propuesto, que deberán considerar para su adopción los efectos y las posibles interacciones resultantes de la aplicación combinada de los distintos requisitos exigidos para los sistemas de inteligencia artificial de alto riesgo, y deberán considerar igualmente el estado de la técnica generalmente reconocido -*the generally acknowledged state of the art*- así como lo reflejado en las normas armonizadas o especificaciones comunes pertinentes.

Como en cualquier otro sistema de gestión de riesgos, conforme igualmente lo exigen las distintas metodologías de referencia, las mejores prácticas aceptadas a nivel internacional en materia de gestión de riesgos, así como los estándares de referencia en otras materias correlacionadas, estos sistemas deben concluir que cualquier riesgo residual asociado a cada peligro así como el riesgo residual global de los sistemas de inteligencia artificial de alto riesgo se consideren aceptables, siempre que el sistema de inteligencia artificial de alto riesgo se utilice de acuerdo con su finalidad prevista o en condiciones de uso indebido razonablemente previsibles.

En relación con el riesgo residual, esto es, aquel que persiste aun después de tomar las medidas necesarias para tratar los riesgos identificados, el Reglamento propuesto impone otra obligación adicional, la de comunicar al usuario los riesgos residuales, entiendo y en

congruencia con el tenor literal del artículo precitado, durante todo el ciclo de vida y por parte del proveedor.

El apartado 4º del artículo 9, establece un conjunto de garantías a seguir para identificar las medidas de gestión de riesgos más adecuadas, en particular, la eliminación o reducción de riesgos en la medida de lo posible mediante un diseño y desarrollo adecuados -es decir, exige un diseño bajo un enfoque de riesgos y la *Security by design*-, la implementación de medidas adecuadas de mitigación y control en relación con los riesgos que no puedan eliminarse en tales supuestos, el suministro de información adecuada conforme a las obligaciones de transparencia reguladas en el artículo 13 en relación con los riesgos y, en su caso, formación de los usuarios.

El precepto indicado exige igualmente que, en la gestión de los riesgos para su eliminación o reducción, se consideren aspectos como el conocimiento técnico, la experiencia, la educación y la capacitación que espera el usuario y el entorno en el que se pretende utilizar el sistema de inteligencia artificial de alto riesgo.

El apartado 5 del artículo 9 exige pruebas previas del sistema para identificar las medidas de gestión de riesgos más adecuadas para su propósito, consistencia y cumplimiento de los requisitos requeridos, las cuales se deberán realizar en cualquier momento durante el proceso de desarrollo y, en todo caso, antes de la comercialización o la puesta en servicio, lo que constituye la aplicación del principio de *Compliance by design*. Las pruebas deberán desarrollarse bajo métricas predefinidas, indicadores clave y de probabilidad adecuados para el propósito del sistema.

Por último, el Reglamento establece que los aspectos regulados en los apartados 1 a 8 del artículo 9 relativo a los sistemas de gestión de riesgos analizados anteriormente deberán formar parte de los procedimientos de gestión de riesgos de las entidades de crédito reguladas en la Directiva 2013/36/UE[572].

[572] Directiva 2013/36/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013, relativa al acceso a la actividad de las entidades de crédito y a la supervisión prudencial de las entidades de crédito y las empresas de inversión, por la que se modifica la Directiva 2002/87/CE y se derogan las Directivas 2006/48/CE y 2006/49/CE. OJ L 176, 27.6.2013. Pp. 338–436

2º) Calidad y gobernanza de los datos

Los conjuntos de datos de entrenamiento, validación y ensayo utilizados por los sistemas de inteligencia artificial de alto riesgo deberán cumplir los criterios de calidad regulados en el artículo 10 del Reglamento propuesto y estarán sujetos a prácticas adecuadas de gestión y gobernanza de datos.

Estas obligaciones se exigen igualmente desde el diseño y concepción del modelo de sistema, sus técnicas y modelos de entrenamiento, y afectará tanto el diseño, como a la recopilación de datos, operaciones de procesamiento, preparación de datos, formulación de supuestos relevante, evaluación previa de disponibilidad, cantidad e idoneidad de los conjuntos de datos necesarios, examen de posibles sesgos o identificación de deficiencias en los datos.

Los apartados 3 y 4 del artículo 10 exigen igualmente que los conjuntos de datos -tanto de formación, validación como de ensayo- deben ser relevantes, representativos, libres de errores y completos, así como disponer de las propiedades estadísticas adecuadas, incluyendo las personas o grupos en los que pretende utilizar el sistema, y tener en cuenta, si lo requiere la finalidad prevista, las características o elementos que sean particulares del entorno geográfico, funcional o de comportamiento concreto en el que se pretende utilizar el sistema, lo que comporta en sí mismo sesgo técnico intrínseco.

Sin embargo, el precepto citado no indica qué debe considerarse por “libres de errores y completos”, por lo que sería deseable alguna matización adicional al respecto por seguridad jurídica.

En relación con este requerimiento, debo significar la habilitación legal específica y especial que contempla el artículo 10.5 del Reglamento propuesto para el tratamiento de datos personales de carácter especial regulados en el artículo 9 del RGPD, con las salvaguardias requeridas, incluyendo la seudonimización y el cifrado, cuando la anonimización no sea posible por afectar significativamente al objetivo pretendido. Sobre este aspecto sería interesante incorporar una referencia a la evaluación de impacto.

3º) Documentación técnica

El Reglamento propuesto exige también la elaboración previa a la comercialización o puesta en servicio de la documentación técnica del sistema de inteligencia artificial de alto riesgo, que deberá estar permanentemente actualizada, conforme establece su artículo 11.

La documentación técnica requerida debe redactarse además con el objetivo de demostrar que el sistema cumple los requisitos requeridos para los sistemas de inteligencia artificial de alto riesgo, proporcionando así la información necesaria a las autoridades nacionales y organismos competentes para evaluar la conformidad del sistema con dichos requerimientos.

La documentación técnica deberá contener la información contemplada en el Anexo IV del Reglamento propuesto, esto es:

a) Descripción general del sistema que incluya:

- Finalidad prevista, la persona o personas que desarrollen el sistema, la fecha y la versión del sistema;

- Cómo interactúa el sistema o cómo se puede utilizar para interactuar con hardware o software que no forme parte del sistema en sí, cuando corresponda;

- Versiones de *software* o *firmware* relevantes y cualquier requisito relacionado con la actualización de la versión;

- Descripción de todas las formas en las que el sistema se comercializa o se pone en servicio;

- Descripción del hardware en el que está previsto que se ejecute el sistema;

- Cuando el sistema sea un componente de productos, fotografías o ilustraciones que muestren características externas, marcado y diseño interno de esos productos;

- Instrucciones de uso para el usuario y, en su caso, instrucciones de instalación.

b) Descripción detallada de los elementos del sistema y del proceso para su desarrollo, incluyendo:

- Métodos y pasos realizados para el desarrollo del sistema, incluido, en su caso, el recurso a sistemas o herramientas previamente entrenados proporcionados por terceros y cómo estos han sido utilizados, integrados o modificados por el proveedor;

- Especificaciones de diseño del sistema que incluyan: La lógica general del sistema y de los algoritmos, las opciones de diseño clave, incluida la justificación y las suposiciones realizadas, también con respecto a las personas o grupos de personas en las que se pretende utilizar el sistema, las principales opciones de clasificación, para qué está diseñado el sistema para optimizar y la relevancia de los diferentes parámetros, las decisiones sobre las posibles compensaciones tomadas en relación con las soluciones técnicas adoptadas para cumplir con los requisitos requeridos;

- Descripción de la arquitectura del sistema que explique cómo los componentes de *software* se complementan o se alimentan entre sí y se integran en el procesamiento general; los recursos computacionales utilizados para desarrollar, entrenar, probar y validar el sistema;

- Cuando proceda, requisitos de datos en términos de hojas de datos que describan las metodologías y técnicas de formación y los conjuntos de datos de formación utilizados, incluida información sobre la procedencia de esos conjuntos de datos, su alcance y características principales; cómo se obtuvieron y seleccionaron los datos; procedimientos de etiquetado -por ejemplo, para aprendizaje supervisado-, metodologías de limpieza de datos -por ejemplo, detección de valores atípicos-;

- Evaluación de las medidas de supervisión humana necesarias de conformidad con los previsto en el artículo 14 del Reglamento propuesto que analizaré con

posterioridad, incluida una evaluación de las medidas técnicas necesarias para facilitar la interpretación de los resultados de los sistemas por parte de los usuarios, de conformidad con lo previsto en el artículo 13.3.d) del Reglamento propuesto;

- En su caso, descripción detallada de los cambios predeterminados en el sistema y su rendimiento, junto con toda la información relevante relacionada con las soluciones técnicas adoptadas para garantizar el cumplimiento continuo del sistema con los requerimientos exigido al mismo;

- Procedimientos de validación y prueba utilizados, incluida información sobre los datos de validación y prueba utilizados y sus principales características, métricas utilizadas para medir la precisión, la solidez, la ciberseguridad y el cumplimiento de otros requisitos relevantes establecidos en los requerimientos exigidos para este tipo de sistemas así como los impactos potencialmente discriminatorios, registros de prueba y todos los informes de prueba fechados y firmados por las personas responsables, incluso con respecto a los cambios predeterminados a que se refiere el apartado anterior.

c) Información detallada sobre el seguimiento, funcionamiento y control del sistema de inteligencia artificial, en particular en relación con: Sus capacidades y limitaciones de rendimiento, incluidos los grados de precisión para personas o grupos de personas específicos en los que se pretende utilizar el sistema y el nivel general esperado de precisión en relación con su propósito previsto; los resultados no deseados previsibles y las fuentes de riesgos para la salud y la seguridad, los derechos fundamentales y la discriminación en vista del propósito previsto del sistema de inteligencia artificial; las medidas de supervisión humana necesarias de conformidad con lo previsto en el artículo 14 del Reglamento propuesto, incluidas las medidas técnicas establecidas para facilitar la interpretación de los resultados de los sistemas por parte de los usuarios; especificaciones sobre los datos de entrada, según corresponda.

d) Descripción detallada del sistema de gestión de riesgos de conformidad con los previsto en el artículo 9 del Reglamento propuesto, anteriormente analizado.

e) Descripción de cualquier cambio realizado en el sistema a lo largo de su ciclo de vida.

f) Lista de las normas armonizadas aplicadas total o parcialmente cuyas referencias se han publicado en el Diario Oficial de la Unión Europea o, cuando no se hayan aplicado dichas normas armonizadas, una descripción detallada de las soluciones adoptadas para cumplir los requisitos exigidos en la Propuesta de Reglamento para sistemas de inteligencia artificial de alto riesgo, incluida una lista de otras normas y especificaciones técnicas pertinentes aplicadas.

g) Copia de la declaración de conformidad de la UE;

h) Descripción detallada del sistema de evaluación del rendimiento del sistema de inteligencia artificial en la fase posterior a la comercialización de conformidad con lo dispuesto en el artículo 61 del Reglamento propuesto, incluido el plan de seguimiento posterior a la comercialización al que hace referencia el apartado tercero de dicho precepto.

Este requerimiento integra la explicabilidad del sistema y demás información determinante base para la posterior exigencia de responsabilidad en el caso de daños. Esta información es determinante para el cumplimiento de otras obligaciones legales por los usuarios, en particular, las de transparencia en materia de privacidad, por lo que sería deseable que sea lo suficientemente completa para permitirle cumplir dichas obligaciones e incorporar, en caso de que el proveedor hubiere llevado a cabo la preceptiva evaluación de impacto, la entrega de la misma.

Conforme se haya regulado este requerimiento en la propuesta, los destinatarios principales de esta información serán las autoridades nacionales y organismos competentes, sin perjuicio de que algunos de los elementos que la integran tengan como destinatario también a los usuarios como, por ejemplo, las instrucciones de uso y, en su caso, de instalación.

En resumen, la documentación técnica requerida es cualitativa y cuantitativamente significativa, rigurosa, abstracta en algunas cuestiones y acorde, entiendo, al riesgo

calificado, pero que considero, podría ser considerado un obstáculo inicial para *startups* que constituyen un motor de innovación actual en la medida que no dispongan de las capacidades y recursos iniciales suficientes para alcanzar todos estos requerimientos en las etapas iniciales de diseño y concepción de sistemas de esta naturaleza.

4º) Registro de eventos y trazabilidad

El Reglamento propuesto exige que los sistemas de inteligencia artificial de alto riesgo integren en su diseño y desarrollo sistemas de registro automático de eventos (logs) mientras estén en funcionamiento, de modo que todo lo que ocurra durante el mismo quede registrado, conforme establece su artículo 12.1.

Esto permite y facilita la *accountability* y facilita la imputación y depuración de responsabilidades, eje conductor de esta investigación y que constituyen principios éticos, así como obligaciones jurídicas de responder en función del hecho y contexto.

Este registro debe garantizar, de un lado, un nivel de trazabilidad adecuado para la finalidad prevista del sistema a lo largo del ciclo de vida del mismo, conforme exige el apartado 2º del artículo precitado, y de otro, permitir supervisar el funcionamiento del sistema en caso de situaciones de riesgo (en el sentido del artículo 3.19 del Reglamento (UE) 2019/1020) o de modificación sustancial, así como facilitar el seguimiento posterior a la comercialización a la que refiere el artículo 61 del Reglamento propuesto.

Por último, el Reglamento propuesto define la información que debe constar y ser extraíble de dicho registro en el caso de los sistemas de inteligencia artificial destinados a ser utilizados para la identificación biométrica remota "en tiempo real" y "posterior" de personas físicas.

En particular, en su artículo 12.4 exige que al menos deberán proporcionar la siguiente información: a) Registro del período de cada uso del sistema -fecha y hora de inicio y fecha y hora de finalización de cada uso-; b) La base de datos de referencia con la que el sistema ha verificado los datos de entrada; c) Los datos de entrada para los que la búsqueda ha dado lugar a una coincidencia; d) La identificación de las personas físicas

implicadas en la verificación de los resultados -referenciadas en el artículo 14.5 del Reglamento propuesto-.

No obstante, no se regula el período de conservación de los registros, lo que sería deseable por seguridad jurídica y a efectos de exigencia de responsabilidad.

5º) Transparencia y suministro de información a los usuarios

De nuevo, el Reglamento propuesto incorpora otra obligación a considerar de inicio en el diseño y desarrollo de sistemas de inteligencia artificial de alto riesgo, de manera que deberán concebirse garantizando un funcionamiento posterior lo suficientemente transparente, de modo que los usuarios puedan interpretar los resultados del sistema y utilizarlos adecuadamente.

La obligación de transparencia e información de halla regulada en el artículo 13 del Reglamento propuesto, que exige que tanto el tipo como el grado de transparencia deberán ser adecuados para lograr el cumplimiento de las múltiples obligaciones adicionales del usuario y del proveedor reguladas en los artículos 16 a 29 del Reglamento propuesto, que serán analizadas más adelante.

Adicionalmente a la obligación de documentar técnicamente determinados aspectos comentada anteriormente, que ya contempla expresamente la inclusión en la misma de las instrucciones de uso para el usuario (y, en su caso, instrucciones de instalación), los sistemas de inteligencia artificial deberá ir acompañados de instrucciones de uso en formato digital apropiado o cualquier otro modo, que incluya información "concisa, completa, correcta y clara", relevante, accesible y comprensible para los usuarios. La información requerida se detalla en el apartado 3 del artículo 13 e incluye:

a) La identidad y los datos de contacto del proveedor y, en su caso, de su representante autorizado;

b) Las características, capacidades y limitaciones de rendimiento del sistema de inteligencia artificial de alto riesgo, que deben incluir:

    - Su finalidad prevista;

- El nivel de precisión, solidez y ciberseguridad esperado a que se refiere el artículo 15 del Reglamento propuesto -que será analizado más adelante-, con respecto al cual se ha probado y validado el sistema, así como cualquier circunstancia conocida y previsible que pueda tener un impacto en ese nivel de precisión esperado, robustez y ciberseguridad;

- Cualquier circunstancia conocida o previsible, relacionada con el uso del sistema de acuerdo con su finalidad prevista o bajo condiciones de uso indebido razonablemente previsibles, que pueda conducir a riesgos para la salud y seguridad o los derechos fundamentales;

- Su rendimiento en lo que respecta a las personas o grupos de personas con las que se pretende utilizar el sistema;

- Cuando proceda, las especificaciones para los datos de entrada, o cualquier otra información pertinente en cuanto a los conjuntos de datos de entrenamiento, validación y prueba utilizados, teniendo en cuenta la finalidad prevista del sistema.

c) Los cambios en el sistema y su rendimiento que hayan sido predeterminados por el proveedor en el momento de la evaluación de conformidad inicial, si los hubiere;

d) Las medidas de supervisión humana reguladas en el artículo 14 del Reglamento propuesto, analizadas a continuación, incluidas las medidas técnicas establecidas para facilitar la interpretación de los resultados de los sistemas de inteligencia artificial por parte de los usuarios;

e) La vida útil prevista del sistema y las medidas de mantenimiento y cuidado necesarias para garantizar el correcto funcionamiento de ese sistema, incluidas las actualizaciones de software.

Este requerimiento es esencial para la posterior exigencia de responsabilidad por los daños que pueda causar el sistema.

En mi opinión, sería deseable que dicha trazabilidad incorporara información sobre las reglas e instrucciones en las que se basen los sistemas para la toma de decisiones.

6º) Supervisión humana

La supervisión humana como principio y norma ética esencial se integra como requerimiento específico en el diseño y desarrollo de sistemas de inteligencia artificial de alto riesgo.

Conforme regula el artículo 14 del Reglamento propuesto, el diseño deberá permitir que los sistemas pueden ser supervisados de manera efectiva por personas físicas mientras el sistema esté en uso, dotándole de un interfaz adecuado con dicha finalidad.

El objetivo de la misma es prevenir y minimizar los riesgos para la salud, la seguridad o los derechos fundamentales que puedan surgir en el uso de este tipo de sistemas, ya sea un uso conforme a su finalidad inicial o en condiciones de uso indebido razonablemente previsible, especialmente cuando estos riesgos persisten a pesar de la aplicación de los requisitos de estos sistemas objeto de análisis en este apartado y regulados en los artículos 8 a 15 del Reglamento propuesto.

La supervisión humana, conforme establece el mismo, deberá garantizarse mediante una o todas las medidas siguientes:

a) Identificada e incorporada, cuando sea técnicamente factible, al sistema de inteligencia artificial de alto riesgo por el proveedor antes de su comercialización o puesta en servicio;

b) Identificadas por el proveedor antes de comercializar o poner en servicio el sistema y que sean adecuadas para ser aplicadas por el usuario.

Adicionalmente, en relación con los sistemas de inteligencia artificial destinados a ser utilizados para la identificación biométrica remota "en tiempo real" y "posterior" de personas físicas, el apartado 5 del artículo 14 del Reglamento propuesto establece que estas medidas deberán garantizar que, además, el usuario no tome ninguna acción o

decisión sobre la base de la identificación resultante del sistema, a menos que ésta haya sido verificada y confirmada por al menos dos personas físicas.

En este sentido, dicha verificación llevada a cabo por dos personas sin cualificación, permitiría cumplir el requerimiento, pero no asegurar absolutamente nada. En consecuencia, sería deseable requerir los conocimientos y cualificación oportuna de las mismas.

Las medidas de supervisión humana previstas en el apartado 3 del artículo 14 del Reglamento propuesto, deberán permitir a las personas a las que se asigne la supervisión humana lo siguiente, según el contexto:

a) Comprender plenamente las capacidades y limitaciones del sistema y poder supervisar debidamente su funcionamiento, de modo que los indicios de anomalías, disfunciones y rendimientos inesperados puedan detectarse y abordarse lo antes posible;

b) Ser consciente de la posible tendencia a confiar automáticamente o en exceso en los resultados producidos por un sistema de inteligencia artificial de alto riesgo, esto es, el denominado "sesgo de automatización", en particular en el caso de los sistemas utilizados para proporcionar información o recomendaciones para las decisiones que deben tomar las personas físicas;

c) Ser capaz de interpretar correctamente los resultados del sistema, teniendo en cuenta, en particular, las características del sistema y las herramientas y métodos de interpretación disponibles;

d) Poder decidir, en cualquier situación concreta, no utilizar el sistema o, de otro modo, ignorar, anular o invertir los resultados del sistema, es decir, la anulación o reversión de sus resultados;

e) Poder intervenir en el funcionamiento del sistema o interrumpir el sistema mediante un botón de "parada" o un procedimiento similar.

En este sentido, sería recomendable una mayor concreción del tipo de intervención humana requerida en la toma de decisiones conforme analicé en los capítulos precedentes, esto es, en un enfoque de supervisión humana o también denominada "Human on the loop" (HOTL), de participación humana o "Human in the loop" (HITL) y/o de control humano o "Human in command" (HIC).

7º) Precisión, robustez y ciberseguridad

Los sistemas de inteligencia artificial se diseñarán y desarrollarán de forma que alcancen, conforme a su finalidad prevista, un nivel adecuado de precisión, robustez y ciberseguridad, y que funcionen de manera consistente conforme a esos aspectos a lo largo de su ciclo de vida.

Los niveles de precisión y las métricas de precisión pertinentes deberán incluirse en las instrucciones de uso precitadas que acompañen a los mismos.

Por lo que se refiere a resiliencia, los sistemas deberán ser resistentes frente a errores, fallos o incoherencias que puedan producirse en el sistema o en el entorno en el que opera el mismo, en particular debido a su interacción con personas físicas u otros sistemas, cuestión que me parece muy relevante, dada la frecuente interacción de los sistemas inteligentes con otras tecnologías y sistemas, que pueden potenciar sus ventajas pero también sus inconvenientes, así como los riesgos asociados a su uso, especialmente de ciberseguridad.

El Reglamento propuesto incluye algunas propuestas para alcanzar la solidez requerida, en particular, la redundancia con planes de respaldo o a prueba de fallos, es decir, lo que en seguridad de sistemas de información denominamos Planes de Recuperación de Desastres -*Disaster Recovery Plan* o DRP por sus siglas en inglés- y planes de continuidad de negocio -*Business Continuity Plan* o BCP por sus siglas en inglés.

El artículo 15.3 del Reglamento propuesto adiciona un requerimiento adicional para los sistemas de inteligencia artificial de alto riesgo que sigan aprendiendo después de su comercialización o puesta en servicio, es decir, de aprendizaje automático y continuo, en particular, la inclusión de medidas de mitigación adecuadas en su diseño en relación con

los resultados sesgados generados en su retroalimentación que se utilizan como entrada para operaciones futuras, conforme ya comenté en su capítulo II al abordar estos riesgos.

Del mismo modo, la resistencia del sistema es igualmente exigida frente a los intentos de terceros no autorizados de alterar su uso o rendimiento aprovechando las vulnerabilidades del sistema.

Y finalmente, el último apartado del precitado artículo establece que las soluciones técnicas destinadas a garantizar la ciberseguridad de los sistemas de inteligencia artificial de alto riesgo serán adecuadas a las circunstancias pertinentes y a los riesgos, es decir, adecuadas al contexto y riesgos a los que puedan ser sometidos, y adiciona que las soluciones técnicas para hacer frente a las vulnerabilidades específicas de la inteligencia artificial incluirán, en su caso, medidas para prevenir y controlar los ataques orientados a la manipulación del conjunto de datos de entrenamiento, es decir, el denominado, "envenenamiento de datos", las entradas diseñadas para hacer que el modelo cometa un error, denominados "ejemplos adversos", o los defectos del propio modelo.

6.8.2. Obligaciones de proveedores y usuarios de sistemas de IA de alto riesgo

Junto a los requisitos precitados, el Reglamento propuesto regula en los artículos 16 a 29 las obligaciones generales y específicas de los proveedores y usuarios de sistemas de alto riesgo y otras partes implicadas, lo cuales serán expuestos a continuación.

A. Obligaciones para proveedores de sistemas de IA de alto riesgo

El artículo 16 del Reglamento propuesto regula las obligaciones generales de los proveedores de sistemas de inteligencia artificial de alto riesgo que se exponen a continuación.

1º) Garantizar que sus sistemas cumplen los requisitos definidos en el anterior apartado.

El Reglamento propuesto coloca al proveedor, sea o no el fabricante del sistema, como garante de que los sistemas que provea cumplan los requisitos definidos en el Capítulo II del Título III (Artículos 8 a 15), analizados en el anterior apartado, con las consecuencias previstas en el mismo que se analizarán posteriormente, de conformidad con lo previsto en el artículo 16.I.a) del Reglamento propuesto.

2º) Disponer de un sistema de gestión de la calidad

Los proveedores deberán disponer de un sistema de gestión de la calidad que garantice el cumplimiento del Reglamento propuesto, conforme a su artículo 17, que deberá documentarse de manera sistemática y ordenada en forma de políticas, procedimientos e instrucciones escritas, y que deberá incluirá al menos los siguientes aspectos:

a) Una estrategia de cumplimiento normativo -*compliance*-, incluido el cumplimiento de los procedimientos de evaluación de la conformidad y los procedimientos para la gestión de modificaciones al sistema inteligente;

b) Técnicas, procedimientos y acciones sistemáticas que se utilizarán para el diseño, control del diseño y verificación del diseño del sistema inteligente;

c) Las técnicas, procedimientos y acciones sistemáticas que se utilizarán para el desarrollo, control de calidad y garantía de calidad del sistema inteligente;

d) Los procedimientos de examen, prueba y validación que deben llevarse a cabo antes, durante y después del desarrollo del sistema inteligente y la frecuencia con la que deben llevarse a cabo;

e) Las especificaciones técnicas, incluidas normas, que se aplicarán y, cuando no se apliquen íntegramente las normas armonizadas pertinentes, los medios que se utilizarán para garantizar que el sistema inteligente cumple los requisitos analizados en el anterior apartado;

f) Los sistemas y procedimientos para la gestión de datos, incluida la recopilación de datos, el análisis de datos, el etiquetado de datos, el almacenamiento de datos, la filtración de datos, la extracción de datos, la agregación de datos, la retención

de datos y cualquier otra operación relacionada con los datos que se realice antes y a efectos de la comercialización o puesta en servicio de los sistemas de IA de alto riesgo;

g) El sistema de gestión de riesgos requerido y mencionado anteriormente, que se haya regulado en el artículo 9;

h) La creación, implementación y mantenimiento de un sistema de seguimiento posterior a la comercialización, de conformidad con lo dispuesto en el artículo 61 del Reglamento propuesto;

i) Los procedimientos relacionados con la notificación de incidentes graves y de mal funcionamiento de conformidad con el artículo 62 del Reglamento propuesto;

j) La gestión de la comunicación con las autoridades competentes, incluidas las sectoriales, que proporcionan o apoyan el acceso a los datos, los organismos notificados, otros operadores, clientes u otras partes interesadas;

k) Los sistemas y procedimientos para el mantenimiento de registros de toda la documentación e información relevante;

l) La gestión de recursos, incluidas medidas relacionadas con la seguridad del suministro;

m) Un marco de rendición de cuentas que establezca las responsabilidades de la dirección y del resto del personal con respecto a todos los aspectos relativos al sistema de gestión enumerados precedentemente.

La dimensión de este requerimiento, especialmente para *startups*, se halla matizada en el apartado 2 del artículo 17 del Reglamento propuesto, al indicar que la implementación de todos estos aspectos será proporcionada al tamaño de la organización del proveedor.

Esta cuestión genera cierta inseguridad jurídica en relación con el grado y proporcionalidad para la implementación de las exigencias precitadas cuando se trate de *startups* en la medida que no se define.

En mi opinión, es necesario mantener el equilibrio y la proporcionalidad para asegurar la retención de talento y la innovación, lo que significa facilitar las cosas a las empresas de reciente creación, pero con cautela para evitar actos desleales y prácticas no deseables por parte de grandes compañías que generen nuevas estructuras, no solo para gestionar nuevos proyectos de inteligencia artificial con recursos específicos, sino para eludir cargas y obligaciones que les resultarían de aplicación en caso de implementar estos proyectos en su seno.

En este sentido, en la práctica profesional estamos empezando a reflexionar con desarrolladores y proveedores de sistemas de inteligencia artificial la necesidad de valorar la futura implementación de sistemas de gestión de cumplimiento o *compliance* global en estas materias, que integren todos los aspectos que se requieran en los marcos reguladores generales como especiales que se exijan de manera integral, así como las estrategias, planes y programas, como las políticas, estándares, procedimientos, requerimientos, medidas y controles corporativos, exigidos para su evaluación y monitorización continua.

En el caso de que los proveedores sean entidades de crédito reguladas por la Directiva 2013/36/UE, la obligación de implantar un sistema de gestión de la calidad se considerará cumplida mediante el cumplimiento de las normas sobre disposiciones, procesos y mecanismos de gobernanza interna con arreglo al artículo 74 de dicha Directiva. En este contexto, se deberán tener en cuenta las normas armonizadas referenciadas en el artículo 40 del presente Reglamento propuesto.

En relación con estas obligaciones examinadas, considero que sería deseable igualmente definir los criterios y frecuencia con la que deben llevarse a cabo los procedimientos de examen, prueba y validación antes, durante y después del desarrollo del sistema inteligente, no dejándolo en manos del proveedor.

Asimismo, en relación con el marco de rendición de cuentas exigido que establezca las responsabilidades de la dirección y del resto del personal, sería interesante incorporar ya los criterios de distribución de la responsabilidad.

3º) Elaborar la documentación técnica del sistema

El Reglamento propuesto traslada esta obligación al proveedor en su artículo 18, conforme a los dispuesto en su artículo 11 y Anexo IV, en los términos que he analizado en el apartado anterior de requerimientos.

Tal y como comenté al abordar dicho requerimiento, en el caso de entidades de crédito reguladas por la Directiva 2013/36/UE mantendrán la documentación técnica como parte de la documentación relativa a la gobernanza interna, las disposiciones, los procesos y los mecanismos, conforme a lo previsto en el artículo 74 de la Directiva precitada.

4º) Conservar los registros generados automáticamente por sus sistemas, cuando estén bajo su control.

El artículo 16 del Reglamento propuesto regula esta obligación del proveedor de conservación de los registros generados por los sistemas cuando estén bajo el control del mismo.

5º) Garantizar que el sistema se somete al correspondiente procedimiento de evaluación de la conformidad, antes de su comercialización o puesta en servicio.

El Reglamento propuesto exige en su artículo 19.1 que los proveedores se aseguren de que los sistemas proveídos por los mismos se sometan al procedimiento de evaluación de la conformidad pertinente antes de su comercialización o puesta en servicio, siguiendo lo establecido al respecto en el artículo 43 del mismo.

Una vez realizada la evaluación de la conformidad y acreditada su conformidad y cumplimiento los requisitos requeridos para estos sistemas en los artículos 8 a 15 del Reglamento propuesto (Capítulo 2 del Título III), el precitado artículo 19.1 exige dos obligaciones asociadas más a los proveedores:

- Redactar una declaración UE de conformidad según lo exigido en el artículo 48 del Reglamento propuesto;

- Colocar el marcado CE de conformidad con lo dispuesto en su artículo 49.

De nuevo, el Reglamento propuesto contempla una previsión específica respecto de los sistemas de inteligencia artificial destinados a ser utilizados para evaluar la solvencia de las personas físicas o establecer su puntaje crediticio (con la excepción de los sistemas de inteligencia artificial puestos en servicio por proveedores de pequeña escala para su propio uso), cuando sean comercializados o puestos en servicio por proveedores que sean entidades de crédito reguladas por la Directiva 2013/36/UE. En estos supuestos la evaluación de la conformidad se llevará a cabo como parte del procedimiento contemplado en los artículos 97 a 101 de dicha Directiva.

6º) Cumplir con las obligaciones de registro

Los artículos 16, 51 y 60 del Reglamento propuesto exigen que los proveedores (o su representante autorizado), registren sus sistemas de inteligencia artificial de alto riesgo regulados en el Anexo III del mismo en la base de datos de la UE, es decir no todos, sino sólo aquellos listados en el precitado anexo, conforme he expuesto anteriormente.

7º) Mantenimiento de registros automáticos

Los proveedores de estos sistemas deberán mantener los registros generados automáticamente por sus sistemas a los que hice alusión anteriormente cuando se hallen bajo su control, tal y como exige el artículo 20 del Reglamento propuesto, ya sea en virtud de un acuerdo contractual con el usuario o por Ley.

Los períodos de conservación serán los adecuados conforme a la finalidad del sistema y las obligaciones legales exigibles.

De nuevo, en el caso de que los proveedores sean entidades de crédito reguladas por la Directiva 2013/36/UE, éstos mantendrán los registros generados automáticamente por sus sistemas de inteligencia artificial de alto riesgo como parte de la documentación exigida, conforme a lo dispuesto en el artículo 74 de aquella.

8º) Adoptar las medidas correctivas necesarias si el sistema no es conforme con los requisitos analizados en el apartado anterior

Si los proveedores del sistema consideran o tienen motivos para considerar que un sistema que hayan introducido en el mercado o puesto en servicio no es conforme con el Reglamento propuesto, de conformidad con lo previsto en el artículo 22 del mismo, tienen la obligación de adoptar inmediatamente las acciones correctivas necesarias para hacerlo conforme, retirarlo o recuperarlo, según proceda.

La cuestión es que es previsible que cuando esto suceda, el proveedor no esté en condiciones de hacerlo, especialmente si afecta a su diseño y desarrollo, por lo que en estos casos deberá optar por retirarlo o recuperarlo según haya sido ya comercializado o puesto en servicio.

Cuando se produzcan estos supuestos, el proveedor tiene la obligación adicional de informar a los distribuidores del sistema en cuestión y, en su caso, al representante autorizado y a los importadores.

9º) Deber de información en caso de riesgo de no conformidad

El artículo 16 del Reglamento propuesto exige al proveedor informar a las autoridades nacionales competentes de los Estados miembros en los que hayan facilitado o puesto en servicio el sistema y, en su caso, al organismo notificado, de la no conformidad y de las medidas correctoras adoptadas.

Si el sistema presenta un riesgo para la seguridad, salud o derechos fundamentales de la persona en el sentido del artículo 65.1 del Reglamento propuesto y es conocido por el proveedor del sistema, éste deberá informar inmediatamente de la no conformidad del mismo y de las medidas correctoras adoptadas a las autoridades nacionales competentes de los Estados miembros en los que haya puesto en circulación el sistema y, en su caso, al organismo notificado que haya expedido un certificado para el sistema de implicado.

10º) Colocar el marcado CE en sus sistemas de IA de alto riesgo para indicar la conformidad con el presente Reglamento.

El Reglamento propuesto regula en su artículo 16 la obligación de los proveedores de este tipo de sistemas de colocar el marcado CE en sus sistemas para indicar la conformidad con el mismo, de conformidad con lo previsto en su artículo 49.

El marcado deberá ubicarse de forma visible, legible e indeleble, salvo cuando no sea posible o no esté justificado debido a la naturaleza del sistema, en cuyo caso se colocará en el embalaje o en la documentación adjunta, según corresponda. El mismo se hallará sujeto a los principios generales establecidos en el artículo 30 del Reglamento (CE) nº 765/2008[573].

El marco irá seguido del número de identificación del organismo notificado responsable de los procedimientos de evaluación de conformidad, debiendo igualmente indicarse en cualquier material promocional que mencione que el sistema cumple los requisitos para el marcado CE.

11º) Demostrar la conformidad del sistema con los requisitos exigidos.

El artículo 16 del Reglamento propuesto establece una obligación general adicional y de carácter proactivo al proveedor, esto es, la de no sólo adoptar medidas para garantizar la conformidad sino de demostrar la misma en relación con los requisitos establecidos en los artículos 8 a 15 del Reglamento propuesto (Capítulo 2 del Título III).

12º) Deber de colaboración con las autoridades competentes

El deber de colaboración se regula en el artículo 23 del Reglamento propuesto en relación con el 16, contemplando dos obligaciones principales:

a) Los proveedores de estos sistemas deberán proporcionar a las autoridades nacionales competentes que lo soliciten, toda la información y documentación necesarias para demostrar la conformidad del sistema con los requisitos

[573] Reglamento (CE) n o 765/2008 del Parlamento Europeo y del Consejo, de 9 de julio de 2008, por el que se establecen los requisitos de acreditación y vigilancia del mercado relativos a la comercialización de los productos y por el que se deroga el Reglamento (CEE) n o 339/93. OJ L 218, 13.8.2008. P. 30-47

establecidos en 8 a 15 del Reglamento propuesto (Capítulo 2 del Título III), y ello en una lengua oficial de la UE determinada por el Estado miembro de que se trate.

b) Los proveedores de estos sistemas deberán igualmente proporcionar acceso a los registros generados automáticamente por los mismos, previa solicitud y "motivada" de la autoridad nacional competente, siempre que dichos registros estén bajo su control, en virtud de acuerdo contractual o por Ley.

13º) Conservación de documentos

El proveedor deberá mantener a disposición de las autoridades nacionales competentes la documentación indicada a continuación, conformidad con lo previsto en el artículo 50 del Reglamento propuesto, y ello durante un plazo de diez años desde que el sistema sea comercializado o puesto en servicio.

La documentación a conservar durante dicho período es la siguiente:

a) La documentación técnica (Artículo 11);

b) La documentación relativa al sistema de gestión de la calidad (Artículo 17);

c) La documentación relativa a los cambios aprobados por los organismos notificados, cuando proceda;

d) Las decisiones y otros documentos emitidos por los organismos notificados, cuando proceda;

e) La declaración UE de conformidad (Artículo 48).

14º) Registro

El proveedor o, en su caso, el representante autorizado, deberá registrar el sistema en la *Base de datos de la UE para sistemas de IA de alto riesgo* regulada en artículo 60 del Reglamento propuesto antes de comercializar o poner en servicio el mismo, conforme a lo dispuesto en su artículo 51.

15º) Monitorización posterior a la comercialización

El artículo 61 del Reglamento propuesto adiciona la obligación de los proveedores de sistemas de inteligencia artificial de alto riesgo de implementar y documentar un sistema de seguimiento posterior a la comercialización, adecuado y proporcionado a las tecnologías de inteligencia artificial y riesgos de los sistemas de alto riesgo implicados.

Estos sistemas deberán recopilar, documentar y analizar de forma activa y sistemática los datos relevantes proporcionados por los usuarios o recopilados a través de otras fuentes sobre el rendimiento de los sistemas a lo largo de toda su vida útil, y deberá permitir al proveedor evaluar el cumplimiento continuo de los sistemas inteligentes de los requisitos analizados en el epígrafe anterior y regulados en los artículos 8 a 15 del Reglamento propuesto (Capítulo 2 del Título III).

El sistema de monitorización deberá estar basado en el plan de seguimiento posterior a la comercialización que forma parte de la documentación técnica a elaborar por el proveedor conforme he expuesto y que se halla regulada en el anexo IV.

El texto propuesto prevé la creación un futuro modelo de plan de seguimiento posterior a la comercialización y el listado de elementos que deben incluirse en el plan por parte de la Comisión mediante actos de ejecución.

Cuando se trate de sistemas de inteligencia artificial de alto riesgo cubierto por los actos jurídicos recogidos en el anexo II -continente del listado la legislación de armonización de la Unión basada en el *Nuevo Marco Legislativo*-, en la medida que ya se haya establecido un sistema y plan de seguimiento posterior a la comercialización en virtud de dicho marco legislativo, los aspectos precitados deberán ser integrados en dicho sistemas y plan conforme proceda.

La implementación y documentación de los sistemas de seguimiento será igualmente también se aplicará a los sistemas de inteligencia artificial destinados a ser utilizados para evaluar la solvencia de las personas físicas o establecer su puntaje crediticio comercializados o puestos en servicio por entidades de crédito reguladas por la Directiva 2013/36/UE, conforme a lo dispuesto en el último inciso del artículo precitado.

16º) Intercambio de información sobre incidentes y mal funcionamiento

Los proveedores de estos sistemas deberán notificar cualquier incidente grave o cualquier mal funcionamiento de dichos sistemas que constituya un incumplimiento derivadas del Derecho de la UE destinadas a proteger los derechos fundamentales, y sólo en estos supuestos, y deberán comunicarlo a las autoridades de vigilancia del mercado de los Estados miembros en los que se haya producido dicho incidente o incumplimiento.

Esta obligación adicional se halla contemplada en el artículo 62 del Reglamento propuesto y constituye una obligación adicional a las previstas en los marcos reguladores de la protección de datos (RGPD y normativas locales de los Estados miembros) así como de ciberseguridad.

La notificación deberá realizarse inmediatamente después de que el proveedor haya tenido conocimiento de la incidencia y haya establecido, bien un vínculo causal entre el sistema y el incidente o mal funcionamiento, o una probabilidad razonable de que exista dicho vínculo y, en todo caso en el plazo máximo de quince (15) días después de que el proveedor haya tenido conocimiento del incidente grave o del mal funcionamiento, lo que comporta la obligaciones inherente de que el proveedor efectué el correspondiente análisis para establecer el posible nexo causal de manera previa en el plazo precitado.

El precepto citado indica el procedimiento a seguir con posterioridad a la misma.

Al igual que respecto de otras obligaciones previamente expuestas, en el caso de sistemas de inteligencia artificial destinados a ser utilizados para evaluar la solvencia de las personas físicas o establecer su puntaje crediticio mediante sistemas comercializados o puestos en servicio por entidades de crédito reguladas por la Directiva 2013/36/UE, así como en el caso de los sistemas de inteligencia artificial de alto riesgo que sean componentes de seguridad de dispositivos o que sean ellos mismos dispositivos cubiertos por el Reglamento (UE) 2017/745 y el Reglamento (UE) 2017/746, la notificación de incidentes graves o de mal funcionamiento se limitará a los que constituyan un incumplimiento de las obligaciones derivadas del Derecho de la UE destinadas a proteger los derechos fundamentales.

B. Obligaciones de fabricante de producto

El marco obligaciones que afecta a los proveedores de sistemas de inteligencia artificial de alto riesgo ha sido analizado anteriormente.

La cuestión es cuando estos sistemas estén relacionados y se comercialicen o pongan en servicio bajo el nombre del fabricante junto con los productos a los que hace referencia la Sección A del Anexo II del Reglamento propuesto, anteriormente analizada, esto es, máquinas, juguetes, embarcaciones de recreo, etc., el fabricante asumirá la responsabilidad de la conformidad del sistema con el Reglamento, quedando sujeto a las mismas obligaciones impuestas por el presente Reglamento al proveedor, de conformidad con lo previsto en su artículo 24. Es obvia la repercusión que ello comporta a efectos de la depuración de las responsabilidades que puedan derivarse por daños asociados a dicho incumplimiento.

Cuando un sistema de inteligencia artificial de alto riesgo que sea un componente de seguridad de un producto regulado por la legislación sectorial que integra el precitado *Nuevo Marco Legislativo -NFL-* no se comercialice o ponga en servicio independientemente del producto, el fabricante del producto final según se define en dicha legislación, debe cumplir con las obligaciones del proveedor establecidas en el Reglamento propuesto.

C. Obligaciones de los representantes autorizados

Los proveedores establecidos fuera de la UE, conforme establece el artículo 25 del Reglamento propuesto, deberán designar a un representante autorizado en la misma mediante mandato escrito y antes de comercializar sus sistemas -en este caso no se hace referencia a puesta a disposición en el precepto que lo contempla-, cuando no sea posible identificar un importador dado que, en tal supuesto, se aplicarán las obligaciones al mismo previstas en el artículo 26 que se exponen a continuación. Sus facultades principales se hallan contemplas en el precitado artículo 25.

D. Obligaciones de los importadores

Los importadores de sistemas de inteligencia artificial de alto riesgo tienen un conjunto de obligaciones previas a la comercialización de los mismos en la UE, conforme regula el artículo 26 del Reglamento propuesto, en particular, las de asegurarse que:

a) El proveedor del sistema ha llevado a cabo el procedimiento de evaluación de la conformidad adecuado;

b) El proveedor ha elaborado la documentación técnica exigida conforme al Anexo IV del Reglamento, anteriormente analizada;

c) El sistema lleva la marca de conformidad CE requerida y va acompañado de la documentación requerida y las instrucciones de uso;

d) Las condiciones de almacenamiento o transporte no comprometan el cumplimiento de los requisitos establecidos en los precitados artículos 8 a 15 del Reglamento propuesto (Capítulo 2 del Título III), mientras el sistema esté bajo su responsabilidad.

Si el importador considera o tiene motivos para considerar que el sistema no es conforme con el Reglamento propuesto, no debe comercializar ese sistema hasta que dicho sistema sea conforme.

Si el sistema presenta un riesgo para la seguridad, salud o derechos fundamentales de la persona en el sentido del artículo 65.1 del Reglamento propuesto el importador deberá comunicarlo al proveedor del sistema y a las autoridades de vigilancia del mercado a tal efecto.

Los importadores deberán indicar en el sistema su nombre, denominación comercial registrada o marca registrada, y la dirección en la que se les puede contactar o, cuando no sea posible, en su embalaje o en la documentación que lo acompañe, según corresponda.

Los importadores también tienen obligaciones de colaboración con las autoridades nacionales competentes. De un lado, deberán proporcionar a las mismas, previa solicitud motivada, toda la información y documentación necesarias para demostrar la conformidad del sistema con los requisitos establecidos en los artículos 8 a 15 del

Reglamento propuesto anteriormente analizados, y ello en un idioma que pueda ser fácilmente comprensible para el autoridad nacional competente, así como acceso a los registros generados automáticamente por el sistema en la medida en que dichos registros estén bajo su control[574] en virtud de un acuerdo contractual con el usuario o por Ley. De otro, también deberán cooperar con dichas autoridades en cualquier acción que adopte la autoridad nacional competente en relación con el sistema en cuestión.

E. Obligaciones de los distribuidores

Los distribuidores de sistemas de inteligencia artificial de alto riesgo tienen un conjunto de obligaciones previas a la puesta a disposición en el mercado de los mismos en la UE, conforme regula el artículo 27 del Reglamento propuesto, en particular, las siguientes:

a) Verificar que el sistema lleve el marco de conformidad CE y va acompañado de la documentación requerida y las instrucciones de uso;

b) Verificar que tanto el proveedor como el importador del sistema, en su caso, han cumplido con las obligaciones establecidas en el Reglamento propuesto.

Si un distribuidor considera o tiene motivos para considerar que un sistema no se ajusta a los requisitos establecidos en artículos 8 a 15 del Reglamento propuesto (Capítulo 2 del Título III), no pondrá el sistema a disposición en el mercado hasta que dicho sistema haya ha sido puesto en conformidad con esos requisitos. Y si el sistema presenta un riesgo para la seguridad, salud o derechos fundamentales de la persona en el sentido del artículo 65.1 del Reglamento propuesto el distribuidor deberá comunicarlo al proveedor o al importador del sistema, según corresponda, a tal efecto, no a las autoridades de vigilancia del mercado, en la medida que esta obligación correspondería a aquéllos y el sistema no habría sido puesto a disposición del mercado.

Los distribuidores deberán asegurarse de que las condiciones de almacenamiento o transporte no comprometan el cumplimiento de los requisitos establecidos en los

[574] El artículo 26.5 del Reglamento propuesto parece contener un error en la medida que al regular esta obligación alude al "proveedor", no al importador.

precitados artículos 8 a 15 del Reglamento propuesto, mientras el sistema esté bajo su responsabilidad, cuando proceda.

A diferencia de los importadores, si un distribuidor considera o tiene motivos para considerar que un sistema que ha puesto a disposición en el mercado no se ajusta a los requisitos establecidos en los artículos 8 a 15 del Reglamento propuesto analizados en el epígrafe anterior, tomará las acciones correctivas necesarias para que el sistema en funcionamiento sea conforme con dichos requisitos, retirarlo o recuperarlo o se asegurará de que el proveedor, el importador o cualquier operador relevante, según corresponda, tome esas acciones correctivas.

En estos supuestos, en la medida que el sistema ya habrá sido puesto a disposición del mercado, si el sistema presenta un riesgo para la seguridad, salud o derechos fundamentales de la persona en el sentido del artículo 65.1 del Reglamento propuesto, el distribuidor deberá comunicarlo inmediatamente a las autoridades nacionales competentes de los Estados miembros en los que haya distribuido el producto a tal efecto, proporcionando detalles, en particular , del incumplimiento y de las acciones correctivas adoptadas.

Los distribuidores también tienen obligaciones de colaboración con las autoridades nacionales competentes. De un lado, deberán proporcionar a las mismas, previa solicitud motivada, toda la información y documentación necesarias para demostrar la conformidad de un sistema de alto riesgo con los requisitos establecidos en los artículos 8 a 15 del Reglamento propuesto. Y, de otro, los distribuidores también deberán cooperar con esa autoridad nacional competente en cualquier medida que adopte la misma.

F. Obligaciones comunes de distribuidores, importadores, usuarios o cualquier otro tercero.

El Reglamento propuesto regula en su artículo 28 determinadas circunstancias en las que tanto el distribuidor, el importador, el usuario como cualquier otro tercero -sin aclarar que relación o vinculación tenga con el sistema implicado- serán considerados “proveedor” a los efectos del Reglamento propuesto y estarán sujetos al cumplimiento de todas las obligaciones generales del mismo previstas en su artículo 16, que he expuesto

anteriormente. Estas circunstancias que motivarían dicha consideración y exigencia, concurran una sola o varias, son las siguientes:

a) Colocar en el mercado o poner en servicio un sistema de inteligencia artificial de alto riesgo con su nombre o marca comercial;

b) Modificar la finalidad prevista de un sistema de inteligencia artificial de alto riesgo ya comercializado o puesto en servicio;

c) Hacer una modificación sustancial al sistema de inteligencia artificial de alto riesgo.

En cualquiera de las últimas dos circunstancias, el proveedor que comercializó o puso en servicio inicialmente el sistema dejará de ser considerado proveedor a los efectos del Reglamento propuesto, conforme establece el artículo 28.2.

Considero que se trata de presunciones *iuris tantum* susceptibles de prueba para desvirtuar la misma, si bien, sería deseable una concreción en el texto final.

G. Obligaciones de los usuarios

El Reglamento propuesto contempla igualmente un conjunto de obligaciones para los usuarios de los sistemas de inteligencia artificial de alto riesgo, sin perjuicio de otras obligaciones que recaigan en los mismos en virtud de la legislación nacional o de la UE, o dimanantes de la propia discrecionalidad del usuario para organizar sus propios recursos y actividades con el fin de aplicar las medidas de supervisión humana indicadas por el proveedor. Conforme establece su artículo 29, los usuarios se hallarán sujetos a las siguientes obligaciones:

a) Utilizar los sistemas de acuerdo con las instrucciones de uso que los acompañan a los sistemas, de conformidad con los apartados 2 y 5 del precitado artículo;

b) Asegurar que los datos de entrada sean relevantes en vista de la finalidad prevista del sistema inteligente de alto riesgo, en la medida que el usuario ejerza el control sobre los datos de entrada;

c) Monitorizar el funcionamiento del sistema sobre la base de sus instrucciones de uso. Cuando los usuarios sean entidades de crédito reguladas por la Directiva 2013/36/UE, la obligación de monitorización precitada se considerará cumplida mediante el cumplimiento de las normas sobre mecanismos, procesos y mecanismos de gobernanza interna de conformidad con el artículo 74 de dicha Directiva;

d) Informar al proveedor o distribuidor y suspender el uso del sistema, cuando el usuario tenga motivos para considerar que el uso de acuerdo con sus instrucciones puede dar lugar a que el sistema presente un riesgo para la seguridad, salud o derechos fundamentales de la persona en el sentido del artículo 65.1 del Reglamento propuesto;

e) Informar al proveedor o distribuidor cuando el usuario detecte cualquier incidente grave o mal funcionamiento en el sentido del artículo 62 del Reglamento propuesto e interrumpir el uso del sistema. Si el usuario no pueda comunicarse con el proveedor, se aplicará, con los ajustes necesarios, el precitado artículo.

f) Mantener los registros generados automáticamente por el sistema, en la medida en que dichos registros estén bajo su control. El Reglamento propuesto también prevé su conservación, estableciendo que se conservarán durante un período que sea apropiado conforme a la finalidad prevista del sistema y las obligaciones legales aplicables en virtud de la legislación nacional o de la Unión. Al igual que respecto de la obligación anterior, cuando los usuarios sean entidades de crédito reguladas por la Directiva 2013/36/UE mantendrán los registros como parte de la documentación relativa a las disposiciones, los procesos y los mecanismos de gobernanza interna de conformidad con el artículo 74 de dicha Directiva.

g) Utilizar la información a proporcionar en virtud del artículo 13 del Reglamento propuesto para cumplir con su obligación de realizar una evaluación de impacto de

la protección de datos en virtud del artículo 35 del RGPD o del artículo 27 de la Directiva (UE) 2016/680[575], cuando proceda.

6.8.3. Autoridades de notificación y órganos notificados

El Reglamento propuesto regula en sus artículos 30 a 39 la designación y funcionamiento de las autoridades de notificación, responsables de establecer y llevar a cabo los procedimientos necesarios para la evaluación, designación y notificación de los organismos de evaluación de conformidad y de su seguimiento, interacción con organismos de evaluación y notificados que verificarán la conformidad del sistema de inteligencia artificial de alto riesgo con arreglo a los procedimientos de evaluación de la conformidad contemplados en el artículo 43 del Reglamento propuesto. Asimismo, en su artículo 46 regula las obligaciones de información de los organismos notificados.

El objeto y alcance limitados de esta investigación impiden abordar con profundidad estos aspectos.

6.8.4. Evaluación de conformidad

Los artículos 43 a 47 del Reglamento propuesto regulan los procedimientos de evaluación de conformidad, si bien, ante el objeto y alcance concretos y limitados de esa investigación, no puedo abordar los mismos desde un punto de vista crítico.

[575] Directiva (UE) 2016/680 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos y por la que se deroga la Decisión Marco 2008/977/JAI del Consejo. OJ L 119, 4.5.2016. Pp. 89-131.

6.8.5. Presunciones de conformidad

El Reglamento propuesto regula una presunción general de cumplimiento de los requisitos previstos en los artículos 8 a 15 del Reglamento propuesto (Capítulo 2 del Título III) de aquellos sistemas de inteligencia artificial de alto riesgo que sean conformes con las normas armonizadas o partes de las mismas cuyas referencias hayan sido publicadas en el Diario Oficial de la UE, de conformidad con lo previsto en su artículo 40.

Del mismo modo, el texto propuesto prevé algunas previsiones específicas, en particular, cuando no existan las precitadas normas armonizadas, cuando la Comisión considere que las normas armonizadas pertinentes son insuficientes o que es necesario abordar problemas específicos de seguridad o derechos fundamentales, la Comisión podrá adoptar especificaciones comunes respecto a los requisitos establecidos en los precitados artículos 8 a 15 del Reglamento propuesto.

Los actos de ejecución que se instrumenten con dichas finalidades deberán adoptarse conforme al procedimiento de examen regulado en el artículo 74.2.

En consecuencia, los sistemas de inteligencia artificial de alto riesgo que sean conformes con las especificaciones comunes indicadas, se presumirán conformes con los requisitos establecidos en los artículos precitados, siempre que dichas especificaciones comunes cubran esos requisitos.

El artículo 41.4 del Reglamento propuesto ya prevé una solución a las situaciones que podrán suscitarse en la práctica en relación con estas cuestiones cuando los proveedores no se ajusten a las especificaciones comunes indicadas, dado que, en estos supuestos, los proveedores deberán justificar debidamente que han adoptado soluciones técnicas al menos equivalentes.

Por último, el Reglamento propuesto también contempla otras presunciones específicas en su artículo 42, en particular, de cumplimiento de determinados requisitos:

a) De un lado, considerando la finalidad prevista, según el artículo precitado, se presumirá que los sistemas de inteligencia artificial de alto riesgo que hayan sido formados y probados con datos relativos al entorno geográfico, conductual y funcional específico en el que se vayan a utilizar cumplen el requisito establecido en el artículo 10.4 del mismo relativo a que los conjuntos de datos de formación, validación y ensayo consideren estos aspectos.

b) De otro, los sistemas de inteligencia artificial de alto riesgo que hayan sido certificados o para los que se haya emitido una declaración de conformidad bajo un esquema de ciberseguridad de conformidad con el Reglamento (UE) 2019/881 del Parlamento Europeo y del Consejo y cuyas referencias se hayan publicado en el Diario Oficial de la UE se presumirán conformes con los requisitos de ciberseguridad establecidos en el artículo 15 del presente Reglamento propuesto, en la medida en que el certificado de ciberseguridad o la declaración de conformidad, o partes de los mismos, cubran dichos requisitos.

6.8.6. Certificados

Los certificados a expedir por los organismos notificados se regulan en el artículo 44 en relación con el Anexo VII del Reglamento propuesto, si bien, como he indicado en el apartado anterior, ante el objeto y alcance concretos y limitados de esa investigación, no puedo abordar los mismos en el marco de la misma.

6.8.7. Declaración UE de conformidad

Conforme se ha referido anteriormente, el proveedor deberá redactar una declaración UE de conformidad por escrito para cada sistema de inteligencia artificial y la mantendrá a disposición de las autoridades nacionales competentes durante diez años después de que el sistema se haya comercializado o puesto en servicio.

La declaración UE de conformidad deberá identificar el sistema de inteligencia artificial para el que se ha elaborado y se entregará una copia de la declaración UE de conformidad a las autoridades nacionales competentes pertinentes que lo soliciten.

La declaración UE de conformidad indicará que el sistema de inteligencia artificial de alto riesgo cumple los requisitos establecidos en los artículos 8 a 15 del Reglamento propuesto y deberá contener la información regulada en el Anexo V, anteriormente citada.

Asimismo, la declaración deberá ser traducida a la lengua/s oficial/es de la UE exigidas por los Estados miembros en los que esté disponible el sistema de inteligencia artificial de alto riesgo.

El artículo 48.3 del Reglamento propuesto prevé una solución cuando los sistemas de inteligencia artificial de alto riesgo estén sujetos a otra legislación de armonización de la UE que también requiera una declaración de conformidad de la UE.

En estos supuestos se redactará una única declaración de conformidad de la UE con respecto a todas las legislaciones de la UE aplicables al sistema que contendrá toda la información necesaria para identificar la legislación de armonización de la Unión a la que se refiere la declaración.

Como he expuesto anteriormente y recoge expresamente el artículo 48.4 del Reglamento propuesto, el proveedor asumirá en la elaboración de la declaración UE de conformidad la responsabilidad del cumplir con los requisitos establecidos en los precitados artículos 8 a 15 del Reglamento propuesto, debiendo igualmente mantenerla actualizada.

El artículo 48.5 del Reglamento propuesto prevé también mecanismos para actualizar el contenido de la declaración para adaptarla al progreso técnico, mediante el otorgamiento de poderes a la Comisión para la adopción de actos delegados.

6.8.8. Marcado CE de conformidad

El marcado CE, conforme ha expuesto al analizar los requisitos de estos sistemas, deberá colocarse de forma visible, legible e indeleble para los sistemas de inteligencia de alto riesgo, salvo que esto no sea posible o no esté justificado debido a la naturaleza del sistema de inteligencia artificial de alto riesgo, en cuyo caso, se colocará en el embalaje o en la documentación adjunta, según corresponda, de conformidad con lo previsto en el artículo 49 del Reglamento propuesto.

6.9. Obligaciones para otros sistemas de inteligencia artificial

El Reglamento propuesto, como he referido anteriormente, regula la prohibición de sistemas de riesgo inaceptable, así como los requisitos y obligaciones de los sistemas de inteligencia artificial de alto riesgo, si bien, no regula los principios, normas éticas u otras obligaciones del resto de sistemas, a excepción de lo previsto en su artículo 52, que regula obligaciones de transparencia e información a proveedores y usuarios, respecto de determinados sistemas de inteligencia artificial, en función de su naturaleza, sean o no de alto riesgo.

Respecto de los sistemas de inteligencia artificial destinados a interactuar con personas físicas, conforme he definido anteriormente, el precepto indicado exige a sus proveedores que se aseguren que dichos sistemas estén diseñados y desarrollados de manera que las personas físicas que interactúen con estos sistemas estén informadas de que están interactuando con un sistema de inteligencia artificial, salvo cuando esto sea obvio por las circunstancias y el contexto de uso.

El Reglamento excluye de esta obligación a los sistemas de inteligencia artificial autorizados por la ley para detectar, prevenir, investigar y perseguir delitos, a menos que esos sistemas estén disponibles para que el público pueda denunciar un delito.

Respecto de los sistemas de inteligencia artificial para el reconocimiento de emociones o de categorización biométrica, conforme he definido igualmente con anterioridad, el apartado 2 del artículo precitado, exige a sus usuarios que informen del funcionamiento del sistema a las personas físicas expuestas al mismo, salvo en el caso de sistemas de inteligencia artificial utilizados para la categorización biométrica, que están permitidos por ley para detectar, prevenir e investigar delitos.

Respecto de los sistemas de inteligencia artificial que generen o manipulen contenido de imagen, audio o video que se parezca apreciablemente a personas, objetos, lugares u otras entidades o eventos existentes y que le parezca falsamente auténtico o veraz a una persona -*deep fakes* o "falso profundo"-, el apartado 3 del precepto indicado exige que los usuarios del sistema deberán revelar que el contenido ha sido generado o manipulado artificialmente, salvo cuando el uso esté autorizado por la ley para detectar, prevenir, investigar y perseguir delitos o sea necesario para el ejercicio del derecho a la libertad de expresión y el derecho a la libertad de las artes y las ciencias garantizados. en la Carta de los Derechos Fundamentales de la UE, y sujeto a las debidas garantías de los derechos y libertades de terceros.

De este modo, el derecho a la libre creación técnica y artística reconocidos como derecho fundamental dentro de la libertad de expresión constituiría una excepción a la obligación de informar sobre su creación o manipulación artificial cuando se trate de una creación técnica y artística en el ejercicio del mismo. El análisis de esta cuestión me parece interesantísima, pero resulta inabordable en el marco del objeto y alcance limitados de esta investigación.

Por otra parte, no se específica cuándo deberá "revelarse" esta información, entendiendo que deberá acompañar al contenido generado y su divulgación, si bien, sería adecuado precisar este aspecto en la revisión de este borrador.

El artículo 52.4 del Reglamento propuesto establece que estas obligaciones adicionales no afectarán a los requisitos y obligaciones analizados en los apartados precedentes, regulados en los artículos 6 a 51 del mismo.

6.10. Medidas de apoyo a la innovación

El Título V regula distintas medidas de apoyo a la innovación en materia de inteligencia artificial:

- Sandbox regulatorio

  En primer lugar, el Reglamento propuesto apuesta en su artículo 53 por los "*sandboxes"* regulatorios al objeto de promover la creación de entornos controlados que faciliten el desarrollo, pruebas y validación de sistemas de inteligencia artificial innovadores, durante tiempo limitado y de manera previa a su comercialización o puesta en servicio.

  La creación y gestión de estos entornos se efectuará con la supervisión y orientación directa de las autoridades competentes, entre las que deberán estar las autoridades en protección de datos, conforme expresamente recoge el artículo precitado, cuando los sistemas impliquen el tratamiento de los mismos.

  El artículo 54.1 del Reglamento propuesto prevé un conjunto de condiciones para permitir el tratamiento de datos personales recopilados para otros fines en el entorno de pruebas con la finalidad de desarrollar ciertos sistemas de inteligencia artificial en interés público.

  No obstante, esta previsión se halla condicionada a lo previsto en su apartado 2, esto es, "en la medida que la legislación de la UE o de los Estados miembros que excluya el tratamiento con fines distintos de los mencionados explícitamente en dicha legislación".

  Las áreas de interés público que identifica el precepto indicado son: a) La prevención, investigación, detección o enjuiciamiento de infracciones penales o la ejecución de sanciones penales, incluida la protección y prevención de amenazas a la seguridad pública, bajo el control y responsabilidad de las autoridades competentes; b) La seguridad pública y salud pública, incluida la prevención, el control y el tratamiento de enfermedades; c) Un alto nivel de protección y mejora de la calidad del medio ambiente.

Las condiciones requeridas para dicho tratamiento incluyen igualmente distintos requerimientos específicos, en particular que: a) Los datos tratados sean necesarios para cumplir con uno o más de los requisitos exigidos por los artículos 8 a 15 del Reglamento propuesto; b) Existencia de mecanismos de seguimiento efectivo y de respuesta ante riesgos; c) Realizarse en un entorno de procesamiento separado, protegido y controlado sólo accesible para personales autorizadas; d) El entorno no debe generar medidas o decisiones que afecten a los interesados; e) Eliminación de los datos tratados una vez que la participación ha terminado o expire el período de conservación de aquéllos; f) Conservación de los registros de tratamiento en estos entornos durante la participación y un año después de su terminación para cumplir las obligaciones de responsabilidad, documentación o legales; g) La descripción del proceso debe formar parte de la documentación técnica; h) Publicación de un breve resumen del proyecto desarrollado en el entorno de *sandbox* regulatorio en las páginas web de las autoridades competentes.

El desarrollo y pruebas en estos entornos estarán igualmente sujetos a análisis de riesgos, de modo que cualquier riesgo significativo para la salud y la seguridad y los derechos fundamentales identificados durante estas tareas exigirá una mitigación inmediata y, en su defecto, a la suspensión del proceso de desarrollo y prueba hasta que se lleve a cabo dicha mitigación.

A los efectos de responsabilidad que constituye el eje principal de esta investigación, el artículo 54.4 prevé expresamente que los participantes en los *sandboxes* regulatorios serán responsables en todo momento por cualquier daño infligido a terceros como resultado de la experimentación en estos entornos, conforme a la legislación de responsabilidad de la UE y de los Estados miembros.

De este modo, la flexibilidad que confieren estos entornos para el desarrollo y testeo previo a la comercialización o puesta en servicio de sistemas no exime de responsabilidad alguna a los participantes por los daños que ocasionen conforme a los marcos de responsabilidad.

Las modalidades y condiciones de funcionamiento de los *sandboxes*, incluidos los criterios de elegibilidad y el procedimiento para la solicitud, selección,

participación y salida del *sandbox*, y los derechos y obligaciones de los participantes, se establecerán en los actos de ejecución previstos en los artículos 53.6 en relación con el 74.2 del Reglamento propuesto.

- Medidas para proveedores y usuarios en pequeña escala

El Reglamento propuesto prevé en su artículo 55 un conjunto de acciones a llevar a cabo por los Estados miembros en relación con proveedores y usuarios en pequeña escala.

Los Estados miembros deberán emprender acciones con la finalidad de proporcionar a los precitados proveedores de pequeña escala y a las empresas emergentes -*startups*- acceso prioritario a los entornos de pruebas precitados en la medida en que cumplan con las condiciones de elegibilidad, de organizar actividades específicas de sensibilización sobre la aplicación del Reglamento adaptadas a las necesidades de los proveedores y usuarios a pequeña escala y, cuando proceda, para establecer un canal de comunicación exclusivo con proveedores y usuarios a pequeña escala y otros innovadores para proporcionar orientación y responder a consultas sobre la aplicación del Reglamento.

Asimismo, el apartado 2 de artículo precitado, indica que se tendrán en cuenta los intereses y necesidades específicos de los proveedores a pequeña escala al establecer las tasas por evaluación de la conformidad, reduciendo las mismas proporcionalmente a su tamaño y tamaño del mercado.

Igualmente, el Reglamento propuesto contempla otras medidas e incentivos en otros apartados del mismo.

De un lado, su artículo 69.4 establece que la Comisión Europea y el Comité Europeo de Inteligencia Artificial deberán tener en cuenta los intereses y necesidades específicas de los proveedores a pequeña escala y de *startups* al fomentar y facilitar la elaboración de los códigos de conducta voluntarios previstos en el mismo para sistemas distintos a los de alto riesgo.

> De otro, su artículo 71.1. prevé la creación de un marco sancionador por parte de los Estados miembros como instrumento coercitivo para el cumplimiento efectivo de sus disposiciones, que deberá tener en cuenta los intereses de los proveedores de pequeña escala y *startups*, así como su viabilidad económica.

A mi juicio todas estas acciones son absolutamente necesarias pero insuficientes a la vista de los requerimientos requeridos en el Reglamento propuesto y el reto que previsiblemente supondrá a nivel cualitativo y cuantitativo para proveedores y usuarios a pequeña escala y *startups*, que constituyen un motor de la innovación a nivel mundial y que no pueden competir en igualdad de armas en el mercado.

Las *startups* se enfrentan a marcos regulatorios cada vez más exigentes en múltiples aspectos, especialmente significativos en países como España, donde deben cumplir, entre otras, con sus obligaciones en materia de *compliance* penal -con todo lo que ello comporta-, en materia de protección de datos o de ciberseguridad, así como otros marcos sectoriales y técnicos como los analizados en este apartado, los cuales obligarían a destinar buena parte de sus recursos iniciales, habitualmente limitados, no tanto a investigar e innovar sino a cumplir dichos marcos, lo que de antemano puede constituir un obstáculo efectivo para la innovación y un aspecto esencial a considerar de inicio para la ubicación física de una nueva empresa, junto con otros más clásicos pero igualmente determinantes, como las ayudas e incentivos públicos o privados y los aspectos fiscales.

Si no disponen de un entorno incentivador, el resultado es que el talento y la innovación se marche fuera de nuestro país y de la propia UE por razones obvias, en ocasiones, ubicando directamente la sede social y operativas en países con mayores incentivos y menos cargas, o siendo participadas o adquiridas por empresas ubicadas en dichos países.

En la medida que el Reglamento propuesto entra en ello, considero que debería concretar la creación de incentivos específicos y la adopción de medidas concretas en lo sucesivo, dado que de otro modo las autoridades nacionales se verán sujetas al marco de requerimientos y obligaciones establecido en el mismo, de modo que cualquier acción incentivadora o flexibilizadora no prevista en aquél que pueda adecuar los niveles de exigencia en determinadas parcelas deberían ser ya previstas en el Reglamento propuesto,

sin perjuicio del valor de las acciones de sensibilización en este sentido a las que alude y las medidas que se promuevan en lo sucesivo en base a sus prescripciones.

A mi juicio, los nuevos marcos deberán contemplar o acompañarse de las condiciones o al menos criterios de selección y participación para optar a dicho acceso prioritario a estos incentivos, así como las medidas concretas de flexibilización de requerimientos y obligaciones manteniendo el necesario equilibrio y consecución de objetivos de seguridad y confianza.

6.11. Gobernanza

El Reglamento propuesto contempla en su Título VI, artículos 56 y siguientes, un marco de gobernanza de la inteligencia artificial, que prevé la creación del Comité Europeo de Inteligencia Artificial, concebido como un órgano de asesoramiento y asistencia a la Comisión para contribuir a las cooperación efectiva de las autoridades nacionales de supervisión y la Comisión, dar soporte a las mismas para garantizar las aplicación coherente del Reglamento propuesto, y coordinar y contribuir a la orientación y el análisis por parte de la Comisión y las autoridades nacionales de supervisión y otras competentes sobre cuestiones emergentes relacionadas con el Reglamento.

Esta fórmula de gobernanza es similar a la prevista en el Reglamento General de Protección de Datos.

El marco propuesto prevé también la designación de las autoridades nacionales competentes para garantizar la aplicación y ejecución del Reglamento, de entre las que deberá designarse una autoridad nacional de supervisión que actuará como autoridad notificante y autoridad de vigilancia del mercado, con posibilidad de que los Estados pueden designar más de una autoridad por razones organizativas y administrativas.

Dentro de sus funciones, su artículo 59.7 prevé que puedan proporcionar orientación y asesoramiento sobre la aplicación del Reglamento incluso a los proveedores a pequeña escala.

El precitado Dictamen conjunto 5/2021, del Comité Europeo de Protección de Datos con el Supervisor Europeo de Protección de Datos sobre el Reglamento propuesto, recoge como ambos organismos consideran que las autoridades de protección de datos deberían ser designadas como autoridades nacionales de control de conformidad con el artículo 59 de la Propuesta.

Asimismo, los artículos 73 y 74 del Reglamento propuesto, regulan el otorgamiento a la Comisión de poderes para adoptar actos delegados en las condiciones recogidas en los mismos, así como la asistencia a la misma de un Comité, si bien, respecto de su composición, funciones y competencias, se estará a lo dispuesto en el Reglamento (UE) no 182/2011 del Parlamento Europeo y del Consejo, de 16 de febrero de 2011, por el que se establecen las normas y los principios generales relativos a las modalidades de control por parte de los Estados miembros del ejercicio de las competencias de ejecución por la Comisión[576].

El citado Dictamen conjunto 5/2021, del Comité Europeo de Protección de Datos con el Supervisor Europeo de Protección de Datos concluye igualmente que, para garantizar su independencia, el futuro Reglamento sobre inteligencia artificial debería dar más autonomía al Comité Europeo de Inteligencia Artificial -EAIB- y garantizar que pueda actuar por iniciativa propia.

6.12. Base de datos para sistemas de inteligencia artificial de alto riesgo independientes

Como he referido anteriormente, el Reglamento propuesto prevé en su Título VII y, en particular, en su artículo 60, la creación de una base de datos de la UE para sistemas de inteligencia artificial de alto riesgo independientes, de cuyo tratamiento será responsable la Comisión Europea, la cual deberá ofrecer a los proveedores apoyo técnico y administrativo adecuado.

[576] OJ L 55, 28.2.2011, Pp. 13-18.

Merece destacar que el título del artículo precitado conserva e incorpora el concepto "autónomo", entiendo que por error, dado que es un concepto evitado en esta propuesta objeto del presente análisis. De hecho, el precepto indicado se titula "Base de datos de la UE para sistemas autónomos de IA de alto riesgo".

El Reglamento propuesto prevé la creación y mantenimiento por parte de la Comisión, en colaboración con los Estados miembros, de una base de datos de la UE que contenga la información sobre los sistemas de inteligencia artificial de alto riesgo relacionados en el Anexo III, esto es los sistemas relacionados con la identificación biométrica y la categorización de personas físicas, la gestión y la operación de infraestructuras críticas, la educación y la formación profesional, el empleo, la gestión de trabajadores y el acceso al autoempleo, el acceso y disfrute de servicios privados esenciales y servicios y beneficios públicos, el cumplimiento de la ley, la gestión de migración, asilo y control de fronteras, la administración de justicia y los procesos democráticos.

La información a registrar será la detallada en el anexo VIII y se hallará accesible al público.

6.13. Aplicación del Reglamento

El artículo 63 del Reglamento propuesto establece los mecanismos de vigilancia del mercado y de control de los sistemas de inteligencia artificial en la UE por parte de las autoridades competentes conforme al mismo.

En el marco de sus actividades, las autoridades de vigilancia del mercado deberán tener acceso completo a datos y documentación, en particular, a los conjuntos de datos de formación, validación y prueba utilizados por el proveedor, para que el artículo 64.1 del Reglamento propuesto prevé distintas opciones, como interfaces de programación de aplicaciones (API) u otros medios técnicos y herramientas adecuados que permitan el acceso a distancia.

El artículo 64.2 faculta a dichas autoridades a acceder al código fuente del sistema de inteligencia artificial cuando sea necesario para evaluar la conformidad del sistema con

los requisitos establecidos en los artículos 8 a 15 del Reglamento propuesto, previa solicitud motivada por parte de dicha autoridad. Ello supone una habilitación legal para acceso al código protegido por los derechos de autor, conforme será analizado más adelante.

Del mismo modo, las autoridades u organismos públicos nacionales encargados de la supervisión o cumplimiento de las obligaciones del Derecho de la UE regulador de la protección de los derechos fundamentales en relación con el uso de sistemas de inteligencia artificial de alto riesgo, se hallarán facultados para solicitar y acceder a cualquier documentación creada o mantenida en virtud de lo regulado en el Reglamento propuesto, cuando el acceso sea necesario para el cumplimiento de sus competencias dentro de su jurisdicción.

Cualquier solicitud de esta naturaleza deberá ser informada a la autoridad de vigilancia.

Si se considerase insuficiente para determinar el incumplimiento, las autoridades u organismo público precitados podrán solicitar motivadamente a la autoridad de vigilancia del mercado la articulación de pruebas en este sentido.

Sobre esta cuestión, el Reglamento establece que en el plazo máximo de 3 meses cada Estado miembro deberá identificar las autoridades y organismos públicos precitados para su publicación en la página web de la autoridad nacional de supervisión.

6.14. Procedimiento para tratamiento de sistemas que presenten riesgos a nivel nacional

El artículo 65.1 establece que los sistemas de inteligencia artificial que presenten un riesgo, sin mayor concreción de su entidad o naturaleza, se entenderán como un producto que presenta un riesgo definido en el artículo 3, punto 19, del precitado Reglamento (UE) 2019/1020[577] sobre vigilancia del mercado y conformidad de los productos, en lo que

[577] Reglamento (UE) 2019/1020 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativo a la vigilancia del mercado y la conformidad de los productos y por el que se modifican la

respecta a los riesgos para la salud o la seguridad o para la protección de los derechos fundamentales de las personas.

En tales supuestos, si la autoridad de vigilancia del mercado del Estado miembro tiene motivos suficientes para considerar que un sistema presenta un riesgo en los términos indicados, deberá a cabo una evaluación del sistema de inteligencia artificial en cuestión en lo que respecta al cumplimiento de todos los requisitos y obligaciones establecidos en el presente Reglamento, en aquello que le resulte de aplicación.

Cuando existan riesgos para la protección de los derechos fundamentales, la autoridad de vigilancia del mercado deberá informar también a las autoridades u organismos públicos nacionales pertinentes precitados, para lo que los operadores implicados deberán cooperar en la medida necesaria con las autoridades de vigilancia del mercado y con las demás autoridades u organismos públicos nacionales.

Si durante dicha evaluación, la autoridad de vigilancia del mercado constate que el sistema no cumple los requisitos y obligaciones establecidos en el Reglamento propuesto, deberá exigir sin demora al agente económico implicado que adopte todas las medidas correctoras adecuadas para que el sistema sea conforme, que lo retire del mercado o que lo recupere en un plazo razonable, de manera proporcionada a la naturaleza del riesgo, según determine. Del mismo deberán comunicar al organismo notificado.

Si el incumplimiento no se limita al territorio nacional de la autoridad de vigilancia del mercado, ésta deberá informará a la Comisión Europea y a los demás Estados miembros de los resultados de la evaluación y de las medidas que haya exigido al agente involucrado.

Conforme expuse anteriormente al analizar los requisitos y obligaciones de los sistemas de alto riesgo y conforme igualmente regula el artículo 65.4 del Reglamento propuesto, el operador involucrado se asegurará de que se adopten todas las medidas correctoras

Directiva 2004/42/CE y los Reglamentos (CE) n.° 765/2008 y (UE) n.° 305/2011. OJ L 169, 25.6.2019. Pp. 1–44

adecuadas en relación con todos los sistemas de inteligencia artificial afectados que haya comercializado en toda la UE.

Si el operador no adopta las medidas correctoras adecuadas en el plazo mencionado precitado, la autoridad de vigilancia del mercado podrá adoptar todas las medidas provisionales adecuadas para prohibir o restringir la comercialización del sistema de inteligencia artificial en su mercado nacional, retirar el producto de ese mercado o recuperarlo. Y dicha autoridad deberá informar sin demora a la Comisión Europa y a los demás Estados miembros de dichas medidas.

Las autoridades de vigilancia del mercado de todos los Estados miembros se asegurarán de que se adopten sin demora las medidas restrictivas adecuadas con respecto al producto en cuestión, como la retirada del producto de su mercado.

El artículo 66 del Reglamento propuesto regula los procedimientos de salvaguardia en relación con las medidas adoptadas por un Estado miembro.

El artículo 67 regula los requerimientos de la autoridad de vigilancia a los operadores para que adopten las medidas adecuadas para garantizar que el sistema no presente un riesgo cuando, tras la evaluación y comprobación de la conformidad, reste un riesgo para la salud o la seguridad de las personas, para el cumplimiento de las obligaciones derivadas del Derecho de la UE o nacional destinadas a proteger los derechos fundamentales o para otros aspectos de la protección del interés público.

Y, por último, el artículo 68 regula los requerimientos de cumplimientos a los proveedores por parte de la autoridad de vigilancia del mercado cuando compruebe el incumplimiento del marcado de conformidad, la falta de elaboración o redacción incorrecta de la declaración UE de conformidad o no identificación del número del organismo notificado participante en los procedimientos de evaluación de conformidad.

6.15. Códigos de conducta

Conforme he expuesto anteriormente, el Reglamento propuesto no regula las obligaciones de los sistemas distintos de los calificados como alto riesgo, con la excepción de las obligaciones generales de transparencia e información, si bien, en su artículo 69, apuesta por el fomento, por parte de la Comisión Europea y los Estados miembros, de los códigos de conducta orientados a fomentar la aplicación voluntaria al resto de sistemas de los requisitos previstos en el mismo para los sistemas de inteligencia artificial de alto riesgo -artículos 8 a 15-, así como otros requisitos relacionados como, por ejemplo, la sostenibilidad medioambiental, la accesibilidad de las personas con discapacidad, la participación de los interesados en el diseño y desarrollo de los sistemas y la diversidad de equipos de desarrollo sobre la base de objetivos claros e indicadores de rendimiento para medir el logro de esos objetivos.

Los códigos de conducta, conforme prevé el apartado 3º del precepto citado, podrán ser elaborados por proveedores individuales de sistemas inteligentes o por organizaciones que los representen o ambos, incluso, a mi juicio muy conveniente, con la participación de los usuarios y las partes interesadas y sus organizaciones representativas.

El Reglamento propuesto apuesta pues por el *soft law* para el resto de sistemas de inteligencia artificial, con lo que discrepo, conforme expuse y argumenté en los capítulos precedentes, en la medida que considero necesario establecer un marco básico y esencial que constituya una base sólida para el desarrollo, aplicación y uso de una inteligencia artificial segura y confiable, garante de los bienes y derechos, especialmente los fundamentales, y que no afecte sino que incluso incentive la competitividad y la innovación, de modo que aspectos como la seguridad, la privacidad, la responsabilidad, la supervisión o el control humano estén garantizados, en atención especialmente a su características o capacidades innatas en su concepción o adquiridas en su utilización mediante técnicas, por ejemplo, de aprendizaje automático y profundo.

6.16. Confidencialidad de la información, secretos comerciales, derechos de propiedad intelectual y cumplimiento

El Reglamento propuesto, prevé en su artículo 70 las obligaciones de confidencialidad y de protección de la información, secretos comerciales y derechos de propiedad intelectual de personas físicas o jurídicas, a las que se hallarán sujetas las autoridades nacionales competentes, los organismos notificados involucrados y la propia Comisión Europea en el desempeño de sus tareas y actividades, tanto en lo relativo a su acceso como intercambio.

Estas prescripciones son importantes en la medida que, en sus actuaciones, las autoridades pueden tener acceso a información protegida por los marcos reguladores del secreto de empresa, las patentes y los derechos de autor.

Del mismo modo, prevé la protección por parte de las mismas de la aplicación efectiva del Reglamento propuesto, así como la integridad de los procedimientos penales y administrativos.

6.17. Sanciones

El Reglamento propuesto prevé la creación de un marco sancionador por parte de los Estados miembros, a los que deriva el establecimiento del mismo, como instrumento coercitivo para el cumplimiento efectivo de sus disposiciones, regulado en sus artículos 71, siguientes y concordantes.

No obstante, el Reglamento propuesto ya recoge un conjunto de infracciones y su sanción correlativa, en el apartado 2 del precitado artículo, cuantificando su sanción en atención a su gravedad.

De este modo, el Reglamento propuesto tipifica las siguientes infracciones como las más graves y que deberán ser sancionadas con multas administrativas de hasta treinta (30)

millones de euros o, si el infractor es una empresa, hasta 6% de su volumen de negocios anual total a nivel mundial durante el ejercicio económico anterior, según resulte el mayor importe:

a) Incumplimiento de la prohibición de las prácticas de inteligencia artificial reguladas en el artículo 5 del Reglamento propuesto, es decir, la puesta en el mercado, puesta en servicio o uso de los sistemas de inteligencia artificial prohibidos en dicho artículo de riesgo inaceptable;

b) Incumplimiento del sistema de inteligencia artificial de los requisitos establecidos en el artículo 10 del Reglamento propuesto, es decir, el incumplimiento de las prácticas de gestión, tratamiento y gobernanza de datos regulados en el mismo.

Asimismo, el incumplimiento de cualquier otro requisito u obligación exigida por el Reglamento propuesto distinto a los enumerados en el anterior párrafo, deberá ser sancionado igualmente con multas administrativas de hasta veinte (20) millones de euros o, si el infractor es una empresa, hasta el 4% de su facturación anual total en todo el mundo durante el ejercicio económico anterior, aplicando el importe que sea más alto.

Por último, el Reglamento propuesto también considera otras infracciones de menor gravedad relativas al suministro de información incorrecta, incompleta o engañosa a los organismos notificados y a las autoridades nacionales competentes en respuesta a una solicitud de la mismas, que deberán ser sancionado con multas administrativas de hasta diez (10) millones de euros o, al igual que las anteriores, si el infractor es una empresa, hasta el 2% de su volumen total de negocio anual en todo el mundo durante el ejercicio económico anterior, aplicando el importe que sea más alto.

Por otra parte, el apartado 6 del artículo 71 del Reglamento establece un conjunto de prescripciones y criterios que deberán ser considerados para decidir la cuantía en cada caso concreto, debiendo considerar junto al contexto, la naturaleza, gravedad y duración de la infracción y de sus consecuencias, si otras autoridades de vigilancia del mercado ya han aplicado multas administrativas al mismo operador por la misma infracción y, por último, el tamaño y la cuota de mercado del operador que comete la infracción

Respecto de la aplicación del régimen sancionador a las autoridades y organismos públicos establecidos en un Estado miembro y en qué medida, el apartado 7 del precepto indicado, se remite a las normas que establezcan los Estados sobre la posibilidad de imponer multas administrativas a aquéllos y qué grado.

En el caso de multas administrativas a instituciones, agencias y organismos de la UE que entren en el ámbito de aplicación del Reglamento propuesto, el artículo 72 del mismo atribuye esa facultad al Supervisor Europeo de Protección de Datos, tipificando en dicho precepto un conjunto de infracciones en la misma línea que las generales precitadas, pero sancionadas con multas sensiblemente inferiores, esto es, 500.000€ las más graves ante incumplimiento de la prohibición de las prácticas de inteligencia artificial reguladas en el artículo 5 del Reglamento propuesto, es decir, la puesta en el mercado, puesta en servicio o uso de los sistemas de inteligencia artificial prohibidos en dicho artículo de riesgo inaceptable o el incumplimiento de las prácticas de gestión, tratamiento y gobernanza de datos regulados en el mismo.

En relación con esta potestad atribuida al Supervisor Europeo de Protección de Datos, esta cuestión debe ser objeto de reflexión y revisión, en la medida que parece no tener demasiado sentido ante la creación de un Comité Europeo de inteligencia artificial.

Del mismo modo, en caso de incumplimiento de un sistema inteligente de cualquier requisito u obligación regulado en el Reglamento distinto de los anteriores, estará sujeto a multas administrativas de hasta 250.000€.

6.18. Otras cuestiones adicionales

Los artículos 75, siguientes y concordantes del Reglamento propuesto modifican múltiples normas comunitarias en congruencia con las disposiciones del mismo.

Por lo que se refiere a su aprobación, entrada en vigor y aplicación, ya prevé una aplicación asíncrona a su entrada en vigor, de modo que se aplicaría dos años después de su entrada en vigor, similar a la entrada en vigor y aplicación del Reglamento General de Protección de Datos (RGPD).

7. Consideraciones finales

La inteligencia artificial está siendo desarrollada, desplegada y aplicada de manera incesante en todo tipo de ámbitos y sectores y lo seguirá siendo a un ritmo de crecimiento exponencial conforme lo haga la tecnología sobre la que sustenta, su relación e interacción con otras, la mayor capacidad de computación, así como la mayor disponibilidad de datos masivos.

Sobre ello ya he anticipado mis reflexiones y posicionamiento sobre distintos aspectos abordados a lo largo de este capítulo.

El Derecho no contempla esta realidad tan compleja y ni mucho menos las soluciones a los retos, riesgos y conflictos que su desarrollo, despliegue, aplicación y uso están generando y que generarán en mayor medida en los próximos años, no sólo por su creciente uso y aplicación, sino también ante su interacción con otras tecnologías, aumento de la capacidad de procesamiento, conectividad, características y capacidades propias y la mayor disponibilidad de datos masivos, que potenciarán igualmente no sólo su crecimiento cuantitativo sino cualitativo, especialmente ante el aumento de su supuesta autonomía, capacidad de autoaprendizaje, independencia y relativa impredecibilidad.

Los marcos jurídicos actuales no están preparados para el potencial disruptivo de la inteligencia artificial y sus aplicaciones.

El Derecho no puede ser creado a la misma velocidad que las realidades que pretende regular.

La autonomía real y plena no es una realidad científica hoy y no debería serlo por razones éticas y de seguridad, pero su peligrosa aproximación requiere ya una reflexión ética y jurídica que no debe postergarse. No podemos partir de una inteligencia artificial simple y débil si queremos mirar hacia el futuro, por lo que debemos considerar también sistemas inteligentes progresivamente más avanzados y con mayor autonomía y capacidades.

Los nuevos marcos reguladores de la inteligencia artificial en el ámbito de la UE deberían centrarse en el ser humano, en sentido individual y colectivo, concibiendo la tecnología

como un medio y no como un fin, y una tecnología dotada de inteligencia por y para la sociedad.

Además deberían partir, a mi juicio, de un enfoque amplio y global considerando todos los retos y riesgos que plantea para todos los sujetos relacionados y para todos los bienes y derechos a proteger, de una combinación de técnicas de *hard* y *soft law* que garanticen la seguridad y el equilibrio entre los intereses y derechos de las distintas partes involucradas, de una convergencia entre ética, seguridad y derecho, así como de premisas incuestionables como que los derechos fundamentales siempre deben prevalecer sobre cualquier tecnología y que debería tener más costes cumplir que investigar.

Y, por último, hay que considerar que los retos y riesgos pueden tener impacto tanto la dimensión física como en la virtual, que constituyen una única realidad, la que vivimos y que, en ocasiones, parece más la que percibimos.

El reconocimiento de una personalidad jurídica electrónica, ya apuntado por el propio Parlamento Europeo en algunas de sus Resoluciones citadas en esta investigación, no es una realidad actual ni una opción prevista a corto y medio plazo, si bien, los imparables avances tecnológicos y el cuestionable objetivo de emular artificialmente la mente humana, posiblemente exigirá revisar en el futuro esta posibilidad.

En este escenario, la UE pretende liderar a nivel internacional la regulación de los marcos éticos vinculantes y de responsabilidad civil por daños de la inteligencia artificial, para lo que su Parlamento aprobó sendas resoluciones el 20 de octubre de 2020 con dicho objetivo que acompañaban sus respectivas propuestas de Reglamento, como norma jurídica de alcance general y eficacia directa, esto es, directamente aplicable en todos los Estados de la UE por cualquier autoridad o particular, sin que sea precisa ninguna norma jurídica de origen interno o nacional que la transponga para completar su eficacia plena, e invocable ante los tribunales nacionales o comunitarios por cualquier particular.

A las propuestas precitadas, se une ahora la nueva propuesta analizada en último apartado de este capítulo, que se presenta como la Ley de inteligencia artificial -*Artificial Intelligence Act*- y que pretende constituir la principal norma reguladora de la UE de la inteligencia artificial, si bien, no la regula en toda su amplitud, limitándose a la

denominada inteligencia artificial "débil", a la prohibición de los sistemas de inteligencia artificial de riesgo actualmente inaceptable y a regular con detalle los sistemas inteligentes considerados de alto riesgo conforme al Reglamento propuesto, sujetando a concretos sistemas no incluidos en los mismos a algunas obligaciones generales de transparencia e información, invitando al resto a someterse a futuros códigos de conducta de adscripción voluntaria.

El objetivo pretendido con esta nueva propuesta no es realmente establecer un marco plenamente uniforme y armonizado en UE para la inteligencia artificial como realidad compleja, sino exclusivamente respecto de los sistemas inteligentes considerados de riesgo inaceptable o alto riesgo conforme a la misma, prohibiendo relativamente los primeros, salvo en los casos excepcionados expresamente previstos, y otros considerados de alto riesgo, sometidos a un amplio conjunto de requisitos y obligaciones.

Esta iniciativa legislativa, a falta de su tramitación y aprobación, podría constituirse en una referencia internacional para los distintos países que tienen en su agenda legislativa el establecimiento de un marco jurídico de inteligencia artificial, si bien, auguro un largo trabajo que exige reflexión sobre qué futuro realmente deseamos en el ámbito de la inteligencia artificial y actuar en consecuencia a nivel legislativo, donde el Parlamento Europeo debe tener especial protagonismo.

Los aspectos más destacables de una primera revisión de este han sido expuestos a lo largo de su análisis, incorporando mis consideraciones y opiniones.

El enfoque de riesgos desde el que se ha construido me parece el adecuado, no tanto el modelo regulador elegido, que no regula la inteligencia artificial de manera horizontal en toda su amplitud, sino que se focaliza en la prohibición de determinados sistemas de riesgo inaceptable y en la regulación detallada de los requerimientos y obligaciones verticales para los sistemas inteligentes de alto riesgo, salvo determinadas obligaciones de transparencia e información para concretos sistemas. Para el resto de sistema apuesta por el sometimiento voluntario a códigos de conducta.

La opción por este modelo comporta la ausencia de principios y normas éticas y jurídicas vinculantes para el resto de los sistemas inteligentes, lo que constituye un riesgo en sí

mismo, como expuse al analizar sus riesgos y retos -ausencia de regulación-, no gestionado en el ámbito de la UE e incongruente con el propio enfoque de riesgos desde el que se ha concebido.

Como he expuesto, los futuros marcos reguladores deberían contemplar una base de principios, normas y obligaciones esenciales que resulten vinculantes para cualquier sistema de inteligencia artificial, cualquiera que sea su nivel de riesgo, sin perjuicio de establecer un marco específico para aquellos sistemas considerados de alto riesgo.

Asimismo, estos marcos no deberían permitir la puesta en funcionamiento y comercialización de sistemas inteligentes cuyo riesgo no haya sido evaluado.

La Propuesta de Reglamento analizada evidencia distintos aspectos que exigirán mayor reflexión y revisión por parte del legislador europeo, algunos expuestos a lo largo de su análisis con mis consideraciones, aunque de manera sucinta, ante el objeto y alcance limitados de esta investigación.

La misma no aborda cuestiones responsabilidad, que precisará su desarrollo en la línea ya iniciada con la Resolución del Parlamento Europeo sobre la misma de 20 de abril de 2020, y no menciona a las personas afectadas por estos sistemas, ni los mecanismos de garantía o tutela de derechos de los mismos.

Constituye un instrumento normativo de alto contenido técnico y de seguridad, especialmente orientado a desarrolladores, fabricantes, proveedores y usuarios de sistemas inteligentes considerados de alto riesgo, pero con aspectos abstractos que precisarán su desarrollo y concreción, cuanto menos, en normas técnicas relacionadas.

La relación de sistemas inteligentes considerados de alto riesgo me parece adecuada, pero es insuficiente, como he expuesto, y muchos de los sistemas incluidos en los mismos se hallan ya regulados en otros marcos reguladores e incluso podrían estar prohibidos por los mismos. Además, la calificación de un sistema inteligente de alto riesgo no puede circunscribirse a los sectores, ámbitos y usos inicialmente identificados u otros que se identifiquen en el futuro dentro de los mismos, sino que la evolución del desarrollo y

aplicación de la inteligencia artificial evidenciará supuestos de alto riesgo no necesariamente pertenecientes a los sectores y ámbitos inicialmente listados.

Desde un enfoque de riesgos, un sistema inteligente deberá clasificarse no sólo en función de un listado objetivo incorporado en una norma o mediante su inclusión en virtud de un acto de ejecución por la Comisión, en el ejercicio de actos delegados, sino que además debería ser calificado, incluido o no formalmente, en atención los riesgos asociados y contexto, considerando las propias características y capacidades de los sistemas como la probabilidad e impacto de aquéllos.

Y del mismo modo, el enfoque regulatorio de la inteligencia artificial debe ser global y desde una perspectiva amplia, y no desde enfoques muy focalizados en los riesgos y retos de privacidad, dado que los riesgos pueden afectar no sólo a personas físicas sino a todo tipo de entidades públicas o privadas o gobiernos, así como no sólo a la salud, la seguridad y los derechos fundamentales, sino a todo tipo de bienes y derechos de distinta naturaleza y clase.

En este sentido, considero que estos marcos deben partir de un enfoque amplio y global, en relación con lo cual, me permito poner como ejemplo instrumentos jurídicos vigentes en el seno de la UE, como el Reglamento (UE) 2019/1020 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativo a la vigilancia del mercado y la conformidad de los productos y por el que se modifican la Directiva 2004/42/CE y los Reglamentos (CE) n.º 765/2008 y (UE) n.º 305/2011[578].

Conforme regula en su artículo 1, el Reglamento citado tiene como objetivo garantizar que solamente se comercialicen en la UE productos conformes que cumplan los requisitos que proporcionan un nivel elevado de protección de intereses públicos, como la salud y la seguridad en general, la salud y la seguridad en el trabajo, la protección de los consumidores, del medio ambiente y la seguridad pública y cualquier otro interés público protegido por dicha legislación.

[578] DOUEL 25.06.2019

La relación, interacción y posibles colisiones entre el Reglamento propuesto y el RGPD parecen claras y sería deseable una mayor armonización entre ambas normas, especialmente para determinar responsabilidades subjetivas en relación con el tratamiento de datos y cumplimiento de las correlativas obligaciones asociadas.

El denominada *Ley de inteligencia artificial* europea es un gran paso hacia a la regulación de la misma orientado a evitar la fragmentación del mercado interior, pero insuficiente a mi juicio, sino se lleva a cabo desde una perspectiva amplia y global que considere aspectos éticos, de seguridad y de responsabilidad, incluyendo a las personas afectadas, que parta de una definición clara de una realidad compleja, dinámica y en constante evolución, y que considere un conjunto de requerimientos y obligaciones esenciales de carácter ético y jurídico a cualquier sistema inteligente, sin perjuicio de las específicas para los sistemas considerados de alto riesgo.

Capítulo V

Responsabilidad civil: Derecho de daños

1. Introducción

De manera previa al análisis de la responsabilidad civil derivada de daños causados por sistemas inteligentes, considero necesario partir de una reflexión previa sobre la definición y elementos que conforman el régimen general de responsabilidad civil para, posteriormente, abordar su posible exigencia a sistemas dotados de inteligencia artificial.

La responsabilidad civil podemos definirla como la obligación de responder por los actos realizados personalmente o por otra persona, con indemnización de los daños y perjuicios producidos a un tercero, sea persona física o jurídica.

La responsabilidad civil comporta principalmente un deber: El de indemnizar los daños y perjuicios causados.

El denominado "ilícito civil" se halla regulado en el artículo 1089 del *Código Civil* español que establece que "Las obligaciones nacen de la Ley, de los contratos y cuasicontratos, y de los actos y omisiones ilícitos o en que intervenga cualquier género de culpa o negligencia", en relación con lo dispuesto en su artículo 1.093, el cual dispone que "Las que se deriven de actos u omisiones en que intervengan culpa o negligencia no penadas por la Ley, quedarán sometidas a las disposiciones del Capítulo II del Título XVI de este Libro", es decir, a los artículos 1902, siguientes y concordantes.

Por su parte, el denominado ilícito penal se halla contemplado en el artículo 1.092 del Código Civil que se remite al Código Penal español, estableciendo que "Las obligaciones civiles que nazcan de los delitos o faltas se regirán por las disposiciones del Código Penal".

Y, por último, la Ley de Enjuiciamiento Criminal española regula en su artículo 100 que, de todo delito o falta nace acción penal para el castigo del culpable, y puede nacer acción civil para la restitución de la cosa, la reparación del daño y la indemnización de los perjuicios causados por el hecho punible, lo que comporta, de un lado, que no se derive necesariamente responsabilidad civil de todos los delitos y, de otro, que todo delito genera obligaciones civiles siempre que se haya producido un daño.

La diferenciación entre ambas regulaciones parece llevarnos a considerar que las consecuencias de ambos ilícitos respectivamente son derivadas de responsabilidades distintas.

La responsabilidad civil es el deber de indemnizar los daños y perjuicios causados a otro por los actos propios o de otra persona, mientras que la responsabilidad penal nace de la calificación y tipificación del acto como delito en el Código Penal, siendo compatible con la responsabilidad civil siempre que se haya causado un daño. En este sentido, la responsabilidad civil derivada de los artículos 109 a 122 del Código Penal español sería la misma que la prevista en los artículos 1902, siguientes y concordantes del Código Civil precitado.

La responsabilidad civil derivada de un hecho delictivo de depurará conjuntamente con la responsabilidad penal en el proceso de esta naturaleza, salvo que la persona perjudicada se reserve el ejercicio de la acción civil.

La responsabilidad derivada del delito será abordada en el capítulo VI de esta investigación, por lo que en los siguientes apartados analizaré la responsabilidad desde una perspectiva estrictamente civil.

La responsabilidad por los daños o perjuicios causados por o mediante un sistema de inteligencia artificial se hallaría inicialmente sometida a este régimen jurídico de responsabilidad civil, sin perjuicio de la responsabilidades de otra naturaleza que pueden derivarse, por ejemplo, de naturaleza penal o incluso administrativa y de responsabilidad patrimonial de las Administraciones públicas, en el caso de sistemas inteligentes utilizados en la prestación de servicios públicos, por ejemplo, en el ámbito de la gestión

de procedimientos administrativos específicos o en la gestión de servicios municipales de *Smart cities*, controles de acceso, tráfico rodado, etc.

Los regímenes clásicos de responsabilidad se construyeron sobre la responsabilidad subjetiva para responder por actos propios y culpables, incluso cuando se debía responder por actuaciones ajenas por culpa *in vigilando, in eligendo* o *in educando*.

La evolución del contexto económico y social motivó una paralela objetivación de la responsabilidad, primero con la presunción de culpabilidad con inversión de carga de la prueba, para posteriormente llegar a una responsabilidad basada exclusivamente en la existencia de un daño al margen de la conducta del sujeto responsable, básicamente ante la dificultad en la investigación y prueba de la culpa, y la rapidez y seguridad que cada vez más exige la sociedad. Y de ahí hasta la conformación y aplicación de la *teoría del riesgo* para construir, en determinados contextos, una responsabilidad basada en el riesgo, de modo que quien pone en el tráfico algo con un riesgo potencial debe soportar las consecuencias dañosas o perjudiciales de su funcionamiento, concurra o no culpa por su parte, máxime si además es el que se beneficia de ello.

En base a esta construcción de la responsabilidad, especialmente la extracontractual, la cuestión es si la inteligencia artificial, su comercialización, puesta en servicio, aplicación y uso constituye una actividad susceptible de clasificación como peligrosa o de riesgo potencial, además de que éste pueda ser especialmente significativo o alto. El Parlamento Europeo considera su funcionamiento y uso una actividad de riesgo.

A estas alturas, considero indudable que potencialmente lo es o puede serlo en función de la naturaleza, capacidades, características y contexto de aplicación y uso del sistema inteligente, y no tanto por los elementos que la integran en sí mismos -*hardware, software* o datos-, que también pueden entrañar altos riesgos potenciales como, por ejemplo, un robot dotado de armamento, una máquina o un coche "autónomo", sino por el contexto y actividad que se lleve a cabo mediante la misma.

Los riesgos potenciales pueden generarse en el marco contractual, desde la contratación, el diseño, el desarrollo, la elaboración del algoritmo y la conversión al lenguaje-máquina, el suministro del sistema, su customización con datos propios confidenciales o secretos,

la introducción de datos, su puesta en servicio, su funcionamiento, su entrenamiento, su aplicación, su uso, su monitorización, su mantenimiento o su ofrecimiento como servicio, así como en el marco extracontractual frente a terceros.

Todo ello anticipa de antemano la complejidad y dificultad consecuente de atribuir la responsabilidad y probar el hecho generador del daño y su nexo causal en función del momento y contexto, máxime ante la pluralidad de sujetos intervinientes durante todo su ciclo de vida.

Siguiendo la doctrina italiana y otros autores citados por Carrascosa[579], como Frazoni y Darío Vergel, tradicionalmente se ha distinguido entre la peligrosidad de la conducta que se incluye en la categoría dominada por la culpa, mientras que la actividad peligrosa se inserta en otra determinada por el elemento objetivo de la actividad misma que es peligrosa en sí o por los medios que adopta.

El concepto de "actividad peligrosa" es un concepto necesariamente relativo y evolutivo que dependerá del estado de la técnica y del avance científico.

A mi juicio, considero incuestionable que determinados sistemas de inteligencia artificial - especialmente por sus características y capacidades, el sector donde operen y la actividad y el uso asociados a los mismos-, deben considerarse como potencialmente peligrosos y de riesgo, y este es el contexto en el que debe construirse la responsabilidad por riesgo de la inteligencia artificial. Y ello sin perjuicio de los demás riesgos asociados a cualquier sistema que pueden potenciarse por la criticidad de sus usos y dependencia. En este sentido, los sistemas de inteligencia artificial tendrán errores y fallos como cualquier otro sistema, del mismo modo que cualquier sistema de ciberseguridad también los tiene y no es absolutamente infranqueable.

Los sistemas de inteligencia artificial comportan riesgos y también incertidumbres inherentes asociadas a su naturaleza, características y capacidades, por lo que sus riesgos,

[579] CARRASCOSA LÓPEZ, V. (2000). *La contratación informática: El nuevo horizonte contractual*. Editorial Comares. 2000. Pp. 283-284. FRAZONI, M. (1988). *Culpa presenta e responsabilitá del debitore*. Cedán, Padova. 1988; y DARÍO VERGEL, S. (1994). Responsabilidad civil derivada de la informática. *Revista Informática y Derecho*. Vol. 4. Aranzadi-UNED. Mérida. 1994.

la incertidumbre y el error debe ser esperado, y además no sólo propios de la misma, sino de las demás tecnologías con las que interactúa y se interrelaciona, incluida la propia conectividad.

De hecho, la Propuesta de Reglamento del Parlamento Europeo y del Consejo, relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial, que acompaña a la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial[580], recoge en su Considerando 2, que el desarrollo de nuevas tecnologías (en este caso, no tan nuevas en cuanto a su existencia, pero si en cuanto a su desarrollo y aplicación), comporta inevitablemente la asunción de riesgos para el propio usuario[581].

El avance de las economías modernas y, supuestamente, el mantenimiento del estado de bienestar, parecen motivar un desarrollo y producción imparable que conlleva riesgos y daños colaterales como, por ejemplo, al medio ambiente, de modo que parece aceptarse el daño inherente y colateral, considerando en muchos supuestos la opción de indemnización en lugar de paralizar aquéllos.

Esta cuestión genera un interesantísimo debate sobre el futuro de la humanidad y del mundo, tal y como lo conocemos, que resulta inabordable tratar en esta investigación, de objeto y alcance específicos.

En paralelo, se ha producido la denominada socialización de la responsabilidad civil, de modo que, por ejemplo, mediante el traslado de los riesgos a un tercero (seguros), los daños sufridos por unos pocos son sufragados por la comunidad asegurada.

[580] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))
[581] "(2) En particular al principio del ciclo de vida de los nuevos productos y servicios, tras haber sido probados previamente, existe cierto nivel de riesgo para el usuario, así como para terceras personas, de que algo no funcione correctamente. Este proceso de prueba y error también es, al mismo tiempo, un factor clave para el progreso técnico, sin el cual la mayor parte de nuestras tecnologías no existirían. Hasta ahora, los riesgos asociados a nuevos productos y servicios se han visto atenuados adecuadamente por la sólida legislación en materia de seguridad de los productos y las normas sobre responsabilidad civil".

Y en relación con todo ello, considero necesario valorar un enfoque híbrido de la responsabilidad derivada de la inteligencia artificial que combine la responsabilidad individual con la colectiva, previa identificación de los grupos de riesgo, de modo que, sin excluir la responsabilidad individual, se aprovechen los beneficios de una responsabilidad colectiva en materia de seguridad y de aseguramiento de un resarcimiento efectivo a las personas afectadas.

En este contexto es en el que considero debemos situar la responsabilidad civil de la inteligencia artificial y la construcción de una responsabilidad por riesgos, pero con todas las cautelas y reservas en aras de mantener el equilibrio entre todos los derechos, bienes e intereses en juego.

Esta fue la línea que siguió el Parlamento Europeo en su Resolución de 20 de octubre de 2020 sobre un régimen de responsabilidad civil en materia de inteligencia artificial, anteriormente citada. No obstante, para su sujeción a un marco de responsabilidad enfocada en el riesgo, debe considerarse su tipología, características, capacidades, sector donde opere, contexto y usos.

La inteligencia artificial plantea importantes retos para el Derecho, conforme analicé en el capítulo II de esta investigación, y, en especial, en materia de responsabilidad, especialmente en relación con algunas de sus características y capacidades, como la autonomía, el autoaprendizaje, la impredecibilidad asociada, la complejidad y la opacidad.

La autonomía permitiría a los sistemas inteligentes actuar "al margen del ser humano", si bien, conforme a los principios y normas éticas más elementales y esenciales, en particular, el control y la supervisión humana continua durante todo su ciclo de vida que, a mi juicio, debería ser jurídicamente vinculante, imperativa e irrenunciable para cualquier sistema inteligente, cualquiera que sea el riesgo inicial o sector donde opere, dicha autonomía nunca debería ser plena, sino relativa, restringida y condicionada.

Las capacidades de autoaprendizaje de los sistemas inteligentes aumentan la autonomía y dificultan el control y la supervisión humana.

Estas capacidades y la complejidad y sofisticación de los sistemas inteligentes comportan enormes retos para la responsabilidad, la cual se ve aumentada con la interacción o integración de la inteligencia artificial con otras tecnologías y sistemas, y su conectividad.

La opacidad, conforme expuse en el capítulo II al analizar sus riesgos y retos, genera la dificultad de comprender su funcionamiento y sus decisiones.

Y todo ello en relación con los múltiples contextos de aplicación y uso, interacción con su entorno y otros sistemas, así como la variedad de agentes que intervienen en su ciclo de vida.

La cuestión esencial a abordar es si los marcos actuales y, en especial, el denominado "derecho de daños" vigente, contempla y puede dar respuestas adecuadas a los daños causados por los sistemas de inteligencia artificial y para el resarcimiento de las personas afectadas o si, por el contrario, el marco actual debería ser revisado y, en su caso, adaptado y complementado, o ser profundamente reformado para afrontar los enormes retos que plantea la inteligencia artificial en esta materia, incluso si debería motivar un nuevo paradigma en materia de responsabilidad civil extracontractual, al igual que en otras materias que también son objeto de reflexión en esta investigación, como la protección, titularidad y contenido de los derechos de propiedad intelectual o industrial respecto de creaciones e invenciones llevadas a cabo por sistemas inteligentes con o sin intervención humana.

El derecho de daños integra el conjunto de reglas dentro de un ordenamiento jurídico que regulan el ejercicio de pretensiones orientadas a la reparación de un daño o a su indemnización, que incluyen las normas generales de responsabilidad civil extracontractual como normas especiales aplicables a sectores o actividades particulares.

Y no opera de manera aislada, sino en combinación con otros instrumentos del ordenamiento jurídico tales como el contrato y las responsabilidades derivadas de su incumplimiento, la regulación administrativa respecto de aspectos como la seguridad de los productos, la regulación procesal, los seguros o los instrumentos penales.

El derecho de daños de los Estados miembros de la UE no está en gran medida armonizado, como destaca el informe *Liability for Artificial Intelligence and other emerging digital technologies*[582] de 2019, con la excepción de la legislación sobre responsabilidad por productos defectuosos conforme a la Directiva 85/374/CE[583], algunos aspectos de la responsabilidad por la infracción de los marcos sobre protección de datos -artículo 82 del Reglamento General de Protección de Datos[584]- y la responsabilidad por infringir la legislación sobre competencia conforme a la Directiva 2014/104/UE[585]. Adicionalmente existe un régimen consolidado que regula los seguros de responsabilidad en relación con los daños causados por el uso de vehículos de motor[586].

Los ordenamientos jurídicos de los distintos Estados miembros no contemplan marcos de responsabilidad específicamente aplicables al funcionamiento y uso de la inteligencia artificial.

Conforme significa Rubí[587], las posiciones iniciales en la última década partían de la autonomía de la que podrían estar dotados los sistemas inteligentes y la dificultad de que concurra su previsibilidad, por lo que se alejaban de una responsabilidad por culpa para proponer la aplicación generalizada de una responsabilidad objetiva por los daños, incluso

[582] *Liability for Artificial Intelligence and other emerging digital technologies.* UE. 2019. Recuperado de: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf. Consultado el 14.12.2020.

[583] Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados miembros en materia de responsabilidad por los daños causados por productos defectuosos (DO L 210 de 7.8.1985, p. 29), modificada por la Directiva 1999/34/CE del Parlamento Europeo y del Consejo, de 10 de mayo de 1999, DO L 141 20 de 4.6.1999.

[584] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos, y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos), DO L 119 de 4.5.2016.

[585] Directiva 2014/104/UE del Parlamento Europeo y del Consejo, de 26 de noviembre de 2014, relativa a determinadas normas por las que se rigen las acciones por daños y perjuicios en virtud del Derecho nacional por infracciones de las disposiciones del Derecho de la competencia de los Estados miembros y de la Unión Europea, DO L 349 de 5.12.2014.

[586] Directiva 2009/103/CE del Parlamento Europeo y del Consejo, de 16 de septiembre de 2009, relativa al seguro de la responsabilidad civil que resulta de la circulación de los vehículos automóviles, así como al control de la obligación de asegurar esta responsabilidad, DO L 263 de 7.10.2009. Pp. 11-31. La Directiva está siendo revisada actualmente, véase la propuesta COM (2018) 336 final.

[587] RUBÍ, A. (2020). Retos de la inteligencia artificial y adaptabilidad del derecho de daños, en CERRILLO I MARTÍNEZ, A. Y PEQUERA POCH, M. (Coord). *Retos jurídicos de la inteligencia artificial*. Aranzadi S.A.U. (Thomson Reuters). Navarra 2020. P. 58.

llegando más allá al proponer una responsabilidad absoluta o barajar incluso la posibilidad la atribución de personalidad jurídica a los sistemas de inteligencia artificial.

Las posiciones actuales mayoritarias que abordaré en su análisis, parecen situarse en la utilidad de los marcos vigentes de responsabilidad extracontractual para su aplicación a los daños causados por sistemas inteligentes, si bien, a mi juicio y anticipando ya algunas de mis consideraciones sobre esta cuestión, con el necesario acompañamiento de una revisión global integradora que complemente e integre adecuadamente los mismos, especialmente a nivel europeo, al objeto de abordar estas cuestiones de una manera armonizada y evitar la fragmentación del mercado único en esta materia.

Según el autor precitado, aunque la inteligencia artificial plantea retos novedosos, el derecho de daños actual dispone de herramientas suficientes para ofrecer respuestas razonables a las situaciones que se puedan plantear, no obstante, en mi opinión, considero que, a pesar de disponer de herramientas y poder ofrecer soluciones en base a las mismas, pueden no ser adecuadas ni unas ni otras en los múltiples contextos que los sistemas de inteligencia artificial pueden plantear, tanto por su tipología, elementos y capacidades, como por el sector donde opere, uso, aplicación y multiplicidad de sujetos participantes durante su ciclo de vida, máxime cuando su nivel de riesgo puede ser cambiante en atención a dichas capacidades y resultar muy diferente durante su funcionamiento al previamente identificado en el momento de su concepción, especialmente con origen en sus funciones autoaprendizaje, entrenamiento, usos y grado de autonomía e impredecibilidad.

Durante el análisis de todas estas cuestiones, abordaré las instituciones clásicas de la responsabilidad civil en nuestro ordenamiento jurídico para valorar su posible aplicación para la depuración de las responsabilidades derivadas de los daños causados por sistemas inteligentes. Asimismo, analizaré la última propuesta reguladora del Parlamento Europeo sobre responsabilidad civil en materia de inteligencia artificial y finalizaré el capítulo con mis consideraciones y opinión personal a modo de reflexión.

Siguiendo un enfoque clásico de la responsabilidad civil, de inicio, diferenciaré entre responsabilidad contractual y extracontractual, si bien, ante el objeto específico de esta investigación, focalizaré especialmente mi análisis en la segunda, relacionada con los

daños que puedan ocasionarse por los sistemas de inteligencia artificial derivados de su funcionamiento, aplicación o uso, sin perjuicio de las consideraciones generales que efectuaré previamente en materia de responsabilidad contractual, de especial interés cuando se produzcan daños derivados del incumplimiento de contratos que puedan celebrarse en relación o con la participación de sistemas de inteligencia artificial, bien en relación con la adquisición o cesión de productos de inteligencia artificial o dotados de la misma, o el suministro de servicios gobernados por la misma, como por ejemplo contratos asistenciales, bancarios, asesoramiento financiero o de compraventa de valores bursátiles mediante este tipo de sistemas, donde ya existe una casuística significativa, variada y de alto impacto[588].

Del mismo modo abordaré la responsabilidad por productos defectuosos en la UE y en España ante los matices diferenciadores y su relevancia en relación con la inteligencia artificial.

2. La responsabilidad civil contractual

2.1. Cuestiones generales

La responsabilidad contractual se haya principalmente regulada en el ordenamiento jurídico español en su Código Civil, en particular, en los artículos 1101, siguientes y concordantes, así como en el Texto Refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias[589] (en lo sucesivo TRLGDCU)

[588] El 6 de mayo de 2010 el mercado de valores estadounidense comenzó la jornada en rojo debido principalmente a la profunda preocupación por la crisis de deuda griega. Sin embargo, a las 14:42h se desplomó repentinamente con extrema rapidez y fuerza ("*flash crash*"), desplomándose más de 600 puntos en cinco minutos y llegando a marcar una pérdida de 1000 puntos en el día. Los índices de acciones y futuros estadounidenses cayeron un 10% en esos minutos con algunas acciones de primera clase cotizando brevemente a un centavo. Veinte minutos después se había recuperado. La posterior investigación emprendida por la SEC determinó que "*fueron los algoritmos de inversión automatizada, y su forma de operar, los que originaron el caos de aquel día*" al generar una inoportuna gran orden de venta. The Economist. 01.10.2010. Recuperado de https://www.economist.com/newsbook/2010/10/01/one-big-bad-trade. Consultado el 15.02.2021.

[589] Real Decreto Legislativo 1/2007, de 16 de noviembre, por el que se aprueba el texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias. BOE n. 287, de 30.11.2007.

española, sin perjuicio, obviamente, de lo previsto en los propios contratos dentro del ámbito dispositivo de las partes.

La responsabilidad contractual deriva del incumplimiento o cumplimiento defectuoso o lesivo de un contrato por una de las partes, consistente en la falta de diligencia o previsión y que genera la obligación de indemnizar al perjudicado.

El artículo 1101 precitado establece que quedan sujetos a indemnización de los daños y perjuicios causados los que, en cumplimiento de sus obligaciones, incurrieren en dolo, negligencia o morosidad, y los que de cualquier modo contravinieren al tenor de aquélla.

Este precepto se halla directamente relacionado con los artículos 1256 y 1258 del mismo cuerpo legal.

El primero establece que la validez y cumplimiento de los contratos no pueden dejarse al arbitrio de uno de los contratantes, mientras que el segundo establece que los contratos se perfeccionan por el mero consentimiento y, desde entonces, obligan no sólo al cumplimiento de lo expresamente pactado, sino también a todas las consecuencias que, según su naturaleza, sean conformes a la buena fe, al uso y a la Ley. Todos estos preceptos constituyen el núcleo regulador de la responsabilidad contractual.

En definitiva, la responsabilidad contractual es aquella que resulta del incumplimiento de una obligación nacida de un contrato, a diferencia de la responsabilidad civil extracontractual que, en su concepción más clásica y como indica Castán Tobeñas[590], se caracteriza por que no existe ningún vínculo obligatorio o relación jurídica preexistente entre el autor del daño y la víctima del mismo.

La responsabilidad contractual se deriva de las obligaciones contraídas por las partes en el contrato como fuente de las mismas.

En consecuencia, las partes deben cumplir con sus respectivas obligaciones contractuales dado que, en caso de incumplimiento de una obligación contractual preexistente por

[590] CASTÁN TOBEÑAS, J. (1988). *Derecho civil español, común y foral*. 14a. Ed., Editorial Reus, Madrid, 1988. T.1V. P. 939.

cualquiera de ellas, ya sea por dolo, culpa o retraso, ésta, la parte que incumplió, quedará obligada a la indemnización de los daños y perjuicios causados.

Y esta nueva obligación de indemnizar derivada del contrato puede nacer del mero incumplimiento culposo del contrato por una de las partes, siempre y cuando no se derive de "caso fortuito" o "fuerza mayor", por lo que no es preciso que exista dolo, engaño o mala fe en el cumplimiento del contrato.

El sistema actual de responsabilidad contractual en España establece que la reparación indemnizatoria de un daño contractual, además de una imputabilidad objetiva de dicho daño a un incumplimiento, exige una imputabilidad subjetiva, es decir, "un juicio de reprochabilidad de la conducta del contratante incumplidor"[591].

La responsabilidad contractual en caso de dolo o mala fe en el cumplimiento de las obligaciones y su justificación es obvia, recibiendo un supuesto trato diferenciado en el artículo 1107 del Código Civil español, que establece una responsabilidad por todos los daños y perjuicios derivados de la falta de cumplimiento de una obligación, a diferencia de la culpa, en la que la responsabilidad se circunscribirá a los daños y perjuicios que sean consecuencia necesaria de su incumplimiento culposo.

El análisis del concepto de dolo en el ámbito de las obligaciones y contratos sería un tema apasionante, pero inabordable en el marco del objeto y alcance específicos de esta investigación, si bien, se ha venido identificando como un aspecto que no ha sido profundizado debidamente en nuestra doctrina y jurisprudencia, por parte de la doctrina más autorizada, como Montes Penadés[592].

Según este autor, partiendo del indicado precepto, el dolo en el cumplimiento es un criterio de imputación que ofrece la característica de hallarse formulado en términos de tal generalidad que lo hacen inaplicable.

---

[591] RODRÍGUEZ-ROSADO, B. (2020). Los sistemas de responsabilidad contractual: Entre la responsabilidad por culpa y la strict liability. *Revista de Derecho Civil*. Vol. I, núm. 4 (Octubre-diciembre, 2014). Estudios. Pp. 155-187. Disponible en: https://www.nreg.es/ojs/index.php/RDC/article/view/103. Consultado el 20.12.2020.

[592] MONTÉS PENADÉS, V.L. (2007). La responsabilidad por dolo en Moreno, J.A. (Coord), *La responsabilidad civil y su problemática actual*. Dykinson 2007. Pp. 719-746.

La gran mayoría de los incumplimientos serían dolosos para la mayor parte de la doctrina, pero la casuística jurisprudencial pone en evidencia que sólo se aplica el régimen especial previsto para los supuestos de dolo en contadas ocasiones. Y siguiendo a este mismo autor precitado, el precepto indicado ha sido criticado severamente por la doctrina e incluso, como cita el mismo, calificado como "enigmático" por Montes, "inaplicable" por Carrasco o “barbaridad” por Pantaleón al considerar que en su tenor literal conduce al absurdo.

Una buena parte de la doctrina se ha ocupado de interpretar y explicar el precitado artículo, unos considerando que parece referirse a un deudor de buena fe culposo, otros considerando la existencia de un territorio “de no buena fe”, no doloso, entre la buena fe y el dolo, y otros, asimilando dolo y mala fe en su tenor literal.

Siguiendo al precitado autor, a los efectos de interpretación de los preceptos indicados, se debe distinguir entre fundamentos subjetivos de la responsabilidad, que serían el dolo y la culpa, y las modalidades objetivas de infracción obligacional, que serían la mora y la "contravención del tenor de la obligación". De este modo los fundamentos subjetivos -dolo o culpa- serían la causa de la infracción obligacional -mora o contravención- y esta, a su vez, sería la causa de los daños.

Conforme a lo establecido en el precitado artículo, el deudor culposo respondería de los daños y perjuicios previstos o que hayan podido prever al tiempo de constituirse la obligación, y sean consecuencia necesaria de su falta de cumplimiento, mientras que el doloso respondería de la totalidad de daños y perjuicios, previstos o no, que se deriven de la falta de cumplimiento de la obligación[593]. Ello comporta un aspecto de indudable importancia para la reparación efectiva del perjudicado.

[593] Respecto a su extensión a la responsabilidad extracontractual, destacar la Sentencia del TS nº 942//942, de 24/11/1995 (Rec. 3105/1994) establece al respecto que: "El art.1107 del Código Civil que establece los límites de la responsabilidad del deudor de buena fe y del deudor por dolo por la falta de cumplimiento de sus obligaciones, entendiéndose por deudor de buena fe al que incumple por culpa, por contraposición al que lo hace con o por dolo, es aplicable, no obstante su ubicación en el Código , también a las obligaciones nacidas de culpa extracontractual y así lo viene señalando la doctrina científica mayoritaria y lo ha establecido esta Sala que en su sentencia de 20 de junio de 1989 afirma que " de acuerdo con el tenor literal del susodicho art.1103, y al margen de la localización sistemática del mismo, éste es un precepto aplicable a toda clase de obligaciones, como en realidad lo son también otros artículos inmediatos del mismo Código (así los arts. 1104, 1106 y 1107) que habitualmente se aplican y proyectan tanto al campo de obligaciones dimanantes de convención o contrato como en el de los que nacen de acto ilícito"; no hay, por ello, obstáculo

No obstante, en congruencia con el objeto y alcance limitados de esta investigación, a continuación, me limitaré a profundizar algo más en mis reflexiones alrededor de la responsabilidad contractual basada en la culpa, de manera previa a abordar la responsabilidad extracontractual.

La culpa se deriva pues de la falta de diligencia y previsión que se presume a la parte que incumple una obligación contractual preexistente y genera la obligación de indemnizar. En este sentido, considero necesario analizar con más detalle esa "falta de diligencia" en relación con el entorno objeto de esta investigación.

Los elementos de la responsabilidad civil contractual basada en la culpa son: a) La existencia de una relación jurídica o contrato entre las partes; b) El incumplimiento total o parcial de alguna de sus obligaciones; c) El incumplimiento derivado de una falta de diligencia o falta de precisión del obligado a ello; d) La relación causa-efecto entre el hecho y el resultado y; e) Generación de un daño o perjuicio reparable y cuantificable.

Los daños comprenden no solo los daños materiales o económicos, en su doble vertiente de daño emergente y lucro cesante, sino también daños morales. Y su indemnización, conforme a lo dispuesto en el artículo 1106 del Código Civil español, incluye no sólo el valor de la pérdida sufrida, sino también el de la ganancia que haya dejado de obtener la parte acreedora de la obligación.

El "incumplimiento culposo" y "relación de causalidad" cobran especial relevancia en el contexto de daños derivados del funcionamiento y uso de sistemas inteligentes, por lo que partiré de los conceptos generales para llegar a mis reflexiones posteriores en relación con las cuestiones analizadas.

El artículo 1104 del Código Civil español establece que la culpa o negligencia del deudor consiste en la omisión de aquella diligencia que exija la naturaleza de la obligación y corresponda a las circunstancias de las personas, del tiempo y del lugar, lo que en el ámbito de la contratación de la adquisición, puesta en funcionamiento y/o uso de sistemas

alguno a que para fijar la cuantía de la indemnización se acuda en las obligaciones extracontractuales al grado de conducta dolosa o culposa observada por el responsable, al igual que se hace en las obligaciones contractuales en aplicación del art.1107 citado."

inteligentes es esencial considerar este contexto, especialmente para mantener el equilibrio de las partes, especialmente cuando no intervengan en el proceso contractual usuarios contratantes cualificados.

Conforme al marco legal vigente en España regulador de la responsabilidad contractual, el incumplimiento "culposo" supone una actuación carente o no ajustada a la diligencia exigible en cada caso concreto en atención a las circunstancias de las personas, del tiempo y del lugar y, en caso de no concreción expresa en el contrato, esta diligencia mínima sería la que correspondería a "un buen padre de familia", conforme lo dispuesto en el artículo precitado.

De manera consecuente, debo abordar dos preguntas esenciales: ¿Cuál es la "diligencia de un padre de familia"? Y, de manera consecuente a dicho concepto y regulación legal ¿Cuál es la diligencia contractual profesional para diseñadores, desarrolladores, fabricantes, proveedores de datos, entrenadores, proveedores de sistemas, operadores o usuarios de sistemas de inteligencia artificial, cuando es con ellos con quién se formaliza el contrato?

La diligencia de un buen padre de familia prevista en el Código Civil español no es la misma que el estándar de diligencia exigible a un profesional o empresario de carácter cualificado, es decir, debemos diferenciar aquella de la diligencia profesional o empresarial, sometida a requerimientos más estrictos.

El propio artículo 1104 del Código Civil refleja la necesidad de distinguir entre ambas, esto es, del deudor normal de la del deudor profesional o empresario, así como entre dos tipos de culpa asociada de carácter normal y la otra de carácter profesional o empresarial, como significa Ramos Herranz[594].

El Tribunal Supremo ha delimitado ambos modelos de diligencia en múltiples pronunciamientos[595], estableciendo que el nivel de exigible a un profesional debe ser siempre superior al de un deudor normal, habiendo de adecuarse a su especialidad, aunque

[594] RAMOS HERRANZ, I. (2006). "El estándar mercantil de diligencia: El ordenado empresario". *Anuario de derecho civil*. ISSN 0210-301X. Vol. 59. Nº 1. 2006. P. 201.
[595] Sentencias del TS de 22 de noviembre de 1971 (Rec. 4974), 29 de abril de 1988 (Rec. 3302), 2 de febrero de 1989 (Rec. 657), 13 de julio de 1987 (Rec. 5488) o 23 de marzo de 1993 (Rec. 2545), entre otras.

no siempre lo ha hecho con la claridad y coherencia deseables, como igualmente significa la autora precitada. Del mismo modo, la doctrina del Tribunal Supremo español ha recogido en algunos de sus pronunciamientos la necesidad de considerar en la apreciación del apartamiento de esa diligencia debida, la naturaleza de la obligación y las circunstancias de las personas, del tiempo y del lugar junto con el sector de actividad o de la vida social en la que el deudor opere o actúe[596].

El estándar civil o común de diligencia exigible es la propia de "un buen padre de familia" con origen en el *paterfamilias* del Derecho Romano, que es la considerada propia del hombre medio. Este modelo se halla recogido tanto en el ámbito del *common law* anglosajón como del *civil law* propio de la tradición jurídica española, si bien, concebida en aquel como la diligencia propia del hombre razonable. En cualquier caso, tanto esta diligencia como la del padre de familia hacen referencia a un mismo concepto, esto es, la diligencia del hombre medio que presta la diligencia propia de un hombre prudente[597].

La diligencia de un buen padre de familia difiere pues de la diligencia profesional y empresarial propia de un "ordenado empresario" dentro del sector de actividad donde opere, con mayores exigencias en base a una cualificación y conocimientos técnicos que no pueden exigirse al padre de familia para el desarrollo de sus actividades no profesionales. En definitiva, los parámetros de esa diligencia son distintos, y dentro de la profesional o empresarial, diferentes para cada sector de actividad.

En consecuencia, en el marco de la inteligencia artificial y los distintos actores intervinientes durante su ciclo de vida, su actividad debe respetar los parámetros de la diligencia profesional y empresarial debida.

En ausencia de marcos reguladores ¿esta diligencia podría estar integrada por el *soft law* o códigos éticos no vinculantes? ¿Cuál sería la diligencia en caso de que una de las partes contratantes pudiese llegar a ser en el futuro un sistema de inteligencia artificial (en caso de reconocérsele en el futuro personalidad jurídica y capacidad de obrar)?

---

[596] Sentencias TS de diciembre de 1971 (Rec. 5232), 20 de mayo de 1993 (Rec. 3718) y 5 de octubre de 1994 (Rec. 7453).
[597] RAMOS HERRANZ, I. (2006). "El estándar mercantil de diligencia: El ordenado empresario". *Anuario de derecho civil*. Op.cit. P. 199

Dicha diligencia, a mi juicio, deberá integrarse no sólo por los conocimientos técnicos presupuestos, sino también por los marcos normativos vigentes, generales y especialmente sectoriales -por ejemplo, normas de seguridad de productos- que puedan afectar al objeto del contrato, pero también por los marcos éticos específicos del sector donde operen las partes -sectoriales-, incluso también los generales y referidos a sistemas de inteligencia artificial -todavía no plenamente consensuados a nivel mundial y no vinculantes legalmente, como he comentado en el capítulo III de esta investigación-, así como los códigos de conducta o de buenas prácticas a los que se halle adherida la parte incumplidora.

Como he anticipado, el artículo 1104 del Código Civil español establece que, cuando el contrato no exprese la diligencia exigible que ha de prestarse en su cumplimiento, se debe exigir la que corresponde a un "buen padre de familia", entendida como la "tipo medio de persona diligente", lo que variará en cada caso, y en contextos comerciales, la profesional o empresarial exigible.

Para Santos Briz, esa "diligencia general debida" responde a un criterio objetivo o abstracto y consiste en la diligencia que dentro de la vida social puede ser exigida en la situación concreta a persona razonable y sensata correspondiente al sector del tráfico, cualificada por la clase de la actividad a enjuiciar. Es decir, la diligencia que puede ser requerida en cada contexto a cada persona razonable y sensata, conforme al sector del tráfico donde opere[598].

Por otra parte, la previsibilidad junto con la evitabilidad son características de la culpa, lo que me lleva a nuevas preguntas que serán abordadas con posterioridad: ¿Podría imputarse la responsabilidad por culpa a un desarrollador, fabricante u operador en caso de sistemas de inteligencia artificial avanzada o fuerte con supuesta autonomía, capacidad de autoaprendizaje e hipotética impredecibilidad en sus decisiones y acciones?

Esa preceptiva medición o graduación de la culpa y más en un contexto como el que integra el objeto de esta investigación, exige la aplicación de las teorías más modernas

[598] SANTOS BRIZ, J. (1993). *La responsabilidad civil, Derecho sustantivo y derecho procesal.* 7ª Edición. Editorial Montecorvo, S.A. Madrid 1993. Pp. 97 y ss.

del arbitrio judicial ante la ausencia de un marco jurídico especial, dejando la apreciación de la culpa y extensión de la responsabilidad a la "sana crítica" de los Tribunales, fundamentada en la exigencia de que se atienda a las circunstancias especiales de caso y resolver en equidad. Por ello, mi posicionamiento sobre la inseguridad jurídica que comporta este mecanismo y la necesidad de nuevos marcos reguladores especiales.

De este modo, son los Tribunales los que caso a caso deberán resolver cuál es la diligencia exigible y la culpa consecuente de la que deba responder la persona o personas responsables -físicas o jurídicas-, lo que comporta inevitablemente una inseguridad jurídica inicial.

La determinación de la diligencia exigible será variable y dependerá de la naturaleza de la obligación y corresponderse al contexto y circunstancias de las personas, del tiempo y del lugar[599].

Este escenario comporta pues una relativa inseguridad jurídica para todas las partes intervinientes en un contrato, especialmente ante tecnologías y sistemas complejos, de difícil comprensión y conocimiento y de reciente integración en nuestra vida diaria, máxime cuando la determinación de la diligencia exigible y la culpa consecuente de la que deba responder el responsable se determinaría *a posteriori* y en el caso de iniciarse un procedimiento judicial.

De hecho, los informes elaborados a instancias del Parlamento Europeo, especialmente en relación con la aplicación práctica y resoluciones judiciales relacionadas con la responsabilidad por productos defectuosos relacionados con la tecnología, evidencian que la experiencia judicial es todavía muy escasa y la inseguridad consecuente, grande, como se abordará en los posteriores apartados.

[599] La jurisprudencia es muy abundante: Sentencia Nº 32/2020, Audiencia Provincial de Barcelona, Sección 11, Rec. 980/2018 de 06 de Marzo de 2020; Auto Tribunal Supremo, Sala de lo Civil, Sección 1, Rec. 4579/2017 de 27 de Mayo de 2020; Sentencia Nº 21/2006, Audiencia Provincial de Cádiz, Sección 5, Rec. 186/2005 de 03 de Febrero de 2006; Sentencia Nº 659/2010, Audiencia Provincial de Valencia, Sección 6, Rec. 724/2010 de 26 de Noviembre de 2010 o Sentencia Tribunal Supremo, Sala de lo Civil, Sección 1, Rec. 1150/1989 de 07 de Enero de 1992, entre otras.

De ahí, la necesidad de marcos reguladores que contribuyan a complementar los actuales en materia de responsabilidad para definir el marco de responsabilidad asociada la inteligencia artificial, sin perjuicio de que acometan, igualmente, la conversión en vinculantes de determinados principios y normas éticas esenciales asociadas a la misma y que regulen la inteligencia artificial y sus riesgos de manera proactiva y preventiva, de modo que se reduzca la necesidad de acudir a marcos reactivos aplicables en supuestos que podrían haber sido ya previstos y evitados en aquellos, mediante el establecimiento de requerimientos y obligaciones claras, con el correlativo marco de consecuencias derivadas de su incumplimiento.

En definitiva, el marco actual comporta incertidumbre e inseguridad jurídica al mercado, por lo que sería deseable una concreción de dicha diligencia exigible y requisitos preceptivos *ab initio*, dejando claras las reglas de juego para todos los *players*, desde la concepción, diseño, comercialización, adquisición, puesta en funcionamiento, entrenamiento, formación, aplicación, uso, mantenimiento y monitorización de sistemas inteligentes.

Y todo ello sin perjuicio de la posibilidad de futuras evaluaciones previas por parte de terceros, como autoridades públicas o laboratorios, que puedan certificar y garantizar *a priori* el nivel de riesgo de los sistemas inteligentes que se introduzcan en el mercado (al menos de los de mayor riesgo), el funcionamiento adecuado de los mismos y el cumplimiento de determinados marcos, y que incluso podrían intervenir posteriormente durante su funcionamiento con objetivo similar.

En cualquier caso, prosiguiendo con mi análisis, la responsabilidad en el ámbito contractual requiere la existencia de una relación de causalidad entre el hecho y el resultado sobre la que se sustente la imputabilidad del deudor y su obligación de reparar el daño.

El ordenamiento jurídico español prevé en materia de responsabilidad contractual una presunción general de que la parte deudora que no cumple una obligación contractual, lo hace porque quiere y es responsable de la falta de cumplimiento, sin necesidad de que la parte acreedora tenga que probar nada más que la existencia de la obligación. En consecuencia, el causante del daño o "deudor" es quién deberá probar su actuación

diligente y que si incumplió no fue por su culpa, con la finalidad de desvirtuar su responsabilidad y obligación de indemnizar.

Es decir, se invierte inicialmente la carga de la prueba *de facto*, en el sentido de que al actor no le corresponderá demostrar la culpa del causante del daño sino que es a éste a quién le corresponde probar su actuar diligente, si bien, lo cierto es que la relación de causalidad entre el hecho y el daño deberá probarse, dado que no se presume ni puede basarse en conjeturas, deducciones, indicios o probabilidades, y esa carga de la prueba corresponderá a la actora de conformidad con lo previsto en el artículo 217 de la Ley de Enjuiciamiento Civil española, quien deberá acreditar que la actuación de la parte demandada fue la que causó el daño generador de la indemnización. Una vez acreditada la relación causal, se presumirá la culpa del causante.

Esta presunción sería inicialmente aplicable a cualquier sistema de responsabilidad en que se base el actor para ejercer su reclamación, incluido el sistema previsto en el TRLGDCU[600] española, que será objeto de análisis posterior, por lo que se refiere a productos defectuosos.

El criterio sobre el que se construye todo este sistema de responsabilidad es el de la integridad de la reparación, si bien, no es un criterio absoluto, en la medida que en determinados supuestos existe la posibilidad de reducir la indemnización que corresponda al perjudicado, por ejemplo, en los casos de concurrencia de culpa de este último, así como en los casos en que el perjudicado tenga el deber de mitigar el daño, legal o contractualmente. Estos aspectos constituirían supuestos de atenuación de la responsabilidad civil.

Además de todo ello, podrían producirse circunstancias de exoneración de la responsabilidad, en particular, el caso fortuito y la fuerza mayor, tanto en el ámbito de la responsabilidad contractual como extracontractual.

[600] Real Decreto Legislativo 1/2007, de 16 de noviembre (Texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias).

Conforme al ordenamiento jurídico español, en particular, tal y como dispone el artículo 1.105 de su Código Civil español[601], nadie podrá responder de aquellos sucesos que no hubieran podido preverse o que previstos fuera inevitables, es decir, causas externas o ajenas como el caso fortuito y la fuerza mayor. Estos supuestos exoneran de responsabilidad a quien incumple el contrato, ya que suponen una quiebra del nexo causal entre la acción y el resultado.

En estos casos se rompería el precitado nexo causal, en la medida que la causa del daño tiene su origen en un supuesto de caso fortuito o de fuerza mayor no imputable al sujeto y que es inevitable e imprevisible, por lo que no podrá imputarse su causa a la acción u omisión humana de dicho sujeto. Ambos supuestos constituyen situaciones imprevisibles o inevitables.

De manera consecuente, me surge una nueva pregunta: Estas circunstancias eximentes de responsabilidad, especialmente la fuerza mayor, ¿resultarían aplicables a un propietario, operador o usuario de un sistema de inteligencia artificial "autónomo" respecto de la supuesta imprevisibilidad asociada al mismo? Abordaré igualmente estas cuestiones con posterioridad.

El caso fortuito se identifica con un suceso que no pudo preverse, pero de haberse previsto se hubiera podido evitar, mientras que la fuerza mayor es un suceso inevitable, se hubiera o no previsto. El caso fortuito constituiría un supuesto de inexistencia de culpa. Del mismo modo, el caso fortuito se produce inicialmente en un contexto interno de la obligación, mientras que la fuerza mayor se produce en un contexto ajeno al círculo de la obligación. El Tribunal Supremo español ha definido estas circunstancias en múltiples sentencias[602].

[601] El Código Civil español contempla el caso fortuito en otros artículos específicos que constituyen aplicaciones concretas de las reglas generales, en particular, en los artículos 1.096, 1.136, 1.575, 1.602, 1.625, 1.744, 1.745, 1.891, 1.896, 1.488. Del mismo modo, contempla la fuerza mayor en otros preceptos que conforme aplicaciones concretas de las reglas generales, en concreto, en los artículos 1.602 y 1.625, 457, 1.777, 1784, 1.905 y 1.908.

[602] La Sentencia del Tribunal Supremo, Nº 23/1998, de 23/01/2004 (Rec. 530/1998) establece que la tendencia jurisprudencial hacia una objetivación de la culpa extracontractual, mediante los mecanismos de la inversión de la carga de la prueba y de la teoría del riesgo, no excluye de manera total y absoluta el esencial elemento psicológico o culpabilístico, como inexcusable ingrediente integrador, atenuado pero no suprimido de la responsabilidad por culpa extracontractual, de tal modo que si de la prueba practicada en

No obstante, ha dictado pronunciamientos muy dispares en situaciones similares, por ejemplo, y en relación con "hechos extraños a la conducción" ante la irrupción o invasión de la calzada de peatones o animales, que podríamos llevarlos a algunos supuestos reales que ya se han planteado en relación con los vehículos "autónomos"[603].

La doctrina mayoritaria se ha pronunciado sobre algunos razonamientos algo confusos del Tribunal Supremo sobre ambos conceptos, y tal y como Ataz López[604] aclara, si el daño era evitable por ser previsible el suceso que lo ocasionó, no puede hablarse de caso fortuito ni de fuerza mayor, con la indudable repercusión que ello puede tener en relación con la responsabilidad por daños causados por sistemas inteligentes.

La cuestión esencial radica en la inevitabilidad del daño, no tanto del suceso que lo causó, de modo que, si éste último era previsible, en función el contexto y circunstancias, el daño pudo haberse evitado.

La fuerza mayor debe consistir en una fuerza superior a todo control y previsión y, para valorar su concurrencia deberá estarse a la norma y razonable previsión que las circunstancias exijan adoptar en cada supuesto concreto o, inevitabilidad de una posibilidad existente.

---

el proceso, con inversión o sin ella, aparece plenamente acreditado que, en la producción del evento dañoso, por muy lamentable que sea, no intervino absolutamente ninguna culpa por parte del demandado o demandados, sino que el mismo fue debido exclusivamente a un imprevisible acaecimiento de fuerza de mayor, ha de excluirse la responsabilidad de los demandados.
La Sentencia del Tribunal Supremo de 31/05/2006 (Rec. 2968/1999) establece que por caso fortuito se entiende todo suceso imposible de prever, o que, previsto, sea inevitable y, por tanto, realizado sin culpa alguna del agente, de manera que el vínculo de causalidad se produce entre el acontecimiento y el daño, sin que en él intervenga como factor apreciable la actividad dolosa o culposa del agente, por lo que para que tal suceso origine exención de responsabilidad es necesario que sea imprevisible e inevitable, y que cuando el acaecimiento dañoso fue debido al incumplimiento del deber relevante de previsibilidad, no puede darse la situación de caso fortuito, debido a que con ese actuar falta la adecuada diligencia por omisión de atención y cuidado requerido con arreglo a las circunstancias del caso, denotando una conducta interfiriente frente al deber de prudencia y cautela exigibles, que como de tal índole es excluyente de la situación de excepción que establece el indicado Art. 1105 ,Código Civil, al implicar la no situación de imprevisibilidad, insufribilidad e irresistibilidad requeridas al efecto.

[603] Sentencias del Tribunal Supremo de 17 de noviembre de 1989 (RJ 7889), de 8 de febrero de 1992 (RJ 1198), 14 de mayo de 2014 (RJ 2729), 4 de febrero de 2015 (RJ 2075) y 11 de febrero de 2016.

[604] ATAZ LÓPEZ, J. (2021). "Caso fortuito y fuerza mayor en el Código Civil y en la jurisprudencia del Tribunal Supremo: La distinción entre ambas nociones", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Ed. Aranzadi-Thomson Reuters. 2021. Pp. 483-484.

A modo de ejemplo e ilustrativo en supuestos específicos más relacionados con la responsabilidad extracontractual, si adquiero un oso de peluche con cámara IP integrada o una aspiradora dotada de idéntico elemento, ¿Qué ocurre si el acceso a dicha cámara una vez conectado el dispositivo a internet es directamente accesible o integra una configuración de seguridad general "de fábrica", conocida por cualquier persona que adquiera dicho dispositivo, sin que obligue o informe al usuario de que debe cambiarla o personalizarla? ¿Sería por sí mismo un defecto del producto? ¿Dónde se exige ese requerimiento? Si alguien accede a este tipo dispositivo y a la cámara a través de internet y puede ver a mi hijo mientras duerme, su dormitorio o toda mi casa, incluyendo obras de arte, presencia o no de personas, etc…. ¿Podría exigir responsabilidad al fabricante o al proveedor? ¿En base a qué marco jurídico? No tan poco previsible, vistas algunas experiencias, la existencia de páginas web donde consultar todos los dispositivos IoT conectados a Internet en el mundo con información de su configuración y la existencia y fácil acceso a tutoriales y herramientas de *hackeo* de los mismos también disponibles en la red, pero, en cualquier caso, en buena medida evitable mediante una adecuada configuración de seguridad.

En este supuesto, nos encontramos ante un nuevo contexto: El precitado *Internet de la Cosas* o *Internet of Things* -IoT por sus siglas en inglés-, donde se produce "la tormenta perfecta". Millones de dispositivos conectados a Internet con puertos abiertos, sin protección o con protección general "de fábrica". Estamos abriendo las puertas a cualquier visitante no deseado. Un ejemplo real lo constituye "Mirai Atack" que expuse en el capítulo II.

Y continuando con mis reflexiones, si el próximo regalo a mis hijos es un dron dotado con cámara que controlo desde el móvil y que se conecta al mismo vía *Bluetooth*, que he adquirido directamente del fabricante, comercializador a su vez del mismo a través de su tienda virtual. ¿Qué ocurre si alguien toma el control a través de mi *app* en mi móvil y lo dirige contra la cara de mi hijo causándole distintos daños? ¿Qué tipo de responsabilidad se derivaría y a quién sería exigible? ¿El fabricante-comercializador sería responsable ante la falta de seguridad de acceso y control del dispositivo? Poco previsible pero seguramente evitable ante una adecuada configuración de seguridad.

La carga de la prueba de estas situaciones eximentes de responsabilidad, ¿correspondería al usuario o a quien la oponga como causa de exoneración de responsabilidad?

Si inicialmente consideramos que la responsabilidad recae en el fabricante o productor, ello significa que la persona a la que se atribuye el incumplimiento, los daños y la consiguiente responsabilidad contractual por los mismos, para poder evitar la misma, debería probar haber obrado con la diligencia exigible para evitar el daño en el ejercicio de su actividad.

Las preguntas consiguientes a realizar son: ¿Dónde están definidos específicamente los compromisos contractuales de seguridad de estos dispositivos? ¿En el contrato, anexos técnicos? En caso de ausencia, ¿existe alguna previsión específica en normativa general o sectorial en materia de seguridad sobre productos o ciberseguridad? La posible respuesta es que, en general, no estarán definidos estos requerimientos y compromisos consecuentes en ninguna de estas fuentes, sin perjuicio de que las nuevas propuestas reguladoras en el ámbito de la UE apuestan por la seguridad de maquinaria, productos y sistemas inteligentes integrados o no en aquellos, por lo que de antemano y salvo futuras disposiciones específicas, deberíamos aplicar los marcos generales de responsabilidad extracontractual o por productos defectuosos y protección del consumidor, en caso de resultar aplicable.

Como expondré a lo largo de este capítulo, si un producto es defectuoso, conforme establecen las directivas europeas vigentes, la Directiva de Máquinas (Directiva 2006/42/CE de 17 de mayo de 2006) o la Directiva de Seguridad General de los Productos (Directiva 2001/95/CE de 3 de diciembre de 2001) que analicé en el capítulo II, se atribuye la responsabilidad inicialmente al productor. No obstante, en la actualidad, no hay criterios de seguridad específicos en estos marcos para los productos con sistemas inteligentes integrados, además de que la Directiva europea sobre responsabilidad por productos defectuosos se refiere a productos y no se extiende a servicios defectuosos, sin perjuicio de las normas en España de transposición y extensión de la misma, en especial, el TRLGDCU, que regula tanto la responsabilidad en caso de productos como servicios defectuosos.

La inteligencia artificial supone un nuevo paradigma que exigirá la revisión de los regímenes tradicionales de responsabilidad.

A modo de ejemplo, un *roboadvisor* inteligente de segunda generación que tome decisiones de gestión sobre nuestro patrimonio o de inversión en relación con nuestra cartera, en caso de pérdidas financieras.

En base a los compromisos contractuales contraídos, se debería determinar si estamos en el marco de un arrendamiento de servicios o un contrato de obra y, en función de ello, depurar las responsabilidades que en su caso concurran, inicialmente frente a su prestador/operador oferente de la operación mediante un sistema inteligente, ante la imposibilidad de hacerlo contra el propio sistema, sin perjuicio que, de concurrir errores o defectos del sistema, sea su productor quien pudiera responder de los mismos, contra el que podría iniciar la vía de repetición hacia responsable inmediato. En este sentido ya existe un procedimiento judicial incoado en un contexto similar en el Tribunal Comercial de Londres, actualmente en tramitación, cuya resolución puede ser muy interesante para la reflexión sobre estas cuestiones.

En relación con el supuesto planteado, significar que en fechas recientes los medios de comunicación de todo el mundo publicaron la caída del primer fondo gestionado por inteligencia artificial en EE.UU.[605], al parecer, por apostar demasiado por las tecnológicas.

Si a continuación seguimos profundizando sobre estas características esenciales de la responsabilidad contractual y las relacionamos con un contexto en el que una de las partes sea el diseñador, fabricante, operador o comercializador de un sistema de inteligencia artificial, o un sistema de inteligencia artificial avanzado o "fuerte", con supuesta autonomía, capacidad de autoaprendizaje y relativa impredecibilidad o incluso, en un hipotético futuro, hasta con personalidad jurídica y capacidad de obrar reconocida, e incluso ante contextos donde las dos partes pudieran ser robots o sistemas dotados de

[605] "El primer fondo gestionado por inteligencia artificial en EEUU cae por apostar demasiado por las tecnológicas". Publicado por El Economista, el 25.05.2021. Recuperado de: El primer fondo gestionado por inteligencia artificial en EEUU cae por apostar demasiado por las tecnológicas - elEconomista.es. Consultado el: 30.05.2021.

inteligencia artificial avanzada dotados de autonomía, personalidad y capacidad, las cuestiones que se nos suscitan son casi infinitas y desde luego apasionantes, pero un análisis jurídico de estas características nos obliga a partir de realidades actuales o potenciales, y no futuribles, y no sé hasta qué punto deseables, como comento en otros capítulos de esta investigación.

Los sistemas de inteligencia artificial son y serán en mayor medida objeto de todo tipo de contratos -diseño, desarrollo, compraventa, integración, formación, entrenamiento, mantenimiento, monitorización, auditoría, alquiler, licencia y/o servicios, entre otros-, y entre los distintos sujetos intervinientes en todo el ciclo de vida de los mismos. Estos contratos podrán o, más bien, deberán integrar información de su funcionamiento, manuales, condiciones de uso, información técnica, normas de seguridad y estándares de adhesión que limiten algunas de sus capacidades, seguros, garantías e incluso acuerdos de nivel de servicio -ANS o SLA por sus siglas en inglés, esto es, *Service Level Agreement*-, que ayudarán a conformar e integrar el marco obligacional vinculante para las partes firmantes, y que deberá adecuarse e integrarse con el marco legal vigente, así como con aquellos marcos éticos, estándares propios de la industria reconocidos y códigos de conducta a los que las partes se hallen adheridas y que deben complementar, a mi juicio, el objeto del contrato y la diligencia debida y seguridad esperada.

En mi opinión, como he manifestado al analizar otros aspectos a lo largo de esta investigación la deseable autorregulación de la industria es insuficiente, en la medida que las mayores exigencias en seguridad exigen también mayores requerimientos de inversión, lo que repercute en el precio de los productos, lo que a su vez puede suponer una desventaja competitiva inicial para los fabricantes sujetos a mayores requerimientos, especialmente ante otros ubicados en jurisdicciones con mecanismos más laxos y permisivos volcados en la innovación, en el fomento de actividad económica, la creación de servicios de valor añadido para el usuario, el crecimiento empresarial y/o en el despliegue y aplicación masiva de la tecnología en todos los ámbitos. Y el incremento de los costes puede afectar directamente a la accesibilidad a la tecnología.

No obstante, la otra lectura de este contexto es que la seguridad y confiabilidad de los productos resultantes también proporciona una ventaja competitiva que ha permitido destacar y posicionar a la UE en distintos mercando en base a dichos atributos.

La autorregulación en dispositivos del Internet de las Cosas -*IoT* por sus siglas en inglés- y la publicación de las buenas prácticas de seguridad asociadas a los mismos se evidenció claramente insuficiente en Reino Unido, lo que motivó la propuesta parlamentaria para la creación de un marco jurídico vinculante que contemple y regule su exigencia.

En este sentido, las propuestas de la UE para regular los principios, valores, normas éticas y obligaciones jurídicas esenciales de la inteligencia artificial y convertirlas en vinculantes conforme a las Resoluciones de 20 de octubre de 2020, son un acierto absoluto como primera aproximación, y considero que constituyen una buena referencia internacional inicial para la regulación de la inteligencia artificial, especialmente para garantizar su confiabilidad y seguridad. La cuestión es si el enfoque y alcance es el adecuado.

La Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[606], conforme se ha expuesto en su análisis durante el capítulo precedente, supone un enfoque algo distinto, muy orientado a la prohibición de determinados sistemas y a una regulación pretendidamente minuciosa de los denominados sistemas de inteligencia artificial de alto riesgo, no sujetando al resto de sistemas inteligentes a ningún tipo de principio, norma u obligación ética o jurídica -a excepción de las obligaciones de transparencia para determinados sistemas-, ni tan siquiera la incuestionable y necesaria supervisión y control humano. Además, como he referido, únicamente se focaliza en la denominada inteligencia artificial "débil".

Para finalizar los aspectos generales sobre la responsabilidad contractual, en congruencia con lo que expondré posteriormente, significar que la prescripción de la acción, el artículo 1964.2 del Código Civil establece el plazo de 5 años desde que pueda exigirse el cumplimiento de la obligación para las acciones personales que no tengan plazo especial de prescripción.

[606] COM (2021) 206 final 2021/0106 (COD)

Por último, el objeto y alcance de esta investigación me impiden abordar distintos aspectos con más profundidad como, por ejemplo, el deber de consejo, información y asesoramiento de la parte contratante más preparada para mantener el equilibrio de las partes, la buena fe, el dolo, la negligencia o la morosidad.

2.2. Responsabilidad por evicción y vicios ocultos.

La responsabilidad por evicción y vicios ocultos debe ser igualmente considerada en relación con los sistemas de inteligencia artificial y no es exclusiva de los contratos de compraventa, si bien, abordaré la misma de manera general para focalizarme, como he anticipado, en los marcos vigentes de responsabilidad extracontractual y de productos defectuosos.

Derivado de la relación contractual establecida, el saneamiento es la obligación que tiene el vendedor de responder ante el comprador cuando se le prive total o parcialmente de la posesión legal, pacífica y útil de la cosa objeto del contrato. Conforme a la doctrina tradicional, las obligaciones del vendedor no se acaban con la entrega de la cosa, sino que el vendedor debe asegurar al comprador la posesión pacífica y útil de la misma.

La responsabilidad del vendedor por evicción y vicios ocultos en los contratos de compra y venta se halla regulada en los artículos 1475 a 1499 del *Código Civil* español. A mi juicio, constituye una garantía contractual de la que se derivan directamente responsabilidades de esta naturaleza para el vendedor de una cosa, en este caso, de un dispositivo, robot o sistema inteligente, que podrían resultar extensibles a su arrendador.

El saneamiento por evicción pretende asegurar la posesión pacífica de la cosa y deriva de la pérdida de la misma en virtud de un derecho preexistente sobre ella por parte de un tercero anterior a la compraventa, así como de la existencia de gravámenes no conocidos y asumidos por el adquirente.

Por su parte, la responsabilidad por vicios ocultos nace por los defectos ocultos de la cosa vendida en el caso de que éstos la hagan impropia para el uso al que destina o si

disminuyen de tal modo dicho uso que, de haberlos conocido el adquirente, no la habría adquirido o habría pagado menos precio. Y todo ello, aunque el vendedor los ignorase.

Esta responsabilidad se derivará igualmente en caso de pérdida de la cosa vendida por efecto de los vicios ocultos, aunque con distinto alcance en función de si el vendedor los conocía o no, si concurrió caso fortuito, culpa del adquirente o el vendedor actuó de mala

Esta responsabilidad también podrá exigirse contractualmente al vendedor de dispositivos, robots y sistemas inteligentes por parte del adquirente que pueda verse afectado por ambas circunstancias, con los efectos previstos en los artículos precitados[607].

A modo de ejemplo, pensemos en un dron con asistencia de vuelo y sistemas de navegación inteligente que desaparece en el horizonte sin destino conocido o el vehículo autónomo que una semana después de su adquisición colisiona frontalmente a 50 km/hora contra las balizas metálicas retráctiles de una calle de circulación provisionalmente restringida y cortada al tráfico.

No obstante, como he referido anteriormente, pretendo centrar mi análisis en la responsabilidad extracontractual.

---

[607] Artículo 1474 CC. En virtud del saneamiento a que se refiere el artículo 1.461, el vendedor responderá al comprador: 1.º De la posesión legal y pacífica de la cosa vendida. 2.º De los vicios o defectos ocultos que tuviere.
Artículo 1475 CC. Tendrá lugar la evicción cuando se prive al comprador, por sentencia firme y en virtud de un derecho anterior a la compra, de todo o parte de la cosa comprada.
El vendedor responderá de la evicción, aunque nada se haya expresado en el contrato.
Los contratantes, sin embargo, podrán aumentar, disminuir o suprimir esta obligación legal del vendedor.
Artículo 1484 CC. El vendedor estará obligado al saneamiento por los defectos ocultos que tuviere la cosa vendida, si la hacen impropia para el uso a que se la destina, o si disminuyen de tal modo este uso que, de haberlos conocido el comprador, no la habría adquirido o habría dado menos precio por ella; pero no será responsable de los defectos manifiestos o que estuvieren a la vista, ni tampoco de los que no lo estén, si el comprador es un perito que, por razón de su oficio o profesión, debía fácilmente conocerlos.
Artículo 1485 CC. El vendedor responde al comprador del saneamiento por los vicios o defectos ocultos de la cosa vendida, aunque los ignorase. Esta disposición no regirá cuando se haya estipulado lo contrario y el vendedor ignorara los vicios o defectos ocultos de lo vendido.
Artículo 1486 CC. En los casos de los dos artículos anteriores, el comprador podrá optar entre desistir del contrato, abonándosele los gastos que pagó, o rebajar una cantidad proporcional del precio, a juicio de peritos. Si el vendedor conocía los vicios o defectos ocultos de la cosa vendida y no los manifestó al comprador, tendrá éste la misma opción y además se le indemnizará de los daños y perjuicios, si optare por la rescisión.
Artículo 1487 CC. Si la cosa vendida se perdiere por efecto de los vicios ocultos, conociéndolos el vendedor, sufrirá éste la pérdida, y deberá restituir el precio y abonar los gastos del contrato, con los daños y perjuicios. Si no los conocía, debe sólo restituir el precio y abonar los gastos del contrato que hubiese pagado el comprador.
Artículo 1488 CC. Si la cosa vendida tenía algún vicio oculto al tiempo de la venta, y se pierde después por caso fortuito o por culpa del comprador, podrá éste reclamar del vendedor el precio que pagó, con la rebaja del valor que la cosa tenía al tiempo de perderse. Si el vendedor obró de mala fe, deberá abonar al comprador los daños e intereses.

2.3. Responsabilidad contractual vs extracontractual

La principal diferencia entre ambas responsabilidades radica en su origen distinto del que derivan sus respectivos regímenes, dado que la responsabilidad contractual presupone la existencia de una relación anterior, que puede ser un contrato o cualquier otra relación jurídica que otorgue un derecho de resarcimiento, mientras que la extracontractual parte de que no existe ningún vínculo obligatorio o relación jurídica preexistente o sólo presupone un daño, con independencia de que exista cualquier relación jurídica preexistente entre las partes.

No obstante, se producen inevitables aspectos coincidentes que atenúan su diferenciación por basarse ambas en el principio general de que quien causa un daño lo debe indemnizar y en su propia finalidad reparadora, ya sea causado por el incumplimiento de una obligación preestablecida o por culpa no relacionada con la existencia de un marco obligacional preexistente. De hecho, se produce la aplicación indistinta de preceptos considerados “comunes”.

Es más, la existencia de un contrato entre las partes no es suficiente para excluir automáticamente la responsabilidad extracontractual, precisándose para esta exclusión que la realización del hecho dañoso se produzca rigurosamente dentro del marco contractual pactado, dado que si se trata de una negligencia ajena a lo que constituye propiamente el objeto del contrato, esta negligencia desplegará sus efectos independientemente y de forma autónoma, por lo que podrían concurrir ambas clases de responsabilidades en una yuxtaposición, que sólo se rompería cuando se den estrictamente los requisitos de una o de otra responsabilidad.

Esto nos lleva a una afirmación práctica de indudable interés para la finalidad reparadora sobre la que se sustentan ambos regímenes y relacionada con el objeto de esta investigación, que no es otra que, en cualquier caso, subsiste la culpa extracontractual completando a la contractual, por cuanto integra todos los elementos orientados al pleno resarcimiento del daño, sin otro límite que dejar indemne al perjudicado.

En la práctica, la jurisprudencia mayoritaria[608] en España se ha decantado por los principios de unidad de culpa civil, finalidad reparatoria y yuxtaposición de ambas responsabilidades, contractual y extracontractual, considerando que lo único vinculante para el juzgador, desde el punto de vista de congruencia, son los hechos de la demanda más que los fundamentos jurídicos relacionados con una u otra responsabilidad, disponiendo de la libertad para encuadrar la conducta del demandado en la culpa contractual o extracontractual, en el marco de sus facultades de aplicación de la norma pertinente ya adecuada.

3. La responsabilidad civil extracontractual

3.1. Cuestiones previas

La responsabilidad extracontractual es aquella que nace como consecuencia de acciones u omisiones imputables a una persona por su culpa o negligencia, que causen daños personales o patrimoniales a otra, que aquélla deberá reparar.

Inicialmente, la responsabilidad extracontractual nace de una relación jurídica entre dos personas entre las que no existe un vínculo contractual previo, como consecuencia de los actos u omisiones imputables a una de ellas por culpa o negligencia, que producen daños de naturaleza personal o patrimonial a la otra, y que conllevan el deber de indemnizar los mismos.

De nuevo, el precepto de referencia es el artículo 1089 del Código Civil español que establece que “las obligaciones nacen de la ley, de los contratos, cuasi contratos y de los actos y omisiones ilícitos o en que intervenga cualquier género de culpa o negligencia". En este caso, estas obligaciones no nacen de un contrato sino de los actos y omisiones ilícitos o en los que intervenga la culpa o la negligencia.

[608] La Sentencia del Tribunal Supremo español nº 135/2009, de 4 de marzo de 2009, recoge la doctrina consolidada y recogida en pronunciamientos previos sobre la diferenciación entre ambos tipos de responsabilidad, contractual y extracontractual y la aplicación de principios como el de unidad de la culpa civil. Ver igualmente la Sentencia del Tribunal Supremo español de 13 de mayo de 2006 (RJ 2006/3497).

Como referí anteriormente, cuando hablamos de "ilícitos" podemos considerar desde una perspectiva amplia tanto ilícitos civiles, penales, administrativos, como de cualquier otra naturaleza, si bien, conforme he indicado anteriormente y en congruencia con el objeto de esta investigación, me circunscribiré en este capítulo al análisis de la responsabilidad civil, dejando el análisis de las nacidas por la comisión de un delito para su tratamiento en el capítulo VI de esta investigación, donde se abordará la responsabilidad penal y algunos de los tipos penales más estrechamente relacionados con el uso de sistemas inteligentes como medio o instrumento para su comisión.

En este sentido, recordar que la culpa o negligencia puede originar responsabilidad penal en caso de hallarse tipificada como tal dicha conducta en el ordenamiento jurídico vigente, así como responsabilidad civil contractual o extracontractual en función de los elementos concurrentes de cada una de ellas, especialmente ante la existencia o no de un vínculo contractual previo.

Como he referido anteriormente, las acciones y omisiones que puedan constituir un ilícito civil en el que intervenga culpa o negligencia y con origen en los compromisos que integran un marco contractual generará responsabilidad contractual.

Aquéllas que se produzcan fuera del marco de un contrato generarán responsabilidad extracontractual, denominada también "responsabilidad aquiliana", la cual tiene su origen en el Derecho romano basado en el *damnum iniuria datum* previsto en la *Lex Aquilia* y que todavía hoy es una institución que utilizamos para resolver conflictos y responsabilidades relacionados con las tecnologías más vanguardistas.

El artículo 1902 del Código Civil español complementa lo dispuesto por su precitado artículo 1089 y establece los elementos que conforman la responsabilidad extracontractual: "El que por acción u omisión causa daño a otro, interviniendo culpa o negligencia, está obligado a reparar el daño causado".

Es decir, deben concurrir una acción o una omisión culposa o negligente, la producción de un resultado dañoso y la existencia de un nexo o relación de causalidad entre la acción u omisión y el daño causado.

La responsabilidad no nace por el incumplimiento de alguna obligación contractual sino por la omisión de un deber de diligencia debida que incumbe a toda persona. Si fuera el incumplimiento de la diligencia debida en un contrato se consideraría incumplimiento del mismo, resultando pues exigible la responsabilidad contractual.

La responsabilidad extracontractual tiene su origen en la omisión de un deber de diligencia y debe existir una relación de causalidad entre la acción y omisión del agente responsable y los daños producidos al perjudicado.

La responsabilidad nace cuando una persona causa -sea por sí misma, por medio de otra de la que responda o por una cosa de su propiedad o que posea-, un daño a otra persona con la que no estaba ligada por un vínculo obligatorio anterior relacionado con el daño causado. De esta conceptualización básica, podríamos extraer una primera aproximación general, respecto de la posibilidad de que el operador, propietario o usuario de un sistema inteligente (o máquina, robot o dispositivo dotado de éste), deba responder de los daños causados a un tercero por medio del mismo, en un contexto inicial de inteligencia artificial "débil".

Los elementos y requisitos exigidos para su concurrencia son tres: La acción u omisión culposa o negligente, el daño y la relación de causalidad.

a) Acción u omisión culposa o negligente

Por lo que se refiere a la existencia de una acción u omisión culposa o negligente, podemos distinguir dos elementos dentro del mismo, uno de carácter objetivo y otro de carácter subjetivo, si bien, la jurisprudencia española mayoritaria hace referencia únicamente al subjetivo, como analizan Muñoz Villarreal y Gallego Corchero[609].

Según estos autores, el elemento objetivo hace referencia a la manifestación externa de la conducta del causante del daño, que puede tratarse tanto de una acción como de una omisión, mientras que el elemento subjetivo hace referencia a "un hacer u

[609] MUÑOZ VILLARREAL, A. Y GALLEGO CORCHERO, V. (2019). Inteligencia artificial e irrupción de una nueva personalidad en nuestro ordenamiento jurídico ante la imputación de responsabilidad a los robots, en Monterrosso Casado, E. (Dir.), *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch 2019. Pp. 86-87.

omitir algo que se encuentra fuera de las normas de cautela y previsión establecidas por el ordenamiento y socialmente aceptadas" y que "se traduce en un *animus* de quien emplea o utiliza el medio productor del resultado"[610].

La jurisprudencia española[611] ha ido modulando la diligencia requerida, considerando que comprende no sólo las prevenciones y cuidados reglamentarios, sino además todos los que la prudencia imponga para evitar el evento dañoso. En relación con su análisis, me remito igualmente a las consideraciones efectuadas en los anteriores epígrafes sobre la diligencia debida, especialmente en el ámbito profesional y empresarial.

De este modo, las conductas con resultado dañoso que se aparten de las normas éticas y de seguridad establecidas por el ordenamiento jurídico vigente, por marcos y normas de conducta sectoriales, por códigos de conducta o de buenas prácticas a los que el responsable se halle adherido y por lo que la propia prudencia imponga para evitar el daño, podrían estar sujetas a esta responsabilidad.

El fundamento de esta responsabilidad tiene un carácter inicialmente subjetivo basado en la culpa o negligencia al causar el daño a resarcir, con una jurisprudencia tendente a aplicar la culpa levísima y a objetivar la responsabilidad en actividades peligrosas -exclusivamente prevista para supuestos específicos-, creando así una responsabilidad *cuasi objetiva* inicial, en la que lo que importa es la responsabilidad

[610] Sentencia del Tribunal Supremo, Sala de lo Civil, de 15.07.1992.

[611] La Sentencia del Tribunal Supremo, Sala de lo Civil, de 29.10.2008, Rec. 942/2003, considera que, para que la responsabilidad extracontractual regulada en el artículo 1.902 del Código Civil sea admitida, se hace preciso la conjunción de los requisitos siguientes: Uno subjetivo, consistente en la existencia de una acción u omisión generadora de una conducta imprudente o negligente atribuible a la persona o entidad contra la que la acción se dirige; otro objetivo, relativo a la realidad de un daño o lesión; y, por último, la relación causal entre el daño y la falta. Del mismo modo, la doctrina jurisprudencial se inclina por la tesis de que no resulta suficiente la diligencia reglamentaria si la realidad fáctica evidencia que las garantías adoptadas para evitar los daños previsibles han resultado ineficaces. Igualmente, se declara que si bien el artículo 1.902 del Código Civil español descansa en un principio básico culpabilista, no debe desconocerse que la diligencia requerida comprende no sólo las prevenciones y cuidados reglamentarios, sino además todos los que la prudencia imponga para evitar el evento dañoso todos los que la prudencia imponga para evitar el evento dañoso, con inversión de la carga de la prueba y presunción de conducta culposa en el agente, así como la aplicación, dentro de prudentes pautas, de la responsabilidad basada en el riesgo, aunque sin erigirla en fundamento único de la obligación a resarcir, no siendo suficiente para la inexistencia de culpa acreditar que se procedió con sujeción a las disposiciones legales que, al no haber ofrecido resultado positivo, revelan su insuficiencia y la falta de algo por prevenir, estando por tanto incompleta la diligencia. Esta Sentencia se sustenta en la jurisprudencia precedente recogida en las Sentencias del Tribunal Supremo, Sala de lo Civil, de 22.04.1987 y de 13.07.1999, entre otras.

existente por haber activado o generado un riesgo, siendo suficiente que exista un daño y un nexo causal con la persona que lo ha producido para que deba resarcir a la persona afectada.

La objetivación se ha construido a nivel jurisprudencial mediante la aplicación de la teoría del riesgo, que principalmente comporta una inversión de la carga de la prueba y un mayor rigor en la diligencia requerida, con el objetivo de proteger a las personas afectadas en el marco de actividades complejas y peligrosas, lo que a su vez se relaciona en la argumentación de su aplicación en supuestos complejos como los relacionados con la inteligencia artificial, en el principio *cuius commoda eius incommoda*[612], que se traduce en que a quien corresponden los beneficios corresponden los inconvenientes

b) Daño

El Código Civil español no establece ningún requerimiento del mismo, si bien, la jurisprudencia[613] ha delimitado el daño indemnizable a través del sistema de responsabilidad civil extracontractual integrado en el mismo, el cual podrá ser no sólo patrimonial sino también extrapatrimonial o moral, si bien, en cualquier caso, el daño deberá ser cierto, aunque podrá ser actual o futuro.

c) Relación de causalidad

Respecto a la relación de causalidad, tradicionalmente se ha planteado la delimitación del concepto de causa y su adecuación para causar el daño, que ha sido

612 https://dpej.rae.es/lema/cuius-commoda-eius-incommoda#:~:text=Gral.,beneficios%2C%20corresponden%20los%20inconvenientes'.

613 Entre otras, citar las siguientes Sentencias: TS, Sala de lo Civil, nº 553/2005, de 07/07/2005, Rec. 296/1999 (Daño patrimonial, daño emergente y lucro cesante); TS, Sala de lo Social, de 12/12/2007, Rec. 25/2007 (Daño moral); TS, Sala de lo Civil, nº 28/2014, de 29/01/2014, Rec. 2509/2011 y TS, Sala de lo Civil, nº 899/2011, de 30/11/2011, Rec. 1692/2010 (Daño permanente o duradero); TS, Sala de lo Contencioso, de 30/10/2012, Rec. 3566/2011, consolidando el criterio jurisprudencial establecido, entre otras, por la TS, Sala de lo Contencioso, nº S/S, de 11/05/2004, Rec. 2191/2000 (Daño continuado o de producción sucesiva); TS, Sala de lo Civil, nº 157/2003, de 21/02/2003, Rec. 2034/1997.

resuelto por la doctrina y jurisprudencia mayoritaria mediante la aplicación de la denominada *teoría de la causalidad adecuada*[614].

De este modo, la responsabilidad tiende jurisprudencialmente a sustentarse sobre el riesgo más que sobre la culpa para su objetivación, y precisamente sobre esta base se están construyendo los futuros marcos reguladores de la responsabilidad por daños derivados del funcionamiento y uso de sistemas de inteligencia artificial, como analizaré con posterioridad.

La jurisprudencia ha ido modulando esta responsabilidad *cuasi objetiva* conforme reflejan distintos autores, entre otros, Gómez-Riesco[615], invirtiendo la carga de la prueba de la culpa, aumentando el rigor del nivel de diligencia exigible, la no exoneración de responsabilidad por la simple inobservancia de los reglamentos de conducta y flexibilizando el nexo de causalidad entre el daño y el hecho que lo genera.

La inteligencia artificial está contribuyendo y contribuirá en mayor medida a reducir los daños en muchos sectores de actividad -diagnóstico médico, tratamiento de enfermedades o conducción por vías públicas-, y especialmente los físicos, pero, por otra parte, también contribuirá a aumentar los mismos en distintos sectores, especialmente en el ámbito emocional o moral y de los derechos fundamentales.

Las responsabilidades de los distintos agentes intervinientes dentro del ciclo de vida de los sistemas inteligentes por los daños derivados de su funcionamiento y uso frente a la persona afectada, es decir, diseñadores, desarrolladores, fabricantes/productores, importadores, comercializadores, distribuidores, formadores, operadores, propietarios y usuarios de sistemas de inteligencia, podrían enmarcarse inicialmente dentro la denominada responsabilidad extracontractual.

De este modo, inicialmente, si el daño se produce por un uso negligente de un sistema inteligente, la responsabilidad dimanante de este sistema de responsabilidad analizado se

[614] Recogida en la Sentencia del Tribunal Supremo español nº 83/2010, de 22 de febrero de 2.010 (Rec.356/2007), F.J. 2.
[615] GÓMEZ-RIESCO TABERNERO DE PAZ, J. (2018). "Los robots y la responsabilidad civil extracontractual", en BARRIO, M. (Dir.). *Derecho de los Robots*. Wolters Kluwer, Madrid, 2018, P. 113.

debería imputar inicialmente a quien lo usaba cuando se produjo la circunstancia causante del daño.

Si fuera consecuencia de un inadecuado diseño, falta de seguridad o defecto en su ensamblaje, integración o construcción la responsabilidad se debería imputar inicialmente al fabricante. Si fuera consecuencia de un mantenimiento deficiente, la responsabilidad se debería imputar inicialmente al propietario o usuario del sistema.

Salvando las distancias, una primera aproximación a casos donde la responsabilidad extracontractual se ha escalado hasta el fabricante, podemos tomar como ejemplo de la misma la vinculación establecida entre compradores de paquetes de tabaco con sus respectivos fabricantes.

En estos supuestos no existe una relación contractual entre los mismos, sí entre la persona afectada y los expendedores que actúan en establecimientos abiertos al público, en nombre y por cuenta propia.

El daño producido excede de las relaciones contractuales establecidas entre ambos, por lo que es escalado al fabricante. El daño en estos supuestos derivaría de una información supuestamente deficiente sobre la peligrosidad del producto imputable al fabricante, importador o distribuidor mayorista. En consecuencia, los daños reclamables pueden ser enmarcados dentro de la responsabilidad extracontractual.

En cualquier caso, cualquier régimen de responsabilidad civil sustentado en la culpa requiere un elemento esencial que es la "imputabilidad", es decir que el agente o persona responsable sea imputable y pueda atribuírsele la culpa o negligencia. Los sistemas de inteligencia artificial carecen actualmente de personalidad jurídica y de la capacidad para ser titulares de derechos y obligaciones, y esta cuestión parece que por el momento ha salido de la agenda del legislador europeo, como he apuntado en anteriores capítulos y comentaré con detalle más adelante más adelante, a la vista de las recientes propuestas reguladoras del Parlamento Europeo y la Comisión.

3.2. Régimen general

El régimen de responsabilidad civil extracontractual vigente en España se sustenta en la existencia de culpa en el agente o persona que causa el daño o perjuicio, agente que debe ser libre y plenamente consciente de sus actos y, en consecuencia, en la posibilidad de imputar la misma y la responsabilidad a quien lo causó.

El régimen basado en la culpa se mantiene incluso en aquellos supuestos en los que para proteger determinados intereses se ha concebido un criterio de imputación objetivo para la exigencia de responsabilidad, puesto que, aunque en éstos se invierta la carga de la prueba para presumir la concurrencia de culpa, el responsable dispondrá de la posibilidad de eximirse de responsabilidad acreditando su "no culpabilidad".

La responsabilidad objetiva o absoluta constituyen sistemas de responsabilidad especiales que se alejan de la culpa y se centran el riesgo, de modo que la responsabilidad se derivaría aunque no hubiera mediado negligencia inicial alguna por parte del responsable. Este tipo de responsabilidad se reserva inicialmente a casos extraordinarios, excepcionales o anormales.

El ordenamiento jurídico español contempla supuestos específicos en los que el grado de objetivación se aplica a su máximo nivel en atención al alto riesgo provocado, de modo que ni tan siquiera se presume la concurrencia de culpa del causante del daño, sino que se le atribuye la responsabilidad directamente por el alto riesgo de daño que genera una actividad concreta, sin posibilidad, con algunos matices, de que el responsable pueda acreditar su diligencia o no culpabilidad para eximirse de su responsabilidad.

En el caso de España, estos regímenes se aplicarían a los daños causados por animales previstos en el artículo 1905 del Código Civil español, los causados en actividades de caza conforme al artículo 33 de la Ley 1/1970, de 4 de abril, de Caza, y los causados en el marco del transporte aéreo, conforme prevé la Ley 48/1960, de 21 de julio, de Navegación Aérea. Algunos de estos regímenes especiales serán abordados más adelante.

No obstante, durante los últimos años, se está debatiendo sobre la posibilidad de aplicar regímenes de responsabilidad objetiva a la inteligencia artificial.

El problema es que el establecimiento general de una responsabilidad objetiva y absoluta, en relación con el desarrollo, explotación, puesta en funcionamiento, aplicación y uso de la inteligencia artificial, especialmente a nivel local por parte de determinados ordenamientos jurídicos, podría tener, de un lado y de manera inmediata, un efecto disuasorio y suponer un obstáculo para la inversión, la innovación, el desarrollo y la aplicación de la misma, así como para el desarrollo tecnológico, económico y empresarial y, de otro y de manera mediata, una obstaculización de la accesibilidad y disponibilidad de tecnologías que pueden mejorar el bienestar social y el mundo en el que vivimos, especialmente ante el incremento de sus costes.

No obstante, en función del enfoque de esta cuestión, también se puede extraer una lectura positiva como indicaba anteriormente, en la medida que la exigencia de este tipo de responsabilidad podría suponer incentivos para la inversión y garantías adicionales por parte de sus productores y proveedores, lo que permitiría la puesta en el mercado de sistemas más seguros y fiables, y mejoraría la necesaria confiabilidad y seguridad en los mismos por parte de la sociedad en general para su despliegue y aplicación.

Por otra parte, no sujetarlas a ningún marco de responsabilidad o regímenes más laxos puede tener un efecto desincentivador para fabricantes, operadores o propietarios de estos sistemas para invertir y adoptar las medidas adecuadas para reducir los riesgos asociados a los mismos y evitar o reducir la causación de daños, así como dificultar enormemente la efectividad de estos regímenes de responsabilidad ante las características intrínsecas y capacidades de estas tecnologías -integradas en su diseño o adquiridas-, especialmente su opacidad, explicabilidad y complejidad técnica- y el resarcimiento efectivo de las personas afectadas.

Actualmente, se está debatiendo en el seno de la UE sobre la conveniencia de utilizar un régimen particular de responsabilidad civil objetiva para supuestos específicos -reservado en la actualidad para actividades de alto riesgo, muy peligrosas o productos defectuosos en atención a aspectos como el elevado coste de prueba y verificación de la negligencia, el nivel de actividad/riesgo o la posibilidad efectiva de socializar las pérdidas-, y de un régimen subjetivo sustentado en la culpa para el resto.

Esta es la línea seguida por el Parlamento Europeo en su reciente propuesta regulatoria de la responsabilidad civil de la inteligencia artificial de 20 de octubre de 2020, objeto de análisis en este capítulo.

Cualquier propuesta reguladora requerirá mantener un equilibrio de todos los intereses precitados en constante tensión y que requerirá revisiones y adaptaciones constantes.

El contexto global actual exige establecer un régimen equilibrado que tenga en cuenta todos los retos, riesgos, derechos e intereses en juego de todas las partes implicadas, personas e industria, que es lo que pretende conseguir la nueva propuesta europea de responsabilidad de la inteligencia artificial precitada.

En cualquier caso, hecho este inciso, pretendo proseguir con mi análisis y reflexión sobre el régimen general de responsabilidad extracontractual sustentado en la culpa.

La culpa exigiría una intención de causar daño o, en su caso, que al menos se hubiera contemplado la posibilidad de que este daño pueda darse -aunque no sea intencionadamente- o, la falta de diligencia en la conducta para su imputación.

En consecuencia, ¿podría imputarse esta responsabilidad al propio sistema de inteligencia artificial? ¿Podría atribuírsele dicha intención o culpa?

Esta es una de las cuestiones más debatidas que se plantean en la actualidad cuando el causante del daño pudiera tratarse de un sistema dotado de inteligencia artificial más avanzada o “fuerte”, con supuesta autonomía, capacidad de aprendizaje y relativa impredecibilidad. ¿Podría ser responsable dicho sistema? ¿Podría ser sujeto de derechos y obligaciones? ¿O es inimputable?

Estas cuestiones serán objeto de análisis en los posteriores apartados, si bien, anticipo mi respuesta en sentido negativo, conforme al marco jurídico vigente y estado de la tecnología actual.

El Parlamento Europeo abordó frontalmente y por primera vez la responsabilidad de los sistemas de inteligencia artificial en la precitada Resolución de 16 de febrero de 2017[616], en la que incorporó sus recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica.

La Resolución parece orientarse en exclusiva a la "robótica" según se desprende de su título y enunciados, si bien, como más adelante analizaré, realmente incluye a la inteligencia artificial de la que pueda estar dotada un robot físico o una máquina virtual.

El Parlamento Europeo abordó el concepto "autonomía" en su Resolución, en referencia a un robot o sistema dotado de inteligencia artificial avanzada, y lo hace, como abordé en los capítulos precedentes, concibiéndola como la capacidad de tomar decisiones y aplicarlas en el mundo exterior, con independencia de todo control o influencia externos y considera que, cuanto más autónomos sean los sistemas y robots, más difícil será considerarlos simples instrumentos en manos de otros agentes como el fabricante, el operador, el propietario, o el usuario.

Es decir, el concepto de autonomía plena de un sistema de inteligencia artificial exigiría la ausencia de control externo y, consiguientemente, parece que debería incluirse entre esos controles el humano, lo que, de inicio, chocaría frontalmente con los principios y normas éticas esenciales sobre las que se pretenden construir los futuros marcos reguladores de la inteligencia artificial y la robótica en la UE, especialmente en materia ética y de responsabilidad objeto de análisis en esta investigación, que exigen la supervisión y control humano en su diseño, despliegue y aplicación.

Como he referido anteriormente, la Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[617], se aparta del concepto de autonomía y se focaliza en la prohibición de determinados sistemas inteligentes y en la regulación pormenorizada de los sistemas de inteligencia artificial que considera de alto riesgo, pero no en el resto, quedando pues

[616] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))
[617] COM (2021) 206 final 2021/0106 (COD)

cualquier otro sistema inteligente al margen de cualquier principio, norma u obligación ética o jurídica vinculante conforme al Reglamento propuesto, ni tan siquiera la supervisión y el control humano.

Es obvio que cuanto mayor libertad de actuación y de decisión tengan estos sistemas, mayor y más rápida será su capacidad de reacción eficiente frente a su entorno y estímulos[618], es decir que a mayor autonomía menor control, y también cierta impredecibilidad. Cuanto menor sea le necesidad de celeridad en su decisión o actuación, menor sería la complejidad requerida para aplicar controles sobre los mismos[619].

La supuesta "autonomía" de un sistema inteligente plantea dos cuestiones principales con especial relevancia en materia de responsabilidad:

a) La naturaleza jurídica de los robots y sistemas dotados de inteligencia artificial avanzada y la posibilidad de incardinarlos en alguna de las categorías existentes en los marcos jurídicos vigentes, incluyendo los de responsabilidad, o si debe crearse una nueva categoría por sus características jurídicas;

b) Si la normativa general vigente sobre responsabilidad es suficiente o si se requerirían normas y principios específicos que aclaren la responsabilidad jurídica de los distintos agentes y su responsabilidad por los actos y omisiones de estos sistemas, cuya causa no pueda atribuirse a un agente humano concreto.

Respecto de la primera de las cuestiones, el Parlamento Europeo propuso en la Resolución precitada de 2017, entre otras posibles opciones a valorar en el futuro, la posibilidad de crear a largo plazo una personalidad jurídica específica para los robots y sistemas dotados de inteligencia artificial, que pudiera permitir en el futuro que los sistemas autónomos más avanzados y complejos puedan ser considerados "personas electrónicas",

[618] NÚÑEZ ZORRILLA, M. C. (2019). *Inteligencia artificial y responsabilidad civil*, Reus Editorial, Madrid 2019, P. 12.

[619] MORIELLO, S. (2006). *Los robots inteligentes tendrán tres niveles de conciencia*. Levante EMV. Megatendencias. 01.01.2006. Recuperado de https://tendencias21.levante-emv.com/los-robots-inteligentes-tendran-tres-niveles-de-conciencia_a832.html#:~:text=Los%20niveles%20reactivo%2C%20deliberativo%20y,fecha%20ninguna%20definici%C3%B3n%20universalmente%20aceptada. Consultado el 28.01.2021.

responsables de reparar los daños que puedan causar, así como reconocer personalidad electrónica en aquellos supuestos en los que estos sistemas tomen decisiones autónomas inteligentes o interactúen con terceros de forma independiente.

Todo ello, conforme a la propuesta del Parlamento Europeo, acompañado de otras medidas como la creación de un registro de robots, la exigencia de un seguro obligatorio de responsabilidad civil o la creación de fondo de compensación, conforme la propia Resolución contempla.

Desde mi modesto punto de vista, entiendo que el estado de la tecnología y los marcos éticos ya altamente consensuados a nivel internacional, como los éticos y jurídicos que pretenden construirse en la UE en relación con la inteligencia artificial, impiden directamente la existencia, por el momento, de este tipo de sistemas, especialmente en la medida que deben estar sujetos al control y supervisión humana en todo momento y deben integrar la seguridad en su concepción, de modo que existiría automaticidad o automatismo pero no una autonomía y libertad plena sino, en el mejor de los casos, restringida o limitada.

Aun así, distintos expertos y autores como Hage[620], consideran que no parece haber razones de peso por las que los sistemas autónomos no deban ser responsables de lo que hacen, si bien, ello no significa que sea lo deseable. Para este experto, lo más eficaz podría ser responsabilizar a los programadores o a los usuarios de los sistemas autónomos desde el punto de vista de la finalidad, "aunque la diferencia entre los humanos y los sistemas autónomos como tales no justifique un tratamiento diferente en lo que respecta a la responsabilidad".

Sin embargo, respecto de la segunda de las cuestiones, el Parlamento Europeo consideraba ya en aquella Resolución que los robots o sistemas dotados de inteligencia artificial avanzada no pueden ser considerados responsables de los actos u omisiones que causen daños a terceros conforme al marco jurídico de responsabilidad actual, como no

[620] HAGE, JAAP. (2017). Theoretical foundations for the responsability of autonomous agents. *Artif Intentell Law* (2017). 31 de agosto de 2017. P. 270

puede ser de otra forma, ante la ausencia de personalidad jurídica o electrónica y su inimputabilidad.

En relación con estas cuestiones, el *Comité Económico y Social Europeo* (CESE) negó igualmente y de manera taxativa cualquier tipo de personalidad jurídica para los robots o la inteligencia artificial en su Dictamen de 31 de mayo, de 2017[621].

De hecho, la propuesta de valorar una futura atribución de personalidad jurídica electrónica a los sistemas basados en inteligencia artificial ha sido omitida o directamente rechazada definitivamente de la agenda regulatoria europea, que se ha evidenciado en varios de los documentos posteriores promovidos por los órganos e instituciones de la UE, especialmente en el informe *Report on Liability for Artificial Intelligence and other emerging technologies*"[622] de 21 de noviembre de 2019, que expresamente descarta esta opción, y ni tan siquiera es considerado como opción en el *Informe de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo, de 19 de febrero de 2020, sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica*[623], ni en las recientes propuestas europeas en materia de regulación de la ética y la responsabilidad civil de la inteligencia artificial, esto es, la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas, y la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial.

El abandono de esta opción, por el momento, se complementa con la nueva propuesta precitada del Parlamento Europeo en materia de responsabilidad, de 20 de octubre de

---

[621] Dictamen del Comité Económico y Social Europeo sobre la "Inteligencia *artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad*" (Dictamen de iniciativa). 2017/C 288/01. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=OJ:C:2017:288:TOC

[622] *Report on Liability for Artificial Intelligence and other emerging technologies*. EU 2019. Disponible en: https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608. Consultado el 07.03.2021.

[623] Informe de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo, de 19 de febrero de 2020, sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica (COM(2020)0064)

2020, que pretende consolidar el derecho de daños vigente, considerando que dispone de instrumentos suficientes para dar respuesta a los retos en materia de responsabilidad civil por daños que plantea la inteligencia artificial, descartando una reforma estructural profunda.

No obstante, conforme he anticipado, la aplicación del régimen general y especial de responsabilidad civil extracontractual para depurar las responsabilidades derivadas de daños causados por sistemas inteligentes puede ser complejo en atención a sus características y capacidades, y se complica aún más cuando nos encontramos ante un sistema de inteligencia artificial más avanzado o "fuerte" que pueda tomar decisiones y ejecutar acciones con "cierta" o supuesta "plena" autonomía y libertad, que aprende de forma autónoma de su experiencia y con posibles comportamientos imprevisibles fruto de sus interacciones entre sus elementos y con el entorno donde opera.

Y más si cabe en los casos de sistemas que pudieran llegar a disponer y desarrollar una inteligencia artificial con cierto grado de consciencia o autorreflexión, con capacidad para calificar la justicia o injusticia, maldad o bondad de su actos, lo que constituye un cualidad humana que, a mi juicio y a la vista del estado actual de la tecnología difícilmente podría atribuirse todavía hoy a un sistema de inteligencia artificial -para algunos imposible como refiero en otros apartados de esta investigación e incluso absolutamente innecesario en muchos casos-, sin perjuicio de que su programación pueda integrar parámetros éticos y morales que le permitan emitir juicios de valor, en contextos concretos, que determinen sus decisiones y acciones.

En definitiva, sistemas con capacidad para tomar decisiones y aplicarlas con independencia y al margen de cualquier control o influencia externa[624], o cuanto menos sin necesidad de limitarse estrictamente en su funcionamiento a las instrucciones incorporadas en su concepción, con comportamientos futuros desconocidos.

La reflexión adicional es la relativa a si los futuros marcos regulatorios de la inteligencia artificial deberían permitir la creación de estos sistemas, entre otras razones por vulnerar

[624] NÚÑEZ ZORRILLA, M.C. (2018). "Los nuevos retos de la Unión Europea en la regulación de la responsabilidad civil por los daños causados por la inteligencia artificial". *Revista Española de Derecho Europeo*, nº 66 (Abril-junio 2018). Editorial Civitas (Thomson Reuters). Navarra 2018.

los principios y normas éticas esenciales que están siendo objeto de un pretendido consenso internacional.

La imprevisibilidad, a mi juicio, vendría dada por la autonomía plena, su entrenamiento, por su capacidad de autoaprendizaje y supuesta libertad en contextos de diversidad de usos y acción, que podrían no ser previstos por sus diseñadores, desarrolladores y fabricantes para adoptar las precauciones oportunas, pero también por su interacción con su entorno, uso y posible manipulación de los mismos por sus usuarios, no previsible en función de diseño y, consecuentemente igualmente imprevista para aquéllos, como del mismo modo lo han reflejado algunos autores[625].

La pérdida de control del diseñador, desarrollador o fabricante consecuente en estos sistemas plantearía adicionalmente la cuestión de si podría imputárseles responsabilidad por los daños causados por los sistemas dentro del margen de autonomía e impredecibilidad otorgado por los mismos a aquellos, especialmente al perder, supuestamente, el control sobre el riesgo, cuanto menos parcialmente, y todo ello reiterando la necesaria reflexión consecuente sobre si la generación de este tipo de sistemas, renunciando al control sobre sus riesgos y, en definitiva, renunciando al control y supervisión humana y seguridad, sería compatible con los marcos éticos actuales y si será compatible con los marcos éticos y jurídicos futuros propuestos a nivel europeo. En mi opinión, no.

A mi juicio, nunca se debería permitir el desarrollo de una inteligencia artificial plenamente autónoma y libre por razones éticas, jurídicas y de seguridad.

A mi opinión, la irrenunciable seguridad, supervisión y control humanos, la consecuente ausencia de una autonomía real -y más en la medida que la inteligencia artificial debería operar bajo instrucciones, parámetros, objetivos, restricciones y limitaciones predefinidas

[625] PALMERINI, E. (2017). Robótica y derecho: Sugerencias, confluencias, evoluciones en el marco de una investigación europea. *Revista de Derecho Privado*, nº 32. Enero-junio 2017. Universidad externado de Colombia, Pp. 73 a 76. Disponible en https://dialnet.unirioja.es/servlet/articulo?codigo=6073053

y programadas-, así como la ausencia de conciencia, impiden actualmente considerar a un sistema de estas características como sujeto de derecho, aunque sí objeto del mismo[626].

Otra cosa distinta serían aquellos sistemas expertos en los que los diseñadores y programadores trasladarían al sistema las normas y los criterios para tomar decisiones imitando el razonamiento humano, con unas capacidades limitadas y predecibles, de modo que la responsabilidad de aquellos parece evidente.

Para finalizar, dentro de estas reflexiones iniciales sobre las características generales de la responsabilidad extracontractual, el ordenamiento jurídico español prevé distintas vías para exigir la responsabilidad por daños y obtener un resarcimiento de los mismos, y al igual que la contractual, la reparación puede producirse de distintas formas[627]:

a) Reparación específica o "in natura". La reparación en forma específica consiste en la exigencia al causante del daño que realice todo lo necesario para restablecer o reponer las cosas al estado en el que se hallaban con anterioridad a la producción del daño.

b) La reparación por equivalente consiste en atribuir a la persona afectada una indemnización en sustitución de los daños y perjuicios sufridos, cuando la reparación en forma específica no sea posible o cuando haya sido la forma elegida por la persona afectada.

c) Y la reparación en especie, mediante la entrega de bienes con valor equivalente al daño sufrido.

[626] En esta misma línea de opinión véase DÍAZ-LIMÓN, J.A. (2016). *Daddy's car:* La inteligencia artificial como herramienta facilitadora de derechos de autor, en *Revista La Propiedad Inmaterial*, nº 22, Universidad Externado de Colombia, 2016. DOI: http://dx.doi.org/10.18601/16571959.n22.06. P. 97.

[627] La jurisprudencia española (Ss STS, 3ª , 11.10.2000, A. 8628 y 11.7.1995, A. 5632) y doctrina coinciden en afirmar que cualquiera de las anteriores formas de reparación están comprendidas en las expresiones "reparar el daño" del artículo 1902 del Código Civil español e "indemnización de los daños y perjuicios" del art. 1101 del mismo. No obstante, fuera de estos casos se plantean problemas terminológicos y sustantivos.

En definitiva, la diversidad de contextos que se pueden plantear en relación con sistemas inteligentes puede limitar verdaderamente las opciones de las que pueda disponer la persona afectada para obtener un resarcimiento efectivo de los daños sufridos.

En conclusión, la inteligencia artificial, sus capacidades, características, desarrollos y usos potenciales sitúa al jurista ante nuevos retos y dificultades para analizar e incardinar en los marcos jurídicos vigentes en materia de responsabilidad, los distintos supuestos que se produzcan relacionados con daños causados por sistemas de inteligencia artificial, si bien, la evolución de los sistemas inteligentes podría enfrentar a los mismos a sistemas más avanzados y complejos, con mayor autonomía, libertad, supuesta impredecibilidad e incluso diseñados al margen de los principios y normas éticas esenciales analizados anteriormente, como el control y la supervisión humana y la seguridad, suponiendo que realmente el avance científico y tecnológico pueda permitir la creación de estos sistemas con todos estos atributos precitados, y que el derecho permita su creación.

3.3. Adecuación del marco de responsabilidad extracontractual a la IA

El marco jurídico actual regulador de la responsabilidad por daños o derecho de daños en España, a mi juicio, dispone ya de mecanismos e instrumentos que permitirían dar respuesta a distintas situaciones que se pueden plantear en materia de responsabilidad civil por el funcionamiento sistemas de inteligencia artificial, aunque no a todas, así como tratar y dar soluciones a muchos de los riesgos que comporta la inteligencia artificial, como así opinan distintos autores como Cerrillo i Martínez y Peguera Poch[628].

No obstante, anticipándome ya a mis reflexiones sobre las propuestas regulatorias europeas que analizaré a continuación, considero que la responsabilidad debería situarse en cualquier caso en la órbita de quienes tienen el control sobre el riesgo -efectivo o potencial- y su materialización en un daño en cada momento dentro del ciclo de vida del sistema inteligente y que pueden renunciar a su adecuada gestión y limitación, total o parcialmente, incluso posiblemente apartándose y vulnerando los principios y normas éticas esenciales que integran los principales marcos objeto de pretendido consenso a

[628] CERRILLO I MARTÍNEZ, A. Y PEQUERA POCH, M. (Coord). (2020). *Retos jurídicos de la inteligencia artificial*. Aranzadi S.A.U. (Thomson Reuters). Navarra 2020. P. 21.

nivel internacional, así como los contemplados en las propuestas reguladoras de la UE en materia de inteligencia artificial, objeto de análisis en esta investigación.

En este sentido y alienado con autores como del precitado Rubí[629], que se sustenta igualmente en autores como Guido Calabresi y Steve Chavell, el derecho debe atribuir responsabilidad civil al sujeto que se halla en una mejor posición para prevenir la causación de daños, lo que a su vez constituye un incentivo para adoptar y aplicar tecnologías más seguras.

Según este mismo autor, los principios de neutralidad tecnológica y de equivalencia funcional sustentan una postura conservacionista de las reglas, instituciones y principios tradicionales que conforman el actual régimen de la responsabilidad civil en España, sin necesidad de acudir a soluciones disruptivas. Precisamente, esta perspectiva fue la acogida inicialmente por el Parlamento Europeo para la elaboración de su Propuesta de Reglamento relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial, de 20 de octubre de 2020.

Sin embargo, considero que esta postura no permite dar una adecuada solución a todos los supuestos que pueden derivarse en función de la tipología, capacidades, características, contexto y concretas aplicaciones de cada sistema inteligente, como he manifestado anteriormente, y así lo significa este autor al expresar sus dudas que de que pueda ofrecer una respuesta única y válida para todos los supuestos.

En mi opinión, cada sistema de inteligencia artificial puede tener unas características y capacidades iniciales singulares, que además pueden ser cambiantes y evolutivas, especialmente ante las capacidades de autoaprendizaje y reprogramación, pero, además, los sectores en los que un sistema sea usado pueden ser muy diversos, las aplicaciones del mismo pueden ser muy dispares y su entrenamiento e interacción con el contexto muy distinta, por lo que se plantearán contextos, situaciones, riesgos y daños en todo su ciclo de vida que podrían requerir soluciones adecuadas no idénticas, por lo que siguiendo mis reflexiones sobre este tema, el sistema actual de responsabilidad civil extracontractual no

[629] RUBÍ, A. (2020). "Retos de la inteligencia artificial y adaptabilidad del derecho de daños", en CERRILLO I MARTÍNEZ, A. Y PEQUERA POCH, M. (Coord). *Retos jurídicos de la inteligencia artificial*. Aranzadi S.A.U. (Thomson Reuters). Navarra 2020. P. 60.

está preparado para dar una solución única a todos los problemas derivados de los sistemas inteligentes actuales, y mucho menos a los futuros más evolucionados y con más capacidades.

Las capacidades y factores a considerar son muy diferentes, entre otras cosas, el sector donde operen, el grado de autonomía y de intervención humana, la concurrencia de negligencia humana[630] en su funcionamiento o uso, el nivel de riesgo, interrelación con otras tecnologías, vulnerabilidades de *software*, la imprevisibilidad de errores en su funcionamiento y conectividad, su funcionamiento estadístico o el sesgo.

Los daños derivados de sistemas inteligentes aplicados a sectores como el de ocio, la domótica, la sanidad, la conducción autónoma, los *robo-advisors* o las noticias automatizadas pueden ser muy diferentes.

Sin embargo, Rubí plantea la utilidad de analizar sectores o tecnologías específicas para valorar el grado de adaptabilidad del "derecho de daños" vigente o para, en su caso, valorar cambios más profundos, si bien, como he referido, este autor sostiene que los niveles de flexibilidad, adaptabilidad y ductilidad del derecho de daños son suficientes para afrontar los retos jurídicos que plantea la causación de daños por sistemas de inteligencia artificial y, en consecuencia, "no hay necesidad de reformas profundas o de un cambio de paradigma jurídico en el sistema de responsabilidad civil extracontractual".

Conforme argumenta el autor citado, el régimen de responsabilidad civil extracontractual español vigente está lo suficientemente evolucionado para proporcionar instrumentos sólidos para afrontar supuestos complejos de causación de daños, entre otros, los criterios de imputación objetiva, la inversión de la carga de la prueba, la solidaridad impropia o la aplicación de responsabilidad *cuasi objetiva* a la que he aludido anteriormente, que reforzarían la posición de la víctima.

Y además de todo ello, en relación con otros instrumentos del ordenamiento jurídico -ya existentes o de nueva creación-, se podrían disponer de mecanismos adicionales para

[630] SELBST, A. D. (2020). Negligence and AI's Human Users. *100 Boston University Law Review* 1315 (2020). Facultad de Derecho de UCLA. Documento de investigación de derecho público Núm. 20-01. Disponible en SSRN: https://ssrn.com/abstract=3350508

compensar a las víctimas, como los seguros de responsabilidad civil, la constitución de fondos o la regulación administrativa de homologación o autorización relacionada con la seguridad de nuevos productos.

Comparto este posicionamiento parcialmente, pero no debemos olvidar que algunos de estos instrumentos son de despliegue o eficacia *a posteriori*, por ejemplo, la interpretación y aplicación *cuasi objetiva* de la responsabilidad a nivel judicial, lo que comporta inseguridad jurídica inicial para los distintos agentes involucrados en el sector de la inteligencia artificial, lo que podría, de un lado, afectar a la innovación, inversión y desarrollo de este sector y, de otro, al despliegue y aplicación efectiva de la inteligencia artificial ante la posible falta de seguridad y confiabilidad para la sociedad en general.

Otros autores apuntan la posibilidad de asignar la responsabilidad a la parte que pueda evitarlos de forma más económica como Calabresi o Panezi[631], de modo que "en ausencia de certeza sobre quién es esa parte o actividad, los costes deberían recaer en la parte o actividad que pueda, con los menores costes de transacción, actuar en el mercado para corregir un error en los derechos, induciendo a la parte que puede evitar los costes sociales de forma más barata a hacerlo".

En base a todo ello y dado que los regímenes vigentes de responsabilidad no darían solución a todos los supuestos que puede plantear la inteligencia artificial por los motivos expuestos, considero que la definición previa de los requerimientos exigibles a la misma y de los marcos de responsabilidad aplicables aportarán transparencia y confianza para todos, para lo que igualmente considero necesario el establecimiento de nuevos marcos jurídicos que complementen adecuadamente los actuales en relación con los sistemas inteligentes y que definan con claridad el régimen aplicable, los requerimientos y obligaciones exigibles y las consecuencias derivadas de su incumplimiento, lo que sin duda evitará y/o minimizará la aplicación de regímenes reactivos para la exigencia de responsabilidades y, en caso de aplicación, se facilitará la imputación e identificación de

[631] PANEZI, A. (2021). "IA: un enfoque ecosistémico para gestionar el riesgo y la incertidumbre", en García Mexia, P. y Pérez Bes, F. (Eds); Panezi. A. (Coord). (2021). *Artificial Intelligence and the Law*. Ed. La Ley (Wolters Kluwer). 2021. Edición electrónica (SMARTECA).

incumplimientos y niveles de diligencia a los efectos de depuración de las responsabilidades.

De este modo, comparto que las reglas e instrumentos esenciales que conforman el marco de responsabilidad extracontractual vigente en España puede ser aplicables para abordar la depuración de responsabilidades de manera eficaz respecto de daños causados por sistemas de inteligencia artificial en determinados contextos, pero no es un marco plenamente actual ni adecuado para abordar la totalidad de supuestos que se pueden suscitar en la práctica.

En consecuencia, estas reglas e instrumentos deberían ser revisados, adaptados, complementados, integrados y predefinidos con absoluta claridad y, en cualquier caso, interpretados, a la vista de las complejidades propias de las tecnologías subyacentes a los sistemas inteligentes y con la finalidad de garantizar un resarcimiento efectivo a la persona afectada.

A modo de ejemplo, según el marco vigente, la carga de la prueba de la relación de causalidad recae en la persona que pretenda la reparación del daño -salvo inversión de la misma-, para lo que un sistema inteligente basado en *software* debería ser capaz de recopilar toda la información relativa a su funcionamiento, uso y contexto en que se tomó una decisión, se llevó a cabo una conducta -acción u omisión- o se produjo un accidente, por lo que si todo ello debe ser registrado, no alterado y disponible por el sistema, se facilitaría la acreditación de las causas del daño. En este sentido, la Propuesta de Reglamento UE de 21 de abril de 2021, como luego referiré, establece como requerimiento de los sistemas inteligentes de alto riesgo la transparencia, información, trazabilidad y registro, lo que contribuiría a ello siempre y cuando se vea complementado con un marco jurídico que prevea la no necesidad de su prueba o la inversión de la carga probatoria *ab initio*, en la medida que esta nueva propuesta no contempla estos aspectos de responsabilidad.

De otro modo, sería muy dificultoso su acreditación, ante la complejidad de los mismos y ausencia de requisitos y obligaciones jurídicas definidas, máxime en caso de aprendizaje profundo, lo que justifica la necesidad de nuevos marcos jurídicos que, de un lado, consideren estas complejidades, que requieran estas exigencias de transparencia,

información, explicabilidad, registro de actividad, rendición de cuentas o auditabilidad y que complementen los regímenes de responsabilidad y, de otro, que regulen nuevos derechos, como el de acceso.

Con este propósito, resultaría de gran valor garantizar a través de una norma sustantiva o procesal -preferiblemente la primera en mi opinión, con la finalidad de dejar claras "las reglas de juego" y propiciar la seguridad jurídica *ab initio*- el derecho de acceso de las víctimas a esta información y/o autoridades competentes, partiendo del hecho de que se exija su registro protegido -se trate o no de datos personales-, disponible, inalterable, accesible y legible en determinadas circunstancias legalmente previstas y por personas legitimadas para ello -víctimas, autoridades competentes o tribunales-.

En este sentido ya tenemos referencias legales en otros ordenamientos jurídicos. Por ejemplo, en Alemania las víctimas de un accidente de tráfico tienen derecho a acceder a la caja negra del vehículo autónomo y en California (EE.UU.), los fabricantes de los vehículos autónomos tienen la obligación de remitir un informe completo de cada colisión a las autoridades públicas[632]. De otro modo, se suscita el problema de probar la relación de causalidad.

Y, es más, a mi juicio, dicho derecho acceso debería de ir acompañado, cuanto menos, de una consecuencia o sanción automática, sencilla y directa para el caso de no ser atendido o no ofrecer la información completa, como puede ser el hecho de considerarse cierto el hecho imputado, sin perjuicio del régimen administrativo sancionador que pueda y, a mi juicio, que deba acompañar a los futuros marcos reguladores por el incumplimiento de obligaciones dimanantes de los mismos.

Este derecho de acceso podría servir de incentivo para que los fabricantes de sistemas inteligentes faciliten la transparencia e información de la inteligencia artificial, atendiendo de este modo a dos de sus principales riesgos, la falta de transparencia y la responsabilidad.

[632] CERRILLO I MARTÍNEZ, A. Y PEQUERA POCH, M. (Coord). "*Retos jurídicos de la inteligencia artificial*". Aranzadi S.A.U. (Thomson Reuters). Navarra 2020. P. 65.

La nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, analizada en el anterior capítulo, establece en este sentido requisitos específicos para los sistemas de inteligencia artificial de alto riesgo, entre otros, la exigencia de análisis de riesgos y de seguridad, la descripción técnica de diseño y funcionamiento, la transparencia y la información, el registro de eventos (logs) y la trazabilidad, además de los requerimientos específicos de transparencia e información para concretos sistemas no considerados de alto riesgo.

Además, un ejemplo de derecho de acceso lo podemos encontrar ya en el ordenamiento jurídico vigente, en particular, en el Reglamento General de Protección de Datos[633], pero con un alcance más específico y concreto.

Asimismo, dado que en determinadas situaciones y ante la complejidad de los sistemas inteligentes, entre otros factores, los marcos vigentes podrían no garantizar el derecho al resarcimiento de las víctimas de los daños, en estos supuestos se deberían plantear otros instrumentos sencillos, por ejemplo y como he referido anteriormente, la inversión de la carga de la prueba respecto de la relación de causalidad.

Esta fue una de las propuestas formuladas por el *Grupo de expertos europeo constituido en materia de responsabilidad y nuevas tecnologías* en su informe *Report on Liability for Artificial Intelligence and other emerging technologies*[634] de 21 de noviembre de 2019, en particular, cuando se valore conjuntamente la concurrencia de los siguientes factores:

a) La probabilidad de que la tecnología en cuestión haya al menos contribuido al daño;

[633] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento General de Protección de Datos o RGPD/GDPR).

[634] "*Report on Liability for Artificial Intelligence and other emerging technologies*". EU 2019. Disponible en: https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608. Consultado el 07.03.2021.

b) La probabilidad de que el daño haya sido causado por la tecnología o por alguna causa dentro de la misma esfera;

c) El riesgo de un defecto conocido dentro de la tecnología, aunque su impacto como causa real no sea evidente por sí mismo;

d) El grado de trazabilidad *ex post* y la inteligibilidad de los procesos subyacentes a la tecnología que pueden haber contribuido a la causa;

e) El grado de accesibilidad *ex post* a los datos recopilados y generados por la tecnología y su comprensión;

f) El tipo y grado del daño potencial realmente causado.

No obstante, considero necesario establecer unas bases más amplias y sólidas de regulación de las posibles acciones u omisiones, que proporcionen seguridad jurídica a todas las partes, atendida la infinita diversidad de tecnologías, capacidades, elementos, aplicaciones, sectores, usos, contextos y demás factores precitados, y bajo este prisma se ha construido la primera propuesta regulatoria europea en materia de responsabilidad civil en materia de inteligencia artificial que será objeto de análisis en posteriores apartados.

De nuevo, considero necesario apostar por propuestas regulatorias adaptativas, evolutivas, "responsive", flexibles y dúctiles, diseñadas sobre tendencias más que sobre novedades, sobre realidades actuales o potenciales, y no sobre cuestiones muy lejos de toda realidad en la actualidad.

En mi opinión, las profecías y vaticinios sobre la evolución y desarrollo de la inteligencia artificial hasta llegar a alcanzar e incluso superar la inteligencia humana y los profundos retos y riesgos que ello podría conllevar incluso para la propia raza humana no deben obviarse, pero tampoco formar parte de la agenda legislativa actual, la cual debe estar centrada en crear un marco sólido sustentado en la ética, la seguridad y la responsabilidad, que vaya adaptándose al desarrollo de la inteligencia artificial.

La aplicación de los regímenes generales o especiales, que luego abordaré, en materia de responsabilidad por daños todavía se complica más, como he referido anteriormente, ante

futuros sistemas dotados de inteligencia artificial más avanzada y "fuerte", con supuesta autonomía y capacidad de autoaprendizaje, sobre los que se ha planteado la posibilidad de dotarlos de personalidad jurídica como una de las posibles opciones para garantizar una respuesta adecuada a las responsabilidades derivadas de su diseño, despliegue o aplicación.

Según algunos expertos, estos sistemas podrían llegar a igualar o superar los niveles humanos de inteligencia[635] en un futuro, creando una inteligencia artificial "fuerte" o incluso la denominada "súper inteligencia" o, capaz de retroalimentarse y perfeccionarse a sí misma y crear otros sistemas inteligentes, ¿quién responderá de los daños causados por la misma?

Conforme he expuesto en anteriores apartados de esta investigación, distintos investigadores consideran que la inteligencia artificial "general" -AGI por sus siglas en inglés-, dentro de décadas será capaz de realizar todas las tareas cognitivas a nivel sobrehumano.

Respecto de las predicciones sobre cuándo se alcanzará esa "súper inteligencia" y si se alcanzará, algunos expertos como Tegmark[636] destacan las discrepancias entre los expertos y considera que simplemente no lo sabemos, puede que llegue en décadas, siglos o nunca.

Y sobre el temor de que la inteligencia artificial pueda volverse malvada o actuar con el propio ser humano, este experto considera que la preocupación real se centra más en que la inteligencia artificial llegue a ser competente y tenga objetivos no alineados con los humanos, y que lo que debemos hacer es planificar con tiempo hasta que llegue para hacer que sea segura.

En este sentido, cobra más vigencia que nunca algunas de las frases más célebres de Irwin Corey: *"Si no enderezamos el rumbo pronto, acabaremos allí donde estamos yendo".*

---

635 NÚÑEZ ZORILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*, Editorial Reus, Madrid 2019. P.13.

636 TEGMARK, MAX (2018). *Vida 3.0: Qué significa ser humano en la era de la inteligencia artificial*. Penguin Random House Grupo Editorial. 2018. P. 62

Es más, todo ello podría desembocar en la denominada "explosión" de la inteligencia y un desafío para la propia humanidad[637].

Algunos científicos, como he referido al abordar los capítulos anteriores, han comenzado a plantear la posibilidad de sistemas robóticos dotados de inteligencia artificial que encuentren un nivel de consciencia semejante al de los humanos, que podrían incluso considerar a los mismos como innecesarios y desembocar en el denominado "punto de singularidad", como hecho futuro probable, aunque a ocurrir en un momento todavía no definido[638].

Otros expertos consideran que la consciencia es una característica exclusivamente humana, una propiedad que únicamente puede derivar del cerebro humano y no de uno electrónico.

No obstante, expertos como Lacruz Mantecón[639], consideran que el cerebro humano es uno de los mecanismos computadores que pueden sustentar una consciencia, pero un cerebro electrónico lo suficientemente complejo también podría hacerlo. Esto podría permitir la emisión de juicios de valor y actuar en consecuencia[640].

Según Marvin Minsky, científico estadounidense cofundador del *Laboratorio de inteligencia artificial y Ciencias de la Computación* del Instituto Tecnológico de Massachusetts (MIT) y uno de los considerados "padres de la inteligencia artificial" junto con John McCarthy, ha venido considerando que, efectivamente, hasta la fecha no se ha diseñado un ordenador que sea consciente de lo que está haciendo, "pero la mayor parte del tiempo, nosotros tampoco lo somos". Y nos ha dejado algunas afirmaciones para la reflexión y ahora más que nunca: "Cuando los ordenadores tomen el control, puede que

[637] TERRONES, A.L. (2018). "Inteligencia artificial y ética de la responsabilidad". *Cuestiones de filosofía*. Vol. 4. Nº 22. Enero-Junio, 2018, Pp. 152 y 153.
[638] GARCÍA-PRIETO, J. (2018). "¿Qué es un robot?". En BARRIO ANDRÉS, M. (Dir.). *Derecho de los robots*. Wolters Kluwer, Madrid 2018. Pp. 31-59.
[639] LACRUZ MANTECÓN, M. L. "Potencialidades de los robots y capacidades de las personas". En ROGEL VIDE, C. (Coor.). *Los robots y el Derecho*. Editorial Reus, Madrid 2018. P. 52.
[640] GÓMEZ-RIESCO, J. (2018). "Los robots y la responsabilidad civil extracontractual". En BARRIO ANDRÉS, M. (Dir.). *Derecho de los Robots*. Wolters Kluwer. Madrid 2018.

no lo recuperemos. Sobreviviremos según su capricho. Con suerte, decidirán mantenernos como mascotas"[641], o "¿Heredarán los robots la Tierra? Sí, pero serán nuestros hijos"[642].

Otros expertos consideran que la consciencia humana podría incorporarse a una máquina en el futuro y no dentro de tanto tiempo. Hay distintos proyectos centrados en la investigación del cerebro y la percepción de modelos de pensamiento neuronal, entre otros el proyecto *Human Brain Project*[643] de la UE -orientado a reproducir tecnológicamente las características del cerebro humano y que ya ha permitido avances en la computación neuromórfica y la creación de herramientas, como las de ayuda para la navegación y guía de la cirugía cerebral en pacientes con epilepsia farmacorresistente-, y el proyecto "BRAIN"[644] de EE.UU.

Asimismo, otros científicos de la talla de Stephen Hawking, al que he hecho referencia en distintos capítulos, han venido advirtiendo de sus peligros, lo que comparto en muchos aspectos y no en base a lo que podamos haber visto en la ciencia-ficción, sino en lo que ya hoy está empezando ocurrir.

Según Hawking, el desarrollo de una inteligencia artificial fuerte significaría el fin de la humanidad, pues evolucionaría a un ritmo que los humanos no podrían seguir. Del mismo modo funcionaría por sí sola y se rediseñaría cada vez más rápido, de modo que los seres humanos, limitados por la lenta evolución biológica, no podrían competir con ella y serían superados.

Durante toda una vida dedicada a la investigación y la divulgación científica, el físico inglés nos dejó también algunas frases para su reflexión y ahora más que nunca: "Los robots podrían llegar a tomar el control y se podrían rediseñar a sí mismos" o "La humanidad tiene un margen de mil años antes de autodestruirse a manos de sus avances científicos y tecnológicos".

El debate ético y jurídico es inevitable y debería plantearse ya para anticiparnos a los problemas, actuar sobre tendencias y no sobre novedades, cuestionarnos de antemano si

[641] Revista "*Life*". Núm. Noviembre de 1970.
[642] Revista "*Scientific American*". Núm. Octubre 1994.
[643] Recuperado de: https://www.humanbrainproject.eu/en/. Consultado el 02.03.2021
[644] Recuperado de: https://braininitiative.nih.gov/. Consultado el 02.03.2021.

debería permitirse la creación de este tipo de sistemas y dejar su gobierno, uso y explotación en manos privadas, así como a la creación de marcos éticos vinculantes y revisión de los marcos jurídicos existentes para regular adecuadamente estos sistemas y su responsabilidad, con los límites y condiciones que se establezcan.

El derecho creado por el hombre se basa y construye sobre sus valores humanos, éticos y morales. La inteligencia artificial creada por el ser humano debería construirse sobre la base de dichos valores.

Un sistema de inteligencia artificial más avanzada de este tipo podría causar daños por decisiones o conductas -acciones u omisiones- autónomas que no serían consecuencia directa de las instrucciones dadas por el diseñador, programador o el fabricante ni por el operador o usuario, sino consecuencia de la supuesta autonomía conferida, grado de libertad otorgado, capacidad de autoaprendizaje, margen de improvisación, procesamiento, impredecibilidad y otras capacidades otorgadas a dicho sistema, por lo que en cualquier caso, este podría no ser considerado un mero instrumento o medio para la realización de la acción causante del daño por un tercero, pero tampoco el sujeto de derecho causante del mismo, conforme al marco jurídico vigente en materia de responsabilidad en la UE y en España, al carecer de personalidad jurídica y no poder ser titular de derechos y obligaciones. De manera consecuente, ¿debemos concluir que en estos contextos no existiría responsable y no se repararía el daño conforme al marco legal vigente?

Por último y para concluir estos aspectos generales de este régimen de responsabilidad de responsabilidad extracontractual, en los supuestos con componentes transfronterizos donde no participan consumidores, se pueden plantear cuestiones relativas al marco de responsabilidad extracontractual a aplicar, en especial, en relación con la posible aplicación del Convenio de la Haya sobre ley aplicable a la responsabilidad por productos, de 2 de octubre de 1973[645], el cual determina el Derecho aplicable a partir de la

[645] Instrumento de ratificación del Convenio sobre la Ley aplicable a la responsabilidad por productos, hecho en La Haya el 2 de octubre de 1973. BOE 25.01.1989. Pp. 2054-2056

combinación de distintos elementos, como el lugar del hecho lesivo, domicilio de la víctima, domicilio del fabricante o productor y lugar de la adquisición del producto.

En conclusión, en mi opinión, el régimen general de responsabilidad extracontractual no dispone de mecanismos adecuados para dar respuesta a todas estas cuestiones que puede plantear la inteligencia artificial en relación con los daños causados derivados de su funcionamiento.

3.4. Sujetos responsables

La regla general sobre la que se sustenta la responsabilidad extracontractual es que el autor material del daño es el sujeto responsable de su reparación, si bien, el Código Civil español contempla supuestos en los que hace responsable del daño a personas o entidades que no lo causaron con el fin de asegurar el resarcimiento total de la víctima *-restitutio in integrum-*, en especial, los supuestos por hecho ajeno que analizaré con posterioridad.

Los sistemas de inteligencia artificial se hayan compuestos por múltiples elementos y tecnologías, incluyendo el *hardware* -carcasas, dispositivos físicos, robots, sistema de impresión o sensores-, *chips*, *software*, algoritmos, datos -de entrada, de entrenamiento y de salida-, bases de datos o los de conectividad, que conforman sistemas con determinadas características y capacidades que pueden incluir su autoaprendizaje o incluso su autoprogramación.

Todo ello involucra a distintos sujetos durante todo el ciclo de vida de un sistema con distinto grado de control sobre el sistema y sus distintos componentes -diseñadores y desarrolladores de *software* y *hardware*, integradores, fabricantes de *hardware*, comercializadores, proveedores de datos, responsables y encargados del tratamiento de datos, operadores de redes de conexión, entrenadores, operadores, propietarios o usuarios-, y la correlativa dificultad para identificar que sujeto está en una mejor posición para prevenir y controlar el riesgo real o potencial asociado al sistema y su materialización en un daño.

El régimen general de responsabilidad civil atribuye la responsabilidad al sujeto que por acción u omisión cause daño a otro, interviniendo culpa o negligencia, por lo que podría recaer en cualquiera o varios de los sujetos identificados anteriormente, aisladamente o en concurrencia de culpas.

Sin embargo, como posteriormente abordaré, el marco jurídico específico que regula la responsabilidad por productos defectuosos a nivel de la UE, la Directiva 85/374/CEE, la atribuye inicialmente al productor, esto es, a la persona que fabrica un producto acabado o que se presente como productor poniendo su nombre, marca o cualquier otro signo distintivo en el producto, considerándose como tal quien fabrica un elemento o componente que integre el sistema.

En cualquiera de los casos, en la práctica, esta concurrencia de sujetos puede plantear serias dificultades a la persona afectada para determinar la responsabilidad individual inicial de entre los sujetos potencialmente responsables.

De inicio, una de las soluciones a estas cuestiones podría ser la atribución de una responsabilidad solidaria para procurar garantizar el derecho de resarcimiento de la persona afectada aplicable en el régimen de responsabilidad extracontractual vigente en España en el caso de pluralidad de agentes con concurrencia causal única o diversa, de conformidad con los artículos 1902 en relación con el 1.137, 1.144 y 1.445 del Código Civil español, conforme se recoge en la jurisprudencia[646].

No obstante, podrían barajarse otras opciones en los futuros marcos que revisen y complementen los regímenes de responsabilidad actuales, en particular y entre otras, la identificación e imputación inicial al sujeto que pueda estar en la mejor posición para controlar el riesgo, conforme indicaba anteriormente y propuso el *Grupo de expertos europeo constituido en materia de responsabilidad y nuevas tecnologías* en su informe precitado anteriormente, que identifica al operador de la tecnología -aquél que controla

[646] Sentencias del Tribunal Supremo de 31.01.2007 (Rec. 236/000) que recoge otras previas del Alto Tribunal de 7.11.2003 y 8.05.2006. A destacar el análisis de la jurisprudencia en materia de solidaridad impropia de Sánchez Gálvez. SANCHEZ GALVEZ, F. (2020). "Jurisprudencia sobre la solidaridad impropia en la responsabilidad civil extracontractual". *Revista Acta Judicial.* N5. Pp. 10-35. Recuperado a partir de https://letradosdejusticia.es/revistaactajudicial/index.php/raj/article/view/40.

el riesgo generado por la tecnología y se beneficia de su funcionamiento- como potencial sujeto responsable.

A mi juicio, la solidaridad me parece una postura adecuada en garantía del perjudicado al margen del posterior derecho de repetición entre los distintos responsables. No obstante, considero que los futuros marcos podrían ir más allá para situar la responsabilidad en la esfera de quien pueda estar en la mejor posición para controlar el riesgo, sea de manera efectiva o potencial, siendo este último escenario en el que podría situarse habitualmente al fabricante/productor, el diseñador y/o programador (por los defectos imputables al mismos) o propio propietario (por ejemplo, en caso de falta de mantenimiento preceptivo) u operador (ante una acción correspondiente al mismo, ejecutada incorrectamente o no ejecutada).

La Propuesta de Reglamento sobre responsabilidad civil derivada de la inteligencia artificial[647] del Parlamento Europeo, que abordaré más adelante, apuesta también por dicha solidaridad.

Adicionalmente a todo ello, conforme al marco jurídico de responsabilidad vigente en España, la persona que sufra daños derivados del funcionamiento o uso de sistemas inteligente podría utilizar el marco general de responsabilidad extracontractual o, en su caso, el específico para productos o servicios defectuosos frente a los sujetos responsables conforme a cada uno de los marcos precitados. Estos marcos podrían responder adecuadamente a distintas situaciones, si bien, el mayor problema en materia de responsabilidad se plantea, como he expuesto, en sistemas más avanzados es los que el daño tenga su origen en circunstancias posteriores, supuestamente, inesperadas, impredecibles y fuera del supuesto control de diseñadores, desarrolladores, fabricantes o incluso de operadores y usuarios, es decir, no en un defecto de fabricación sino en un funcionamiento o decisión autónoma.

El Derecho español, conforme he referido y destaca igualmente Núñez Zorrilla[648], parte de una responsabilidad de la persona basada en los juicios de previsibilidad y de

[647] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial
[648] NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Op. cit. P. 23.

evitabilidad, de modo que el agente causante del daño deberá responder sólo cuando el daño causado era predecible en el momento de la puesta en servicio del sistema en cuestión, y sólo en la medida que efectivamente hubiera podido ser predecible, podría ser evitado.

Los sistemas de inteligencia artificial más avanzada o fuerte podrían ir más allá de las órdenes e instrucciones de sus creadores, operadores y usuarios en el marco de esa autonomía e independencia otorgada, y realizar acciones no previsibles para éstos y que causen daños a terceros.

No obstante, en mi opinión, considero que en cualquier caso, estos sistemas necesariamente actuarían con un margen de autonomía, libertad, improvisación o imprevisibilidad "conferido" o "permitido" cuanto menos[649] por alguien -incluso con capacidad para reprogramarse a sí mismos u otros sistemas-, conferidos por su creador, esto es, su diseñador, programador o fabricante/productor, que son los que determinan su programación, objetivos, permisos, autorizaciones, instrucciones, funciones, criterios, atributos, capacidades, condiciones, restricciones y limitaciones y que deberían mitigar los riesgos potenciales asociados de los sistemas asociados a todo ello -al menos los razonablemente previsibles-, por lo que la responsabilidad por daños derivados de decisiones y acciones al margen de las órdenes o instrucciones dadas, cuanto menos de buena parte de ellas, debería orientarse hacia los mismos, por ser quienes tenían los conocimientos técnicos y profesionales, así como el control efectivo o potencial sobre el riesgo en origen -no sobre el hecho posterior dimanante de la inexistencia de dichos controles y la materialización del riesgo-, y no lo gestionaron adecuadamente o dispusieron de todos los medios a su alcance para prevenir el riesgo potencial razonablemente previsible (o permitir su adecuada gestión por el operador o usuario) y su materialización.

[649] No debería permitirse la creación de sistemas plenamente autónomos sin sometimiento al control y supervisión humana, como he referido en los apartados anteriores, en la medida que atenta contra los principios y normas éticas esenciales más extendidos y consensuados a nivel internacional y que se recogen en la *Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas* de octubre de 2020.

En consecuencia, por dichos motivos, considero que no debería considerarse la concurrencia de estos factores como una causa de plena exención de responsabilidad que afirman autores como la precitada Núñez Zorilla[650].

Según la misma, conforme al régimen tradicional de responsabilidad del fabricante de la que hablaré posteriormente, la imprevisibilidad, concebida como riesgos desconocidos para el estado de conocimientos de la ciencia, sería una causa de exención de la misma -cuestión que evidencia la necesidad de que los nuevos marcos que revisen y complementen los regímenes actuales de responsabilidad consideren los avances tecnológicos-, si bien, esta supuesta exención general es una cuestión sobre la que estoy parcialmente de acuerdo por razones de seguridad jurídica conforme al marco actual y en base a los argumentos anteriormente expuestos, dado que, de otro modo, comportaría una exención total de responsabilidad al fabricante por un hecho dañoso que no podría ser imputable por quedar, supuestamente, al margen de su control.

En mi opinión, los sistemas de inteligencia artificial deben incorporar en su arquitectura y programación la ética, la seguridad y el cumplimiento normativo en su diseño, lo que comporta establecer niveles de autorización, seguridad y controles durante todo su ciclo de vida acordes a las características y capacidades de los mismos, especialmente autonomía, libertad, autoaprendizaje, autoprogramación y margen de improvisación, ya sean decisiones o acciones adoptadas por el propio sistema o por la orden o instrucción de su operador o usuario. Es decir, "líneas rojas" impuestas en diseño y concepción, no sólo exigibles por las buenas praxis de la industria, sino por marcos éticos, legales y de seguridad claramente definidos y precisos, y a su vez adaptables a cada contexto, que nunca deben traspasarse si queremos una inteligencia artificial segura y bajo el control y supervisión humana.

Es más, la responsabilidad de fabricantes, diseñadores y programadores ante la posible falta de integración efectiva de un sistema de prevención de determinas conductas y sus consecuencias, podría construirse en los necesarios nuevos marcos reguladores y con las debidas distancias, como la responsabilidad penal de las personas jurídicas en el ámbito

[650] NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*, Reus Editorial, Madrid 2019, P. 19.

penal -*Compliance penal*-, por los delitos que comenten las personas a su cargo, de la que podrían ser exoneradas o verla atenuada si establecieron un sistema de prevención de delitos y lo implementaron de forma efectiva, acompañado de su correspondiente seguimiento y formación a los agentes implicados.

En consecuencia, la implementación efectiva de sistemas de gestión y prevención de riesgos en el diseño, desarrollo y producción de sistemas inteligentes deberían identificar los riesgos potenciales, obviamente conforme al estado de la tecnología, y establecer las medidas adecuadas para su mitigación, al objeto de evitar el daño y, en caso de producirse, valorar su eficacia como causa limitadora o eximente de responsabilidad.

Sin duda es una cuestión controvertida que exige una amplia y profunda reflexión de los marcos actuales de responsabilidad aplicables a la inteligencia artificial y la base para la definición de los futuros marcos reguladores.

La casuística puede ser muy variada, el contexto complejo -considerando tipología, características, sector, uso y aplicaciones, entre otros factores- y las cuantías indemnizatorias por daños podrían resultar astronómicas, por lo que lógicamente existe preocupación e interés también por las empresas tecnológicas por clarificar el marco, especialmente excluyendo o, en cualquier caso, limitando su responsabilidad.

Como he apuntado anteriormente y analizaré posteriormente con más detalle al abordar la personalidad jurídica de los robots en el capítulo VII de esta investigación, algunos expertos y autores han planteado, al abordar la responsabilidad por daños, la posibilidad de otorgar una personalidad jurídica a sistemas de inteligencia artificial y robots dotados de la misma, similar a la que se otorga a las sociedades mercantiles, para poder imputarles la responsabilidad civil, si bien, no comparto en la actualidad esta opción, conforme he expuesto anteriormente y abordaré con más detalle posteriormente, siendo además una cuestión, por el momento y como he referido, rechazada igualmente por el Comité Económico y Social Europeo en su precitado Dictamen de 31.05.2017, considerando que "*la comparación con la responsabilidad limitada de las sociedades no es válida, puesto que el responsable en última instancia es siempre una persona física*".

La cuestión esencial sobre la que debemos reflexionar con más profundidad es la conceptualización de los robots y los sistemas dotados de inteligencia artificial avanzada que construyamos hoy y en el futuro, esto es, como sujeto o como objeto de derecho, especialmente ante unas capacidades que necesariamente deben estar dirigidas con un fin legítimo y lícito, ser seguras y sujetas en todo momento al control y supervisión humana. A mi juicio, hoy sólo podemos considerar la segunda opción.

Actualmente, se está trabajando en proyectos para la generación de sistemas de inteligencia artificial "fuerte" pero no son todavía una realidad desplegada, si bien, en un primera fase, nos podremos encontrar ante sistemas dotados de inteligencia artificial más avanzada que no dispongan de una autonomía real o plena ni, desde luego, inteligencia y conciencia plena -lo que según muchos no ocurrirá nunca conforme al estado de la tecnología actual, tal y como he puesto de relieve en otros capítulos de esta investigación-, lo que de inicio y entre otros factores, imposibilitaría además valorar la posibilidad de atribuirle la consideración de sujeto de derechos y obligaciones y, por tanto, abrir la posibilidad de imputarle responsabilidad directa por los daños que ocasione.

Ante la imposibilidad de reconocer personalidad jurídica a los mismos, de imputarles responsabilidad por culpa y de considerarlos sujeto de derechos y obligaciones, el sistema clásico de responsabilidad civil actual, como he razonado, puede ofrecer algunas soluciones adecuadas en la actualidad en relación con la responsabilidad civil por el funcionamiento de los robots y sistemas dotados de inteligencia artificial, pero no a todas, sin perjuicio de considerarlos objeto de Derecho, aunque con cierta autonomía, pero siempre relativa y restringida conforme al estado de la técnica actual y normas éticas esenciales, lo cual espero sea igualmente un requerimiento jurídico taxativo en los futuros marcos reguladores, con el objetivo de garantizar la seguridad y un control y supervisión efectiva a los sistemas de inteligencia artificial.

Los futuros sistemas dotados de inteligencia artificial calificable como "fuerte" serían inimputables desde el punto de vista jurídico y, consecuentemente, no podría imputársele responsabilidad por hechos propios ni resultarían obligados al resarcimiento de las personas que sufran daños derivados de su funcionamiento.

En mi opinión, si finalmente consideramos los sistemas dotados de inteligencia artificial más avanzada como meros productos, bienes u objetos y hasta la disponibilidad de un nuevo marco de responsabilidad construida sobre el riesgo, la problemática relativa a la responsabilidad debería abordarse, en función del contexto, desde la perspectiva general de la responsabilidad civil extracontractual -en la que aquellos serían meros instrumentos o medios para la causación del daño creados por una persona física o jurídica y puesta en manos de otra para su comercialización y/o uso-, así como desde la legislación de protección del consumidor y de reparación de productos y servicios defectuosos (TRLGDCU) cuando resulte aplicable, en cuyo caso, su fundamento se situaría en la responsabilidad por riesgos o en la responsabilidad objetiva, conforme el propio Parlamento Europeo recoge en su Resolución de 16 de febrero de 2017 sobre robótica. La responsabilidad debería recaer en la órbita de quién/es incumple/n un deber de diligencia y está/n en mejor posición para evitar o controlar el riesgo. Estos aspectos serán analizados en los siguientes apartados.

En consecuencia, ante la multiplicidad de tipologías, capacidades, finalidades, contextos, aplicaciones, usos y sujetos intervinientes durante su ciclo de vida, es necesario crear un nuevo marco jurídico que revise, actualice, complemente, integre y clarifique el régimen de responsabilidad por daños causados por la inteligencia artificial.

3.5. Responsabilidad por hechos ajenos y otros supuestos

Siguiendo mis reflexiones sobre el régimen de responsabilidad civil por daños vigente en España y su posible aplicación a los daños causados por sistemas de inteligencia artificial, considero oportuno profundizar y reflexionar sobre otras posibles soluciones disponibles en el sistema general de responsabilidad civil extracontractual así como, en especial, en la denominada responsabilidad civil por hechos ajenos, por falta de cuidado de las cosas o daños provocados por animales, conforme a lo dispuesto en el Código Civil español, es decir, respondiendo por agentes que no son directamente imputables por no poder actuar de modo culposo, como ocurriría con los padres o tutores respecto de los hijos menores o de las personas con capacidad judicialmente modificada.

En este sentido, si consideramos que un sistema dotado de inteligencia artificial avanzada o "fuerte" dispondría, supuestamente, de autonomía en cuanto actuación y/o movilidad, pero sin inteligencia ni conciencia ni personalidad jurídica, o si lo consideramos un mero objeto inanimado, podríamos plantearnos la posibilidad de aplicar el régimen de responsabilidad por hecho ajeno precitado, pudiendo considerarse al mismo un "*cuasi-agente* con una responsabilidad reducida", conforme proponen algunos autores, como Asaro y Gómez-Riesco[651].

Navas Navarro[652] considera, en relación con los sistemas expertos, que concebidos como un "auxiliar" más que un "instrumento", en base a su autonomía de la que disponen que, según la misma, podría hacerlos equiparables a la propia de un ser humano en cuanto a toma de decisiones con trascendencia jurídica, podría incardinarse en el ámbito de la responsabilidad indirecta o por hecho de un auxiliar, regulada en los artículos 1101 y 1904 del Código Civil español, considerando igualmente que, en el caso de obligaciones de hacer o de dar en el ámbito contractual en cuyo cumplimiento se empleen sistemas expertos, se debería articular una responsabilidad objetiva.

Y, del mismo modo, si pudiéramos sostener una interpretación extensiva del marco vigente, podríamos incluso valorar la posible aplicación de la responsabilidad por daños causados por animales, supuesto en el que responden las personas que tenían que cuidar de los mismos y evitar que causen el daño, la cual, además, es objetiva.

Vázquez de Castro considera que, si consideramos asimilables estos supuestos a la responsabilidad de los poseedores o usuarios de animales prevista en el artículo 1905 del Código Civil español, podría imputarse la responsabilidad al poseedor o propietario del mismo, derivada de la propiedad o titularidad del sistema o dispositivo dotado de inteligencia artificial o de los derechos de propiedad intelectual o industrial sobre el mismo[653].

---

[651] GÓMEZ-RIESCO, J. (2018). "Robots y la responsabilidad civil extracontractual" en BARRIO ANDRÉS, M. (Dir.), *Derecho de los Robots*. Wolters Kluwer, Madrid 2018. P. 117.

[652] NAVAS NAVARRO, S. (2017). "Derecho e inteligencia artificial desde el diseño. Aproximaciones", en Navas Navarro, S. et al. *Inteligencia artificial, tecnología, derecho*. Valencia, Tirant lo Blanch, 2017. P.44.

[653] VÁZQUEZ DE CASTRO, E. (2020). "Aproximación a la responsabilidad derivada de los riesgos de la inteligencia artificial en Europa". En Solar Cayón, J.I. *Dimensiones éticas y jurídicas de la inteligencia*

Es más, si adicionalmente a la responsabilidad objetiva precitada, analizamos el marco en España que regula la tenencia de animales considerados de riesgo o "peligrosos", se sumarían obligaciones adicionales como la de tener la oportuna licencia, registrarlos y disponer de un seguro de responsabilidad civil por daños a terceros con una cuantía mínima garantizada.

Pues bien, dejando a un lado la posibilidad de interpretaciones extensivas por asimilación, quizás este régimen específico podría ser una referencia a considerar para la creación de los futuros marcos reguladores de la responsabilidad de los sistemas de inteligencia artificial en función de su nivel de riesgo, de modo que se conciba como una responsabilidad objetiva por riesgos con la obligación de identificarlos, registrarlos y de establecer determinadas medidas de aseguramiento del derecho a un resarcimiento efectivo de la persona afectada, por ejemplo, mediante los seguros, especialmente en atención a su nivel de riesgo.

En apoyo de esta argumentación, citar a Botella Hermosa[654] que también se planteó estas reflexiones en relación con los robots.

Sin embargo, la dificultad radica en incardinar a los sistemas dotados de inteligencia artificial en los supuestos expresamente contemplados por el ordenamiento jurídico vigente, en la medida que no son personas ni animales, ni disponen de la capacidad de ser titulares de derechos u obligaciones, tengan o no lo capacidad de ejercerlas, que tampoco.

La responsabilidad por hecho ajeno se contempla en el artículo 1903 del Código Civil español, si bien, la legislación española contempla otros supuestos como la responsabilidad patrimonial de la Administración Pública por daños causados por sus funcionarios[655], la responsabilidad como miembro indeterminado de un grupo en la legislación sobre caza[656], la responsabilidad del importador de un producto defectuoso[657]

artificial en el marco del Estado de Derecho. Cuadernos de la Cátedra de Democracia y Derechos Humanos. Universidad de Alcalá: Defensor del Pueblo. 2020. P. 246.

[654] BOTELLO HERMOSA, P. (2020). "La responsabilidad civil extracontractual de los daños originados por robots a terceros: ¿Por qué no una ley española sobre el régimen jurídico de la tenencia y uso de robots?", en BELLO JANEIRO, D. (Coord.) *Nuevas tecnologías y responsabilidad civil*. Editorial Reus 2020. P. 323.

[655] Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público. BOE 02.10.2015.

[656] Ley 1/1970, de 4 de abril, de caza. BOE 06.04.1970.

[657] Real Decreto Legislativo 1/2007, de 16 de noviembre (Texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias). BOE 30.11.2007

o la responsabilidad de la aseguradora en relación con el seguro obligatorio para la circulación de vehículos a motor[658].

El artículo 1903 del Código Civil precitado viene referido a los actos u omisiones de las personas de las que se debe responder, y los sistemas dotados de inteligencia artificial, aunque sea avanzada o "fuerte", no son considerados "personas", por el momento.

Además, dicho precepto relaciona una serie de supuestos en los que unas personas responden por otras, considerando que el fundamento de la responsabilidad en estos supuestos es subjetivo, por culpa *in educando, in eligendo* o *in vigilando*, y se configura como una responsabilidad directa en relación con quien ha causado el daño (el menor o el trabajador) y no subsidiaria.

Las cuestiones que se suscitan para valorar la aplicación de este precepto son, de un lado, considerar si este precepto sería aplicable a fabricantes, diseñadores, desarrolladores, entrenadores, formadores, operadores, propietarios o usuarios de robots o sistemas dotados de inteligencia artificial respecto de los daños causados por éstos y, de otro, si sería aplicable este marco especial a estos contextos en los que se halle involucrado un sistema inteligente o dicho precepto establece un *numerus clausus*.

Respecto de la primera de las cuestiones, la relación de jerarquía o subordinación exigida a nivel subjetivo por el precepto podría permitir valorar la aplicación a todos estos sujetos, si bien, partiendo en todo caso, como he referido, de la ausencia actual de personalidad jurídica y consciencia de los sistemas inteligentes, sin perjuicio del posible reconocimiento de aspectos como su interactividad y cierto grado de autonomía y libertad.

Respecto de la segunda de las cuestiones, la doctrina no es pacífica conforme analizó García Varela[659]. Según el mismo, de un lado, De Ángel Yágüez y la mayoría de la doctrina consideran que los supuestos relacionados son taxativos y no admiten ampliación, es decir, constituyen un *numerus clausus*. De otro, Díez-Picazo y Gullón Ballesteros consideran que en todos los supuestos contemplados en este precepto existen

[658] Real Decreto Legislativo 8/2004, de 29 de octubre, por el que se aprueba el Texto Refundido de la Ley sobre responsabilidad civil y seguro en la circulación de vehículos a motor. BOE 05.11.2004.
[659] GARCÍA VARELA, R. (2008). "La responsabilidad por hecho ajeno". *Diario La Ley*, Nº 6874, Sección Columna, 1 de febrero de 2008, Año XXIX, Ref. D-27, Editorial LA LEY.

unas relaciones jurídicas determinantes de un nexo de jerarquía o subordinación que permite dar a otro órdenes o instrucciones, por lo que para extender el principio de la responsabilidad por hecho ajeno a otras situaciones distintas de las contempladas específicamente en el precepto, debe concurrir dicho contexto, indicando como ejemplo la relación amo-criado, utilizando la propia terminología del Código Civil español. Ambos autores consideran que este ejemplo planteado por los mismos debe entenderse dentro de ámbito de aplicación del precepto, aunque no esté expresamente contemplado en el mismo, al igual que cuando el daño se ha producido por la actividad de una empresa a la que otra la había encargado un trabajo.

Sin embargo, como refiere Vázquez de Castro, otra parte de la doctrina interpreta el 1903 del Código Civil de un modo distinto, entre otros, López Sánchez y Ossorio Serrano, considerando que el mismo contempla un *numerus apertus*[660] de supuestos para imputar la responsabilidad al empresario.

Conforme destaca el precitado García Varela en su análisis de esta cuestión, otros autores como Lacruz, consideran que la enumeración de este artículo 1903 es muy concreta y precisa y, en consecuencia, hace suponer que, en principio, no cabe admitir la aplicación por analogía, no pudiendo sostener una interpretación o aplicación extensiva de un precepto que impone responsabilidad. Otros autores, como Díaz Alabart, consideran que los responsables por hechos ajenos mencionados no están enumerados de forma exhaustiva y es factible, por tanto, la aplicación analógica de la norma.

En cualquier caso, este clausulado hace referencia a una relación de jerarquía o subordinación entre personas, no entre persona y máquina, por lo que no considero viable una interpretación analógica y extensiva, y no considero que la responsabilidad extracontractual de los sistemas de inteligencia artificial pueda incardinarse en este

[660] Ver el análisis doctrinal efectuado por Gómez Calle, en especial, de López Sánchez y Ossorio Serrano. GÓMEZ CALLE, F. (2008) "Los sujetos de la responsabilidad civil: La responsabilidad por hecho ajeno". En REGLERO CAMPOS, L.F. (Coord.). *Tratado de Responsabilidad Civil.* Vol. I. 2008 (Parte General). Tomo I. Pp. 1068-1069. Asimismo, respecto de las actividades que pueden ser englobadas dentro del artículo analizado, serán incluidas todas las actividades concordantes, accesorias, preliminares o necesarias para el servicio de que el empresario se esté beneficiando, siguiendo a Diez- Picazo y Gullón. DIEZ PICAZO, L. Y GULLÓN, A. (2001). *Sistemas de Derechos Civil*. Vol. II. Tecnos. Madrid. 2001. P. 568.

régimen, para lo que se exigiría una modificación del régimen especial de responsabilidad previsto en el Código Civil español para su expresa inclusión.

La posible interpretación analógica de este precepto podría abrir la vía para incardinar la responsabilidad por daños de robots y sistemas dotados de inteligencia artificial en este tipo de supuestos, si bien, mi postura es contraria a la misma sin una necesaria intervención legislativa, en la medida que, de un lado, no considero adecuada una interpretación extensiva de una norma que pretende imponer una responsabilidad y, de otros, estos sistemas carecen de personalidad jurídica y capacidad de obrar, no pudiendo ser titulares de derechos ni de obligaciones.

La cuestión relativa a la atribución de la responsabilidad a los sistemas inteligentes y su atribución al empresario no es nueva, en la medida que ya se planteó incluso judicialmente en 1.984 en el Asunto EE.UU. vs Athlone Indus Inc[661], en el que se pretendía que el robot fuese responsable del daño, siendo resuelto desfavorablemente al considerar que los robots no pueden ser demandados.

García Teruel[662] analiza y cita a parte de la doctrina internacional como O'Sullivan, Van den Hoven, Chen y Burgess, que consideran que deben ser estos sistemas en los que no haya intervención humana, los que respondan de los daños causados ante la dificultad de atribuirla a una persona determinada.

Por otra parte, el artículo 1904 del Código Civil español habla de dependientes y docentes en centros escolares, el 1905 de animales, el 1.906 de heredades de caza, el 1907 de edificios, el 1908 de explosión de máquinas, la inflamación de sustancias explosivas, humos excesivos, caída de árboles, emanaciones de cloacas o depósitos de materias infectantes, el 1909 de defectos constructivos y el 1910 de cosas que se arrojen o caigan

[661] Recuperado de: https://casetext.com/case/united-states-v-athlone-industries-inc?__cf_chl_jschl_tk__=pmd_wUmjFE2OYOuq46oDO5tysIRK46Cb22gNlka9x_UpbhY-1634214650-0-gqNtZGzNAmWjcnBszQiR. Consultado el 14.12.2020.

[662] GARCÍA TERUEL, R.M. (2021). "El derecho de daños ante la inteligencia artificial y el machine learning: una aproximación desde las recomendaciones del Parlamento Europeo y del Grupo de Expertos de la Comisión Europea", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Tomo II. Ed. Aranzadi-Thomson Reuters. 2021. Pp. 1.018.

de una casa. Sin duda, supuestos más propios del contexto civil en el que se aprobó esta norma hace dos siglos.

En relación con estos supuestos específicos de responsabilidad por daños causados por animales o heredades de caza, o por cosas inanimadas por su falta de reparación o defecto de construcción, al hilo de los argumentos comentados anteriormente, podría barajarse una interpretación extensiva inicial para considerar forzadamente la inclusión de los sistemas dotados de inteligencia artificial más avanzada en este régimen especial.

No puedo compartir una interpretación en este sentido y que considero inviable así como la consecuente aplicación por analogía de estos preceptos para determinar la responsabilidad de un sistema de inteligencia artificial, la cual considero extremadamente forzada y choca frontalmente con la doctrina precitada que considera estos supuestos como un *numerus clausus*, y más, cuando comporta la imputación de una obligación o responsabilidad, lo que requeriría una modificación de *lege ferenda*, conforme he manifestado anteriormente.

El artículo 1907 del Código Civil español aborda la responsabilidad del propietario del edificio por los daños que puedan resultar de la ruina total o parcial del mismo con origen en una falta de reparaciones necesarias. La aplicación de este precepto a un sistema inteligente que, por supuesto ni puede ser "propietario" ni es el "edificio" se me hace difícil de imaginar. Un sistema de esta naturaleza podría tener encomendada la monitorización y mantenimiento de determinados aspectos de un edificio, si bien, la responsabilidad recaería en su propietario.

Del mismo modo, el artículo 1908 del Código Civil español, en relación con lo dispuesto en el 1909, contempla distintos supuestos de responsabilidad también del propietario en caso de explosiones de máquinas, la inflamación de sustancias explosivas, humos nocivos para personas o propiedades, caída de árboles o emanaciones de cloacas o depósitos de materiales infectantes. En todos estos casos la responsabilidad recae en el propietario sin perjuicio de que los sistemas inteligentes puedan ser un medio o instrumento orientado a su gestión, monitorización, seguridad o mantenimiento, pero en modo alguno susceptibles de imputación de responsabilidad.

El artículo 1910 del Código Civil español regula de manera general y sucinta un régimen específico de responsabilidad objetiva del cabeza de familia que habita una casa o parte de ella, el cual será responsable de los daños causados por las cosas que se arrojaren o cayeren de la misma. Podría relacionarse con el uso de sistemas inteligentes, pero nunca podría imputársele responsabilidad al propio sistema. García Teruel[663] ha analizado recientemente todos estos supuestos regulados en los artículos 1907 a 1910 con similares conclusiones.

Las reflexiones realizadas sobre todos estos supuestos y su relación con sistemas inteligentes me han suscitado cuestiones apasionantes, cuyo análisis excede del objeto y alcance limitados de esta investigación, como el juego del consentimiento de la persona afectada o la asunción por su parte del riesgo y de sufrir, en consecuencia, un daño como causa de exoneración de responsabilidad, así como otras extrapoladas al ámbito penal, como la fuerza mayor, la culpa de la víctima, el estado de necesidad u otras en relación con distintos supuestos. En definitiva, ¿la aceptación de una persona afectada del riesgo asociado a un sistema de inteligencia artificial y, en consecuencia, de los daños derivados de su materialización, permitiría eximir de responsabilidad al sujeto responsable?

Asimismo, existen otros regímenes especiales que *objetivizan* la responsabilidad y la basan en el riesgo y no en la culpa conforme he comentado anteriormente, especialmente en relación con los daños derivados de actividades de caza conforme al artículo 33 de la Ley 1/1970, de 4 de abril, de Caza, y los causados en el marco del transporte aéreo, conforme prevé la Ley 48/1960, de 21 de julio, de Navegación Aérea.

Los mismos resultan inaplicables a los sistemas de inteligencia artificial como sujeto imputable de la responsabilidad, sin perjuicio de que ejemplifiquen la responsabilidad basada en el riesgo que considero, es el enfoque más adecuado para la conformación futura de una solución integral sobre la que deberían construirse los nuevos marcos

[663] GARCÍA TERUEL, R.M. (2021). "El derecho de daños ante la inteligencia artificial y el machine learning: una aproximación desde las recomendaciones del Parlamento Europeo y del Grupo de Expertos de la Comisión Europea", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Tomo II. Ed. Aranzadi-Thomson Reuters. 2021. P. 1.026.

reguladores en materia de responsabilidad por daños causados por la inteligencia artificial.

La responsabilidad objetiva especial es abordada de manera diversa en los distintos ordenamientos jurídicos en el ámbito europeo e internacional.

Los tribunales en países como Austria tienden a aplicar por analogía los marcos de responsabilidad objetiva, aunque con mucha prudencia y reserva, lo que permite su aplicación a supuestos no expresamente previstos por el legislador. Sin embargo, países como Alemania, Italia o Suiza rechazan su aplicación de manera extensiva a cualquier supuesto no expresamente contemplado en el ordenamiento jurídico.

El Derecho francés regula una responsabilidad general por los hechos de las cosas que están bajo la propia guarda. El Derecho inglés contempla la responsabilidad objetiva para un número limitado de supuestos, por ejemplo, en los daños causados por animales. Asimismo, sujeta al sistema de responsabilidad civil general los accidentes de tráfico que, en la mayoría de países europeos, se consideran sujetos a sistemas de responsabilidad objetiva.

Por todo ello, al igual que otros autores[664], considero necesario revisar y adaptar el Código Civil español -al igual que la Directiva 85/374/CEE en materia de responsabilidad por productos defectuosos y el TRLGDCU española conforme analizaré con posterioridad- a la sociedad digital actual y nuevas exigencias de nuestros tiempos, en paralelo a las iniciativas reguladoras de la inteligencia artificial, especialmente y entre otros aspectos, los supuestos específicos de responsabilidad civil relacionados, lo que podría permitir la inclusión expresa y formal en los mismos de la responsabilidad por daños derivados por sistemas dotados de inteligencia artificial, especialmente la avanzada o "fuerte", sin necesidad, por el momento, de valorar la concurrencia de su personalidad jurídica.

La Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial de 20

[664] DIAZ ALABART, S. (2018) "Robots y responsabilidad civil*"* y SEOANE SPIEGELBERG, J.L. (2019) "La responsabilidad civil tras 130 años de vigencia del Código Civil". Diario La Ley, Nº 9537, Sección Tribuna, Wolters Kluwer 16 de diciembre de 2019.

de octubre de 2020, recoge supuestos de responsabilidad civil por hechos ajenos que serán analizados posteriormente.

3.6. Prescripción de las acciones

El plazo para ejercitar la acción de responsabilidad civil extracontractual general para la reclamación de los daños y perjuicios sufridos es de un año, de conformidad con lo previsto en el artículo 1902 del Código Civil español en relación con su artículo 1.968, sin perjuicio de los regímenes especiales.

3.7. Reflexiones globales

En base al análisis general y consideraciones realizadas en relación con los distintos aspectos tratados, el marco vigente en España para la determinación de la responsabilidad civil aplicable a los sistemas de inteligencia artificial en caso de daños causados por los mismos, estaría conformado por el régimen previsto en el Código Civil español para la responsabilidad contractual y extracontractual, así como por los marcos específicos de responsabilidad regulados en la TRLGDCU -que analizaré a continuación-, cuando resulten de aplicación en relación con contratos con consumidores y usuarios y responsabilidad civil por bienes o servicios defectuosos. Asimismo, se aplican regímenes especiales en determinados sectores como, por ejemplo, el del automóvil, que serán objeto de tratamiento posterior.

En la medida que el objeto y alcance de esta investigación se ha focalizado especialmente en la responsabilidad por daños derivada de sistemas inteligentes, un análisis más profundo y estratégico del sistema de responsabilidad civil extracontractual actual en España, basado en tendencias de lo que será novedad y realidad mañana, nos enfrenta directamente y a corto plazo a una de las primeras cuestiones abordadas en los párrafos precedentes pero esencial para la definición de los futuros marcos regulativos: Los sistemas de inteligencia artificial supuestamente "autónomos" (de manera absoluta o relativa), independientes pero sin consciencia o *pseudoconsciencia*, con capacidad de autoaprendizaje -autodidactas-, susceptibles de entrenamiento y relativa impredecibilidad

en sus decisiones y acciones conforme a su experiencia e interacciones entre sus elementos, datos, situaciones y contexto.

La variedad de aplicaciones, sectores, contextos y usos en los que operarían determinados sistemas o robots dotados de estas capacidades comportan riesgos adicionales que previsiblemente no podrán ser anticipados en su totalidad por sus diseñadores, programadores y fabricantes, a los efectos de adoptar las previsiones y medidas oportunas. Y, además, la posible interacción de los propios operadores y usuarios podría llevar a su alteración, manipulación y consecuente decisión y actuación no prevista en su concepción y desconocida para el diseñador, desarrollador y fabricante[665].

Nos encontraríamos en estos supuestos ante sistemas de inteligencia artificial capaces de tomar decisiones y ejecutarlas supuestamente sin sujeción plena a controles externos, es decir, sin una supervisión y control humano total sobre las mismas ni sobre sus riesgos, ante la independencia conferida en su diseño y concepción.

En resumen, agentes sin personalidad jurídica y sin consciencia pero que podrían afectar a nuestra vida, bienes y derechos e incluso atentar contra todo ello.

Estos sistemas supondrían la posible ausencia de control consecuente por parte de diseñadores, programadores, propietarios o fabricantes.

Además, como he referido, no existe actualmente un marco regulador que establezca los requerimientos y obligaciones éticas, jurídicas y de seguridad de los sistemas inteligentes y, los marcos actuales de responsabilidad, en mi opinión, no proporcionan las herramientas adecuadas de inicio para dar seguridad jurídica y solución a todos los contextos de responsabilidad que pueden plantearse durante todo el ciclo de vida de un sistema inteligente, sea “débil” o “fuerte”, y en relación con los múltiples sujetos que intervengan en su ciclo de vida.

[665] NÚÑEZ ZORILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Editorial Reus, Madrid 2019. P.12.

4. La responsabilidad civil por productos defectuosos.

4.1. Cuestiones generales

4.1.1. Marco jurídico en la UE

La responsabilidad civil por productos defectuosos se reguló a nivel europeo en la Directiva 85/374/CEE del Consejo, de 25 de julio de 1985[666], que fue complementada posteriormente con las Directivas 92/59/CEE del Consejo, de 29 de junio de 1992, relativa a la seguridad general de los productos[667] y la posterior Directiva 2001/95/CE del Parlamento Europeo y del Consejo, de 3 de diciembre de 2001, relativa a la seguridad general de los productos[668].

Con posterioridad fue reformada al objeto de reforzar la protección del consumidor mediante la Directiva 1999/34/CEE[669] que, entre otros aspectos, modificó su artículo 2, estableciendo que, a los efectos de dicha Directiva, se debe entender por "producto" cualquier bien mueble, aun cuando esté incorporado a otro bien mueble o a un bien inmueble y que también se entenderá por "producto" la electricidad.

La responsabilidad por productos defectuosos no está totalmente armonizada en el seno de la UE, en la medida que existen diferencias en la aplicación de la misma y los Estados miembros siguen conservando vías alternativas de indemnización, como expresamente

[666] Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados miembros en materia de responsabilidad por los daños causados por productos defectuosos. DOUEL 7 agosto 1985

[667] Directiva 92/59/CEE del Consejo, de 29 de junio de 1992, relativa a la seguridad general de los productos. OJ L 228, 11.8.1992. Pp. 24-32.

[668] Directiva 2001/95/CE del Parlamento Europeo y del Consejo, de 3 de diciembre de 2001, relativa a la seguridad general de los productos. Diario Oficial n° L 011 de 15/01/2002. Pp. 4-17.

[669] Directiva 1999/34/CE del Parlamento Europeo y del Consejo, de 10 de mayo de 1999, por la que se modifica la Directiva 85/374/CEE del Consejo relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados miembros en materia de responsabilidad por los daños causados por productos defectuosos. OJ L 141, 4.6.1999.

recoge el informe *Liability for Artificial Intelligence and other emerging digital technologies*[670] de 2019.

Conforme recoge el precitado informe, sólo la responsabilidad estricta de los productores por productos defectuosos está armonizada en la UE por la Directiva indicada, si bien, los distintos Estados miembros han establecido distintos marcos específicos sobre aspectos no objeto de armonización, lo que exige ajustes tanto en los regímenes de responsabilidad nacionales como de la UE por lo que se refiere a la inteligencia artificial y otras tecnologías emergentes

Los intereses protegidos por este marco europeo de responsabilidad por productos defectuosos se orientan principalmente a la vida y la salud y a los bienes de los consumidores, no otros, lo que de antemano supone un marco muy limitado que excluye otros bienes, derechos y sujetos.

La Directiva 85/374/CEE se elaboró sobre la base del principio de neutralidad tecnológica, considerándose a fecha actual por las instituciones europeas un instrumento eficaz que contribuye a mejorar la protección de los consumidores, la innovación y la seguridad de los productos, conforme se recoge en las últimas resoluciones del Parlamento Europeo en materia de inteligencia artificial que están siendo objeto de análisis en esta investigación.

No obstante, como significa el informe precitado y abordaré con mayor detalle más adelante, algunos de los conceptos esenciales en los que se basa el régimen de responsabilidad contemplado en el mismo, resultan hoy actualmente inadecuados para afrontar los riesgos potenciales de las nuevas tecnologías, entre otras, la inteligencia artificial.

Los aspectos clave sobre los que se construyó este régimen de responsabilidad fueron concebidos en base a productos y modelos de negocio tradicionales, esto es, "objetos materiales que son puestos en el mercado por una acción única de su productor, tras la

[670] *Liability for Artificial Intelligence and other emerging digital technologies.* UE. 2019. Recuperado de: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf. Consultado el 14.12.2020.

cual éste no mantiene el control sobre el producto". Sin embargo, la evolución tecnológica requiere revisar los conceptos de producto, defecto y productor/fabricante.

Respecto al concepto de "producto", la inteligencia artificial cuestiona la diferenciación tradicional entre producto y servicio, ante la simbiosis que presentan habitualmente estos sistemas entre ambos conceptos y la imposibilidad de separarlos y otorgarles tratamientos distintos.

Del mismo modo, conforme he anticipado, se plantea si el *software* encajaría en el concepto jurídico de producto o de componente de producto, y la respuesta debería ser diferente para el *software* integrado o no integrado, incluyendo sus actualizaciones, con la problemática consecuente para las actualizaciones efectuadas desde fuera de la UE.

Respecto al concepto "defecto" se determina sobre la base de las expectativas de seguridad de un consumidor medio teniendo en cuenta todas las circunstancias pertinentes.

Sin embargo, las características -por ejemplo, la complejidad y opacidad- y las capacidades de las que pueden estar dotados los sistemas inteligentes -especialmente autoaprendizaje e incluso autoprogramación, así como su relativa impredecibilidad e interconexión con otras tecnologías y sistemas- plantean dificultades para identificar el defecto y probarlo. ¿Las desviaciones imprevisibles podrían considerarse un defecto? Y si lo fueran ¿puede alegar el productor como causa de exención el estado de la técnica?

De inicio, la Directiva 85/374/CEE se focaliza en el momento de la puesta en circulación del producto para la determinación de la responsabilidad del productor, lo que excluiría las actualizaciones o mejoras llevadas a cabo por el mismo, y tampoco prevé ninguna obligación específica de control de los productos después de su circulación, ante la evidente necesidad de establecer controles durante todo el ciclo de vida de los sistemas inteligentes.

De este modo, se evidencia la insuficiencia de la Directiva para acometer los riesgos y retos que plantea la inteligencia artificial y para abordar las cuestiones de responsabilidad

por daños que puede generar en función de las características, capacidades, sector, contexto y usos de la misma.

A mayor abundamiento, en los sistemas inteligentes aparecen otros agentes que pueden intervenir en distinto grado durante todo su ciclo de vida, especialmente si el funcionamiento del sistema requiere datos proporcionados por terceros o recogidos por el propio sistema, posee capacidades de autoaprendizaje, es susceptible de entrenamiento y formación por terceros, permite ciertos ajustes para su personalización y uso por el propio operador/ usuario o requiere una monitorización permanente, conforme a la exigencia que, por el momento, contempla la nueva Propuesta de Reglamento de inteligencia artificial de 21 de abril de 2021, que fue analizada en el capítulo anterior.

Estos agentes o sujetos, incluyen al diseñador, al desarrollador, al productor -que puede mantener cierto grado de control sobre el sistema inteligente mediante desarrollos, parches y actualizaciones de manera posterior a su puesta en el mercado, pero limitado y no exclusivo-, fabricante de componentes, ensamblador, proveedor de datos, entrenador, comercializador, operador o usuario.

Este nuevo paradigma que supone la inteligencia artificial diluye los roles y responsabilidades y el control sobre el riesgo ante la intervención de múltiples actores que intervienen en su diseño, funcionamiento y uso durante todo su ciclo de vida.

Además de todo ello, la mayoría de los Estados miembros, como España, han adoptado como limitación de responsabilidad, la exención por riesgo de desarrollo, que evita la responsabilidad al productor si el estado de la ciencia y los conocimientos técnicos en el momento de la puesta en circulación del producto no permitían al mismo poder identificar la existencia del defecto.

Otra de las cuestiones que se plantean en el informe anteriormente citado, es que el régimen de la Directiva 85/374/CEE protege la vida y la salud, así como la propiedad de los consumidores, si bien, en relación con esto último, no parece claro ni totalmente resuelto si cubre los daños a los datos, ya que éstos pueden no ser un "bien" en el sentido del artículo 9 de la misma. A mi juicio, lo es, y en opinión de autores como Vázquez de

Castro[671], el daño indemnizable incluye los datos, como así incluso se recoge en otros marcos jurídicos, como el RGPD, que considera la destrucción de los datos de la víctima como un daño indemnizable. Del mismo modo, el informe precitado *Liability for Artificial Intelligence and other emerging digital technologies*[672] de 2019, concluye que "la destrucción de los datos de la víctima debe considerarse un daño, indemnizable en condiciones específicas. En algunos países es discutible que la destrucción de datos pueda considerarse una pérdida patrimonial, dado que la noción de propiedad se limita a los objetos corpóreos y excluye los intangibles[673].

En cualquier caso, a pesar de estas cuestiones, la Comisión Europea ha apostado por este régimen de responsabilidad de productos defectuosos como el mecanismo aplicable para depurar las responsabilidades por los daños causados por la inteligencia artificial y su funcionamiento.

La Directiva 85/374/CEE se complementó con otras normas europeas.

La Directiva 98/6/CE sobre los precios por unidad obliga a los comerciantes a indicar el precio de venta y el precio por unidad de medida a fin de mejorar y simplificar las comparaciones de precios y de cantidades entre los productos comercializados. Por otra parte, la precitada Directiva 1999/44/CE establece garantías para los productos adquiridos por los consumidores, para lo cual exige a los comerciantes que venden bienes de consumo en la Unión que subsanen los defectos existentes en el momento de la entrega que se manifiesten en el plazo de dos años.

---

[671] VÁZQUEZ DE CASTRO, E. (2020). "Aproximación a la responsabilidad derivada de los riesgos de la inteligencia artificial en Europa". En SOLAR CAYÓN, J.I. (2020). *Dimensiones éticas y jurídicas de la inteligencia artificial en el marco del Estado de Derecho.* Cuadernos de la Cátedra de Democracia y Derechos Humanos. Op. cit. P. 258.

[672] *Liability for Artificial Intelligence and other emerging digital technologies.* UE. 2019. Recuperado de: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf. Consultado el 14.12.2020.

[673] A modo de ejemplo, el artículo 90 del BGB alemán, según el cual una "cosa", por definición, debe ser corpórea, en comparación con el artículo 285 del ABGB austriaco, que no establece tal limitación, de modo que las "cosas" también pueden ser inmateriales.

La UE dispone igualmente de distintas normas dirigidas a la protección del consumidor como la también precitada Directiva 2001/95/CE[674], que establece un sistema de seguridad general de los productos en cuya virtud, cualquier producto de consumo que haya sido comercializado tiene que respetar determinadas normas en lo relativo al suministro de información a los consumidores, las medidas para evitar riesgos para la seguridad, el control de la seguridad del producto y la trazabilidad, y ello aunque no esté regulado por la normativa de un sector específico.

Esta Directiva fue incorporada al ordenamiento jurídico español mediante el Real Decreto 1801/2003, de 26 de diciembre, sobre seguridad general de los productos[675], del que me permito destacar su apuesta por la autorregulación mediante códigos de buenas prácticas para garantizar la seguridad general de los productos, sometidos a una aprobación por los órganos administrativos competentes en la que se valore su utilidad como instrumento al servicio de la seguridad general de los productos.

La Directiva 2001/95/CE relativa a la seguridad general de los productos, regula también la seguridad que deben tener algunos productos relacionados con la tecnología y la inteligencia artificial, la cual se ha ido complementado con Decisiones de la Comisión con finalidad integradora respecto de los requisitos de seguridad de determinados productos y para actualizar algunas normas técnicas de seguridad ante su modificación o nueva creación.

La Comisión Europa, mediante Decisión de Ejecución (UE) 2019/1698[676] procedió a la posterior publicación en el Diario Oficial de la Unión Europea de las referencias de las normas europeas sobre productos redactadas en apoyo de la Directiva 2001/95/CE, que incorpora en sus anexos, con derogación de las diversas decisiones adoptadas por la misma hasta entonces.

[674] Directiva 2001/95/CE del Parlamento Europeo y del Consejo, de 3 de diciembre de 2001, relativa a la seguridad general de los productos (Texto pertinente a efectos del EEE). Diario Oficial n° L 011 de 15/01/2002. Pp. 4-17.
[675] Real Decreto 1801/2003, de 26 de diciembre, sobre seguridad general de los productos. BOE 10 enero 2004.
[676] Decisión de Ejecución (UE) 2019/1698 de la Comisión de 9 de octubre de 2019 relativa a las normas europeas sobre productos redactadas en apoyo de la Directiva 2001/95/CE del Parlamento Europeo y del Consejo relativa a la seguridad general de los productos. DOUEL 10.10.2019.

Las normas técnicas referenciadas incluían las relativas a tecnologías y sistemas de información como, por ejemplo, aparatos de audio, video y aparatos electrónicos análogos, con remisión a sus requisitos de seguridad establecidos en la norma técnica IEC 60065:2001 y los equipos de tecnología de la información-Seguridad, con remisión a sus requisitos generales establecidos en las normas técnicas IEC 60950-1:2005 (versión modificada); EN 60950-1:2006/A12:2011. Estas normas han sido derogadas con posterioridad.

La Directiva precitada fue posteriormente modificada por distintos instrumentos, especialmente por el Reglamento (CE) nº 765/2008 del Parlamento Europeo y del Consejo, de 9 de julio de 2008, por el que se establecen los requisitos de acreditación y vigilancia del mercado relativos a la comercialización de los productos y por el que se deroga el Reglamento (CEE) nº 339/93[677], el cual ha sido modificado y complementada adicionalmente aquélla por el Reglamento (UE) 2019/1020 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativo a la vigilancia del mercado y la conformidad de los productos y por el que se modifican la Directiva 2004/42/CE y los Reglamentos (CE) n.º 765/2008 y (UE) n.º 305/2011[678].

No obstante, junto a la Directiva precitada y de los marcos sectoriales, la UE ha regulado otros aspectos a través de distintos instrumentos, entre otros, los siguientes: Reglamento (UE) 2018/302 del Parlamento Europeo y del Consejo de 28 de febrero de 2018 sobre medidas destinadas a impedir el bloqueo geográfico injustificado y otras formas de discriminación por razón de la nacionalidad, del lugar de residencia o del lugar de establecimiento de los clientes en el mercado interior y por el que se modifican los Reglamentos (CE) 2006/2004 y (UE) 2017/2394 y la Directiva 2009/22/CE, el Reglamento (UE) 2017/2394 del Parlamento Europeo y del Consejo de 12 de diciembre de 2017 sobre la cooperación entre las autoridades nacionales responsables de la aplicación de la legislación en materia de protección de los consumidores y por el que se deroga el Reglamento (CE) n. o 2006/2004, la Directiva 2011/83/UE del Parlamento Europeo y del Consejo de 25 de octubre de 2011 sobre los derechos de los consumidores, por la que se modifican la Directiva 93/13/CEE del Consejo y la Directiva 1999/44/CE

[677] DOUE Nº 218, de 13 de agosto de 2008. Pp. 30 a 47.
[678] DOUEL 25 junio 2019

del Parlamento Europeo y del Consejo y se derogan la Directiva 85/577/CEE del Consejo y la Directiva 97/7/CE del Parlamento Europeo y del Consejo, la Directiva 2005/29/CE del Parlamento Europeo y del Consejo, de 11 de mayo de 2005, relativa a las prácticas comerciales, desleales de las empresas en sus relaciones con los consumidores en el mercado interior, que modifica la Directiva 84/450/CEE del Consejo, las Directivas 97/7/CE, 98/27/CE y 2002/65/CE del Parlamento Europeo y del Consejo y el Reglamento (CE) nº 2006/2004 del Parlamento Europeo y del Consejo (Directiva sobre las prácticas comerciales desleales), la Directiva 93/13/CCE del Consejo, de 5 de abril de 1993, sobre las cláusulas abusivas en los contratos celebrados con consumidores, la Directiva 1999/44/CE del Parlamento Europeo y del Consejo, de 25 de mayo de 1999, sobre determinados aspectos de la venta y las garantías de los bienes de consumo, la Directiva 98/6/CE del Parlamento Europeo y del Consejo, de 16 de febrero de 1998, relativa a la protección de los consumidores en materia de indicación de los precios de los productos ofrecidos a los consumidores, la Directiva 2006/114/CE del Parlamento Europeo y del Consejo de 12 de diciembre de 2006 sobre publicidad engañosa y publicidad comparativa, la Directiva 2011/83/UE del Parlamento Europeo y del Consejo, de 25 de octubre de 2011, sobre los derechos de los consumidores, por la que se modifican la Directiva 93/13/CEE del Consejo y la Directiva 1999/44/CE del Parlamento Europeo y del Consejo y se derogan la Directiva 85/577/CEE del Consejo y la Directiva 97/7/CE del Parlamento Europeo y del Consejo, la Directiva 2009/22/CE del Parlamento europeo y del Consejo de 23 de abril de 2009 relativa a las acciones de cesación en materia de protección de los intereses de los consumidores.

Los requisitos de seguridad de los explosivos de uso civil y productos similares se recogen en las Directivas 93/15/CEE, 2008/43/CE y 2004/57/CE, y en la Decisión 2004/388/CE, recientemente refundidas todas ellas en la Directiva 2014/28/UE sobre explosivos con fines civiles y en la Directiva 2013/29/UE sobre artículos pirotécnicos. Los requisitos de seguridad de los juguetes se establecen en la Directiva 2009/48/CE.

A todo ello, debemos adicionar las Decisiones 93/683/CEE y 93/580/CEE en las que se instauró un sistema europeo para la investigación de los accidentes domésticos y de ocio -*Ehlass* por sus siglas en inglés-, que reúne información sobre accidentes en el hogar y accidentes ocurridos durante las actividades de ocio, así como un sistema comunitario de intercambio de información sobre los riesgos que comportan determinados productos para

la salud o la seguridad de los consumidores (excepto los productos farmacéuticos y los productos destinados a ser utilizados por profesionales).

En conclusión, la seguridad de los productos se haya regulada con amplitud por la UE a través de marcos generales y sectoriales.

Sin embargo, la responsabilidad civil por productos defectuosos en la UE pasó a ser regulada principalmente en la Directiva 85/374/CEE precitada, pero limitándose exclusivamente a los mismos, sin incluir servicios. Sobre esta cuestión, al hilo de la simbiosis precitada que se produce en la inteligencia artificial entre producto y servicio, se suscitan algunas interesantes reflexiones más específicas, inabordables en el marco de esta investigación, en especial, sobre si el suministro de sistemas de inteligencia artificial basados en *software* como servicio o bajo un modelo *cloud* deben considerarse un servicio o un producto, conforme luego abordaré al analizar la transposición de la misma en España.

Conforme he anticipado, el régimen de responsabilidad por producto defectuoso previsto en la Directiva 85/374/CEE plantea distintas carencias actuales, en especial en relación con su posible aplicación a daños causados por la inteligencia artificial y su funcionamiento, por lo que continuación expongo algunas de las principales cuestiones que plantea:

a) Concepto de producto. La cuestión inicial que se plantea es si el *software* que se suministre digitalmente y no integrado en un dispositivo de *hardware* puede ser considerado "producto" según la *Directiva 85/374/CEE sobre responsabilidad por los daños causados por productos defectuosos*. Según la misma, se considera "producto" cualquier bien mueble, excepto las materias primas agrícolas y los productos de la caza, aun cuando está incorporado a otro bien mueble o a uno inmueble.

   La solución actual a esta cuestión ha venido de la mano de interpretaciones extensivas de la norma, lo que no deja de generar cierta inseguridad jurídica. No obstante, esta cuestión puede plantear posiciones enfrentadas, especialmente

cuando la inteligencia artificial se ofrezca como servicio basado en *software*, infraestructura o plataforma para su uso, por ejemplo, mediante servicios *cloud*.

La nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[679], establece un nuevo concepto de sistema de inteligencia artificial que lo asocia exclusivamente al *software*.

En particular, el artículo 3 del Reglamento propuesto define "Sistema de inteligencia artificial" (sistema de IA) como el software desarrollado con una o más de las técnicas y enfoques enumerados en su Anexo I (aprendizaje automático, sistemas expertos, etc.) que puede, para un conjunto dado de objetivos definidos por el ser humano, generar resultados como contenido, predicciones, recomendaciones o decisiones que influyen en los entornos con los que interactúan.

Adicionalmente a lo anterior, las normas de clasificación para los sistemas de alto riesgo recogida en su artículo 6, establece en su apartado primero, que el sistema inteligente podrá ser clasificado en distintos contextos como tal tanto si se comercializa o pone en servicio independientemente de los productos que acompañe, es decir, como componente de estos o de manera independiente, constituyendo el producto en sí mismo.

En definitiva, parece que la nueva propuesta contribuye a sostener la consideración del sistema inteligente exclusivamente basado en *software* como producto en sí mismo a los efectos de aplicación de esta Directiva.

b) Concepto de defecto. La Directiva 85/374/CEE no distingue entre defectos de fabricación, diseño o información, considerando que un producto es defectuoso cuando no ofrezca la seguridad a la que una persona tiene legítimamente derecho.

[679] COM (2021) 206 final 2021/0106

Esto comporta en la práctica que, para considerar si estamos ante una falta de esa seguridad a la que la persona afectada tiene derecho, de nuevo, son los jueces y tribunales los que tienen que hacer un balance de *riesgo-utilidad,* lo que comporta, siguiendo a Rubí[680], que el diseño de un producto será defectuoso si es posible identificar un diseño alternativo que hubiera contribuido a evitar un accidente, si los daños esperables que se hubieran evitado con el diseño más seguro fueran superiores a los costes asociados a la adopción de aquél.

De nuevo nos encontraríamos ante un escenario no definido por el momento *ab initio,* con la inseguridad jurídica consecuente para todos los agentes involucrados y para la sociedad en general, lo que puede impactar en la seguridad, confiabilidad, innovación e inversión en inteligencia artificial y su comercialización.

Además, el problema adicional que todo ello plantea es que ese criterio posterior de *riesgo-utilidad* no será previsiblemente suficiente en el caso de sistemas inteligentes en los que el comportamiento del individuo puede ser menos relevante y donde las capacidades de autonomía y autoaprendizaje del sistema son las que pueden determinar si, en el momento del daño, éste funcionaba de un modo razonablemente seguro.

La nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021 precitada regula de manera general en su artículo 15 los requisitos de precisión, robustez y ciberseguridad de los sistemas de inteligencia artificial considerados de alto riesgo, pero no respecto del resto.

El precitado artículo exige una seguridad adecuada al riesgo y contexto durante todo el ciclo de vida ¿Constituiría pues este precepto el marco de seguridad a la que una persona tiene legítimamente derecho al que alude la Directiva precitada en el ámbito de estos sistemas? ¿Y respecto del resto de sistemas inteligentes de nivel medio o

[680] RUBÍ, A. (2020). "Retos de la inteligencia artificial y adaptabilidad del derecho de daños". En Cerrillo i Martínez, A. y Pequera Poch, M. (Coord). *Retos jurídicos de la inteligencia artificial*. Aranzadi S.A.U. (Thomson Reuters). Navarra 2020. P. 67.

bajo, o simplemente de riesgo no clasificado? ¿Cuál sería el marco de seguridad de éstos últimos?

Respecto de sistemas de inteligencia artificial de alto riesgo y hasta la modificación de la Directiva precitada sobre responsabilidad por productos defectuosos, considero que el precepto indicado establece el marco de referencia de seguridad general, en la línea de la seguridad exigida en materia de protección de datos por el RGPD.

c) Causas de exoneración de responsabilidad. Las características y posibles capacidades de las que disponga un sistema inteligente hacen que el mismo pueda ir actualizándose progresivamente, ya sea de forma controlada por el desarrollador y fabricante o de forma autónoma por parte del propio sistema en virtud de sus características, capacidad, conectividad, experiencia, interacción con su entorno y aprendizaje. Esto podría afectar a dos de las causas de exoneración reguladas en la Directiva 85/374/CEE.

De un lado, la Directiva precitada prevé la posibilidad de exención de responsabilidad del fabricante si acredita que, teniendo en cuenta las circunstancias, es probable que el defecto que causó el daño no existiera en el momento en que puso el producto en circulación o que este defecto apareciera más tarde.

Y de otro, también prevé la posibilidad de exención si prueba que en el momento en que el producto fue puesto en circulación, el estado de los conocimientos científicos y técnicos no permitía descubrir la existencia del defecto (excepción de riesgos de desarrollo).

La apreciación de la concurrencia de cualquiera de estas causas en base a una interpretación estricta y literal de la norma -que por otra parte, considero sería la adecuada en el marco de un contexto obligacional y punitivo-, puede impedir la finalidad proteccionista y de resarcimiento efectivo de la persona que sufrió los daños conforme a este marco, especialmente ante los cambios acontecidos e imprevisibilidad concurrente. Posteriormente abordaré con mayor profundidad estas cuestiones.

Como expongo a continuación, la precitada Directiva 85/374/CEE del Consejo precitada fue transpuesta al ordenamiento jurídico español por Ley 22/1994, de 6 de julio, de responsabilidad civil por los daños causados por productos defectuosos, la cual fue derogada el 1 de diciembre de 2007 por el apartado 4 de la Disposición Derogatoria Única del R.D. Legislativo 1/2007, de 16 de noviembre, por el que se aprueba el Texto Refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias, quedando integrado en este texto refundido la responsabilidad por daños causados por productos defectuosos.

4.1.2. Marco jurídico en España

La responsabilidad por productos defectuosos en España se regula principalmente de manera conjunta con la protección del consumidor, en particular, en el TRLGDCU. Y se complementó con otras normas en aspectos concretos, en particular, el Real Decreto 1801/2003, de 26 de diciembre, sobre seguridad general de los productos precitado, el Real Decreto 1828/1999, de 3 de diciembre, por el que se aprueba el Reglamento del Registro de Condiciones Generales de la Contratación, el Real Decreto 820/1990, de 22 de junio, por el que se prohíbe la fabricación y comercialización de los productos de apariencia engañosa que pongan en peligro la salud o seguridad de los consumidores.

La Directiva 85/374/CEE del Consejo precitada fue transpuesta al ordenamiento jurídico español por Ley 22/1994, de 6 de julio, de responsabilidad civil por los daños causados por productos defectuosos.

Ni el ámbito subjetivo de tutela de la Directiva ni el objetivo coincidía con la entonces vigente Ley 26/1984, de 19 de julio, General para la Defensa de los Consumidores y Usuarios. La Ley 22/1994 pasó a regular un régimen de responsabilidad objetiva -no absoluta-, en la que los sujetos protegidos eran los perjudicados por el producto defectuoso, tuvieran o no la condición de consumidores.

Esta responsabilidad objetiva se venía ya aplicando ya en otros países como EE.UU. a través de su *Restatement (Second) Torts* de 1965.

La Ley 22/1994, de 6 de julio, de responsabilidad civil por los daños causados por productos defectuosos fue derogada el 1 de diciembre de 2007 por el apartado 4 de la disposición derogatoria única del R.D. Legislativo 1/2007, de 16 de noviembre, por el que se aprueba el texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias (TRLGDCU), quedando integrado en este texto refundido la responsabilidad por daños causados por productos defectuosos.

La primera cuestión que plantea este marco regulador es la integración de las directrices de la Directiva precitada en esta normativa orientada específicamente a la protección de consumidores y usuarios, restringiendo y reservando inicialmente la misma a los sujetos que tengan dicha condición.

Conforme al artículo 3 del TRLGDCU y a los efectos de esta ley, "sin perjuicio de lo dispuesto expresamente en sus Libros Tercero (Responsabilidad civil por bienes o servicios defectuosos) y Cuarto", son consumidores o usuarios las personas físicas que actúen con un propósito ajeno a su actividad comercial, empresarial, oficio o profesión, lo que de inicio excluye de su ámbito de protección a profesionales y empresas que actúen en el ámbito de sus actividades comerciales, profesionales o empresariales, si bien, estarían protegidas cuando actúen en el ámbito particular, personal o doméstico.

Del mismo modo, se consideran igualmente consumidores a efectos de esta norma las personas jurídicas y las entidades sin personalidad jurídica que actúen sin ánimo de lucro en un ámbito ajeno a una actividad comercial o empresarial.

El "Libro Tercero" al que alude el precepto indicado regula la responsabilidad civil por bienes y servicios defectuosos, cuando la *Directiva 85/374/CEE* se circunscribe a productos defectuosos (no servicios), y no habla de "consumidores y usuarios" sino directamente de "perjudicados" sin especificar su condición. En particular, el artículo 128.I del TRLGDCU establece que "Todo perjudicado tiene derecho a ser indemnizado en los términos establecidos en este Libro por los daños o perjuicios causados por los bienes o servicios". De este modo, conforme al precepto indicado incluiría a profesionales y empresarios, además de consumidores.

Sin embargo, el artículo 129.1 TRLGDCU matiza algunos aspectos al regular su ámbito de aplicación y establece que, el régimen de responsabilidad previsto en este Libro comprende los daños personales, incluida la muerte, y los daños materiales, siempre que éstos afecten a bienes o servicios objetivamente destinados al uso o consumo privados y en tal concepto hayan sido utilizados principalmente por el perjudicado.

Conforme apunté al abordar el régimen general de responsabilidad civil contractual y extracontractual previsto en el Código Civil español, las cuestiones que comportan los sistemas inteligentes dificultan igualmente la plena adecuación de los marcos europeos y nacionales en materia de responsabilidad por productos defectuosos para resolver los distintos y múltiples contextos que se pueden plantear en relación con la responsabilidad por daños derivados del funcionamiento y uso de la inteligencia artificial.

La normativa española de responsabilidad por productos o servicios defectuosos se aplica a ambos, esto es, a productos y servicios, mientras que la normativa europea sobre seguridad de productos no resulta inicialmente aplicable a servicios y puede plantear dificultades iniciales, como analicé anteriormente, para incardinar en este concepto los sistemas inteligentes basados estrictamente en *software*[681].

En cualquier caso, ambos marcos, español y europeo, no están enfocados a tecnologías complejas (o conjunto de las mismas) o emergentes, dinámicas y evolutivas, en despliegue constante en todo tipo de ámbitos y sectores, como lo es la inteligencia artificial, que evolucionan, interactúan con el contexto y se retroalimentan y, que pueden crear nuevos riesgos en momentos posteriores al de su fabricación y durante todo su ciclo de vida, lo que los podría hacer supuesta y relativamente imprevisibles en dicho momento.

De ahí una de las preocupaciones de la Comisión Europea sobre la necesidad, de un lado, de generar un marco de normas específicas en materia de seguridad y posible adaptación

[681] En relación con la aplicación de la Directiva europea y la legislación española de transposición, la doctrina española se ha posicionado tradicionalmente de manera diferente. A favor de su aplicación, significar SOLER MATUTES, P. (2004). *El contrato para la elaboración de programas de ordenador*. Pamplona, 2004. P. 266. En contra, BAUZÁ REILLY, M. (1994). Responsabilidad civil en materia informática, en *Actualidad Informática Aranzadi*, nº 12, julio 1994, P. 13; RIBAS ALEJANDRO, J. (1994). Informática y responsabilidad civil. *Actualidad Informática Aranzadi*, nº 12, julio 1994. Pp. 6-11. Ver también LÓPEZ-TARRUELLA, A. (2006). *Contratos internacionales de software*. Tirant lo Blach. Valencia. 2006. P. 405.

de la Directiva 85/374/CEE, como refleja en su *Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica*[682] de 19.02.2020, y, de otro de asegurar un marco jurídico adecuado de responsabilidad que proporcione seguridad a todas las partes implicadas y garantice el resarcimiento efectivo de la persona afectada.

Durante los próximos apartados analizaré el marco vigente en España en materia de responsabilidad por productos y servicios defectuosos.

4.2. Ámbito de aplicación y otros aspectos

De inicio, como he referido anteriormente, la normativa española en la que se incardina esté régimen de responsabilidad, está principalmente orientada a la protección del consumidor, si bien, la problemática relativa a la responsabilidad derivada de los daños causados por robots o sistemas dotados de inteligencia artificial integra un espectro más amplio, tanto por lo que se refiere al objeto de protección como a los sujetos relacionados con la misma, que debería ser cualquier persona, física o jurídica, consumidor o no.

El Texto Refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias, aprobado mediante Real Decreto Legislativo 1/2007, de 16 de noviembre (TRLGDCU), en su Libro Tercero, regula la responsabilidad civil por bienes o servicios defectuosos, tras refundir en este texto la antigua Ley 22/1994, de 6 de julio, de Responsabilidad Civil por los Daños causados por Productos Defectuosos. Esté régimen reconoce el derecho de indemnización a cualquier perjudicado.

[682] COM/2020/64 final. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52020DC0064.

4.2.1. Régimen de responsabilidad

La responsabilidad por productos defectuosos prevista en ese marco jurídico se presume objetiva, a pesar de que no se haga mención expresa a dicha naturaleza, de conformidad con lo previsto en el artículo 139 del TRLGDCU, dado que no se debe probar la negligencia del productor y no se permite la exoneración del mismo, aunque demuestre su actuación diligente.

Sin embargo, la persona afectada deberá probar el defecto, el daño y la relación de causalidad entre ambos, conforme a lo previsto en el precitado artículo 139, lo que constituye uno de los aspectos que mayores dificultades pueden plantear a la persona afectada en la práctica para conseguir un resarcimiento efectivo, ante las características de la inteligencia artificial, especialmente su complejidad y opacidad, y sus capacidades, especialmente su capacidad de autoaprendizaje, entrenamiento e incluso autoprogramación.

Como más adelante expondré, algunos autores lo califican como un régimen de responsabilidad *cuasi objetiva*, finalísticamente subjetivo o no plenamente objetivo, y ello en base a esa necesidad de probar el defecto, el daño y la relación de causalidad entre ambos por el perjudicado, así como de las particulares previsiones del TRLGDCU relativas a limitación o exoneración de responsabilidad en determinadas circunstancias cuando el daño causado sea debido conjuntamente a un defecto del producto y a la culpa del perjudicado o de una persona de la que este deba responder civilmente.

La argumentación para la consideración de que se trata de una responsabilidad más próxima a la subjetiva o por culpa se basa en que, el perjudicado que pretenda obtener la reparación de los daños causados tendrá probar el defecto, el daño y la relación de causalidad entre ambos, por lo que realmente no estaríamos ante una responsabilidad totalmente objetiva o absoluta, en la que el mero hecho de sufrir un daño implique la existencia de una responsabilidad, en la medida que el perjudicado debe probar que el daño ha sido debido al producto defectuoso. Por ello, a mi juicio, lo considero un régimen

de responsabilidad *cuasi objetiva* con tendencia a su *objetivización* plena en su aplicación jurisprudencial[683].

El objeto y alcance limitados de esta investigación me impiden profundizar en esta conceptualización y discusión doctrinal, sin perjuicio de su análisis en relación con la inteligencia artificial y su posible adecuación.

A modo de ejemplo, en sistemas inteligentes que integren autoaprendizaje y técnicas de *Machine Learning* y *Deep Learning*, puede resultar extremadamente complicado demostrar el defecto y la relación de causalidad sin los registros de actividad pertinentes activados y la configuración pertinente de control y seguridad, incluso para su desarrollador o productor, que el nuevo Reglamento propuesto de 21 de abril de 2021 pretende exigir respecto de sistemas inteligentes de alto riesgo.

4.2.2. Concepto de productor

El artículo 135 de TRLGDCU establece que los productores serán responsables de los daños causados por los defectos de los productos que, respectivamente fabriquen o importen, considerando productor a los efectos de aplicación de este régimen de responsabilidad, conforme a lo dispuesto en su artículo 138 en relación con su artículo 5, al fabricante del bien o al prestador del servicio o su intermediario, o al importador del bien o servicio en el territorio de la UE, así como a cualquier persona que se presente como tal al indicar en el bien, ya sea en el envase, el envoltorio o cualquier otro elemento de protección o presentación, o servicio, su nombre, marca u otro signo distintivo.

Asimismo, será considerado productor a dichos efectos, al fabricante o importador no sólo del producto terminado, sino también de cualquier elemento integrado en un producto terminado o de la materia prima.

[683] Citar, entre otras, la Sentencia del Tribunal Supremo, Sala 1º, de 14.09.2018.

Y, por último, conforme a lo dispuesto en su artículo 138.2, si el productor no puede ser identificado, será considerado como tal el proveedor del producto, salvo que, en el plazo de tres meses, indique al perjudicado la identidad del productor o de quien le hubiera suministrado o facilitado a él dicho producto, aplicando la misma regla en el caso de un producto importado, si el producto no indica el nombre del importador, aun cuando se indique el nombre del fabricante.

4.2.3. Ámbito de protección y aplicación

En este contexto, su artículo 128.1 establece que "*Todo perjudicado tiene derecho a ser indemnizado en los términos establecidos en este Libro por los daños o perjuicios causados por los bienes o servicios*".

Según el precepto indicado parece brindar su protección a cualquier perjudicado, cualquiera que sea su condición, si bien, una interpretación sistemática y teleológica podría hacer pensar que dicha protección y derecho pretende restringirse al perjudicado que tenga la consideración de consumidor, si bien, no es el caso conforme al tenor literal de la norma.

Conforme recogen autores como Vázquez de Castro, "la responsabilidad civil especial por productos defectuosos está diseñada para la protección de los consumidores y usuarios víctimas de los daños de productos que están a su disposición en el mercado"[684].

La legislación española sobre responsabilidad aplicable a los daños causados por productos defectuosos antes de la Directiva 85/374, en particular, la Ley 26/1984, de 19 de julio, para la defensa de consumidores y usuarios, se orientaba a los consumidores y usuarios en sentido estricto conforme a la misma, esto es, como destinatarios finales de bienes y servicios para uso personal o familiar, de modo que cuando la víctima del daño

[684] VÁZQUEZ DE CASTRO, E. (2020). "Aproximación a la responsabilidad derivada de los riesgos de la inteligencia artificial en Europa". En Solar Cayón, J.I. *Dimensiones éticas y jurídicas de la inteligencia artificial en el marco del Estado de Derecho*. Cuadernos de la Cátedra de Democracia y Derechos Humanos. Universidad de Alcalá: Defensor del Pueblo. 2020. P. 239.

no era un consumidor o usuario a tales efectos, los daños no eran indemnizables conforme a aquella, y debían reclamarse conforme a las normas generales de responsabilidad[685].

La promulgación de la Ley 22/1994, de 6 de julio, de responsabilidad civil por los daños causados por productos defectuosos, de transposición de la Directiva europea, supuso la protección de cualquier tercero perjudicado, consumidor o no[686]. De este modo, la protección se extendió a cualquier perjudicado, sea o no parte en una relación de consumo, que pueda sufrir un daño, no quedando restringida a la existencia de una relación jurídica entre el dañado y el fabricante del producto -*bystanders*- y extendiéndose a las personas jurídicas[687].

Del mismo modo, la norma española también regula, como he indicado anteriormente, la responsabilidad por daños causados no sólo por bienes, como la *Directiva 85/374/CEE*, sino también por servicios.

Su ámbito de protección se regula en su artículo 129. En particular, su apartado primero establece que el régimen de responsabilidad civil que regula ese Libro Tercero comprende "los daños personales, incluida la muerte, y los daños materiales, siempre que éstos afecten a bienes o servicios objetivamente destinados al uso o consumo privados y en tal concepto hayan sido utilizados principalmente por el perjudicado", por lo que quedarían fuera del ámbito de protección de su régimen de responsabilidad los daños causados en bienes y servicios destinados a un uso profesional, empresarial o comercial, o si, aun siendo de uso mixto o alternativo -ya empresarial, ya personal, doméstico, familiar o privado-, han sido utilizados por el perjudicado preferentemente con tales fines profesionales[688].

---

685 PARRA LUCÁN, M. A. (2011). *La protección del consumidor frente a los daños. Responsabilidad civil del fabricante y del prestador de servicios.* Editorial Reus. Madrid 2011. P. 37.

686 Conforme significa Lasarte Álvarez, "las disposiciones aplicables parten de un principio universalista, en virtud del cual, debe protegerse a todo sujeto perjudicado". En LASARTE ÁLVAREZ, C. (2013). *Manual sobre protección de consumidores y usuarios*. 5ª Ed. Dykinson. Madrid 2013. P. 238.

687 PARRA LUCÁN, M. A. (2011). *La protección del consumidor frente a los daños. Responsabilidad civil del fabricante y del prestador de servicios.* Editorial Reus. Madrid 2011. P. 41.

688 La exclusión de los daños en bienes empresariales o comerciales no ha sido algo tradicionalmente pacífico en la doctrina (Zurita Martín, Parra Lucán y Marín López) y tampoco en la jurisprudencia, que en distintas resoluciones ha prescindido de la exigencia legal del destino y el uso privados que ha de tener el objeto dañado, entre otras, Sentencias de la Audiencia Provincial de Madrid de 9 de octubre de 2007 (Rec. 253/2007), Audiencia Provincial de Barcelona de 29 de enero de 2007 (Rec. 302/2006) y 13 de mayo de

En relación con la exclusión de los daños a cosas no destinadas o utilizadas para uso o consumo privado, considero necesario destacar la Sentencia del Tribunal de Justicia de la Comunidad Europea (TJCE) de 4 de junio de 2009, en el Asunto C-285/08, en la que se resuelve la cuestión prejudicial planteada sobre el particular por la Corte de Casación francesa en relación al art. 9 de la Directiva 85/374/CEE. La sentencia abre la vía sobre su posible inclusión de daños a cosas destinadas o utilizadas para uso profesional o empresarial.

El TJCE estimó que "puesto que la armonización realizada por la Directiva 85/374 no incluye la reparación de los daños causados a una cosa destinada al uso profesional y utilizada para tal uso, dicha Directiva no impide que un Estado miembro establezca a este respecto un régimen de responsabilidad similar al que ella instaura". En consecuencia, considerando compatible un régimen nacional de responsabilidad objetiva con el igualmente previsto en la Directiva 85/374/CEE para los daños no cubiertos, el TJCE consideró que esta Directiva "debe interpretarse en el sentido de que no se opone a la interpretación de un Derecho nacional o a la aplicación de una jurisprudencia interna reiterada según las cuales el perjudicado puede solicitar la reparación de los daños causados a una cosa destinada al uso profesional y utilizada para tal uso aportando únicamente la prueba del daño, del defecto del producto y de la relación causal entre dicho defecto y el daño".

4.2.4. Producto defectuoso

Por otra parte, el régimen establecido en este marco regulador objeto de análisis está dirigido a productos defectuosos (también servicios como analizaré más adelante), considerando como tal, siguiendo lo dispuesto en el artículo 137.1 del TRLGDCU, aquel que no ofrezca la seguridad que cabría legítimamente esperar, teniendo en cuenta todas

2008 (Rec. 424/2007) y Audiencia Provincial de Jaén de 27 de abril de 2009 (Rec. 110/2009). Recuperado de GUTIÉRREZ SANTIAGO, P. (2012). "El 'daño' en la responsabilidad civil por productos defectuosos (Régimen jurídico de sus clases, cobertura y limitaciones en la legislación de consumo española, a la luz del cuarto Informe de la Comisión Europea de 8 de septiembre de 2011 sobre la Directiva 85/374/CEE)". *Diario La Ley*, Nº 7859, Sección Doctrina, 16 de mayo de 2012, Año XXXIII, Ref. D-201, Editorial LA LEY.

las circunstancias y, especialmente, su presentación, el uso razonablemente previsible del mismo y el momento de su puesta en circulación. Asimismo, conforme al apartado segundo del precitado precepto, en cualquier caso, se considerará que un producto es defectuoso si no ofrece la seguridad normalmente ofrecida por los demás ejemplares de la misma serie.

Según analiza García Teruel[689], podría considerarse producto defectuoso aquél que no tiene suficientes instrucciones o información o, cuando no se hicieron las pruebas de seguridad pertinentes previas a su puesta en circulación. La autora considera igualmente un producto de estas características, siguiendo a Brüggemeier, en caso de defecto posterior a supuesta en el tráfico, cuando el productor no advierte de un defecto potencial del producto aparecido de forma sobrevenida.

Y, según este marco regulador, conforme a lo dispuesto en su artículo 136 del TRLGDCU en relación con el 6, se considera "producto" cualquier bien "mueble", aun cuando esté unido o incorporado a otro bien mueble o inmueble, y poniendo todo ello en relación con el artículo 335 del Código Civil español, según el cual se consideran bienes muebles los susceptibles de apropiación no comprendidos entre los bienes inmuebles y, en general, todos los que se pueden transportar de un punto a otro sin menoscabo de la cosa inmueble a que estuvieren unidos.

En consecuencia y en relación con el objeto de esta investigación, podrían tener cabida dentro del concepto de producto máquinas, robots, vehículos autónomos, dispositivos, drones u otros bienes muebles dotados de inteligencia artificial. La cuestión apuntada que podría suscitarse es en relación a la consideración de sistemas inteligentes basados exclusivamente en *software* como producto a los presentes efectos. Abordaré está cuestión con más profundidad más adelante.

En definitiva, considerando que concurran los elementos requeridos por este marco regulador, considerando a un robot o sistema dotado de inteligencia artificial como un

[689] GARCÍA TERUEL, R.M. (2021). "El derecho de daños ante la inteligencia artificial y el machine learning: una aproximación desde las recomendaciones del Parlamento Europeo y del Grupo de Expertos de la Comisión Europea", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Tomo II. Ed. Aranzadi-Thomson Reuters. 2021. P. 1.028.

bien mueble y producto a los efectos de este marco regulador, y considerando que el mismo sea defectuoso conforme al mismo, esto es, restringiendo dicho defecto a que no disponga de la seguridad esperada a nivel físico y/o lógico (digital), este marco especial de responsabilidad resultaría de aplicación, para la depuración de la responsabilidad derivada de los daños causados por el sistema inteligente al perjudicado. A modo de ejemplo, un robot asistencial proporcionado a una persona con movilidad reducida, donde podría disponerse del mismo por el perjudicado como producto o como servicio, en función del modelo de comercialización.

La falta de seguridad esperada podría responder a muchas de las situaciones que podrían plantearse en la práctica bajo una visión amplia de este concepto, tanto para la persona, como para sus bienes, derechos e intereses, y tanto a nivel lógico como físico, siempre que comporten o pueda asociarse en su valoración e interpretación a dicha falta de seguridad. A juicio de autores como el precitado Vázquez de Castro, en función del contexto, los defectos pueden incluir defectos de manipulación, de conservación o de información[690].

Consecuentemente, este régimen no resultaría de aplicación a la responsabilidad civil por daños causados al margen de cualquier defecto así considerado por este marco normativo, lo que supone otro de los aspectos que dificultan su aplicación generalizada para depurar la responsabilidad por daños causados por sistemas inteligentes.

Sobre el carácter defectuoso del producto, la doctrina, como señala Gómez-Riesco[691], como he referido anteriormente, se integraría por dos elementos que se encuentran íntimamente relacionados entre sí: "que el producto resulte más peligroso para el adquirente de lo que un hombre razonable pudiera esperar de él dadas sus características y que el daño quede fuera de lo razonable[692]".

[690] VÁZQUEZ DE CASTRO, E. (2020). "Aproximación a la responsabilidad derivada de los riesgos de la inteligencia artificial en Europa". Op.cit. P. 248.
[691] GÓMEZ-RIESCO, J. (2018). "Los robots y la responsabilidad civil". En Barrio Andrés, M. (Coord.), *Derecho de los Robots*. Wolkers Kluwer. 2018, P.121.
[692] DÍEZ-PICAZO, L. (2011). *Fundamentos del Derecho Civil Patrimonial. V. La Responsabilidad Civil Extracontractual*. Editorial Civitas-Thomson Reuters. Navarra 2011. P. 478.

Sin embargo, respecto de sistemas de inteligencia artificial avanzada y "fuerte", con supuesta plena autonomía, capacidad de autoaprendizaje y de relativa improvisación e impredecibilidad, ¿Cuál es la expectativa de seguridad que el fabricante puede o debe garantizar? ¿Cuál es la seguridad que legítimamente cabe esperar teniendo en cuenta factores como el uso razonablemente previsto del mismo? ¿El fabricante no puede garantizar ningún grado de seguridad ni prever posibles usos acordes a las capacidades y libertades con las que ha dotado al sistema? ¿Debería prever restricciones o limitaciones a las capacidades de las que dote un sistema inteligente? ¿En base a qué exigencia, personal, corporativa, sectorial o de la industria, ética o legal?

La nueva Propuesta de Reglamento sobre inteligencia artificial del Parlamento Europeo y del Consejo, de 21 de abril de 2021, regula expresamente, aunque de manera genérica, la identificación, análisis y gestión del riesgo, así como las obligaciones de ciberseguridad de los sistemas inteligentes considerados conforme al mismo de alto riesgo, no respecto de otros, conformé abordé en el capítulo II de esta investigación. La ciberseguridad exigida es la ciberseguridad adecuada al contexto y nivel de riesgo.

Según algunos autores, como Núñez Zorilla[693], en el marco de los robots inteligentes el fabricante no puede ofrecer ninguna expectativa de seguridad, no puede garantizar una seguridad esperable, se consideran por sí inseguros e inciertos en el modo de operar. No comparto esta opinión.

No podemos permitir poner en el mercado un producto inseguro, menos en sectores o aplicaciones de alto riesgo y mucho menos robots o sistemas de inteligencia artificial sin expectativa de seguridad alguna, cualquiera que sea el uso, finalidad o sector para la que fue inicialmente diseñado o concebido, ante la supuesta posibilidad de actuar al margen de éstas.

A mi juicio no lo permiten ya los marcos éticos más consensuados a nivel internacional que exigen su seguridad, fiabilidad, control y supervisión humana, tampoco los marcos de seguridad generales de los productos y tampoco los marcos de privacidad respecto de los derechos y libertades que pretenden garantizar, y no lo deberían permitir los nuevos

[693] NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Op. Cit. P. 24.

marcos éticos y jurídicos propuestos en el seno de la UE que pretenden regular la inteligencia artificial.

La ausencia actual de un marco regulatorio, de marcos éticos plenamente consensuados y vinculantes y la evolución potencial de la inteligencia artificial evidencian la urgencia de acometer su regulación sin demora, incluyendo la regulación de estos estos aspectos relativos a la responsabilidad.

El fabricante debe prever usos adecuados, los usos indebidos, cuanto menos los razonablemente previsibles, así como las decisiones o actuaciones permitidas en base a la autonomía conferida, incluyendo riesgos asociados potenciales a todo ello, en función de las capacidades y libertades de las que dote a su producto o sistema, de modo que integre controles y medidas de seguridad durante su ciclo de vida dirigidas a asegurar que cualquier comportamiento no previsto se vea automáticamente restringido o, cuanto menos, sujeto a autorización previa humana o con activación de mecanismos de control, alerta, revisión y/o restricción, integrando junto a la seguridad exigida, los requerimientos éticos y, esperemos que en breve, jurídicos, de control y supervisión humana en todo el ciclo de vida de los mismos, desde el diseño y concepción hasta su aplicación y funcionamiento.

Conforme a lo previsto en la nueva Propuesta de Reglamento sobre inteligencia artificial precitada, de 21 de abril de 2021, en particular y exclusivamente respecto de los sistemas inteligentes de alto riesgo, se debe garantizar la monitorización y seguridad durante todo el ciclo de vida (Artículo 15), deben preverse tanto los riesgos conocidos como previsibles en los sistemas de gestión de riesgos (Artículo 9) y evaluar todos los riesgos que podrían surgir cuando el sistema se utilice conforme a su finalidad prevista, pero también cuando se le dé un uso indebido, cuanto menos, razonablemente previsible.

Asimismo, conforme al Reglamento propuesto, la información preceptiva sobre el sistema deberá incluir las circunstancias conocidas o previsibles que podrían afectar al nivel de precisión, solidez y ciberseguridad esperado, incluyendo a mi juicio las asociadas al grado de autonomía conferida, así como cualquier circunstancia conocida o previsible, asociada a la utilización del sistema conforme a su finalidad prevista o a un uso indebido

razonablemente previsible, que pueda dar lugar a riesgos para la salud y la seguridad o los derechos fundamentales (Artículo 13).

Y del mismo modo, la supervisión humana exigida tendrá como principal objetivo prevenir o reducir al mínimo los riesgos para la salud, la seguridad o los derechos fundamentales que pueden surgir cuando un sistema inteligente de alto riesgo se utilice conforme a su finalidad prevista, cuando funcione conforme al posible grado de autonomía conferida, o cuando se le da un uso indebido razonablemente previsible.

De este modo, el concepto "uso indebido razonablemente previsible" cobra especial protagonismo en el marco de la seguridad y la gestión de riesgo, pero de indudable repercusión a efectos de responsabilidad, en función del régimen específico por el que finalmente se opte en los futuros marcos reguladores de la misma para la inteligencia artificial.

El artículo 3 del Reglamento propuesto precitado define "Uso indebido razonablemente previsible" como la utilización de un sistema de inteligencia artificial de un modo que no corresponde a su finalidad prevista, pero que puede derivarse de un comportamiento humano o una interacción con otros sistemas razonablemente previsible.

Dejando a un lado este inciso y prosiguiendo con mi análisis del TRLGDCU, a mi juicio, el productor o fabricante debe garantizar unas expectativas mínimas de seguridad en este tipo de sistemas, como sin duda lo deben de hacer en un vehículo a motor o aeronave, dotados o no de sistemas de inteligencia artificial y autonomía.

En estos aspectos radica otro de los principales obstáculos para que el régimen de responsabilidad por productos defectuosos proporcione una adecuada respuesta a los distintos contextos de daños causados por sistemas de inteligencia artificial, especialmente los más avanzados, con mayor autonomía y mayores capacidades.

Además, debemos igualmente considerar las opciones de adaptación y customización de algunos sistemas inteligentes y de su posible instrucción o modificación posterior por su operador o usuarios para el seguimiento de instrucciones o realización de funciones inicialmente no previstas. Deberían establecerse controles sobre todo ello para permitir la

atribución de la responsabilidad, especialmente ante modificaciones inadecuadas únicamente atribuible al operador o usuario.

Otra cuestión adicional sería la capacidad de aprendizaje de los sistemas inteligentes para lo que podrían establecerse límites en su diseño y concepción, de modo que únicamente aprendan tareas, tomen decisiones y lleven a cabo actuaciones contempladas en su programación según instrucciones y condiciones predefinidas, conforme a lo que se denomina "*code as law*", es decir, el propio código fuente incorporado en su diseño y programación contendría restricciones y controles a esa supuesta autonomía y libertad, lo que sería un postura diligente y prudente ante los riesgos evidenciados, por parte de quién tendría una mejor posición de cara a la gestión del riesgo en el ciclo de vista de estos sistemas, esto es, diseñadores, desarrolladores y fabricantes.

Y, por último, todo ello debería ir acompañado de información previa, clara y transparente del productor sobre los posibles riesgos asociados a las capacidades con las que dote sus productos inteligentes que, en una futura revisión de los marcos vigentes de responsabilidad, no debería tener eficacia eximente, sin perjuicio de que pueda tener eficacia atenuante, a fin de garantizar un resarcimiento efectivo.

La precitada Propuesta de Reglamento de 21 de abril de 2021 regula nuevas obligaciones de información muy detalladas.

4.3. Sujetos responsables

Si concurren los elementos precitados, la responsabilidad prevista en este marco regulador podrá exigirse al productor[694] del robot o sistema dotado de inteligencia

[694] Según el artículo 138 de la LGDCU, a los presentes efectos se considerará productor, además del definido en el artículo 5, el fabricante o importador en la UE de un producto terminado, de cualquier elemento integrado en un producto terminado o de una materia prima. Si el productor no puede ser identificado, será considerado como tal el proveedor del producto, a menos que, dentro del plazo de tres meses, indique al dañado o perjudicado la identidad del productor o de quien le hubiera suministrado o facilitado a él dicho producto. La misma regla será de aplicación en el caso de un producto importado, si el producto no indica el nombre del importador, aun cuando se indique el nombre del fabricante. Por otra parte, en su artículo 5, establece que, sin perjuicio de lo dispuesto en el precitado artículo 138, a efectos de lo dispuesto en esta norma se considera productor al fabricante del bien o al prestador del servicio o su

artificial de conformidad con lo previsto en sus artículos 138 y 5 del TRLGDCU, con posibilidad de atribuir dicha condición al fabricante o importador en la Unión Europea[695], así como al propio proveedor del producto en caso de que el productor no pueda ser identificado -salvo que lo identifique en el plazo de tres meses la identidad de productor o de que quien le hubiere suministrado el producto- o cuando el proveedor haya suministrado el producto a sabiendas de la existencia del defecto, conforme a lo dispuesto en el artículo 146 del mismo.

La cuestión es que puede concurrir la responsabilidad de distintos agentes durante el ciclo de vida de sistema inteligente como producto o servicio, de modo que el criterio establecido por este marco específico es la solidaridad entre todos ellos, lo que constituye una opción garantista del perjudicado y para la obtención de un resarcimiento efectivo.

El artículo 132 del TRLGDCU establece que "las personas responsables del mismo daño por aplicación de este libro lo serán solidariamente ante los perjudicados", de modo que, el que hubiera respondido ante el perjudicado, tendrá derecho a repetir frente a los otros responsables, según su participación en la causación del daño. Esta previsión se encuentra igualmente contemplada en el marco general sobre obligaciones previsto en el artículo 1144 del Código Civil español[696].

intermediario, o al importador del bien o servicio en el territorio de la UE, así como a cualquier persona que se presente como tal al indicar en el bien, ya sea en el envase, el envoltorio o cualquier otro elemento de protección o presentación, o servicio su nombre, marca u otro signo distintivo.

[695] En este sentido debe significarse la práctica judicial española sustentada en una visión proteccionista de la persona afectada. La Sala Primera del Tribunal Supremo, al amparo de las reglas generales de responsabilidad, no ha tenido ninguna duda en condenar a vendedores o distribuidores de productos defectuosos en supuestos en los que, sin embargo, había quedado acreditado que el defecto tenía su origen en el proceso de fabricación. Teóricamente, con el régimen procedente de la Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos, no resultaría posible alcanzar fallos en este sentido cuando el fabricante estuviera identificado. La práctica jurisprudencial evidencia que el Tribunal Supremo no se resiste a condenar al suministrador de un producto, al amparo de las reglas de responsabilidad por culpa. Lo que sucede es que, inevitablemente, la terminología que utiliza el Tribunal Supremo para referirse a la culpa es que el producto no ofrecía la seguridad que legítimamente cabe esperar, es decir, la terminología de un régimen legal de responsabilidad sin culpa en la que responde el productor, pero no el suministrador, por el carácter defectuoso del producto. STS de 25 de noviembre de 2013 (RJ 7827).

[696] Artículo 1144 CC: El acreedor puede dirigirse contra cualquiera de los deudores solidarios o contra todos ellos simultáneamente. Las reclamaciones entabladas contra uno no serán obstáculo para las que posteriormente se dirijan contra los demás, mientras no resulte cobrada la deuda por completo.

En base al mismo, en caso de aplicabilidad del régimen previsto en el TRLGDCU al supuesto planteado en materia de daños causados por un sistema inteligente, comportaría la responsabilidad solidaria del productor del hardware que integre dicho sistema, del productor de sensores y componentes electrónicos del mismo, del productor del sistema inteligente y del *software* sobre el que se sustenta, así como del proveedor del producto final.

El informe *Liability for Artificial Intelligence and other emerging digital technologies*[697] de 2019 incorpora la solidaridad como uno de los mecanismos que deberían contemplar los futuros marcos regulatorios que complementen los regímenes vigentes de responsabilidad, y la Propuesta de Reglamento sobre responsabilidad civil incorporada a la Resolución del Parlamento Europeo de 20 de octubre de 2020, así lo incluye como luego abordaré en su análisis.

La dificultad que puede plantear determinar quién de los proveedores es responsable cuando, en el mejor de los casos, se pueda identificar el conjunto de los que han participado en el proceso de fabricación del producto defectuoso, también podría resolverse por la teoría del *Market share liability*- o por cuota de mercado- utilizada en EE.UU., mediante la responsabilidad entre los fabricantes de acuerdo con su participación en el mercado del producto que da lugar al daño, si bien, la misma no se haya prevista en el marco europeo[698], el cual se basa en la responsabilidad solidaria de los mismos. Distintos autores como Robert Guillén[699], en el ámbito de la impresión 3D y 4D, no ve obstáculo en su aplicación ponderada con independencia de la solidaridad, en función de lo que interese proteger de inicio, esto es, al perjudicado o al productor participante no responsable, si bien, en la práctica, considero que la solidaridad ofrece mayores facilitades al perjudicado.

[697] *Liability for Artificial Intelligence and other emerging digital technologies.* UE. 2019. Recuperado de: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf. Consultado el 14.12.2020.
[698] NAVAS NAVARRO, S. (2016). "Smart Robots y otras máquinas inteligentes en nuestra vida cotidiana". *Revista CESCO Digital*. Nº 20. 2016. Pp. 105-106.
[699] ROBERT GUILLÉN, S. (2017). "Impresoras 3D y 4D", en Navas Navarro, S. (Dir.). *Inteligencia artificial, tecnología, derecho*. Valencia, Tirant lo Blanch, 2017. P. 229.

4.4. Carga de la prueba

Por lo que se refiere a la prueba, corresponderá a la persona perjudicada que pretenda obtener la reparación de los daños sufridos, probar el defecto, el daño y la relación de causalidad entre ambos, de conformidad con lo previsto en su artículo 139 del TRLGDCU.

En la práctica esto puede plantear enormes dificultades para la persona afectada como he referido, especialmente en caso de sistemas inteligentes más complejos y con mayores capacidades, especialmente aprendizaje profundo, en la medida que le exigirá recabar de inicio informes técnicos forenses dirigidos a acreditar esa falta de la seguridad esperada (física y/o virtual), el daño y la relación de causalidad, por lo que, como he expuesto anteriormente, constituye otro de los aspectos que evidencia la falta de adecuación de este régimen para resolver las distintas situaciones de responsabilidad por daños derivados del funcionamiento de sistemas inteligentes.

El precitado informe *Liability for Artificial Intelligence and other emerging digital technologies* recoge en su resumen ejecutivo el aumento de las dificultades de prueba y la necesidad de garantizar el derecho a la facilitación de la prueba, concluyendo la necesidad de que los marcos que complementen los regímenes de responsabilidad en el futuro contemplen funciones de registro -que ya incorpora la nueva propuesta de Reglamento de 21 de abril de 2021-, y ello estableciendo mecanismos sencillos consecuentes a la ausencia de información registrada o negación de acceso a los datos registrados, como el despliegue de una presunción *iuris tantum* de que se cumple la condición de responsabilidad que demostraría la información que falta, así como la directamente la inversión de la carga de la prueba en general cuando existan dificultades o costes desproporcionados para establecer el nivel de seguridad pertinente o para demostrar que no se ha cumplido el nivel de seguridad exigido.

La precitada Propuesta de Reglamento sobre inteligencia artificial de 21 de abril de 2021, analizada en el capítulo precedente, contempla obligaciones específicas de registro de eventos o logs, entre otras obligaciones de transparencia y explicabilidad, lo que facilitará la prueba, en particular, de la relación de causalidad.

4.5. Supuestos de limitación o exoneración de responsabilidad

El precitado marco, como también he anticipado anteriormente, contempla determinados supuestos de exoneración de la responsabilidad del productor en su artículo 140 del TRLGDCU, en el caso de que pruebe su concurrencia.

En particular, la primera causa es la relativa a no haber puesto circulación el producto. La cuestión es si un sistema de inteligencia artificial causa daños en entornos de pruebas (sea o no en entornos físicos o virtuales reales o con datos reales), si podría acogerse a esta causa de exoneración.

Para García Teruel[700], el productor se exoneraría de responsabilidad simplemente alegando que no había comercializado un vehículo autónomo inteligente, como ocurrió en un accidente de Uber en Arizona, lo que no impediría que la víctima pudiese ejercer su pretensión a través de la responsabilidad extracontractual analizada anteriormente conforme al artículo 1902 del Código Civil español.

Sin embargo, no comparto completamente esta tesis, en la medida que la cuestión radica en que debemos de entender por "puesto en circulación" y si deberíamos acudir a una interpretación amplia y proteccionista de dicho concepto en el ámbito tecnológico para asociarlo a su mera puesta en el tráfico o a disposición de sus usuarios, desvinculándolo de que sea estrictamente para su comercialización inmediata, incluyendo otros objetivos como la investigación, experimentación en entornos reales, en entornos de prueba con sujetos reales, el testeo y/o la validación previa a la comercialización.

De poder sostener este argumento, a mi juicio, un vehículo autónomo "gobernado" o "gestionado" por un sistema inteligente, que pueda estar circulando en fase de pruebas o piloto por vías abiertas o restringidas, como está ocurriendo o se ha solicitado ya en

[700] GARCÍA TERUEL, R.M. (2021). "El derecho de daños ante la inteligencia artificial y el machine learning: una aproximación desde las recomendaciones del Parlamento Europeo y del Grupo de Expertos de la Comisión Europea", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Tomo II. Ed. Aranzadi-Thomson Reuters. 2021. P. 1.030.

distintos países europeos, en la medida que está circulando y operando de modo efectivo en entornos, situaciones, con personas y/o con datos que pueden ser reales, podría considerarse ya en circulación aunque con las finalidades indicadas, y por consiguiente, aplicar este régimen específico de responsabilidad, siempre y cuando concurran los demás elementos requeridos por el mismo.

La segunda y tercera causa de exoneración de responsabilidad son las relativas a si, en base a las circunstancias del caso, es posible presumir que el defecto no existía en el momento en que se puso en circulación el producto o que el estado de los conocimientos científicos y técnicos existentes en el momento de la puesta en circulación no permitía apreciar la existencia del defecto, es decir, el denominado "riesgo de desarrollo"[701].

Estas dos causas de exoneración pueden ser de gran relevancia en caso de robots y sistemas dotados de inteligencia artificial, especialmente la más avanzada, con mayores capacidades, capacidad de autoaprendizaje e interacción con su entorno, en la medida que esa retroalimentación y aprendizaje pueden ser la causa de ese defecto de seguridad, no existente y no previsto en origen, lo que permitiría al productor exonerarse de responsabilidad.

El precitado informe *Liability for Artificial Intelligence and other emerging digital technologies* de 2019 considera que el productor debería ser estrictamente responsable de los defectos de la tecnología, incluso si dichos defectos aparecen después de la puesta en circulación del producto, siempre que el productor siga controlando las actualizaciones o mejoras de la tecnología. Según el grupo de expertos, el productor no debería poder eximirse se la misma alegando el riesgo de desarrollo.

La prueba de la concurrencia o no de estas causas resulta compleja, especialmente en función de las características y capacidades de cada sistema inteligente. En este sentido, la precitada García Teruel[702] considera que sería difícil justificar una exoneración por

[701] Camacho Clavijo considera que la exoneración de responsabilidad del fabricante podrá ser alegada cuando el productor pruebe que el defecto no era reconocible a la luz del conjunto de conocimientos que se encontraban disponibles en el ámbito científico y técnico en el momento de su circulación. Ver CAMACHO CLAVIJO, S. (2017). "La subjetividad "cyborg". En NAVAS NAVARRO, S. (Dir.). *Inteligencia artificial, tecnología, derecho*. Editorial Tirant lo Blanch. Valencia. Pp 254-255.

[702] GARCÍA TERUEL, R.M. (2021). "El derecho de daños ante la inteligencia artificial y el machine learning: una aproximación desde las recomendaciones del Parlamento Europeo y del Grupo de Expertos de la

riesgos de desarrollo, por cuanto que considera que deberían ser totalmente desconocidos e indetectables de acuerdo con el estado actual de la técnica, de modo que un robot dotado con inteligencia artificial y capacidad de autoaprendizaje cause un daño es hoy algo detectable y conocido. La autora considera que el defecto del producto también existiría cuando, "siendo previsible que el robot aprendiera una acción dañosa, el productor no informara de dicho riesgo en el momento de su comercialización (defecto de las instrucciones o información) o posteriormente (defecto post marketing), así como cuando no se hicieran las pruebas de seguridad pertinentes para descartar posibles daños causados por la IA".

No comparto plenamente esta opinión. En primer lugar, como luego expondré en mis consideraciones y conclusiones, reiterar que es necesario la revisión de estos marcos específicos de responsabilidad para adecuarlos a la realidad compleja que supone la inteligencia artificial, como recogía el informe precitado.

En segundo lugar, las capacidades de las que puede estar dotado un sistema inteligente puede incluir su autoaprendizaje y profundo, su entrenamiento y formación posterior, un grado determinado de autonomía, interacción con su entorno e incluso autoprogramación, así como los contextos de uso que pueden ser muy distintos para los que fue concebido o instrucciones del propio usuario, lo que imposibilita un conocimiento y detección de todos los posibles riesgos futuros o de posibles defectos relativos a un previsible uso indebido, lo que podría llevar al productor a exonerarse de responsabilidad, sin perjuicio de que pueda y deba informar de manera general sobre dichas capacidades y riesgos antes o después de su comercialización, no debiendo permitirse que una mera declaración genérica podría servirle según esta postura para eximirse de responsabilidad. De ahí, que para determinados sistemas con mayores capacidad y autonomía, debiera exigirse una responsabilidad absoluta por los riesgos asociados, dado que se podrían producirse situaciones de desamparo de los perjudicados.

Tal y como he referido anteriormente, la nueva Propuesta de Reglamento sobre inteligencia artificial del Parlamento Europeo y del Consejo, de 21 de abril de 2021, en

---

Comisión Europea", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Tomo II. Ed. Aranzadi-Thomson Reuters. 2021. P. 1.031

relación con los sistemas inteligentes de alto riesgo y entre otros aspectos, exige que: a) Se garantice la monitorización y seguridad durante todo el ciclo de vida del sistema; b) Que se prevean tanto los riesgos conocidos como previsibles en los sistemas de gestión de riesgos; c) Que se evalúen todos los riesgos que podrían surgir cuando el sistema de se utilice conforme a su finalidad prevista pero también cuando se le dé un uso indebido razonablemente previsible; d) Que la información preceptiva sobre el sistema incluya las circunstancias conocidas o previsibles que podrían afectar al nivel de precisión, solidez y ciberseguridad esperado, así como cualquier circunstancia conocida o previsible asociada a la utilización del sistema conforme a su finalidad prevista o a un uso indebido razonablemente previsible, que pueda dar lugar a riesgos para la salud y la seguridad o los derechos fundamentales y; e) Que se garantice la supervisión humana para prevenir o reducir al mínimo los riesgos para la salud, la seguridad o los derechos fundamentales que puedan surgir cuando un sistema inteligente de alto riesgo se utilice conforme a su finalidad prevista o cuando se le da un uso indebido razonablemente previsible.

No obstante, el Reglamento propuesto, como he referido anteriormente, circunscribe este uso indebido razonablemente previsible a la utilización de un sistema inteligente de un modo que no corresponda a su finalidad prevista, pero que pueda derivarse de un comportamiento humano o una interacción con otros sistemas razonablemente previsible.

Y, en tercer lugar, las necesarias pruebas previas, en el caso de sistemas complejos y opacos, considero que no podrían detectar todos los usos indebidos previsibles, riesgos e hipotéticos defectos concretos, dada la diversidad de capacidades y contextos de uso.

En consecuencia, en estos contextos, a mi juicio, una posible solución a estas cuestiones, con el objetivo de garantizar el resarcimiento efectiva del perjudicado, sería mediante el establecimiento de un régimen de responsabilidad objetiva absoluta para determinados sistemas más avanzados y con mayor autonomía y capacidades, que fije la misma en la órbita del productor o del operador, sin perjuicio de los matices coherentes en función de los distintos agentes que puedan intervenir en el ciclo de vida del sistema inteligentes, como formadores-entrenadores o el propio propietario o usuario, conforme abordaré en mi consideraciones y conclusiones finales.

El resto de causas previstas en el artículo precitado son que el producto no hubiere sido fabricado para la venta o cualquier otra forma de distribución con finalidad económica, ni fabricado, importado, suministrado o distribuido en el marco de una actividad profesional o empresarial, o que el defecto se debió a que el producto fue desarrollado conforme a normas imperativas existentes en el momento de su concepción.

Por último, este régimen prevé la exoneración o reducción de la responsabilidad en caso de concurrencia con la culpa del perjudicado o de una persona de la que éste deba responder civilmente, conforme a lo previsto en su artículo 145 TRLGDCU.

4.6. Límites temporales y de cuantía.

El régimen de responsabilidad por productos defectuosos tiene límites temporales que comportan la extinción de la responsabilidad, en particular, un plazo de preclusión de diez (10) años desde que se puso en circulación el producto conforme a su artículo 144 del TRLGDCU, lo que significa que, los derechos que asistan a la persona afectada (perjudicado) no podrían ser ejercitados transcurrido dicho plazo.

Asimismo, este régimen limita cuantitativamente la responsabilidad del productor, en la medida que se deducirá una franquicia de 500 euros de la cuantía de la indemnización de los daños materiales, y la responsabilidad civil global del productor por muerte y lesiones personales causadas por productos idénticos que presenten el mismo defecto, tendrá como límite máximo la cuantía de 63.106.270,96 euros.

4.7. La inteligencia artificial como servicio -AIaaS-

Por último, debo significar los supuestos en los que la puesta a disposición de un usuario-consumidor de un sistema de inteligencia artificial pueda considerarse como un servicio, esto es, *AI as a Service* o *AIaaS*-, en cuyo caso podría resultar de aplicación el régimen de responsabilidad previsto en los artículos 147 a 149 del TRLGDCU español.

En los últimos años la tecnología ha revolucionado los modelos de negocio y en todos los sectores, desde los más vanguardistas a los más clásicos en decadencia, como el video o la música con la irrupción de Netflix o Spotify. Asimismo, ha permitido reenfocar la tecnología como producto, a la tecnología como servicio, por ejemplo, con la irrupción del *Cloud Computing* y sus modelos SaaS (*Software as a Service*), IaaS (*Infraestructure as a Service*) o PasS (*Platform as a service*).

Los robots y sistemas dotados de inteligencia artificial avanzada pueden ser puestos en el tráfico bajo la modalidad de servicios, por ejemplo, robots asistenciales o *chatbots* de seguimiento para personas que requieran una atención especial domiciliaria o cabinas médicas inteligentes que efectúan un diagnóstico inicial completo a la persona que accede a su interior, para activar los protocolos sanitarios que correspondan en función del mismo.

Esta conversión permitiría, de un lado, exigir la responsabilidad contractual y extracontractual que proceda y, en especial, este marco normativo específico de responsabilidad, cuando resulte de aplicación en estos contextos, especialmente en atención a su ámbito objetivo y subjetivo.

La TRLGDCU recoge en sus artículos 147 a 149 un régimen general y especial de responsabilidad en caso de prestación de servicios a consumidores y usuarios, de modo que, parece de inicio exclusivamente acotado a nivel subjetivo a los mismos.

Conforme al régimen general previsto en los mismos, en especial, en su artículo 147, los prestadores de servicios serán responsables de los daños y perjuicios causados a los consumidores y usuarios, salvo que prueben que han cumplido las exigencias y requisitos reglamentariamente establecidos y los demás cuidados y diligencias que exige la naturaleza del servicio, lo que comporta la inversión de la carga de la prueba. A falta de previsión legal que contenga esas exigencias y requisitos, serán exigibles la diligencia y cuidados que la naturaleza del servicio exija, tomando como referencia las fuentes generales del contenido de la relación jurídica obligatoria, de conformidad con el 1258 del Código Civil, conforme a la buena fe, los usos y la ley.

Asimismo, se prevé un régimen especial, en el que se responderá de los daños originados en el correcto uso de los servicios, cuando por su propia naturaleza, o por estar así reglamentariamente establecido, incluyan necesariamente la garantía de niveles determinados de eficacia o seguridad, en condiciones objetivas de determinación, y supongan controles técnicos, profesionales o sistemáticos de calidad, hasta llegar en debidas condiciones al consumidor y usuario.

Esta responsabilidad se derivaría en caso de no respetar no solo la seguridad determinada sino también de su eficacia predeterminada, es decir, su capacidad para producir el efecto pretendido.

Conforme recoge este marco normativo especial, se consideran sometidos a este régimen de responsabilidad los servicios sanitarios, los de reparación y mantenimiento de electrodomésticos, ascensores y vehículos de motor, servicios de rehabilitación y reparación de viviendas, servicios de revisión, instalación o similares de gas y electricidad y los relativos a medios de transporte, es decir, servicios para los que ya se han desarrollado robots y sistemas dotados de inteligencia artificial, pero no se regulan en el mismo otros donde los riesgos son especialmente significativos, por ejemplo, los servicios asistenciales.

En consecuencia, es otra muestra más de la necesidad de adaptar el derecho vigente a las nuevas realidades, en especial, al objeto de contemplar también, expresamente, la responsabilidad del prestador frente a cualquier perjudicado, tenga o no la condición de consumidor.

Y, en el marco de servicios, la responsabilidad prevista en el TRLGDCU también se halla limitada cuantitativamente, al igual que la responsabilidad por productos.

Este régimen específico contemplado para los servicios podría ser un marco mucho más adecuado en los supuestos de servicios que puedan ser habitualmente gestionados, operados o prestados por un sistema inteligente, en la medida que se atribuye la misma al prestador del servicio, no al productor del sistema y sus elementos, y no limita la responsabilidad a un defecto con origen en la falta de la seguridad previsible, sino que lo amplía en los términos indicados, respondiendo de los daños y perjuicios causados, salvo

que el prestador pruebe que fue diligente, esto es, que cumplió las exigencias y requisitos reglamentariamente establecidos y los demás cuidados y diligencias que exigiera la naturaleza del servicio.

La cuestión es que este régimen específico requeriría la definición legal de las exigencias y requisitos de dichos servicios y funcionamiento de estos sistemas para la prestación de sus utilidades, por lo que de nuevo, se hace necesario un nuevo marco regulador que complemente estos regímenes, cuanto menos de manera específica para sistemas inteligentes y que dé cobertura no sólo a consumidores sino a usuarios de estos sistemas, sean consumidores actuantes en su esfera personal o profesionales y empresas.

4.8. Plazo de prescripción

A diferencia de la responsabilidad civil extracontractual, en la que el plazo de prescripción de la acción es de un año, la acción de reparación de los daños y perjuicios por productos defectuosos, conforme a lo previsto en el artículo 143 del TRLGDCU prescribirá a los tres años a contar desde la fecha en que el perjudicado sufrió el perjuicio, ya sea por defecto del producto o por el daño que dicho defecto le ocasionó, siempre que se conozca al responsable de dicho perjuicio.

Por su parte, la acción del que hubiese pagado la indemnización contra todos los demás responsables del daño prescribirá al año, a contar desde el día del pago de la indemnización.

4.9. Reflexiones globales

Los marcos reguladores de la responsabilidad civil por productos defectuosos vigentes en la UE y España no son adecuados para resolver todos los problemas que pueden plantear los sistemas de inteligencia artificial, especialmente ante su complejidad y posible opacidad, distintos grados de autonomía, mayores capacidades de autoaprendizaje,

pluralidad de contextos y agentes durante su ciclo de vida, pérdida o merma de la influencia control y dirección sobre el sistema o problemas de imputación.

Los sistemas inteligentes podrían tomar decisiones ante situaciones previstas o no previstas en su programación y actuar en función del contexto en base a su autoaprendizaje, interacción y experiencia.

Las situaciones y contextos de los que se pueden derivar daños son múltiples, en los que el diseñador, desarrollador, fabricante, proveedor de datos o del sistema, entrenador, propietario, operador y/o el usuario, entre otros agentes, pueden tener una participación absolutamente irrelevante, o ser partícipes o protagonistas de los mismos, ya sea por su concepción, uso, entrenamiento, interacción, falta de actuación requerida o de corrección, entre otros acciones, por lo que, en consecuencia, el grado de responsabilidad del diseñador, desarrollador, fabricante o proveedor es muy distinta, especialmente ante su previsión y diligencia, de modo que los niveles de diligencia y responsabilidad serán distintos en función del contexto específico.

La doctrina es cada vez más proclive al establecimiento de una responsabilidad objetiva por riesgo para productores y operadores de un sistema de inteligencia artificial que no esté limitado en cuanto al tipo de daños y perjuicios indemnizables, lo que proporcionaría los instrumentos más adecuados para garantizar un resarcimiento efectivo. Entre otros autores partidarios de esta opción, citar a Ebers[703] o Navas Navarro[704], sin perjuicio de la acción de repetición. Ante la pluralidad de los posibles sujetos involucrados durante todo el ciclo de vida, sería aplicable la solidaria como instrumento para facilitar al perjudicado su acción y resarcimiento efectivo.

Conforme recoge el informe precitado *Liability for Artificial Intelligence and other emerging digital technologies* de 2019, constituye una solución deseable a contemplar

[703] EBERS, M. (2016). "La utilización de agentes electrónicos inteligentes en el tráfico jurídico jurídico; ¿Necesitamos reglas especiales en el Derecho de la responsabilidad civil?", en *Indret: Revista para el Análisis del Derecho*. N. 3, 2016. Pp 15-16.
[704] NAVAS NAVARRO, S. (2017). "Derecho e inteligencia artificial desde el diseño. Aproximaciones", en Navas Navarro, S. et al. *Inteligencia artificial, tecnología, derecho*. Valencia, Tirant lo Blanch, 2017. P.44.

por los futuros marcos regulatorios, como así la recoge la Propuesta de Reglamento sobre responsabilidad civil de 20 de octubre de 2020.

El informe determina que se debería repercutir el daño a aquellos sujetos cuyo comportamiento jurídicamente reprochable causó el daño, o que se beneficiaron de la actividad que causó el daño, o que controlaban el riesgo de que se materializará o que eran los que de manera más económica podrían evitar los costos. Estos aspectos han sido criticados por algunos autores como el precitado Vazquez de Castro, en la medida que se mezclan y combinan posicionamientos, a juicio del mismo, sin definir unos criterios de imputación objetiva "sobre la base de imputar la responsabilidad a quien crea un riesgo que va más allá de los riesgos generales de la vida o que incrementa esos riesgos"[705].

Algunos autores como Ebers, consideran la necesidad de aplicar una responsabilidad solidaria entre productores y operadores en determinados contextos y uso, ante la dificultad de su discernimiento. Otros consideran que no es aceptable el riesgo de no tener el control y que el sistema tenga el poder de decisión, como la precitada Díaz Alabart[706].

A modo de síntesis, las principales razones por las que este sistema de responsabilidad no sería adecuado para proporcionar respuestas adecuadas a los distintos supuestos de daños causados por sistemas inteligentes, tanto las ya apuntadas como las que se exponen a continuación, podrían agruparse del siguiente modo:

a) Régimen de responsabilidad de ámbito de aplicación inicialmente restringido y con ciertas dificultas probatorias para el perjudicado, lo que limita su aplicación inicial a los distintos supuestos que se habitualmente pueden plantearse, sin perjuicio de la interpretación y aplicación posterior jurisprudencial, con la consiguiente inseguridad jurídica.

   La cobertura de este régimen específico analizado se limita, en relación con productos defectuosos a los daños personales, incluida la muerte, y los daños materiales causados, siempre y cuando éstos afecten a bienes y servicios destinados

[705] VÁZQUEZ DE CASTRO, E. (2020). "Aproximación a la responsabilidad derivada de los riesgos de la inteligencia artificial en Europa". Op.cit. P. 241
[706] DÍAZ ALABART, S. (2018). "Robótica y Responsabilidad Civil". Ponencia impartida en la Real Academia de Jurisprudencia y Legislación el 31 de mayo de 2018.

a un uso personal o consumo privado, y en tal concepto hayan sido utilizados principalmente por el perjudicado, lo que dejaría inicialmente fuera de su ámbito de protección los que afecten a bienes y servicios utilizados para un uso profesional, empresarial o comercial, así como los causados a bienes del perjudicado que hayan sido utilizados preferentemente en tal concepto, sin perjuicio de la doctrina emanada del TJUE en algunos asuntos, conforme analicé anteriormente.

La cobertura de este régimen de responsabilidad se halla inicialmente acotado subjetivamente en materia de servicios "defectuosos", en la medida que ya no se habla de daños y perjuicios a perjudicados sino a consumidores y usuarios. De modo que, en estos contextos, quedarían excluidos inicialmente de este régimen los servicios asociados a la inteligencia artificial y robóticos adquiridos por empresarios y profesionales para fines relacionados con sus actividades industriales, comerciales y profesionales, salvo en los supuestos en que puedan ser considerados "consumidores y usuarios" conforme a lo previsto en el TRLGDCU, los cuales constituyen una categoría principal y potencial de clientes para la industria de la inteligencia artificial, que se vería exonerada de su responsabilidad como proveedor de servicios en este contexto, a pesar de que posiblemente sea su principal fuente de ingresos y beneficios dentro de su volumen de negocio.

b) Régimen de responsabilidad basado en un criterio de imputación calificado por distintos autores precitados como "finalísticamente subjetivo", de modo que el fabricante quedaría exonerado de responsabilidad demostrando que, en el momento de la puesta en circulación del sistema, el estado de la ciencia y la tecnología no le permitía conocer la existencia del defecto, aun obrando con la máxima diligencia.

Es decir, podría constituir una causa de exclusión que sería esgrimida por los fabricantes de manera habitual sobre la base de los "riesgos de desarrollo", máxime ante las capacidades de las que se halle dotado el sistema, especialmente autonomía, autoaprendizaje, interacción con su entorno e impredecibilidad, así como en base a la posibles opciones de adaptación y customización de algunos sistemas inteligentes o de su posible modificación posterior por su operador, entrenador o usuarios para funciones inicialmente no previstas en el momento de su puesta en circulación.

c) Régimen de responsabilidad restringido a productos defectuosos en el ámbito de la UE y extendido de manera restringida a servicios "defectuosos" en España. Conforme al mismo, se consideran productos defectuosos aquellos que no ofrezcan la seguridad que cabría legítimamente esperar, es decir, se asocia el carácter defectuoso a su falta de seguridad.

La aplicación de este sistema especial de responsabilidad tiene un ámbito de aplicación restringido y circunscrito a los daños producidos por sistemas de inteligencia artificial "no seguros", lo que inicialmente podría dejar fuera errores y defectos en su diseño o fabricación (programación, funcionales, etc.) conforme a su tenor literal, salvo que se asocien a dicha falta de seguridad que, en mi opinión, se podrá producir en la mayoría de los casos.

Los sistemas que podrían verse involucrados podrían ser tanto aquellos que tengan cierto grado de autonomía como aquellos que carezcan de la misma, en los que no concurriría esa imposibilidad eximente de garantizar la seguridad conforme a su imprevisibilidad.

Este régimen de responsabilidad no resultaría inicialmente de aplicación para el resto de supuestos en los que no concurra ese "defecto", especialmente aquellos en los que no existiera ningún defecto, ningún vicio oculto, ni instrucciones incorrectas, sino un sistema de inteligencia artificial más avanzado o incluso "fuerte", correctamente construido inicialmente, que procesa y razona lógicamente con cierto grado de autonomía, que toma decisiones y realiza acciones de manera supuestamente independiente, informada y relativa o totalmente imprevisible para su diseñador o desarrollador.

Y reitero que "inicialmente", dado que, en mi opinión, la seguridad y el control humano, exige restringir al máximo en su diseño y fabricación dicha autonomía y sus capacidades de decisión, acción e improvisación, así como dotarlos de supervisión y control humano y herramientas al efecto durante todo su ciclo de vida, conforme a los marcos éticos objeto de pretendido consenso internacional y que

constituye una exigencia para los nuevos marcos reguladores propuestos en materia de inteligencia artificial, conforme recogen las propuestas europeas de octubre de 2020 y de abril de 2021, si bien, circunscritas a sistemas inteligentes de alto riesgo.

Si partimos de las propuestas reguladoras precitadas a nivel ético y jurídico, considero indiscutible que la *Ethics, security & compliance by design* debería impedir poner en el mercado un sistema inteligente de alto riesgo con unas características y capacidades cuyos riesgos no hayan sido previamente analizados y valorados por su productor o fabricante, con aplicación de los controles y medidas necesarias para restringir y controlar capacidades, reducir el nivel de riesgo -real o potencial razonablemente esperable- al máximo y llevarlo a un nivel de riesgo resultante o residual aceptable debidamente informado, y durante todo su ciclo de vida.

Si garantizamos un enfoque de riesgos, seguridad y otros principios éticos y normas en todo el ciclo de vida de estos sistemas, cualquiera que sea su riesgo inicial y empezando en su diseño y concepción, los productores y fabricantes no deberían poner en circulación sistemas de inteligencia artificial que no garanticen un nivel de riesgo gestionado aceptable, dado que de otro modo, la falta de los controles, restricciones y medidas de seguridad adecuados para reducir en origen (en su diseño) el riesgo inherente a los sistemas de inteligencia artificial en atención, en especial, en atención a su naturaleza, características y capacidades, debería poder conllevar la exigencia directa de responsabilidad conforme a este régimen especial, en base a la posibilidad de reducir y controlar, cuanto menos en parte, el riesgo inherente asociado a los mismos, por ser quienes estaban en la mejor posición de hacerlo.

d) Imposibilidad de reproche hacia el productor o fabricante ante determinados sistemas, según algunos autores, ante su imposibilidad de ofrecer y garantizar una expectativa de seguridad que pueda ser esperable en sistemas de inteligencia artificial más avanzados o “fuertes”, por considerarlos potencialmente inseguros y no ser posible la previsibilidad y anticipación por parte del productor o fabricante en base a los infinitos contextos en los que pueda operar un sistema de ese tipo en su interacción con el contexto donde opere y de sus reacciones.

Es decir, partiríamos de sistemas caracterizados por la inseguridad y peligrosidad por lo que, a mi juicio, de antemano y consecuentemente, ni tan siquiera deberían llegar al mercado sin sujeción a un marco de seguridad mínimo, a pesar de los beneficios que pudieran suponer para el ser humano, y configurando una responsabilidad absoluta hacia el propio fabricante y/o operador que pone el riesgo en el mercado.

Sobre este aspecto, me remito a mi posicionamiento sobre esta cuestión repetidamente expuesto en mis reflexiones previas, dado que el productor o fabricante debe establecer mecanismos de seguridad que limiten de manera efectiva y reduzcan al máximo posible los riesgos -reales o potenciales- derivados de la autonomía, capacidad de autoaprendizaje, libertad e impredecibilidad restringida que confiera a los sistemas que fabrique y ponga en el mercado, incluso los relacionados con usos indebidos, especialmente los razonablemente previsibles, y ello mediante controles y medidas de seguridad preventivas, detectivas, reactivas y evolutivas frente a los riesgos derivados de dichas capacidades durante todo el ciclo de vida de los mismos, con sujeción al control y supervisión humana y dotación de herramientas para posibilitarla, y. en cualquier caso, partiendo de dos cuestiones clave: De un lado, que nunca podría evitarse cierto nivel de riesgo que, en cualquier caso, debería llevarse a un nivel aceptable e inherente a este tipo de sistemas y, de otro, que nunca debería ponerse en el mercado un sistema de alto riesgo o de riesgo inaceptable sin los controles y mecanismos de supervisión adecuados durante todo su ciclo de vida.

e) Sistema de responsabilidad *cuasi objetiva* que exigiría que la persona perjudicada que pretenda obtener la reparación deba probar, de un lado, el "defecto" del producto, es decir, que no ofrecía la seguridad esperable en base a los distintos criterios precitados, de otro probar el daño y, por último, la relación de causalidad entre ambos.

Es obvia la complejidad de acreditar todo ello como expuse al analizar tanto el régimen general de responsabilidad extracontractual cono el especial en materia de productos defectuosos, especialmente ante las características propias de estos sistemas, su complejidad y posible falta de transparencia, opacidad y explicabilidad.

f) Régimen de responsabilidad que no impide la puesta en circulación de un producto peligroso de antemano, cuya falta de seguridad es esperada por el fabricante y que, supuestamente, también debería ser esperada por el consumidor, especialmente el error, conforme he expuesto en otros apartados de esta investigación.

g) Régimen de responsabilidad que inicialmente sólo alcanza a los daños personales (corporales) y a los morales derivados de la muerte, quedando excluidos el resto de daños morales que sufra el perjudicado, sus parientes o allegados que no deriven de la muerte, que pueden ser especialmente significativos en los casos de sistemas de inteligencia artificial más avanzada utilizados en el cuidado y asistencia de personas, ante la profunda relación e interacción con el humano y los graves daños derivados de los mismos, que podrían impactar gravemente contra la dignidad, intimidad, autoestima y sentimiento de las personas.

Las características de este sistema de responsabilidad comportan algunas ventajas para el perjudicado, en especial, su carácter parcialmente absoluto en determinados aspectos, en la medida que no pueden prevalecer y operar las cláusulas de exoneración o limitación de la responsabilidad que hayan podido incorporarse en contratos o pedidos.

Y también es solidaria, directa y automática, si bien, como he referido, la persona afectada que pretenda obtener la reparación de los daños sufridos tendrá que probar el defecto, el daño y la relación de causalidad entre ambos, lo que comporta que no se trate de una responsabilidad totalmente objetiva en la que el mero hecho de sufrir un daño implique la existencia de una responsabilidad, pues hay que probar que el daño ha sido debido al defecto en el producto o servicio.

Además, comporta algunos aspectos de especial relevancia en relación con sistemas dotados de inteligencia artificial complejos, como he expuesto, en particular, la ausencia de responsabilidad del productor de una parte o de un elemento integrante de un producto o sistema terminado, si demuestra que el defecto es imputable a la concepción del producto al que ha sido incorporado o a las instrucciones dadas por productor o fabricante del producto, cuestión que se suscitará con toda seguridad en el caso de sistemas complejos compuestos por *hardware*, *software*, algoritmos, datos, etc, especialmente en los casos en que el diseño y funcionamiento del *hardware* asociado al sistema inteligente

sea especialmente relevante para la obtención el resultado esperado, por ejemplo, robots de precisión quirúrgica, exoesqueletos o prótesis inteligentes para asistencia a la movilidad de personas con diversidad funcional, marcapasos, estimuladores vagales o dispensadores de insulina inteligentes, etc.

5. Responsabilidad civil extracontractual vs responsabilidad civil del fabricante por producto defectuoso

El régimen de responsabilidad civil extracontractual sustentado en la culpa sería el régimen general a aplicar por defecto en el caso de daños derivados de sistemas de inteligencia artificial. No obstante, este régimen convive con otro especial anteriormente analizado, el de la responsabilidad civil "objetiva" del fabricante por producto defectuoso, así como otros sectoriales, abordados en este capítulo.

Conforme destaca Vázquez de Castro[707], "habrá supuestos en los que encaje bien el régimen jurídico de la responsabilidad por productos defectuosos, pero habrá otros en los que la víctima del daño podrá optar por otro régimen de responsabilidad que mejor se ajuste a las circunstancias del caso y las características del daño ocasionado".

De hecho, conforme vayan aumentando los niveles de autonomía de los sistemas inteligentes y aumenten sus capacidades, entre otras, las de autoaprendizaje, el poder de control sobre los dispositivos o sistemas y sobre sus riesgos se irá desplazando del usuario al fabricante del mismo o de alguno de sus componentes, por ser quién estaría en mejor disposición de controlar el riesgo -real o potencial- en su diseño y concepción, por lo que podrían resultar de aplicación las normas de responsabilidad del fabricante de directa aplicación en España y la UE, las cuales fueron creadas en otro contexto muy diferente al actual y, aunque han sido modificadas desde su aprobación, considero inevitable su revisión, adecuación y complemento, al objeto de adaptarlas a los nuevos retos que plantea la tecnología y, en particular, de la inteligencia artificial.

[707] VÁZQUEZ DE CASTRO, E. (2020). "Aproximación a la responsabilidad derivada de los riesgos de la inteligencia artificial en Europa". Op.cit. P. 236.

Los diversos contextos, sistemas, capacidad, características, ciclo de vida, agentes involucrados y casuística no permiten de inicio generalizar un régimen único de responsabilidad civil por daños para todo, como destacan autores como el precitado Vázquez de Castro[708].

Sin embargo, ambos regímenes evidencian sus limitaciones para poder dar solución a los distintos contextos que pueden generarse en el ámbito del funcionamiento y uso de sistemas inteligentes, especialmente ante sistemas más avanzados y fuertes, lo que a mi juicio requiere una revisión de los mismos y nuevos marcos que complementen y actualicen los mismos ante una realidad tan compleja como lo es la inteligencia artificial.

El informe *Liability for Artificial Intelligence and other emerging digital technologies* citado anteriormente, concluye a favor de la coexistencia y superposición de los distintos regímenes de responsabilidad en materia contractual, extracontractual, por culpa, objetiva por riesgos y por productos defectuosos, ante las complejidades, contextos y diversidad de agentes, ante la imposibilidad de que exista un régimen único según el mismo. Asimismo, este mismo informe concluye sobre la necesidad de llevar a cabo "ciertos ajustes en los regímenes de responsabilidad nacionales y de la UE" y, en especial, que "es necesario considerar las adaptaciones y modificaciones de los regímenes de responsabilidad existentes".

La responsabilidad solidaria y el futuro establecimiento de una responsabilidad objetiva absoluta en la órbita de los productores de sistemas inteligentes más avanzados, con mayor autonomía o, en cualquier caso y con independencia de su categorización como "débil" o "fuerte", de alto riesgo, contribuirá a garantizar el resarcimiento efectivo del perjudicado, sin perjuicio de su acompañamiento de otras medidas complementarias, como la exigencia de seguros.

[708] VÁZQUEZ DE CASTRO, E. (2020). "Aproximación a la responsabilidad derivada de los riesgos de la inteligencia artificial en Europa". Op.cit. P. 238.

6. La responsabilidad del usuario

La responsabilidad del usuario depende, entre otros factores, de si el uso del sistema inteligente por parte del mismo se lleva a cabo en calidad de empresario o profesional, o en el ámbito privado, personal o doméstico.

En primer lugar, el usuario que adquiere y utiliza el sistema inteligente en el marco de su actividad empresarial (empresarios) o profesional (profesionales) con ánimo de lucro es responsable de los daños que cause a la persona que sufra un daño derivado de dicho uso, si bien, deberíamos diferenciar si dicho el daño se produce en un contexto de relación contractual entre la persona afectada y el empresario o profesional, en cuyo caso, podrá exigírsele una responsabilidad contractual, o en un contexto ajeno a cualquier relación contractual previa, en cuyo caso le resultaría exigible una responsabilidad extracontractual.

Si nos encontráramos ante un sistema de inteligencia artificial con mayores capacidades, más avanzada o "fuerte", algunos autores[709] valoran la posibilidad de aplicar una responsabilidad vicaria o por hecho ajeno cuando no exista una relación contractual entre la persona afectada y el empresario o profesional, y de una responsabilidad contractual indirecta cuando entre ambos exista un contrato, y todo ello sobre la base de lo dispuesto en el artículo 1903.4º del Código Civil español, que regula la responsabilidad del empresario por los daños cometidos por sus empleados o dependientes a su servicio.

No comparto esta opinión en el ámbito español por los motivos expuestos al analizar otros apartados dentro de esta investigación, conforme al interpretación de los preceptos que la regulan, en la medida que los sistemas inteligentes actuales carecen de la condición jurídica de persona, por lo que ni pueden considerarse "empleados" ni "dependientes", por lo que no se hayan incluidos en los supuestos expresamente previstos en el precepto indicado. En todo caso, constituirán objetos, medios o recursos a disposición del empresario o profesional, bien para gestionar su organización, desarrollar su actividad empresarial o profesional o prestar servicios a sus clientes.

[709] NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Op. Cit. P. 74.

El informe *Liability for Artificial Intelligence and other emerging digital technologies,* citado anteriormente, recoge en una de sus conclusiones que, si "el daño es causado por una tecnología autónoma utilizada de manera funcionalmente equivalente al empleo de auxiliares humanos, la responsabilidad del operador por hacer uso de la tecnología debería corresponder al régimen de responsabilidad vicaria de un principal por tales auxiliares".

En este sentido, el propio informe citado destaca como, en la actualidad, distintos Estados miembros prevén en sus ordenamientos jurídicos la responsabilidad indirecta por la conducta del auxiliar con diverso alcance[710].

La responsabilidad por los daños ocasionados por estos medios empresariales recaerá en el empresario o profesional propietario o poseedor de los mismos por cualquier título.

La cuestión es la naturaleza de dicha responsabilidad en función del origen de la conducta, esto es, si concurre la culpa o negligencia del empresario o únicamente se considera imputable al sistema autónomo.

A priori, como propietario o usuario del mismo, debería responder hasta donde alcance su culpa y negligencia, pero en la medida que la conducta dañina tenga su origen exclusivamente en el sistema que actuó de manera autónoma e independiente a las instrucciones del empresario, será al productor o fabricante del mismo el que debería responder frente a cualquier tercero en base a los regímenes analizados.

La cuestión es si el empresario o profesional -no meramente usuario, sino operador-, debiera igualmente responder por ello.

[710] El artículo 1315 del ABGB austríaco establece que "Quien, para la dirección de sus asuntos, se sirva de una persona no apta o de una persona peligrosa a sabiendas, es responsable de los daños que dicha persona cause a otra en esa calidad". El artículo 831 del BGB alemán, establece que el principal puede excusarse "cuando el principal ha actuado con la debida diligencia en la selección del agente y -en la medida en que tiene que proporcionar equipos o herramientas o tiene que supervisar el cumplimiento de las obligaciones- ha actuado con la debida diligencia en dicha provisión y supervisión, o cuando el daño se habría producido incluso si se hubiera ejercido dicha diligencia". El artículo 429 del Código Civil polaco, establece que el principal es responsable de la conducta ilícita (pero no necesariamente culpable) del agente, a menos que el principal haya elegido cuidadosamente al agente o haya elegido a un agente profesional, y el artículo 430 del Código Civil polaco, hace al principal estrictamente responsable de la conducta culpable del agente si éste es un subordinado del principal.

En mi opinión, si el sistema fue adquirido con todas sus capacidades y riesgos asociados a las mismas -conocidos o no- para ser utilizado como usuario para la gestión de su organización y actividades, debiera quedar al margen de una responsabilidad recayente en el productor o fabricante, por resultar igualmente parte perjudicada por sus acciones, aunque lo sea indirectamente.

Ahora bien, si el sistema fue adquirido para proporcionar servicios a terceros y lucrarse directamente de ello, el empresario o profesional que opera el mismo con dicha finalidad, con sus capacidades y sus riesgos asociados, este debería igualmente responder ante la persona afectada en el mismo plano que el productor o fabricante de manera *cuasi objetiva* y solidaria, para lo que debería igualmente disponer de los oportunos y necesarios seguros, y en caso de exigencia de responsabilidad asumir la parte de responsabilidad que le corresponda en la cadena en función de su nivel de control de riesgo, a salvo de su acción de repetición contra el fabricante respecto del resarcimiento que anticipe o abone en exceso a su nivel de responsabilidad exigible a aquél.

En segundo lugar, el usuario del sistema inteligente que adquiere el mismo para su propio uso privativo, personal o doméstico, es decir, sin finalidad empresarial o profesional, ya sea propietario, licenciatario, poseedor o cesionario del mismo por cualquier título (alquiler, *renting*, *leasing*, etc), mientras se halle en posesión del mismo, debería ser responsable civil de los daños que genere a terceros la cosa que posea o use cuando tenga su origen en su acción u omisión a través de la misma.

Si los daños tienen su origen en la conducta autónoma e impredecible sobre la que el usuario no puedo ejercer ningún control razonable para evitarla, considero que concurriría en el mismo una situación de exención de responsabilidad, en todo caso, por caso fortuito o fuerza mayor, recayendo toda la responsabilidad en el fabricante. Si tuvo la opción de evitarla y no lo hizo cuando disponía de un control sobre el riesgo, debería responder del mismo directamente.

El informe precitado establece en sus conclusiones que, de un lado, que una persona que maneja una tecnología debería estar sujeta a una responsabilidad estricta por los daños resultantes de su funcionamiento, y, de otro, que una persona que utiliza una tecnología con cierto grado de autonomía no debería ser menos responsable del daño resultante que

si dicho daño hubiera sido causado por un auxiliar humano. El argumento esgrimido me parece muy razonable, considerando que el uso de la asistencia de una máquina autónoma y con capacidad de autoaprendizaje no debería tratarse de forma diferente al empleo de un auxiliar humano, si dicha asistencia provoca un daño a un tercero, conforme a un principio de equivalencia funcional. Además, determina que "los fabricantes de productos o contenidos digitales que incorporan tecnología digital emergente deben ser responsables de los daños causados por los defectos de sus productos, incluso si el defecto fue causado por los cambios realizados en el producto bajo el control del productor después de su comercialización, especialmente mediante actualizaciones o mejoras, significando que no debería aplicarse una exención de responsabilidad por riesgo de desarrollo. Asimismo, el informe concluye que la responsabilidad estricta debería recaer en la persona que controla el riesgo relacionado con el funcionamiento de la tecnología y que se beneficia de su funcionamiento, en clara referencia a su operador.

El propietario para uso doméstico o particular debería responder por los daños que cause el sistema del que es titular en el mismo plano objetivo que el fabricante frente a terceros, sin perjuicio de su derecho de repetición frente a éste. No obstante, los futuros seguros sobre los robots o sistemas inteligentes avanzados exigible a sus fabricantes deberían tener como beneficiario no sólo el fabricante sino los propietarios de los sistemas a los que se comercialice, correlacionado con un registro de sistemas inteligentes y sus propietarios-operadores.

Si el propietario lo cede a un tercero, debería concurrir una responsabilidad solidaria a partes iguales entre el propietario y el usuario poseedor del sistema, en base al control

La persona perjudicada podría dirigir su acción frente a cualquiera de los dos. No obstante, el que responda dispondrá de la acción de repetición frente al corresponsable para recuperar lo pagado conforme a lo dispuesto en los artículos 1144 y 1145 del *Código Civil* español.

La responsabilidad de ambos debería ser objetiva y por riesgo, en la medida que quien crea un riesgo, aunque su conducta inicial sea lícita y aunque emplee la diligencia debida en la misma, debe soportar las consecuencias derivadas de su actuar peligroso del que se beneficia. La construcción de la exigencia de esta responsabilidad, partiría del principio

regulado en el artículo 5:101 de los *Principios de Derecho Europeo de la Responsabilidad Civil*[711]. En su artículo 5:102 establece la posibilidad de que las autoridades nacionales contemplen otros supuestos de responsabilidad objetiva por actividades peligrosas, pudiendo ser aplicado el principio de responsabilidad objetiva analógicamente a otras situaciones que originen un riesgo de daño. No obstante, no se trataría de una responsabilidad definida *ab initio* con la consiguiente inseguridad jurídica en caso de su reclamación vía judicial.

Estos principios permitirían la responsabilidad del agente bajo estos parámetros, aunque el demandado haya ejercido el máximo cuidado, tanto desde el punto de vista del estándar objetivo como subjetivo, respondiendo en consecuencia por la fuente de peligro que se encuentra dentro de su esfera[712].

En caso de hurto, robo o apropiación indebida, debería responder exclusivamente el usuario autor del ilícito penal.

7. Reflexiones finales en materia de responsabilidad

7.1. Generales.

A falta de un marco regulador que revise, actualice, complemente y/o modifique el marco jurídico vigente y que pueda regular la inteligencia artificial, en caso de daños causados por un sistema de inteligencia artificial se deberá aplicar el régimen general de responsabilidad civil contractual o extracontractual previstos en el Código Civil español, en su caso, así como el marco especial de responsabilidad por productos defectuosos previsto en el TRLGDCU cuando resulte de aplicación y resulte el más adecuado.

El régimen de responsabilidad extracontractual basado en la culpa y en la acreditación de la misma resulta inadecuado ante una realidad tecnológica tan compleja como la

[711] Recuperado de: http://www.egtl.org/PETLSpanish.html. Consultado el 02.01.2021.
[712] NÚÑEZ ZORILLA, M.C. (2018). "Los nuevos retos de la Unión Europea en la regulación de la responsabilidad civil por los daños causados por la inteligencia artificial". *Revista Española de Derecho Europeo*, Nº. 66, Editorial Civitas (Thomson Reuters Aranzadi). Navarra 2018. Pp. 9-53.

inteligencia artificial, dado que no puede proporcionar soluciones adecuadas a los distintos supuestos que se puedan plantear en los que se vean involucrados sistemas inteligentes. No obstante, la construcción doctrinal y jurisprudencial de una responsabilidad *cuasi objetiva* podría proporcionar soluciones en la práctica a algunos de los supuestos que se puede producir en el ámbito de la responsabilidad por daños causados por sistemas inteligentes, hasta la disposición de un marco regulador adecuado.

En cualquier caso, serán compatibles las acciones por responsabilidad contractual o extracontractual con las derivadas de productos defectuosos reguladas en España en la TRLGDCU, en los casos en los que concurran los requisitos de las mismas.

Conforme se ha expuesto, se trataría de un concurso de normas para la exigencia de responsabilidad derivada del funcionamiento y/o uso de sistemas inteligentes, de modo que habrá supuestos en los que el contexto determinará la conveniencia de utilizar uno y otro.

De este modo, si existe un vínculo contractual entre el creador del sistema o *software* subyacente defectuoso y el perjudicado por los defectos del mismo, podrían ejercitarse ambas acciones y, del mismo modo, de no existir dicho vínculo contractual entre los mismos, podrían ejercitarse las acciones reguladas en la TRLGDCU en caso de producto defectuoso y las de responsabilidad extracontractual reguladas en el artículo 1902 del Código Civil español.

La facultad de ejercitar junto con la acción por producto defectuoso, las de responsabilidad contractual o extracontractual, obedece a la necesidad de protección y fortalecimiento de la posición del perjudicado, de modo que sea íntegra y efectivamente resarcido por los daños causados.

Dicha previsión se hallaba ya prevista en el artículo 15 de la Ley 22/1994, posteriormente derogada, y que ha pasado a regularse en el artículo 128.II del TRLGDCU que establece que “Las acciones reconocidas en este libro no afectan a otros derechos que el perjudicado pueda tener a ser indemnizado por daños y perjuicios, incluidos los morales, como consecuencia de la responsabilidad contractual, fundada en la falta de conformidad de los

bienes o servicios o en cualquier otra causa de incumplimiento o cumplimiento defectuoso del contrato, o de la responsabilidad extracontractual a que hubiere lugar".

Conforme fue objeto de análisis, la responsabilidad por producto o servicio defectuoso no cubre todos los daños que puedan causarse a la persona perjudicada, cuyos ámbitos de aplicación -objetivo y subjetivo- y de indemnización se hallan restringidos en función del contexto.

La cuestión es si estas acciones pueden interponerse alternativa o cumulativamente.

Tradicionalmente la doctrina ha considerado que el perjudicado podrá optar por la opción más favorable, tanto en términos sustantivos -especialmente en relación con la indemnización susceptible de reclamación- como procesales -prescripción o prueba-, y podrían instarse cumulativamente en una misma demanda varias acciones por distintos tipos de responsabilidad.

Esta solución fue propuesta por distintos autores como Jiménez Liébana o Moreno Quesada y Soler Matutes[713], y refrendada por la jurisprudencia sucesiva, especialmente ante la doctrina consolidada del Tribunal Supremo español sobre la unidad de responsabilidad civil o "unidad de culpa", en el sentido de que cuando un mismo hecho sea susceptible de generar culpa contractual y extracontractual, podrán ejercitarse las acciones derivadas de las mismas de manera alternativa o subsidiaria, o incluso proporcionando el supuesto fáctico al juzgador para que sea éste quien aplique la norma más adecuada y beneficiosa para el perjudicado.

En cualquier caso, como he referido, todos estos marcos pueden resultar, a mi juicio, inadecuados e insuficientes ante la propia naturaleza de la realidad compleja de la que pueden derivarse los daños y los diversos contextos en los que se pueden producir dentro del ciclo de vida de un sistema inteligente, su tipología, características y capacidades del mismo -especialmente su grado de autonomía, autoaprendizaje, interacción con su entorno, autoprogramación o impredecibilidad-, como del sector donde opere, uso,

[713] SOLER MATUTES, P. (2006). "El contrato de desarrollo de software. La responsabilidad de las partes", en Soler Matutes, P. (Dir.), *Manual de Gestión y Contratación Informática*. Editorial Aranzadi. 2006. Pp. 716-718.

finalidad, aplicación y sujetos que puedan intervenir en su causación durante su ciclo de vida, dados los elementos requeridos por estos regímenes de responsabilidad, limitaciones en su ámbito de aplicación restringido (productos defectuosos), dificultad de prueba, causas de exoneración y demás aspectos analizados.

En este sentido, este posicionamiento es compartido por autores como Ramón Fernández[714].

El sistema común de responsabilidad civil y el sistema de responsabilidad por productos defectuosos no pueden proporcionar una solución adecuada actual a todos los supuestos que pueden plantearse en la práctica en relación con los daños causados por sistemas de inteligencia artificial.

El marco regulador actual puede dar soluciones a algunas de las situaciones que se pueden plantear en materia de responsabilidad por daños causados por robots o sistemas dotados de inteligencia artificial, pero no a todas.

La preocupación de los legisladores es creciente y prueba de ello son las iniciativas del Parlamento Europeo objeto de análisis en esta investigación, especialmente en materia de responsabilidad civil, orientadas a motivar los análisis y reflexiones necesarias para la creación de un marco regulador de la inteligencia artificial que se integre, revise, actualice, complemente y, en su caso modifique, algunos de los regímenes actuales de responsabilidad civil en el ámbito de la UE, al objeto de proporcionar soluciones adecuadas y armonizadas a las nuevas realidades y necesidades evidenciadas en esta materia.

En relación con todo ello, considero que no debemos "reinventar la rueda", sino que se pueden y deben tomar como referencia y punto de partida inicial para dichos análisis y reflexiones los marcos actuales de responsabilidad civil por riesgos, de responsabilidad

[714] RAMÓN FERNÁNDEZ, F. (2019). "Robótica, inteligencia artificial y seguridad: ¿Cómo encajar la responsabilidad civil?". *La Ley Digital* 9365. Wolters Kluwer 2019. Pp. 1-13.

cuasi objetiva o de responsabilidad objetiva como, por ejemplo, la responsabilidad civil por daños corporales derivados de la circulación de vehículos a motor o de aeronaves.

Los enfoques para la creación de nuevos marcos de responsabilidad en el ámbito de la inteligencia artificial, a mi juicio, deben sustentarse en primer lugar, en el principio de precaución, ante la existencia de riesgos potenciales inherentes conforme a su naturaleza, características y posibles capacidades y la incertidumbre técnico-científica sobre su producción, así como, en segundo lugar, en un enfoque de riesgos. De este modo, la mera creación de un riesgo y daño potencial debería comportar el nacimiento de la obligación de responder por el productor, fabricante, propietario u operador según el caso, en el supuesto de que efectivamente se materialice el daño, con mecanismos adicionales para asegurar un resarcimiento efectivo del daño, en especial, de un seguro obligatorio a concertar por los sujetos potencialmente responsables y, en su caso, de un fondo de compensación que responda en caso de ausencia o insuficiencia del mismo.

Asimismo, en tercer lugar, considero que estos nuevos marcos de responsabilidad en el ámbito de la inteligencia artificial deben sustentarse igualmente en los principios y normas éticas esenciales que están siendo objeto de un pretendido consenso mayoritario a nivel internacional y que los futuros marcos jurídicos deben convertir en vinculantes, siguiendo la técnica utilizada para la elaboración de las sendas propuestas de Reglamento en materia ética y de responsabilidad que acompañaron a las Resoluciones del Parlamento Europeo de 20 de octubre de 2020, así como en la propuesta de Reglamento de 21 de abril de 2021.

Este es uno de los principales motivos por los que abordé esta investigación desde un enfoque global, ético, jurídico y de seguridad para abordar uno de sus principales retos que constituye su eje central, como lo es la responsabilidad.

La construcción de estos marcos sobre estos principios y normas éticas esenciales garantizarán una inteligencia artificial segura, robusta, fiable, transparente, trazable, explicable, que rinda cuentas, auditable, responsable, respetuosa del marco legal y de los derechos y libertades fundamentales y sometidas al control y supervisión humana durante todo su ciclo de vida.

Corresponde al ser humano definir su futuro, estableciendo los objetivos y definiendo las estrategias y medios para conseguirlo.

En cualquier caso, cualquier marco regulador de la inteligencia artificial debería partir de una visión global de sus posibles retos y riesgos y no circunscribirse a los más comunes o previsibles, considerar todos los intereses, bienes, derechos y sujetos que pueden verse afectados por los mismos, así como considerar las distintas tipologías de sistemas inteligentes -en base a capacidades y nivel de riesgo-, su grado de autonomía, independencia, capacidad de aprendizaje e interacción, control humano, predictibilidad, decisiones y acciones para las que este diseñado y autorizado, sectores, aplicaciones, posibles usos, etc.

Del mismo modo, los futuros marcos reguladores deberán definir el marco de requerimientos y obligaciones jurídicas de los sistemas inteligentes, especialmente los de mayor riesgo, al objeto de actuar como marcos preventivos y proactivos, reduciendo la activación de marcos reactivos una vez ya materializado el riesgo, como son los de responsabilidad, administrativos-sancionadores y, en su caso, penales o, en caso de activación, facilitando la depuración de responsabilidades y el resarcimiento efectivo de las víctimas.

Cuanto mayor sea la autonomía de los sistemas y, a mi juicio, también cuando mayor sea la capacidad de autoaprendizaje del sistema inteligente o incluso de autoprogramación, más complicado será aplicar los marcos reactivos existentes de responsabilidad civil, tanto contractual como extracontractual, entre otros factores, en la medida que la responsabilidad de los distintos agentes se diluye. Así lo destacan distintos autores como Santos González[715] respecto del primero de los aspectos, que afirma que cuanto más aumenta la autonomía, la responsabilidad se diluye entre los distintos actores como programadores, fabricantes, operadores, compradores, propietarios o usuarios.

En particular, en relación con el posicionamiento de este autor frente a los robots, entiendo que dotados de inteligencia artificial, Santos González considera que “las

[715] SANTOS GONZÁLEZ, M.J. (2017). “Regulación legal de la robótica y la inteligencia artificial: Retos del futuro”. *Revista Jurídica de la Universidad de León*. Nº 4. 2017. Op.cit. P. 38.

normas tradicionales de responsabilidad civil no son suficientes para generar responsabilidad jurídica por los daños ocasionados por el robot, ya que no permiten determinar la parte que ha de hacerse cargo de la indemnización, ni exigir a dicha parte que repare el daño ocasionado".

El Parlamento Europeo ya distinguía dos regímenes de responsabilidad distintos en la Resolución del Parlamento Europeo de 16 de febrero de 2017 sobre normas de Derecho civil sobre robótica.

De un lado, un sistema de responsabilidad objetiva, que únicamente exigiría probar que se ha producido un daño o perjuicio y el establecimiento de un nexo causal entre el funcionamiento perjudicial del sistema inteligente y los daños y perjuicios causados a la persona afectada. Y de otro, un sistema de responsabilidad sustentada en el riesgo que se focalizaría en la persona que tiene el control del riesgo potencial y es capaz, en determinadas circunstancias, de minimizarlo y gestionar el impacto negativo.

Asimismo, el Parlamento Europeo propuso una solución híbrida compuesta por instrumentos legislativos y no legislativos, esto es, de *hard* y *soft law*, sin que se pudiera limitar el tipo y alcance de daños y perjuicios objeto de compensación, ni limitar la naturaleza de la compensación, que la Directiva 85/374/CEE si permitía a los Estados, en particular, limitar la responsabilidad por daños personales[716].

Algunos autores[717] se han venido planteando incluso la posibilidad de establecer un sistema de indemnización, compensación o de garantía de indemnizaciones prescindiendo de la responsabilidad y centrado en la reparación del daño causado sin tener que analizar a quién sea imputable, como en el que se inspira la Ley nº 85/677, de 5 de julio de 1985 en Francia, también conocida como *Ley Badinter*, sobre la indemnización de víctimas de accidentes de circulación y que configura un sistema de seguridad indiscriminada a favor de las víctimas de circulación, al margen de la culpa o cualquier otra consideración.

[716] BADILLO ARIAS, J.A. (2019). "Responsabilidad civil y aseguramiento obligatorio de los robots". En MONTERROSSO CASADO, E. (Dir.). *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch 2019. P. 49.
[717] BARRIO ANDRÉS, M. (2018). *Derecho de los Robots*. Wolters Kluwer. Madrid 2018. P. 124.

Un sistema de este tipo permitiría que, una vez producido el daño, éste deba ser reparado, bien mediante la existencia de un seguro obligatorio concertado por el fabricante, propietario, operador o usuario o, en caso de no ser cubierto por el mismo, mediante un fondo de compensación, similar a la figura actual del Fondo de Compensación de Seguros en España.

La dotación de este fondo podría venir de las contribuciones obligatorias de fabricante, propietarios u operadores, exacciones fiscales indirectas a añadir al precio del producto o servicio u otras aportaciones.

La precitada Resolución del Parlamento Europeo de 16 de febrero de 2017 sobre normas de Derecho civil sobre robótica, que será objeto de análisis más detallado en el siguiente apartado, realiza un enfoque híbrido, considerando, de un lado, la necesidad de mantener un sistema de responsabilidad que en la actualidad debe recaer en personas y no en un sistema dotado de inteligencia artificial, según recoge en el apartado 56 *in fine* de la misma y, de otro, el establecimiento de un conjunto de medidas complementarias para garantizar el resarcimiento efectivo del perjudicado como los seguros y fondos de compensación.

Conforme expondré con mayor detalle a continuación -ya lo fue en relación con los aspectos éticos y de seguridad-, el Parlamento Europeo evidenció su preocupación por la multiplicidad de actores que podían intervenir en la producción de un robot o sistema dotado de inteligencia artificial, así como en su posterior entrenamiento y aprendizaje, así como la variedad de usos y aplicaciones que el propietario o usuario pueda darle, coincidentes o no con las inicialmente pensadas por su productor o fabricante.

En cualquier caso, el contexto que plantean los sistemas dotados de inteligencia artificial, conforme ya reflejaba esta Resolución, supone una enorme complejidad para la determinación de la persona o personas responsable/s en todo el ciclo de vida de un sistema, ya sea por defecto en diseño o fabricación, control directo o indirecto del mismo o, en base a las distintas modalidades de responsabilidades comentadas, esto es, por responsabilidad objetiva o por riesgo.

Adicionalmente, comporta la dificultad de distribuir la responsabilidad entre los mismos -fabricante, diseñador, desarrollador, propietario, formador, entrenador, operador, usuario- o simplemente exigirla solidariamente a cualquiera de ellos, sin perjuicio de su repetición a los demás.

En este sentido considero que la perspectiva más proteccionista para los perjudicados sería considerar una responsabilidad solidaria, como expuse anteriormente, sin perjuicio de repetirse entre los responsables conforme a su grado de responsabilidad, control o distribución del riesgo. En esta línea se han posicionado algunos autores como Ebers[718] y es por la que se decanta la Propuesta de Reglamento del Parlamento Europeo sobre responsabilidad civil en materia de inteligencia artificial de 20 de octubre de 2020[719], que analizaré con detenimiento en el próximo apartado. Además, disponemos ya de marcos jurídicos de referencia en el entorno europeo que regulan y exigen esta responsabilidad en relación con su objeto, como el artículo 82 del RGPD.

Asimismo, el Parlamento Europeo consideraba en su precitada Resolución de 2017 que la responsabilidad debería ser también proporcional al nivel de instrucciones impartidas al sistema inteligente y a su grado de autonomía, de forma que cuanto mayor sea la capacidad de aprendizaje o autonomía o cuanto más larga haya sido la formación, mayor debiera ser la responsabilidad de su formador.

Del mismo modo, el Parlamento Europeo también destacaba que, para determinar y atribuir la responsabilidad de los daños o perjuicios causados por un robot o sistema dotado de inteligencia artificial, deberían diferenciarse las competencias adquiridas a través de su formación de las dependientes estrictamente de su capacidad de aprender de forma autónoma. Y en adición al sistema de responsabilidad por daños que se defina, ya sea objetiva o basada en riesgos, se debería completar, conforme el Parlamento Europeo adicionaba, por un seguro obligatorio, así como otras, como los fondos de compensación.

[718] EBERS, M. (2016). "La utilización de agentes electrónicos inteligentes en el tráfico jurídico: ¿Necesitamos reglas especiales en el Derecho de la responsabilidad civil?". En *InDret: Revista para el Análisis del Derecho*, nº 3, 2016, P. 16.
[719] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial. Op.cit.

El problema, de nuevo, lo plantearán los sistemas dotados de inteligencia artificial más avanzados, con cierto grado de autonomía, independencia y capacidad de autoaprendizaje e interacción con su entorno, con decisiones y acciones impredecibles, al menos relativamente. ¿Quién tendrá el control total o parcial sobre los mismos? ¿Quién ejercerá esa supervisión humana? ¿Cómo valorar el riesgo y posibles responsabilidades por parte de fabricantes o aseguradoras?

No obstante, la imposibilidad de acotar y valorar previamente el riesgo o su desconocimiento del que podrían responder podría comportar, de un lado, que los fabricantes vean desincentivada su actividad innovadora y desarrolladora y que pueda desistir de fabricar determinados tipos de robots o sistemas dotados de inteligencia artificial o de seguir invirtiendo en I+D+I -repercutiendo negativamente en la innovación y la capacidad competitiva de las empresas europeas- y, de otro, que las compañías aseguradoras se nieguen a concertar este tipo de seguros o a establecer primas muy elevadas al objeto de proteger sus intereses empresariales.

En este sentido, algunos autores, como el precitado Gómez-Riesco, considerando el posicionamiento del Parlamento Europeo en la Resolución precitada y alineándose con el mismo, consideran preciso el establecimiento de algún límite cuantitativo a las indemnizaciones para permitir la evaluación previa del riesgo, tanto por fabricantes como por aseguradoras, que de un lado tendría seguridad para dar cobertura a los mismos así como para su posterior reclamación al responsable último de los daños mediante acción subrogatoria.

n este sentido, lo considero razonable, si bien, no deberían operar estas limitaciones en todos los contextos, debiendo cubrirse en su integridad por todos los responsables, especialmente en casos donde concurre una culpa evidente.

Adicionalmente, el Parlamento Europeo propuso en su Resolución de 2017 como complemento del seguro obligatorio, la constitución de un fondo de compensación que garantice la reparación de los daños o perjuicios causados por un robot o sistema dotado de inteligencia artificial avanzada ante la ausencia de un seguro o no cobertura por éste, del que además podría beneficiarse el fabricante, programador, operador, propietario o usuario, ante la posibilidad de ver limitada su responsabilidad si contribuyen al mismo.

Realmente el Parlamento parecía hacer referencia a dos tipos de fondos de compensación y complementarios, como destaca Badillo Arias[720].

Uno similar al existente en automóviles que se haría cargo de cubrir los daños producidos por los robots inteligentes que no tienen seguro de responsabilidad civil a pesar de estar obligados a ello. Se trataría de un patrimonio o fondo que aseguraría el derecho de resarcimiento efectivo de la víctima en el caso de que no sea posible hacerlo mediante la instauración de un sistema de responsabilidad objetiva y un seguro asociado.

Y otro, un fondo de compensación distinto que compense los daños ocasionados por los robots que permita limitar la responsabilidad civil de los agentes intervinientes como el fabricante, programador, propietario o usuario en el caso de que contribuyan a su dotación (o si suscriben un seguro conjunto que garantice el resarcimiento de la víctima) y, en consecuencia, de las compañías aseguradoras.

Considero que el juego de todos estos mecanismos adicionales y complementarios a un sistema de responsabilidad objetiva o por riesgos, asegurarían un resarcimiento íntegro y efectivo de la víctima, y deben ser considerados en los futuros marcos reguladores que deben revisar e integrarse con los marcos preexistentes de responsabilidad, con las adaptaciones que sean necesarias.

Estas garantías reparadoras adicionales constituyen la denominada "socialización del riesgo", susceptible de crítica desde distintas perspectivas, pero que sin duda puede contribuir a proporcionar seguridad jurídica a todas las partes interesadas y evitar obstáculos a la innovación y el avance tecnológico en beneficio de la sociedad.

En este sentido, el Parlamento Europeo solicitaba en dicha Resolución de 2017 la presentación de una Directiva relativa a las normas de legislación de civil en materia de robótica, siguiendo las recomendaciones recogidas en su anexo. La elección de este instrumento no parece la más adecuada por la demora en su eficacia ante el período de transposición y desde un punto de vista proteccionista del ciudadano, como así también

[720] BADILLO ARIAS, J.A. (2019). "Responsabilidad civil y aseguramiento obligatorio de los robots". En MONTERROSSO CASADO, E. (Dir.). *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch 2019. P. 59.

lo consideran autores como Díaz Alabart[721], dejando a un lado las cuestiones de armonización si no se regula con claridad y amplitud su objeto, de modo que no pueda provocar transposiciones asíncronas y de diferente alcance y contenido por los Estados miembros.

En relación con todo lo expuesto anteriormente y anticipándome ya a su análisis, el Parlamento Europeo en su Resolución de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial[722], incluye una Propuesta de Reglamento relativa a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial, en la que aborda distintas cuestiones tratadas en aquella Resolución y en la que distingue dos regímenes de responsabilidad, uno objetivo y otro subjetivo, conforme expondré más adelante. Esta propuesta se construyó sobre el principio de precaución y enfoque de riesgos precitado.

En resumen, a la vista del marco jurídico vigente, la legislación sobre productos defectuosos, en particular, a nivel europeo la Directiva 85/374/CEE, y a nivel español la TRLGDCU, que transpone la anterior y amplia su alcance, serán de aplicación en los supuestos específicos contemplados por la misma de los que se deriva la responsabilidad del fabricante/productor, en los que podrían encajar dispositivos dotados con sistemas de inteligencia artificial y a los propios sistemas con mayor o menor autonomía que puedan encajar en la definición de "producto" y, en el caso de la norma española, también sería susceptibles de aplicación a determinados servicios.

La persona afectada se encontraría ante un sistema de responsabilidad más que objetiva, *cuasi objetiva*, como he referido y destacan algunos autores como Zurita Martín[723], en la medida que, junto a la prueba del daño efectivo y la relación de causalidad entre éste y la acción u omisión causante, se exige a la persona afectada la acreditación del defecto del

[721] DÍAZ ALABART, S. (2018). *Robots y responsabilidad civil*. Editorial Reus, Madrid 2018. P. 59.
[722] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))
[723] ZURITA MARTÍN, I. (2021). "Gestión de riesgos y responsabilidad civil de los robots", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón.* Tomo III. Ed. Aranzadi-Thomson Reuters. 2021. P. 2013.

producto, calificable como tal conforme a los marcos precitados y no en cualquier supuesto.

Del mismo modo, en éstos como en el resto de supuestos, se debe valorar también la aplicación de las normas generales de responsabilidad civil contractual o extracontractual para depurar las responsabilidades por los daños causados por parte del agente causante de los mismos, entre los que podría hallarse desde el fabricante, el proveedor, el diseñador, el desarrollador, el propietario, el formador o entrenador, el operador o el usuario.

En relación con los sujetos responsables, siguiendo las propuestas del *Grupo de Expertos en Responsabilidad Civil y Nuevas Tecnologías de la Comisión Europea* en su informe precitado de 21 de noviembre de 2019 -*Liability for Artificial Intelligence and other emerging digital technologies*-, se diferenciarían dos categorías focalizadas en el fabricante o productor y el operador, pero nunca asociándolo a un único individuo en el caso de sistemas inteligentes, distinguiendo los que operan con el sistema y no lo utilizan (*Back-end*), que pueden actualizar el *software*, introducir mejoras, revisar y/o supervisar, y los que operan con el sistema y lo utilizan (*Front-end*), que recaería en el propietario, el usuario o el poseedor del sistema inteligente.

La cuestión es que los futuros marcos reguladores de la responsabilidad no podrán predeterminar el sujeto responsable y la causalidad en todas las situaciones, sino que será *ad hoc*, en atención al contexto y en base a un marco que defina adecuadamente la responsabilidad por los daños derivados del funcionamiento y uso de los sistemas inteligentes.

7.2. Personalidad jurídica de los sistemas inteligentes avanzados

El Parlamento Europeo planteó valorar, como posibilidad a explorar en su precitada Resolución del Parlamento Europeo de 16 de febrero de 2017 sobre normas de Derecho civil sobre robótica, el posible otorgamiento futuro de personalidad jurídica propia y específica a los sistemas de inteligencia artificial avanzada o fuerte, sin perjuicio de otras

medidas, como la creación de un registro de este tipo de sistemas y robots dotados con los mismos, constitución de un fondo, seguro obligatorio y establecimiento de límites de indemnización en caso de daños materiales.

Sobre la posibilidad de valorar en el futuro el otorgamiento de una personalidad jurídica a los robots o sistemas dotados de inteligencia artificial, ya me he pronunciado sobre ello en los apartados precedentes de esta investigación y que abordaré con más detenimiento en el siguiente capítulo en relación con los robots, posibilidad que considero inviable a la vista del contexto tecnológico actual y marco jurídico vigente, y más todavía ante el previsible marco ético y jurídico futuro en la UE, siendo una opción que ha sido desechada por el momento por el legislador europeo.

Si en el futuro se optara por atribuir una responsabilidad jurídica a los sistemas de inteligencia artificial más avanzados o "fuertes", con cierto grado de autonomía en sus decisiones y conductas, quedaría por resolver cómo atribuirle la responsabilidad a un agente sin consciencia, sin emociones, sin remordimientos, sin capacidad para distinguir entre lo justo e injusto, entre lo bueno y lo malo y que, además, aun otorgándole plena autonomía, la tendría sin libertad plena, en la medida que debería actuar conforme a instrucciones, objetivos, funciones, parámetros, controles y restricciones predefinidas previamente -con independencia de que pueda incorporar instrucciones generadas por el mismo en el marco de su capacidad de autoapredizaje y autoprogramación-, y sujeto a un poder y control humano superior y a la amenaza de desconexión -tampoco percibida como tal, sino como mera consecuencia asociada un hecho previo definido-, cuando su actuación se aleje de lo esperado. Y ello mediante controles de suspensión o desactivación, salvo que en su diseño se incumplieran los principios y normas éticas esenciales que conforman los distintos marcos con mayor consenso a nivel internacional y, en caso de aprobación, de los futuros marcos jurídicos reguladores de los mismos en el ámbito de la UE.

Si sólo el ser que es libre puede actuar con voluntad real y regir sus decisiones y acciones por su consciencia o voluntad, la libertad se basa en la voluntad, por lo que autores como

Núñez Zorilla[724] o Encabo Vera[725] entienden que no se puede hablar de libertad en los actos no voluntarios -los que procederían de un sistema de inteligencia artificial más avanzada o fuerte- condicionados por el conocimiento por parte de la inteligencia artificial de la posibilidad de ser desconectada o inhabilitada-.

En este sentido, no comparto totalmente estas afirmaciones, porque si bien coincido en la ausencia de autonomía y libertad plena por los motivos a los que he aludido anteriormente, es decir, la actuación conforme a instrucciones, objetivos, funciones, parámetros, controles y restricciones predefinidas previamente y sujetos a control y supervisión humana, lo cierto es que las personas pueden actuar libremente en el mundo físico y virtual sin perjuicio de que sean conscientes o no de que su actuación pueda ser ilegítima, ilícita o incluso delictiva, con las consecuencias que ello pudiese conllevar conforme a normas, no solo de suspensión o privación temporal o definitiva de derechos y libertades sino incluso, en algunos países, la extinción de la misma a través de la pena de muerte.

En el caso de sistemas inteligentes, estas consecuencias serían la predefinidas en el código fuente.

Y otra cosa distinta a todo ello, es la inexistencia de consciencia en sus actos e inimputabilidad, que imposibilita hacerla responsables de sus actos.

En consecuencia, no podría atribuirse la culpa, ni responsabilidad, ni personalidad a este tipo de sistemas, cuanto menos por el momento, conforme he abordado anteriormente.

No obstante, como también he expuesto, la posibilidad de crear en el futuro sistemas de inteligencia artificial no sólo con plena autonomía -a mi juicio imposible de considerar a la vista de los marcos éticos y legales que deberán regular esta realidad y posibilidad- y supuesta consciencia, podría suponer un nuevo contexto que podría exigir la reevaluación de las bases y criterios para la atribución de la personalidad jurídica y la responsabilidad.

[724] NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Op. Cit. P. 36.
[725] ENCABO VERA, M.A. (2012). *Derechos de la personalidad*. Marcial Pons. Madrid 2012. P. 75.

7.3. Responsabilidad de entes sin personalidad jurídica

¿Sería una incongruencia el propio tenor literal del título de este apartado?

En mi opinión, considero necesario reflexionar sobre otra opción distinta a la posibilidad de dotar a los sistemas inteligentes más avanzados de personalidad jurídica, como solución a las cuestiones de responsabilidad por daños.

En concreto, me planteo la posibilidad de atribuir "responsabilidad" a los robots y sistemas dotados de inteligencia artificial avanzada o fuerte, sin autonomía real plena sino restringida y sin consciencia, si bien como "agentes", pero bajo la categoría de *entes sin personalidad jurídica*, pero con obligaciones y responsabilidades específicas atribuidas por el ordenamiento jurídico.

Es decir, considerarlos *entes sin personalidad jurídica* que no responderían conforme a los marcos de responsabilidad actuales, pero sobre los que podrían recaer determinadas obligaciones legales en materia de responsabilidad, conforme a los futuros nuevos marcos reguladores de la misma, incluyendo la indemnización con cargo a la persona, entidad o al fondo y/o seguro que tengan asociado a su activación y puesta en funcionamiento, con sujeción, en su caso, a los límites cuantitativos que se establezcan respecto de daños materiales. Detrás de los mismos, estarían dichos medios para asegurar el resarcimiento efectivo a proveer y dotar por las personas físicas o jurídicas que deban de responder de su puesta en circulación y uso.

Del mismo modo, este instrumento podría facilitar la distribución de la responsabilidad sobre todos los sujetos que pudieren ser también corresponsables junto al fabricante.

Esta figura no es nueva, ha sido ya utilizada por nuestro ordenamiento jurídico.

En particular y a modo de ilustrativo, en materia de protección de datos personales, el Reglamento de desarrollo de la derogada Ley Orgánica 15/1999, de 13 de diciembre, de

protección de datos de carácter personal, aprobado mediante Real Decreto 1720/2007, de 21 de diciembre, recogía en sus artículos 5.1.i) y 5.1.q) la posibilidad de que los *entes sin personalidad jurídica* que actúen en el tráfico como sujetos diferenciados pudieran ser considerados, respectivamente, responsables y encargados del tratamiento a los efectos de esta marco regulativo y a sus obligaciones.

Asimismo, a nivel procesal, la Ley de Enjuiciamiento Civil española, prevé expresamente en su artículo 51.2 que "Los entes sin personalidad podrán ser demandados en el domicilio de sus gestores o en cualquier lugar en que desarrollen su actividad". En su artículo 6.1.5º establece que podrán ser parte en los procesos ante los tribunales civiles "las entidades sin personalidad jurídica a las que la ley reconozca capacidad para ser parte". Y, por último, respecto de la comparecencia en juicio y representación, en su artículo 7.6. establece que las entidades sin personalidad a que se refiere el artículo anterior comparecerán en juicio por medio de las personas a quienes la ley, en cada caso, atribuya la representación en juicio de dichas entidades.

En otros ámbitos como el fiscal o el administrativo, las entidades sin personalidad jurídica también tienen determinadas obligaciones, desde la necesaria relación electrónica con la Administración hasta las tributarias, por ejemplo, la obtención de un número de identificación fiscal y su registro.

En consecuencia y para superar las controversias sobre el posible reconocimiento futuro de personalidad jurídica a los robots y sistemas dotados de inteligencia artificial y, consecuente capacidad jurídica y posible capacidad de obrar, su reconocimiento como ente sin personalidad jurídica podría ser una de las posibles soluciones, con la creación de un estatus jurídico propio, la regulación detallada de sus obligaciones y responsabilidades derivadas de su funcionamiento, decisiones, acciones y omisiones, y su registro obligatorio, sin perjuicio de la limitación cuantitativa de las responsabilidades y el establecimiento de seguros obligatorios y fondos de compensación para garantizar el pleno resarcimiento de los daños causados por los mismos.

7.4. Revisión del régimen especial de responsabilidad por productos defectuosos.

En congruencia con lo expuesto, si bien la normativa de transposición española contempla expresamente su aplicación a productos y a servicios defectuosos, a diferencia de la Directiva 85/374/CEE que los circunscribe a los primeros, considero que el marco analizado de responsabilidad del productor de sistemas dotados de inteligencia artificial -especialmente si es avanzada o fuerte-, debe revisarse íntegramente ante sus carencias evidentes para poder resultar de aplicación efectiva a los diversos supuestos de responsabilidad causada por el funcionamiento de sistemas de inteligencia artificial y garantizar el derecho de resarcimiento de las personas y entidades afectadas.

El régimen de responsabilidad por producto defectuoso, como he expuesto, se haya tradicionalmente orientado a la protección del consumidor, tanto a nivel europeo como español, de hecho, la Directiva europea sobre productos defectuosos fue finalmente transpuesta en España a través de una norma elaborada para la protección de los consumidores.

Los marcos específicos de responsabilidad deben considerar y tener como finalidad no sólo la protección del consumidor o usuario sino también los intereses empresariales, profesionales o del sector público como damnificados por el funcionamiento de los sistemas inteligentes, y contemplar no sólo productos sino servicios.

La revisión de este sistema de responsabilidad debería suplir las carencias analizadas en mis reflexiones y, entre otros aspectos, incluir los daños morales de distinta naturaleza, especialmente ante sistemas de inteligencia artificial avanzada o fuerte utilizados en el cuidado y asistencia de personas a los que me he referido en su análisis.

Conforme a este sistema de responsabilidad analizado, salvo en los supuestos en los que sí sería plenamente aplicable este régimen, el productor podría quedar exonerado de los daños que no pudieran haber sido anticipados y que tendrían su origen en la autonomía y capacidad de decisión y actuación conferida al sistema de inteligencia artificial, el cual, como he expuesto anteriormente, tampoco sería imputable, entre otras razones, al carecer de personalidad jurídica.

La posible modificación de este sistema en el futuro y su plena objetivización, considero que requeriría valorar la plena atribución de la responsabilidad al productor o fabricante y valorar la eliminación o condicionamiento como causa eximente del desconocimiento por éste de los posibles daños futuros, aunque con matices, por considerar que debió identificar, gestionar e informar adecuadamente de los riesgos, incluso los derivados de usos indebidos, especialmente los razonablemente previsibles o asociados a las capacidades y características de las que dotó al sistema.

Sin embargo, este posicionamiento ha generado rechazo en parte de la doctrina, en la medida que podría constituir un obstáculo para la innovación, desarrollo y aplicación de la inteligencia artificial, partiendo de que la gestión del riesgo y su responsabilidad consecuente debería atribuirse a quién en cada momento se halle en la mejor posición para gestionarlo y, también, siguiendo a autores como Solé i Feliu[726], para soportarlo en términos de racionalidad económica.

Es incuestionable la complejidad para productores y fabricantes afrontar, de un lado, la inversión y anticipación mediante controles y medidas de seguridad efectivas y, de otro, los costes posteriores de los procedimientos e indemnizaciones, si bien, es la parte que se lucra con ello y posiblemente, entre las cualidades de sus productos y factores que determinan la adquisición de sus sistemas con mayor riesgo, posiblemente sea el grado de autonomía, capacidad de autoaprendizaje y de resolución de problemas exclusivamente por el sistema sin intervención humana, de manera eficaz y eficiente, con el consiguiente ahorro de costes, por lo que podría ser razonable exigir medidas y controles coherentes en base a dicho argumento para evitar decisiones o comportamientos lesivos, seguramente de muy difícil predicción en su materialización concreta, pero si genérica y razonablemente previsibles sin la adopción de control o medida alguna que los restrinja en su ciclo de vida.

Asimismo, otro de los argumentos en contra de la pretensión de imputar una responsabilidad objetiva directa y exclusivamente al productor o fabricante de esta naturaleza es la posibilidad de que pueda tener un efecto desincentivador del desarrollo e

[726] SOLÉ I FELIU, J. (1997). *El concepto de defecto del producto en la responsabilidad civil del fabricante*. Editorial Tirant lo Blanch. Biblioteca Jurídica Cuatrecasas. Valencia, 1997. Pp 504-505.

innovación en inteligencia artificial, que obviamente interesa y beneficia a los usuarios y a la sociedad en general, y el consiguiente, aumento de los costes de aplicación de estas tecnologías, de indudable utilidad social.

De hecho, algunos autores como Solé i Feliu[727] y Núñez Zorrilla[728] consideran que la atribución al productor o fabricante de toda la responsabilidad por los daños derivados de riesgos desconocidos equivaldría a gravarle con una responsabilidad no simplemente objetiva sino absoluta.

Y además de todo ello, el desconocimiento de los riesgos, su indefinición, imposibilidad de estimación plena y su posible gravedad, podría comportar la dificultad de su aseguramiento por terceros y, en caso de su aseguramiento, la aplicación de primas por cuantías muy elevadas que comportaría la elevación del precio de la tecnología y, en consecuencia, afectaría al derecho de acceso a la misma, considerando el mismo un derecho cuasi fundamental, como instrumento necesario para el ejercicio, en muchos casos, imprescindible, de derechos, entre otros, los fundamentales.

No obstante, a mi juicio, la insuficiencia del sistema de responsabilidad por productos defectuosos considero que podría abordarse igualmente mediante la creación de un sistema de responsabilidad objetiva, ajena a la culpa o negligencia del productor o fabricante, y que vaya más allá, esto es, un sistema de responsabilidad objetiva derivado del hecho de generar un riesgo o peligro para la sociedad, de modo que sea irrelevante aspectos como que el fabricante haya actuado con diligencia en su fabricación, que haya puesto todas las medidas a su alcance para evitar el daño, que conforme al estado de la técnica no hubiera podido conocer el defecto en su construcción o que tenga su origen en la autonomía, capacidad de autoaprendizaje e independencia conferida al sistema construido. Y ello, sin perjuicio de los casos de corresponsabilidad o inexistencia de la misma por tener su origen en el usuario u otros sujetos, especialmente ante usos específicos o su propio entrenamiento por el mismo.

[727] SOLÉ I FELIU, J. (1997). *El concepto de defecto del producto en la responsabilidad civil del fabricante*. Op.cit. Pp. 506-508.
[728] NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Op. Cit. P. 33.

Todo ello debería ir acompañado de límites en las indemnizaciones respecto de determinados daños que permita asimismo su aseguramiento, con identificación del riesgo real y potencial para todas las partes implicadas, incluyendo aseguradoras, que permita primas razonables y asumibles.

Adicionalmente, debería regularse la responsabilidad del propietario/operador/usuario cuando sea exigible al mismo y no al fabricante o productor, por ejemplo, en el caso de vehículos "autónomos", de modo que también se acotase la misma respecto de determinados daños a nivel cuantitativo posibilitando la asunción del riesgo por traslación por las aseguradoras bajo primeras razonables.

Del mismo modo, se hace igualmente necesario actualizar a nivel de la UE, la Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos.

Esta necesidad no es nueva y no se ha evidenciado exclusivamente en relación con los dispositivos y sistemas dotados de inteligencia artificial.

Ha sido puesta de manifiesto reiteradamente durante estos últimos años por la propia Comisión Europea, en especial, a través del *Informe de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo* sobre la aplicación de esta Directiva, de fecha 07.05.2018[729].

La Comisión Europea puso de relieve en este informe los nuevos retos a los que nos enfrentamos en la actualidad y que serán más pronunciados en el futuro, especialmente los relacionados con la digitalización, el internet de las cosas, la inteligencia artificial y la ciberseguridad y, desde mi óptica, todavía mayores ante la interacción de todo ello entre sí y el tratamiento de datos masivos.

La inteligencia artificial es considerada por la Comisión Europea como una de las tecnologías más importantes del siglo XXI, destacando en este sentido su Comunicación

[729] COM/2018/246 final. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52018DC0246

anteriormente citada bajo el título *Maximising the benefits of Artificial Intelligence*[730] (Maximizar los beneficios de la inteligencia artificial), en la que, entre otros aspectos, aborda las dificultades legales asociadas a la misma, destacando que para beneficiarse de la misma por parte de la sociedad en general, empresas y ciudadanos, es necesario garantizar la seguridad y la responsabilidad de los productos en caso de producirse daños.

La Comisión Europea ya sometió a una evaluación[731] la Directiva, en particular en 2018, con el propósito de determinar su efectividad, preguntándose cuestiones como, por ejemplo, si aborda de forma adecuada las dificultades que plantean los dispositivos cada vez más autónomos y la ciberseguridad.

La evaluación realizada demostró, a juicio de la Comisión que, pese a que los productos son mucho más complejos en la actualidad que en 1985, la Directiva relativa a la responsabilidad por productos defectuosos sigue siendo una herramienta adecuada. Es obvio a mi juicio, a la vista del análisis realizado, que no lo es para depurar la responsabilidad por daños causados por un sistema de inteligencia artificial ante sus múltiples tipologías y en los múltiples contextos que se puedan plantear, por lo que requiere su revisión.

No obstante, debo destacar que la evaluación externa llevada a cabo y de la que se derivó el informe precitado, tuvo un objeto y alcance limitados por lo que se refiere a las nuevas tecnologías, y la propia Comisión puso de relieve que la falta de información sobre asuntos judiciales concretos, las reclamaciones de los consumidores o la correspondiente experiencia práctica de las partes interesadas hacía imposible alcanzar una conclusión definitiva.

De hecho, el estudio realizado únicamente pudo identificar una causa judicial en la que el tema tratado tenía que ver concretamente con las tecnologías digitales emergentes, en

[730] Comunicación de la Comisión, *Maximising the Benefits of Artificial Intelligence for Europe* («Maximizar los beneficios de la inteligencia artificial para Europa», documento en inglés), COM(2018)237).
[731] Documento de trabajo SWD (2018)157 de los servicios de la Comisión sobre la evaluación de la Directiva adjunto.

particular, con una unidad de almacenamiento de datos en Bulgaria. (Asunto de Bulgaria n.º 20942/2012).

Todo ello evidencia y refuerza la necesidad de reflexionar sobre la revisión de dicha Directiva.

Como he manifiestado en diversos apartados de esta investigación, el Derecho debe ser proactivo y no meramente reactivo, con la finalidad de anticiparse a los desafíos y problemas a los que deberá prever y dar una respuesta efectiva y adecuada.

Los desafíos y problemas relacionados con sistemas de inteligencia artificial se están produciendo ya y su crecimiento será exponencial como consecuencia del despliegue y aplicación masiva de la inteligencia artificial en todos los ámbitos y sectores.

Es por ello, que se precisará de un nuevo marco regulador de la misma y una revisión de los marcos de responsabilidad existentes para adecuarlos a las nuevas realidades (algunas no tan nuevas como he expuesto, pero todavía no materializadas en gran medida pero de aplicación potencial), para de este modo disponer de mecanismos ágiles y efectivos para la protección de los derechos e intereses de todas las partes implicadas y, en especial, en materia de responsabilidad, para garantizar un íntegro y efectivo resarcimiento de las personas que sufran daños derivados del funcionamiento de sistemas inteligentes.

En este sentido, el propio informe precitado concluye que, considerando las características de estas tecnologías, especialmente su complejidad y autonomía, la Comisión Europea deberá dar seguimiento a todas las preguntas todavía sin respuesta, y algunas de estas características pueden cuestionar si el actual marco en materia de responsabilidad por productos es adecuado para garantizar una compensación eficaz a los consumidores y estabilidad de inversión a las empresas. La respuesta deberíamos encontrarla en la Resolución del Parlamento Europeo de 20 de octubre de 2020, sobre un régimen de responsabilidad civil en materia de inteligencia artificial[732].

[732] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))

El informe precitado incide en el concepto de responsabilidad objetiva de los productores, de modo que si un producto es defectuoso y provoca lesiones personales o daños materiales por importe superior a 500€ en un bien destinado principalmente al uso o consumo privados, los productores serán responsables independientemente de si han cometido o no una falta, considerando que un producto es defectuoso si no ofrece la seguridad a la que tiene derecho una persona.

Asimismo, el informe destaca también que esta seguridad debe tener en cuenta todas las circunstancias, incluida la presentación del producto, el uso que razonablemente pudiera esperarse del mismo y el momento en que el producto se puso en circulación, incorporando el siguiente ejemplo, que me parece muy oportuno para asociarlo a los supuestos que pueden producirse en relación con los vehículos "autónomos":

"Mientras conducía su vehículo, un conductor tuvo que esquivar un obstáculo que apareció inesperadamente. Al salirse de la carretera su vehículo comenzó a vibrar fuertemente. Los sensores del airbag consideraron que se trataba de un accidente y se activaron. Uno de los *airbags* laterales golpeó al conductor en el cuello, comprimiéndole una arteria, con lo que le provocó un infarto. Los tribunales trataron de determinar si el productor había calculado con exactitud el riesgo de mal funcionamiento de los sensores. La reclamación fue rechazada en dos instancias, y posteriormente invalidada en una instancia superior. Finalmente, el caso se resolvió extrajudicialmente."

Si se tratara de un vehículo autónomo que pudiera llegar a activar los *airbags* en otros supuestos distintos a los inicialmente parametrizados en su diseño y causara idénticas lesiones en virtud de su autonomía ¿deberíamos considerar que el riesgo de actuación imprevista y consecuente mal funcionamiento no debería haber sido igualmente considerado y limitado mediante controles y medidas por el productor? ¿o no le resultaría imputable responsabilidad alguna al mismo por esta vía?

En mi opinión, las reflexiones alrededor de este supuesto evidencian la insuficiencia y carencias del sistema de responsabilidad por productos defectuosos en el marco de los sistemas inteligentes.

En relación con la evaluación de la Directiva llevado a cabo y, en particular, sobre la posibilidad de exoneración de responsabilidad del productor, destacar que Francia, España, Finlandia, Luxemburgo y Hungría adoptaron en su transposición la denominada "excepción por riesgo de desarrollo" a la que he hecho referencia en mi análisis y que se recoge en el artículo 15, apartado 1, letra b) de la Directiva, por la que un productor puede resultar exonerado de responsabilidad si el estado de los conocimientos científicos y técnicos en el momento en que el producto se puso en circulación no permitía descubrir el defecto.

Sin embargo, sólo Finlandia y Luxemburgo lo aplican a todos los sectores, mientras que Hungría excluye los productos farmacéuticos, España excluye los medicamentos, alimentos o productos alimentarios y Francia excluye los productos procedentes del cuerpo humano.

A mi juicio, los futuros marcos de responsabilidad que revisen y complementen los vigentes en materia de responsabilidad por productos defectuosos, entre otras cuestiones, deberían exigir una reflexión previa sobre la ineficacia de cualquier causa de exclusión de la responsabilidad en el caso sistemas de inteligencia artificial dotados de autonomía y capacidad de autoaprendizaje sin controles, restricciones y/o medidas compensatorias para la toma de decisiones o realización de acciones al margen de las instrucciones, finalidades y restricciones previamente definidas en su diseño, así como en caso de operar sin control y supervisión humana en su diseño y durante todo su ciclo de vida.

La no aplicación a los sistemas de inteligencia artificial nos llevaría a la responsabilidad absoluta por riesgos precitada anteriormente.

Por último, el informe citado evidencia igualmente que la Comisión Europa es consciente de que los problemas y retos a los que nos enfrentamos hoy, inmersos en una nueva revolución tecnológica, difieren notablemente de los que existían en el mundo predominantemente analógico de 1.985, especialmente ante la interconexión, digitalización, autonomía e inteligencia de los sistemas, por lo que destaca que la

necesidad de una respuesta coherente y global ante los nuevos retos, ya recogida en la iniciativa sobre inteligencia artificial anteriormente citada[733].

El propio informe de la Comisión refleja como su eficacia se ve obstaculizada por conceptos como "producto", "productor", "defecto", "daño" o la "carga de la prueba", significando igualmente la complejidad de esta última cuando se ven involucradas tecnologías digitales emergentes.

Del mismo modo reitera la necesidad de que, de un lado, en última instancia, el productor sea responsable del producto que pone en circulación, de otro, que los perjudicados puedan demostrar que ha habido un daño causado por un defecto y, por último, que ambas partes sepan qué esperar de los productos, en lo que a seguridad se refiere, mediante la definición previa de un marco de seguridad claro.

Esto último no se haya definido claramente en la actualidad, como he abordado en distintos apartados de este informe, lo que exigiría la definición de un marco global, ético y jurídico que exija seguridad, y un marco específico de ésta para los sistemas inteligentes, en lo que ya está trabajando ENISA y la UE como he expuesto en el capítulo II de esta investigación.

En este sentido, la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial *-Artificial Intelligence Act-*, de 21 de abril de 2021[734], exige determinados requerimientos de ciberseguridad y robustez, entre otros, pero exclusivamente para los sistemas de inteligencia considerados de alto riesgo conforme al mismo y no respecto del resto. La seguridad requerida sería la seguridad adecuada a los riesgos y contexto en este tipo de sistemas.

Con todos estos objetivos formalizados en el informe precitado, la Comisión Europea puso en marcha un grupo de expertos en materia de responsabilidad con el propósito de

[733] Comunicación de la Comisión, *Maximising the Benefits of Artificial Intelligence for Europe* («Maximizar los beneficios de la inteligencia artificial para Europa», documento en inglés), COM(2018)237).
[734] COM (2021) 206 final 2021/0106 (COD)

analizar detalladamente el efecto de estas cuestiones no resueltas. El grupo de expertos se conformó por dos subgrupos con dos objetivos y configuraciones asociadas.

El primero con la finalidad de ayudar a la Comisión en la interpretación, aplicación y posible actualización de la Directiva, especialmente a la vista de la evolución tanto de la jurisprudencia nacional como de la UE, las implicaciones de las tecnologías nuevas y emergentes y cualquier otro avance en materia de responsabilidad por productos. El mismo se compuso por representantes de los Estados miembros, la industria, organizaciones de consumidores, la sociedad civil y el sector académico.

El segundo con el objetivo de evaluar si el régimen general de responsabilidad es "adecuado para facilitar la asimilación de las nuevas tecnologías al promover la estabilidad de las inversiones y la confianza de los consumidores"[735]. El mismo se compuso únicamente por expertos académicos y profesionales independientes.

El informe concluyó destacando su objetivo final, que no es otro que establecer un marco favorable y fiable en materia de responsabilidad por productos defectuosos que promueva la innovación, el empleo y el crecimiento, si bien, protegiendo a los consumidores, pero también, la seguridad del público en general. Y esto es precisamente lo que necesitamos en materia de inteligencia artificial, y que proteja a personas (cualquiera que sea su condición), empresas y Administraciones Públicas como usuarias de la misma.

7.5. Otras posibles soluciones y alternativas

A continuación, profundizaré sobre algunas de las consideraciones efectuadas y otras adicionales en relación con las posibles soluciones y alternativas expuestas, ante la insuficiencia del régimen general de responsabilidad extracontractual y en materia de productos defectuosos en España.

735 En el *Staff working Document on Liability for emerging digital technologies* (Documento de trabajo de los servicios de la Comisión sobre la responsabilidad por las tecnologías digitales emergentes, documento en inglés) (SWD (2018) 137) ya se hace hincapié en algunos de los temas que se debatirán en esta configuración.

De inicio, considero que el productor o fabricante es quién conoce mejor el sistema, el que lo dota de sus capacidades, atributos, autorizaciones, restricciones, limitaciones y seguridad, por lo que consiguientemente, a mi juicio, es quién estaría en la mejor posición para controlar los riesgos asociados al mismo, acordes a las capacidades y atributos otorgados, y para establecer mecanismos, medidas y controles para asegurar su seguridad y el cumplimiento regulativo, incluyendo el control y la supervisión humana.

Y precisamente esas capacidades y atributos, y sus ventajas asociadas -como automatización, agilidad, eficacia, eficiencia o ahorro de costes- son lo que puede hacer diferencial el sistema inteligente como producto o servicio, crear una ventaja competitiva y constituir el factor determinante de adquisición del mismo en lugar de otros que pueden coexistir en el mercado.

La responsabilidad del productor o fabricante, en mi opinión, debe pues dimensionarse a las capacidades y atributos de su producto o servicio, en definitiva, a sus características y a los riesgos asociados al mismo y a estas, en la medida que ello debería ser un incentivo para la creación de productos más innovadores, pero también más seguros y, en definitiva, más beneficiosos para la sociedad en general y, en consecuencia, más atractivos para su adquisición y consiguiente beneficio.

La posibilidad de que se pueda dotar a estos sistemas de mayores capacidades en su diseño y construcción que les permitan evolucionar en sus capacidades de procesamiento y aprendizaje, adaptarse a los diferentes contextos, modificar su comportamiento o poder tener cierta capacidad de autoprogramarse, comporta necesariamente un incremento de los riesgos potenciales asociados a los mismos y, en consecuencia, de la responsabilidad de la que debe responder su productor, el cual deberá extremar su diligencia en todo el proceso, anticipándose a los posibles problemas mediante adecuados procedimientos de aseguramiento de la calidad, seguridad, conformidad regulatoria, análisis de riesgos y evaluaciones de impacto en la tecnología que desarrolle, de los cuales deberán derivarse e implementarse controles y medidas de seguridad adecuadas, a mi juicio, no sólo de naturaleza meramente preventiva o precautoria, sino incluso detectivas -dentro de una monitorización y seguimiento, automatizado o no-, correctivas (o reactivas) y, en su caso, hasta incluso evolutivas durante todo su ciclo de vida, en la medida que determinados sistemas esenciales para la vida o continuidad de un negocio y red, podrían verse

afectadas por cambios, por ejemplo en la tecnología, que pudieren conllevar su imposibilidad de interoperar, suspender funciones o incluso sus funciones, lo que exigiría revisiones y mantenimiento.

De este modo, considero que este enfoque de la responsabilidad aseguraría la ética -*Ethics by design*-, la seguridad -*Security by design*- y la conformidad regulatoria -*Compliance by design*- en el diseño y concepción de los sistemas de inteligencia artificial y, debería igualmente asegurarse en todo su ciclo de vida en la medida que sea controlable total o parcialmente por el producto o fabricante.

Sin duda, la aprobación de las propuestas regulatorias de la UE objeto de esta investigación en materia ética, jurídica y de responsabilidad de la inteligencia artificial contribuiría de manera definitiva a todo ello mediante su conversión en una exigencia legal, si bien, debe revisarse su enfoque, objeto y alcance conforme a las reflexiones y consideraciones que he efectuado a lo largo de esta investigación.

La seguridad debería estar garantizada al máximo nivel posible, tanto desde el momento de su activación como con posterioridad, y la gestión de los riesgos se situarían no tanto en el sistema, objeto o ente sin personalidad jurídica sino a todas las personas que estarían en mejor disposición para controlar y minimizar los riesgos y gestionar su materialización e impacto negativo. De inicio, esta persona, en mi opinión, sería principalmente el productor o fabricante, de un lado, y el propietario, operador o usuario, de otro, en función del contexto, pero en modo alguno el propio sistema inteligente. Además, el propio producto debe evidenciar un nivel de diligencia adecuado a los dispositivos, sistemas y tecnología que fabrica, asumiendo igualmente responsabilidad en su elección y vigilancia de los diseños, desarrollos y tecnologías más adecuadas, como de los componentes y datos que integra en aquellos.

Desde una óptica proteccionista de los intereses de las personas afectadas por el funcionamiento de sistemas con niveles de riesgo significativo o alto, la responsabilidad objetiva en estos supuestos debería ser imputada al productor o fabricante en base a estas consideraciones, que sería responsable de los daños causados por el sistema de inteligencia artificial con independencia de la imprevisibilidad y del conocimiento en el

momento de su puesta en circulación en el mercado de los posibles o potenciales riesgos y daños futuros.

Se trataría pues de un sistema de responsabilidad objetivo y, a mi juicio, *cuasi absoluta* o absoluta derivado del hecho de generar un riesgo o peligro para la sociedad, conforme al cual sería irrelevante que el fabricante probara el desconocimiento de los riesgos de diseño y desarrollo en el momento de la puesta en circulación y activación de los sistemas, según el estado de los conocimientos científicos y técnicos en ese momento para exonerarse de responsabilidad. Y todo ello, como he referido, al margen de las responsabilidades concurrentes o responsabilidad exclusiva que pueda recaer en otros agentes, en especial, al propietario, operador o usuario del sistema inteligente. De este modo, la responsabilidad recaería igualmente en quién se beneficia del riesgo creado.

Como he comentado al analizar esta cuestión anteriormente, distintos autores como los citados Solé i Feliu y Núñez Zorrilla consideran que este tipo de sistema equivaldría a gravar al productor o fabricante con una responsabilidad no simplemente objetiva sino absoluta, lo que podría conllevar distintas desventajas para el empresario y para el propio acceso del ciudadano a la tecnología como también apuntaba.

Sin embargo, considero que todo ello constituiría un factor esencial para motivar el desarrollo de una innovación segura por parte del sector empresarial y una necesidad de mejora continua, esfuerzo e inversión constante para mejorar las medidas y el perfeccionamiento de los sistemas para reducir los riesgos inherentes, potenciales o eventuales de los mismos y sus capacidades, aunque ello podría conllevar razonablemente y desde el enfoque empresarial, el incremento de costes para el empresario y correlativo incremento del precio de estos sistemas y de los posibles servicios de monitorización o mantenimiento asociados a los mismos que, junto con los riesgos adicionales de indemnización asociada, pudiera a su vez conllevar a su inaccesibilidad para la sociedad en general.

No obstante, una posible solución a este último problema, como he referido, podría ser el establecimiento de un sistema de responsabilidad limitada, es decir, que al menos se limite o palie el impacto de este sistema de responsabilidad para el empresario, en particular, estableciendo una limitación máxima en las cuantías indemnizatorias a abonar en

concepto de daños materiales por los productores o fabricantes, nunca por daños de otra naturaleza, lo que facilitaría la suscripción de seguros que proporcionen las coberturas adecuadas con terceros a precios adecuados ante la acotación del riesgo.

En definitiva, se trataría de encontrar un punto de equilibrio entre todos los intereses en juego, de modo que, de un lado, no se desincentive el desarrollo, la innovación, la inversión y el desarrollo empresarial y, de otro, se garantice la seguridad, la confianza, la accesibilidad, los usos y los beneficios de la tecnología por parte de la sociedad en general, como destacan autores como Ercilla García[736].

Esta opción es por la que parece apostar la UE en sus documentos de trabajo, informes y propuestas reguladoras, con distintos matices, en especial, en su reciente propuesta regulatoria sobre responsabilidad civil derivada de funcionamiento de sistemas de inteligencia artificial, que será objeto de análisis en el posterior apartado.

Incluso podría diseñarse un sistema de responsabilidad de responsabilidad limitada podría ser extensible y de aplicación a diseñadores, desarrolladores, propietarios o usuarios, máxime cuando los sujetos responsables puedan haber contribuido de manera efectiva a un fondo de compensación o bien suscriban un seguro que garantice la compensación de los daños causados por un sistema de este tipo, conforme a lo propuesto en la Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica, precitada anteriormente.

Del mismo modo, podría concebirse como un régimen de responsabilidad objetiva en la que la persona afectada únicamente debería probar el daño y la relación de causalidad entre éste y la conducta del sistema, y en la que al productor o fabricante correspondería, en su caso, la acreditación de cualquier causa que pudiera eximirle de la misma, como podría ser, la actuación correcta del sistema inteligente, la fuerza mayor, la culpa del perjudicado -según establece el artículo 145 del TRLGDCU[737]-, no haber puesto en circulación el sistema inteligente -según el artículo 140.1º.a) del TRLGDCU- o que el

[736] ERCILLA GARCÍA, J. (2018). *Normas de Derecho Civil y Robótica. Robots inteligentes, personalidad jurídica, responsabilidad civil y regulación*. Aranzadi. Navarra 2018. P. 114.
[737] Real Decreto Legislativo 1/2007, de 16 de noviembre (Texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias). BOE 30.11.2007.

sistema no haya sido fabricado para la venta o cualquier otra forma de distribución con finalidad económica, ni fabricado, importado, suministrado o distribuido en el marco de una actividad profesional o empresarial -según el artículo 140.1º.c) del TRLGDCU-.

No obstante, la carga de la prueba se debería ajustar al principio de disponibilidad probatoria, en la medida que es el propio sistema inteligente y, teóricamente, su productor o fabricante, son quienes están en mejor posición para para acreditar los hechos y circunstancias, en base a las exigencias éticas de transparencia, trazabilidad, responsabilidad y *accountability*.

La información previa suministrada por el productor o fabricante a usuarios podría tener una relevancia determinante a la hora de valorar la culpa del perjudicado como causa eximente y de reducir o excluir la responsabilidad del productor o fabricante, donde se debería advertir expresamente de los peligros asociados a sus capacidades y usos recomendados (o permitidos) y no recomendados (o no permitidos), especialmente en contextos o finalidades ajenas para las que el sistema fue inicialmente concebido.

En este sentido, la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021[738], establece importantes aspectos relacionados con estas cuestiones, al regular los requisitos y obligaciones de los sistemas de inteligencia artificial considerados de alto riesgo, no respecto del resto.

Entre otros requisitos, exige establecer, implementar, documentar y mantener un sistema de gestión de riesgos, el cual deberá contemplar la identificación y análisis de los riesgos conocidos pero también los "previsibles" asociados con cada sistema de inteligencia artificial de alto riesgo, la valoración y evaluación de los riesgos que pueden surgir cuando el sistema se utilice de acuerdo con su propósito previsto pero también en condiciones de uso indebido razonablemente previsible, la evaluación de otros riesgos que puedan surgir sobre la base del análisis de los datos recopilados del sistema de seguimiento posterior a

[738] COM (2021) 206 final 2021/0106 (COD)

la comercialización mencionado en el artículo y la adopción de las medidas adecuadas de gestión de riesgos conforme a lo anterior.

Es decir, exige considerar los riesgos conocidos, pero también los previsibles, evaluar y gestionar los riesgos tanto cuando el sistema se utilice conforme a su finalidad, pero también en condiciones de uso indebido razonablemente previsible al que hice referencia anteriormente, y también de los riesgos dimanantes de su seguimiento tras su comercialización y la adopción de las medidas adecuadas.

La nueva propuesta define dicho concepto de "uso indebido razonablemente previsible", como expuse anteriormente, como el uso de un sistema de inteligencia artificial de una manera que no está de acuerdo con su propósito previsto, pero que puede resultar de un comportamiento humano razonablemente previsible o de la interacción con otros sistemas. El problema es que introduce en su definición un nuevo concepto jurídico indeterminado "comportamiento humano razonablemente previsible".

De este modo, se exige un esfuerzo adicional a los fabricantes en la evaluación de riesgos para considerar aquellos riesgos potenciales derivados de posibles usos alejados de su finalidad prevista pero que pueden derivarse de un comportamiento humano razonablemente previsible o de la interacción con otros sistemas, que pueden estar dotados o no de inteligencia artificial.

La cuestión será aclarar ese concepto jurídico indeterminado de "comportamiento humano razonablemente previsible", que puede incluir comportamientos ilegítimos, pero también ilícitos desde un punto de vista civil, administrativo, penal o de cualquier otra naturaleza.

A mi juicio, como considero comparten autores como Argyri Panezi[739], la incertidumbre y cierta imprevisibilidad forma de parte de la inteligencia artificial y no debe tratarse como un error o defecto -que por otra parte es inherente a toda tecnología-, si bien, comporta un riesgo que debe ser consecuentemente considerado y para el que deben

[739] PANEZI, A. (2021). "IA: un enfoque ecosistémico para gestionar el riesgo y la incertidumbre", en García Mexia, P. y Pérez Bes, F. (Eds); Panezi. A. (Coord). (2021). *Artificial Intelligence and the Law*. Ed. La Ley (Wolters Kluwer). 2021. Edición electronica.

adecuarse los marcos de responsabilidad actuales para gestionarlo adecuadamente al objeto de garantizar un resarcimiento íntegro y efectivo.

Por otra parte, la precitada Propuesta de Reglamento de 2021 exige que las medidas adoptadas sean tales (para sistemas de alto riesgo) que cualquier riesgo residual asociado a cada peligro, así como el riesgo residual global se consideren aceptables, siempre que el sistema de inteligencia artificial de alto riesgo se utilice de acuerdo con su finalidad prevista o en condiciones de uso indebido razonablemente previsibles. Y además exige que estos riesgos residuales se comuniquen al usuario.

Del mismo modo, el Reglamento propuesto exige que la información incluya los resultados no deseados previsibles y las fuentes de riesgos para la salud y la seguridad, los derechos fundamentales y la discriminación en vista del propósito previsto del sistema de inteligencia artificial, así como una descripción detallada del sistema de gestión de riesgos.

Y, por último, adicionalmente, su artículo 13 regula los requisitos de transparencia y suministro de información a los usuarios, que incluye la de comunicar cualquier circunstancia conocida o previsible, relacionada con el uso del sistema de acuerdo con su finalidad prevista o bajo condiciones de uso indebido razonablemente previsibles, que pueda conducir a riesgos para la salud y seguridad o los derechos fundamentales.

Al hilo de mi argumentación sobre la posible construcción de esta responsabilidad del fabricante y sus causas eximentes, la deficiente o inexistente información podría conllevar la reducción o inaplicación del uso indebido o incorrecto del sistema por el usuario como causa de exoneración de responsabilidad, especialmente cuando el uso indebido o incorrecto tuviera su origen en una falta de información que hubiera motivado un actuar distinto por parte del usuario, cuestión que ha sido analizada por algunos autores en relación con la protección del consumidor[740].

[740] PARRA LUCÁN, M.A. (2011). *La protección del consumidor frente a los daños. Responsabilidad civil del fabricante y del prestador de servicios.* Colección Derecho del Consumo. Editorial Reus. Madrid, 2011. Pp. 137 y 138.

En el marco de mis reflexiones sobre este posible escenario, el productor o fabricante siempre debería responder por los daños del sistema, haya o no informado en mayor o menor medida sobre los riesgos, sin perjuicio de la concurrencia de culpa o culpa exclusiva del usuario al realizar una conducta o uso en un contexto o propósito inadecuados. Y ello por cuanto que es el fabricante quién dota de autonomía y demás capacidades al sistema bajo unas instrucciones, parámetros, condiciones, autorizaciones y restricciones preconcebidas, unas que podrían ser alteradas por el propio sistema en base a su autoaprendizaje y facultades de autoprogramación y otras contra las que nunca podría operar y funcionar el sistema. ¿O permitiríamos el uso de este tipo de sistemas con dichas capacidades para gestionar la seguridad de nuestras redes, sistemas de gestión empresariales, sistemas de regulación del tráfico, de conducción o vuelo autónomo, el cuidado de nuestros mayores, enfermos, la dirección de una intervención quirúrgica o incluso nuestra propia vida?

En cualquier caso, en caso de que los futuros marcos no impidan la aplicación de esta causa de exoneración de responsabilidad, dada la diversidad y complejidad de los sistemas inteligentes y contextos de uso para según qué tipo de usuarios, la eficacia de la información previa para la exoneración de la responsabilidad debería ser de aplicación muy restringida para supuestos evidentes. En cualquier caso, la mera información previa nunca debería poder legitimar la distribución en el mercado de sistemas que entrañen riesgo para personas, cosas e instalaciones, ni conocidos ni informados, y que carezcan de expectativa general de seguridad.

En este sentido, me permito poner como ejemplo los servicios de una red social. ¿Puede ponerse en el tráfico una red social que no garantice seguridad y privacidad a sus usuarios, ni tan siquiera expectativa alguna? Es clara la utilidad social de este tipo de redes, pero obviamente la respuesta es negativa, a pesar de algunos episodios escandalosos que han salido a la luz pública durante los últimos años, que han motivado actuaciones con cierta determinación por parte de algunas de las autoridades competentes y que requieren y requerirán mayor intervención, precisamente y entre otros factores, ante la adición e integración de la inteligencia artificial en su gestión.

Otros autores, como Jimeno Muñoz[741], considera determinantes, entre otros aspectos, las advertencias de los riesgos y las precauciones de uso que el fabricante ofrezca citando la Sentencia nº 1087/2008 del Tribunal Supremo español, Sala 1ª, de 21.11.2008.

No obstante, otros autores[742], al abordar la responsabilidad del productor o fabricante, consideran que ese deber de información de éstos tiene actualmente un límite, que sería, de nuevo, la inexistencia de obligación para el mismo de advertir sobre riesgos o daños totalmente imprevisibles conforme al estado de la técnica en el momento de la puesta en el tráfico del sistema, sin perjuicio de la obligación de informar de todo ello tanto pronto haya tenido conocimiento del riesgo concreto.

De este modo, consideran que la obligación de informar sólo puede proyectarse sobre riesgos previsibles, en la medida que no podrían advertir sobre lo desconocido. En este sentido me permito citar la doctrina recogida en la Sentencia del Tribunal Supremo español, Sala 1ª, de 3 de diciembre de 1997, en virtud de la cual un sistema inteligente podría considerarse defectuoso debido a la inadecuación o inexistencia de instrucciones o advertencias conforme al estado de la ciencia y la técnica.

El protagonismo de los creadores de sistemas inteligentes es crucial ante estos posibles marcos de responsabilidad, en la medida que deberá asegurarse ya en el proceso de diseño y programación del mismo que el sistema no pueda aprender pautas, tomar decisiones o realizar acciones al margen de las instrucciones dadas que le puedan llevar a causar daños o incurrir en actos ilícitos, en la medida que puedan y deban preverse, sin perjuicio del principio general de que nadie debería responder de eventos que no hubieran podido preverse o que, previstos, fueran inevitables.

[741] JIMENO MUÑOZ, J. (2017). *La responsabilidad civil en el ámbito de los ciberriesgos*. Fundación Mapfre, Madrid 2017. P.130.
[742] SALVADOR CODERCH, P. S. Y RAMÓN GONZÁLEZ, S. (2008). "Defectos de producto". En SALVADOR CODERCH, P. S. Y GÓMEZ POMAR, F. (Coords). *Tratado de responsabilidad civil del fabricante*. Capítulo IV. Editorial Thomson Civitas. Navarra 2008. P. 196. En este mismo sentido, NÚÑEZ ZORILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Op. cit. P. 73.

En cualquier caso, la construcción de un régimen de responsabilidad adecuado para los sistemas de inteligencia artificial, dada la dimensión cuantitativa y cualitativa de los daños potenciales, debería ir acompañado, como he expuesto previamente, de un seguro obligatorio de responsabilidad civil a concertar por el productor o fabricante, al igual que se hace, por ejemplo con vehículos a motor o con un perro de raza peligrosa en base a un enfoque de riesgos, al objeto de garantizar el derecho al resarcimiento efectivo de las personas afectadas por aquéllos en caso de conductas dañinas (imprevisibles o no), junto con otras medidas garantes de un resarcimiento integral y efectivo, como los sistemas de compensación.

Algunos autores como Badillo Arias[743], proponen un sistema de responsabilidad objetiva atribuible al propietario, arrendatario o al detentador de un sistema inteligente robótico, "el cuál debería indemnizar los daños producidos al tercero, salvo que haya alguna causa de exoneración de esta responsabilidad, como la culpa exclusiva de la víctima o la fuerza mayor", en términos similares a lo previsto en el artículo 1105 del Código Civil español.

De este modo, una vez indemnizado al tercero por el responsable civil directo (la entidad aseguradora en caso de mediar un seguro), éste podría repetir contra el que finalmente fuere responsable dentro de la cadena como fabricante, programador o entrenador. Se trataría pues un sistema de atribución de responsabilidad inspirado en el vigente en España en relación con los daños personales derivados de la circulación de vehículos regulado en la Ley de responsabilidad civil y seguro en la circulación de vehículos a motor[744]. De este modo, las víctimas serían indemnizadas salvo concurrencia de causa eximente, y las distintas responsabilidades concurrentes en la cadena se depurarían en vía de regreso.

Desde un punto de vista empresarial, el seguro como complemento adicional para asegurar el derecho de resarcimiento de la persona afectada constituirá un coste adicional que inevitablemente repercutirá en el precio de la venta o cesión del sistema inteligente, por lo que en la línea comentada, deberían además limitarse las cuantías de la

[743] BADILLO ARIAS, J.A. (2019). "Responsabilidad civil y aseguramiento obligatorio de los robots". En MONTERROSSO CASADO, E. (Dir.). *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch 2019. P. 47.
[744] RDLeg. 8/2004, de 29 de octubre (Texto Refundido de la Ley sobre responsabilidad civil y seguro en la circulación de vehículos a motor)

responsabilidad patrimonial, en aras, de un lado, de no desincentivar la inversión, la actividad empresarial, el desarrollo y la innovación tecnológica y, de otro, de permitir la existencia en el mercado de este tipo de contratos de seguro y a precios adecuados, dado que de otro modo podrían no ser productos rentables para las compañías aseguradoras, dado el elevado riesgo potencial o, en caso de serlo, podrían ser comercializados con primas desproporcionadas, especialmente para empresas emergentes y *startups*, que constituyen uno de los motores del desarrollo y la innovación tecnológica a nivel mundial.

En este contexto, considero que estos regímenes no deberían limitar los daños de carácter no patrimonial.

Todo ello podría comportar que, en determinados casos, el valor de los daños a indemnizar sea muy superior a los límites prefijados, lo que desampararía a las víctimas que, en cualquiera de los casos, verían afectado su derecho al resarcimiento.

Esta cuestión podría resolverse paralelamente mediante los precitados fondos de compensación que pudiesen indemnizar, en caso de ausencia de seguro o de insuficiencia de éste, completando el mismo, el cual podría constituirse mediante exacciones fiscales, como señalan distintos autores[745], por parte de quien crea el sistema, los dota de sus concretas capacidades y autonomía, crea igualmente el riesgo y es quién mayor control tiene sobre el mismo, en definitiva, el productor o fabricante. Todo ello sin perjuicio de las aportaciones voluntarias de otros sujetos que podrían verse involucrados en materia de responsabilidad, que podrían ver limitada su responsabilidad en caso de haberlas efectuado. El informe *Liability for Artificial Intelligence and other emerging digital technologies*[746] de 2019, recoge expresamente en sus conclusiones la opción del establecimiento de fondos de compensación para proteger a las víctimas en el caso de que, conforme a las normas de responsabilidad aplicables, sus reclamaciones no puedan ser satisfechas.

[745] Ver Diaz Alabart y NÚÑEZ Zorilla: DÍAZ ALABART, S. (2018). *Robots y responsabilidad civil*. Editorial Reus. Madrid 2018. Pp. 83-88; NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Op. Cit. P. 73.
[746] *Liability for Artificial Intelligence and other emerging digital technologies.* UE. 2019. Recuperado de: https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2020/01-09/AI-report_EN.pdf. Consultado el 14.12.2020.

Por lo que se refiere a su aseguramiento y el establecimiento de límites, debería abrirse un diálogo con el sector asegurador para la generación de productos adecuados para los fines pretendidos a precios asumibles. No obstante, algunos autores, como la precitada Núñez Zorrilla, propone que no se limite la indemnización, que se amplíe el concepto de productor lo máximo posible, que extienda el plazo de prescripción y la aplicación del principio de beneficencia cuando un sistema inteligente cause daños para evitar otros mayores. Es decir, la posible creación de un sistema de responsabilidad objetiva absoluta pero con ciertas limitaciones en la línea de la Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica, que facilitaría el establecimiento de esta medida de traslado del riesgo a un tercero solvente (aseguradora) y aseguraría el resarcimiento de la persona afectada, complementado con un fondo de compensación para garantizar la reparación de los daños en los casos de ausencia de una cobertura del seguro obligatorio.

En cualquier caso, los futuros marcos jurídicos que complemente los regímenes de responsabilidad vigentes en relación con la inteligencia artificial, deberán ser acompañados por la regulación de algunos aspectos relacionados con la misma ante sus riesgos y retos, al objeto de definir marcos claros de requisitos y obligaciones éticas y jurídicas que deban cumplir los mismos bajo un enfoque proactivo, al objeto de evitar o minimizar de materialización de sus riesgos y, de manera consecuente, de precisar acudir a mecanismos reactivos. Cualquier sistema inteligente deberá reunir y cumplir con los requerimientos generales exigidos por los marcos generales vigentes, así como los específicos que se construyan en materia ética, jurídica y de seguridad.

Si no pueden ponerse en circulación vehículos a motor sin frenos o airbags que pongan en peligro al propietario, usuario o terceros (ni en breve, posiblemente, tampoco vehículos que no lleven sistemas de detección automática de obstáculos y frenado), si no pueden circular aeronaves por encima de zonas con determinada densidad de población por los riesgos de seguridad y molestias que comporta, si tenemos un perro de raza peligrosa y salimos a la calle no podemos poner en peligro a terceros y debemos llevarlo atado y con bozal, si no se pueden poner en el mercado vacunas que salvan vidas hasta que se demuestre su seguridad y eficacia, de manera consecuente, un sistema y, además, supuestamente, inteligente, no puede ponerse en el mercado sin garantizar que sea seguro en su funcionamiento, tanto en su comportamiento conforme a las instrucciones y

finalidades preconcebidas en su diseño, como en su comportamiento conforme a las instrucciones adquiridas y autogeneradas por el mismo en el marco de las capacidades de las que fue dotado, es decir, no puede lanzarse al mercado un sistema sin las medidas de seguridad y controles necesarios para gestionar uno de los principales riesgos de seguridad asociado a su grado de autonomía y capacidades, que es su impredecibilidad, especialmente grave en función de su finalidad y contexto de aplicación y uso. Y tampoco que comporte riesgos potenciales por usos indebidos razonablemente posibles o, cuanto menos, previsibles.

La doctrina está abordando en los últimos años los posibles regímenes complementados o alternativos a los vigentes.

Ercilla García[747] propone un sistema de responsabilidad en cascada, esto es, *culpa in curando, culpa in faciendo, culpa in educando, culpa in codificando* y culpa del sistema inteligente, si bien, si el acto dañoso es consecuencia de una decisión autónoma en sentido propio, y el actuar se aparta de la lógica humana a pesar de que el sistema ha operado conforme a los principios lógicos concebidos por el mismos, propone hablar más exactamente de *culpa in singularitatem*.

Algunos autores como Núñez Zorilla[748], proponen la obligación del fabricante de llevar a cabo un seguimiento del sistema inteligente una vez puesto en circulación, para comprobar su uso potencial más allá del apropiado, de modo que la obligación del productor o fabricante de prevenir, identificar y evitar riesgos podría extenderse después de su comercialización, ya sea bajo una modalidad remota, ya sea continua o eventual -que debería ofrecer las garantías necesarias ante la posible afectación para la privacidad de los usuarios y terceros relacionados con el sistema-, o incluso física eventualmente, especialmente a través de revisiones oficiales obligatorias de forma periódica.

[747] ERCILLA GARCÍA, J. (2018). "Aproximación a una Personalidad Jurídica Específica para los robots". *Revista Aranzadi de derecho y nuevas tecnologías*. Nº 47. 2018. Pp. 21-22.
[748] NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Op. Cit. P. 72.

La nueva Propuesta de Reglamento de abril de 2021, conforme he analizado en el anterior capítulo, ya contempla esa obligación de seguimiento y monitorización de los riesgos posteriormente a su comercialización durante su ciclo de vida.

Esto conllevaría un análisis de riesgo constante y actualizado que permitiría ampliar los análisis de riesgos y evaluación de impacto previos a su puesta en el tráfico, tanto de origen interno -del sistema y su evolución de sus capacidades y conductas-, como externo -especialmente evolución del estado de la técnica-, adoptar y proponer medidas, controles y parches de seguridad, y proporcionar más información y advertencias de sus consecuencias dañosas, especialmente a la vista de los usos potenciales y nuevos riesgos detectados, que podrían conllevar desde la propuesta de su suspensión a su retirada del tráfico. Todo ello podría llevarse a cabo mediante verificaciones periódicas.

Por último, la atribución de una responsabilidad directa absoluta al fabricante sin posibilidad de exoneración por los riesgos del desarrollo y consiguiente litigiosidad y complejidad de su prueba, evitaría también este problema, dado que se consideraría responsable con independencia si pudo conocer o no los riesgos, con lo que no sería necesario presentar ningún tipo de prueba con este objetivo.

La exigencia de este tipo de responsabilidad objetiva absoluta debe ser un elemento también motivador para que no se pongan en el mercado productos de utilidad social que comporten riesgos desconocidos, sin expectativas generales de seguridad y sin medidas de control alguno, y todo ello en base a la preceptiva seguridad en el tráfico.

De este modo, sólo concibo una inteligencia artificial al servicio de la humanidad, que satisfaga necesidades, solucione problemas y mejore nuestra vida y nuestro mundo, en la que se asegure el control y la supervisión humana y la seguridad durante todo su ciclo de vida, es decir el dominio de la raza humana sobre entes artificiales.

Este enfoque de responsabilidad chocaría frontalmente además contra cualquier postura futura de otorgar personalidad jurídica a los sistemas de inteligencia artificial que pondría en riesgo el objetivo reparador y preventivo de la responsabilidad civil, se podría fomentar el uso indebido y plantearía posibles conflictos morales.

7.6. Hacia un nuevo marco normativo

El marco jurídico vigente en materia de responsabilidad civil extracontractual y de responsabilidad derivada de productos defectuosos en España y en la UE, como he referido, no es adecuado para resolver todos los supuestos que los sistemas de inteligencia artificial pueden plantear ya y mucho menos en lo sucesivo ante sistemas más avanzados, por lo que es necesario revisar los mismos para complementarlos y, en su caso, llevar a cabo las modificaciones y actualizaciones necesarias para conformar un marco sólido, adecuado e integrado que proporcione la seguridad necesaria a todas las partes interesadas.

Conforme al marco jurídico actual y, en general, cuando el daño se produzca por la operación o uso negligente del sistema, la responsabilidad recaería inicialmente en quien la operaba o usaba cuando se causó el mismo.

Cuando el daño tenga su origen en un diseño inadecuado o defecto la responsabilidad recaería inicialmente en el fabricante, sin perjuicio de la posible repetición de responsabilidad frente a diseñadores, desarrolladores o ensambladores externos.

De los errores de programación debería responder el programador, si bien, de hallarse integrado en el fabricante, debería responder éste.

Cuando responda a un entrenamiento o formación deficiente o inadecuada, la responsabilidad debería recaer en el formador.

Cuando se deba a un deficiente mantenimiento del sistema inteligente, inicialmente la responsabilidad debería recaer en el propietario o usuario del mismo, en función de los compromisos contraídos entre los mismos.

La autonomía en las decisiones o acciones diluye las responsabilidades.

Si el sistema inteligente puede incardinarse como producto a los efectos de la normativa sobre responsabilidad de productos defectuosos europea o española (o servicio en el caso de esta última), la responsabilidad se situaría inicialmente en la órbita del fabricante. En otro caso, se aplicarían las normas generales sobre responsabilidad civil contractual o

extracontractual, siendo inicialmente responsable, en función del contexto, el fabricante, el proveedor, el distribuidor, el operador, el propietario, el usuario o cualquier otro agente involucrado.

La insuficiencia del marco regulador vigente y la dificultad de incardinar y resolver los diversos supuestos que puedan suscitarse en relación con sistemas inteligentes en base a los mismos -especialmente sistemas dotados de inteligencia artificial más avanzada o "fuerte", con supuesta autonomía y capacidad de aprendizaje, requiere una profunda reflexión global sobre un nuevo marco, complementario, ampliatorio, modificativo y/o sustitutivo en algunos casos de los existentes, y tanto desde una perspectiva y enfoque ético y jurídico, como de seguridad y de riesgos.

En atención a la tipología y capacidades del sistema inteligente, sector, finalidad, aplicación o nivel de riesgo, estos marcos podrían sustentarse en sistemas de responsabilidad civil objetiva o por riesgos, en todo caso solidaria y con limitación cuantitativa de la indemnización, y complementada con seguros obligatorios y fondos de compensación para responder en caso de ausencia o falta de cobertura de aquéllos, con previsión de ciertos incentivos de limitación de responsabilidad para su dotación por los potenciales responsables. Incluso se podría plantear una responsabilidad objetiva absoluta.

Estos marcos deberán concebirse desde el equilibrio entre seguridad, confiabilidad, innovación, inversión, economía, despliegue y aplicación de la inteligencia artificial, y deben abordase desde una perspectiva global, general y estratégica, de modo que no se aborde la responsabilidad de la inteligencia artificial desde casos o sectores concretos, en la medida que los contextos, características, capacidades, tipologías, multiplicidad de sujetos intervinientes o ciclo de vida podría llevar a una conceptualización sesgada, errónea y no exportable a otros, sino que deben ser marcos adaptables a la evolución del desarrollo y aplicación de la inteligencia artificial. Algunos autores como Vázquez de

Castro[749] opinan que un enfoque basado en una casuística concreta o sectorial llevaría a una regulación fragmentada.

Conforme al marco jurídico vigente en materia de responsabilidad contractual, extracontractual y responsabilidad por productos defectuosos, tal y como se ha analizado anteriormente, no puede considerarse responsable a un sistema de inteligencia artificial, el cual carece de personalidad jurídica.

Otra cosa sería considerar estos sistemas entes sin personalidad jurídica, entre otras opciones objeto de reflexión, pero con prerrogativas adecuadas a su tipología, capacidades, grado de autonomía, impredecibilidad, uso o aplicaciones previstas, y de cuyos daños responderían los seguros o fondos asociados o las personas detrás de los mismos.

Por todo ello, ante la insuficiencia de estos marcos para dar soluciones adecuadas a los diversos conflictos que ya se están generando y que se generarán el futuro, y la imposibilidad de imputar la responsabilidad jurídica al propio sistema inteligente, los nuevos marcos propuestos deben dar una respuesta adecuada a todo ello. En el próximo apartado se analizará el sistema definido en la propuesta del Parlamento Europeo de 20 de octubre de 2020.

8. La propuesta europea

8.1. Introducción

Durante el tratamiento de los apartados anteriores, he puesto de relieve la dificultad de aplicación e insuficiencia de los sistemas comunes y especiales de responsabilidad civil

[749] VÁZQUEZ DE CASTRO, E. (2020). "Aproximación a la responsabilidad derivada de los riesgos de la inteligencia artificial en Europa". En Solar Cayón, J.I. *Dimensiones éticas y jurídicas de la inteligencia artificial en el marco del Estado de Derecho.* Cuadernos de la Cátedra de Democracia y Derechos Humanos. Universidad de Alcalá: Defensor del Pueblo. 2020. P. 235.

vigentes para responder a los contextos actuales y venideros en materia de daños causados por sistemas de inteligencia artificial.

La atribución de la responsabilidad al productor puede suponer cuantiosas indemnizaciones derivadas de los daños causados por los sistemas inteligentes producidos por aquél, si bien, exclusivamente en los supuestos en los que sea aplicable el marco vigente.

La utilización del sistema de responsabilidad civil extracontractual comporta igualmente la necesidad de construir la culpa o negligencia del responsable en la que intervendría un sistema de inteligencia artificial, con supuesta autonomía, nunca plena conforme analicé anteriormente, que en cualquier caso debería ser restringida y segura mediante la adopción de controles y medidas en su diseño que impidan una verdadera libertad y autonomía en sus decisiones y actuaciones y, en consecuencia su plena impredecibilidad, especialmente mediante instrucciones matrices, "líneas maestras", criterios, objetivos, restricciones e incluso mecanismos reactivos, de modo que se garantice el pretendido control y supervisión humana en todo el ciclo de vida.

El contexto analizado ha provocado un necesario debate en el seno de la UE sobre las posibles soluciones a estas cuestiones, entre las que se ha barajado desde un inicio y desde su comité de expertos en inteligencia artificial, la posibilidad de imputar la responsabilidad directamente a determinados sistemas de inteligencia artificial, mediante la atribución de una personalidad jurídica propia y diferenciada, opción que parece definitivamente desechada, por el momento, de la agenda legislativa actual en la UE.

En los próximos apartados pretendo analizar el posicionamiento de la UE en esta materia que ha culminado con la Resolución del Parlamento Europeo, de 20 de octubre de 2020 sobre un régimen de responsabilidad civil en materia de inteligencia artificial[750], que ya incorpora una Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial.

[750] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))

8.2. Antecedentes

El Parlamento Europeo abordó frontalmente la responsabilidad de los sistemas de inteligencia artificial en una primera Resolución de 16 de febrero de 2017[751], que ha sido objeto de análisis parcial en distintos apartados de esta investigación, en la que incorporó sus recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica.

La Resolución, como he expuesto al analizar otros aspectos, parece referirse en exclusiva, conforme al tenor literal de su propio título y enunciados, a los denominados "robots", si bien, realmente aborda la inteligencia artificial avanzada dotada de autonomía, de la que igualmente puede estar dotado un robot físico o una máquina virtual.

A continuación, analizaré los aspectos abordados en dicha Resolución en materia responsabilidad.

El Parlamento Europeo no abordó en dicha Resolución el concepto "robot" pero si define el concepto "autonomía" referida a un robot como la capacidad de tomar decisiones y aplicarlas en el mundo exterior, con independencia de todo control o influencia externos y considera que, cuanto más autónomos sean los robots, más difícil será considerarlos simples instrumentos en manos de otros agentes como el fabricante, el operador, el propietario, o el usuario.

Es decir, el concepto de autonomía real de un sistema de inteligencia artificial, conforme igualmente expuse al analizar esta definición, exigiría la ausencia de control externo, incluido el humano, lo que, de inicio, chocaría frontalmente con los principios esenciales y normas éticas y jurídicas sobre las que se pretenden construir los marcos reguladores de la inteligencia artificial y robótica en la UE, especialmente su seguridad y la supervisión y control humanos.

[751] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))

El precitado concepto plantea dos cuestiones iniciales y principales para el Parlamento Europeo: a) La naturaleza jurídica de los robots y la posibilidad de incardinarlos en alguna de las categorías existentes o si debe crearse una nueva categoría por sus características jurídicas; b) Si la normativa general sobre responsabilidad es suficiente o si se requerirían normas y principios específicos que aclaren la responsabilidad jurídica de los distintos agentes y su responsabilidad por los actos y omisiones de los robots cuya causa no pueda atribuirse a un agente humano concreto.

Respecto de la primera de las cuestiones, el Parlamento Europeo consideró, como una de las posibles opciones a valorar en el futuro, la posibilidad de crear a largo plazo una personalidad jurídica específica para los robots que permitiera que los robots autónomos más complejos puedan ser considerados personas electrónicas responsables de reparar los daños que puedan causar, así como reconocer personalidad electrónica en aquellos supuestos en los que los robots tomen decisiones autónomas inteligentes o interactúen con terceros de forma independiente.

Respecto de la segunda de las cuestiones el Parlamento Europeo consideró que los robots no pueden ser considerados responsables de los actos u omisiones que causen daños a terceros conforme al marco jurídico de responsabilidad actual.

En relación con estas cuestiones, como ya expuse al comentar estas opciones en los anteriores apartados, el Comité Económico y Social Europeo (CESE) negó igualmente cualquier tipo de personalidad jurídica para los robots o la inteligencia artificial en su Dictamen de 31 de mayo, de 2017[752].

Según el Parlamento Europeo, el marco jurídico vigente permitiría inicialmente atribuir la acción u omisión de un robot inteligente a un agente humano concreto, en particular al fabricante, al operador, al propietario o al usuario, partiendo de que dicho agente podía haber previsto y evitado el comportamiento del robot que ocasionó los daños. Asimismo, el Parlamento Europeo consideró que los fabricantes e incluso los operadores, los

[752] Dictamen del Comité Económico y Social Europeo sobre la "Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad", de 31 de mayo, de 2017

propietarios o los usuarios podrían ser considerados objetivamente responsables de los actos u omisiones de un robot.

De este modo, se aplicarían a los daños causados por los robots o la inteligencia artificial, de un lado, el régimen de responsabilidad por daños causados por productos defectuosos cuando resulte de aplicación y respondiendo el fabricante de su mal funcionamiento o defecto y, en los demás supuestos, el régimen general de responsabilidad extracontractual por daños analizado anteriormente, respondiendo el usuario de un producto ante cualquier conducta que cause un daño con origen en el mismo.

No obstante, como he expuesto, estos sistemas plantean dificultades para dar respuesta adecuada a la amplísima casuística que se puede producir en la práctica en atención a factores como la tipología de los sistemas, características, capacidades, sector, contexto, uso, finalidad o sujetos intervinientes.

Además, todo ello se complicaría cuando realmente nos pudiésemos encontrar ante un sistema de inteligencia artificial que pueda tomar decisiones verdaderamente autónomas (no meramente automatizadas o automáticas), y más si pudiera llegar a desarrollar una inteligencia artificial con conciencia o autorreflexión -lo que todavía no se ha producido-, con capacidad para calificar la justicia o injusticia, maldad o bondad de su actos, lo que constituye un cualidad humana que difícilmente podría atribuirse a un sistema de inteligencia artificial, sin perjuicio de que su programación pueda integrar parámetros éticos y morales que le permitan emitir juicios de valor en contextos concretos que determinen sus decisiones y acciones.

Esta hipótesis es lo que, en mi opinión, podría motivar una reflexión futura más profunda y debate sobre la posible atribución de una personalidad jurídica a un sistema de inteligencia artificial más avanzado o “fuerte”.

A mi juicio, la ausencia de una autonomía real y plena en la actualidad -especialmente en la medida que la inteligencia artificial opera bajo instrucciones programadas, finalidades y restricciones predefinidas y sujeción al control humano- y la ausencia de conciencia

impiden actualmente considerarla sujeto de derecho, aunque sí objeto del mismo. En la misma línea, destacar a autores como Díaz-Limón[753].

En este contexto, el Parlamento consideró en la Resolución precitada que los futuros análisis sobre responsabilidad deberán considerar algunos aspectos clave como la autonomía plena, el conocimiento, la voluntad y el control sobre la decisión o acto realizado.

De hecho, el Parlamento Europeo instó a la Comisión en esta Resolución para que presentara una propuesta de marco normativo -*hard law*- sobre los aspectos jurídicos relacionados con el desarrollo y el uso de la robótica y la inteligencia artificial, junto con instrumentos no legislativos -*soft law*- como directrices y códigos de conducta, similares a los que ya adjuntó como anexo a la precitada Resolución.

Asimismo, en esta Resolución, el Parlamento Europeo ya se remitió a la futura evaluación que efectúe la Comisión Europea para determinar si deberá aplicarse un enfoque de responsabilidad objetiva o de riesgos.

Según el mismo, la responsabilidad objetiva únicamente exigiría probar que se ha producido un daño o perjuicio y el establecimiento de un nexo causal entre el funcionamiento perjudicial del sistema inteligente y los daños o perjuicios causados a la persona que los haya sufrido.

Sin embargo, un enfoque de riesgos no se centraría en la persona que pueda actuar negligentemente como personalmente responsable, sino en la persona que tenga la capacidad de control sobre el riesgo, pudiendo en determinadas circunstancias minimizar el riesgo y gestionar el impacto negativo.

La Resolución objeto de análisis también reflejó las consideraciones del Parlamento respecto de la imputación, graduación y distribución de la responsabilidad, determinando que, una vez que las partes responsables hayan sido identificadas, "dicha responsabilidad

[753] DÍAZ-LIMÓN, J.A. (2016). "Daddy's car: La inteligencia artificial como herramienta facilitadora de derechos de autor", en *Revista La Propiedad Inmaterial*, nº 22, Universidad Externado de Colombia, 2016. DOI: http://dx.doi.org/10.18601/16571959.n22.06. P. 97.

debería ser proporcional al nivel real de las instrucciones impartidas a los sistemas inteligentes y a su grado de autonomía, de forma que cuanto mayor sea la capacidad de aprendizaje o la autonomía y cuanto más larga haya sido la 'formación' del robot, mayor debiera ser la responsabilidad de su formador".

En este sentido, la Resolución estableció que, al determinar la persona responsable de los daños o perjuicios causados por un sistema inteligente, las competencias adquiridas a través de la formación del mismo no deberían confundirse con las competencias estrictamente dependientes de su capacidad de aprender de modo autónomo.

En cualquier caso, la Resolución incorporó en su apartado 56ª la posición entonces del Parlamento Europeo respecto de la posibilidad de imputar la responsabilidad a un robot o sistema dotado de inteligencia artificial, concluyendo que, en la actualidad, la responsabilidad debe recaer en un humano, y no en un robot.

A partir de aquí, la Resolución recoge algunas posibles soluciones para la determinación de la responsabilidad en caso de daños o perjuicios por robots o sistemas de inteligencia artificial.

La primera para "robots cada vez más autónomos", la exigencia de un seguro obligatorio como el que se exige a los automóviles, que debería tener en cuenta todas las responsabilidades potenciales en la cadena.

La segunda y complementaria a la anterior, al igual que con el seguro de vehículos de motor, considerar la creación de un fondo que garantice la reparación de daños en los casos de ausencia o insuficiencia de una cobertura de seguro, para lo que Parlamento Europea pide al sector de los seguros que desarrolle nuevos productos y tipos de ofertas adaptados a la evolución de la robótica.

En este sentido, el Parlamento Europeo volvió a remitirse a la Comisión para que fuera ésta la que explorara, analizara y considerara todas las posibles soluciones jurídicas, proponiendo, entre otras:

- Establecer un régimen de seguro obligatorio -en los casos en que sea pertinente y necesario para categorías específicas de robots-, que cubra los posibles daños y

perjuicios causados por sus robots, similar al existente para los automóviles. Los obligados a su suscripción serían los fabricantes o los propietarios de robots.

- Establecer un fondo de compensación que garantice la reparación de los daños o perjuicios causados por un robot ante la ausencia del seguro.

- Permitir que el fabricante, el programador, el propietario o el usuario puedan beneficiarse de un régimen de responsabilidad limitada si contribuyen a un fondo de compensación o bien si suscriben conjuntamente un seguro que garantice la compensación de daños o perjuicios causados por un robot.

  Esta propuesta entró en contradicción con el Considerando 52 de la precitada Resolución que indicaba que cualquier instrumento legislativo que se adopte sobre aspectos jurídicos relacionados con el desarrollo y el uso de la robótica y la inteligencia artificial no deben, en modo alguno, limitar el tipo o el alcance de los daños y perjuicios que puedan ser objeto de compensación, ni tampoco limitar la naturaleza de dicha compensación, por el único motivo de que los daños y perjuicios hayan sido causados por un agente no perteneciente a la especie humana.

  Es decir, de un lado se propone la posibilidad de limitar la responsabilidad y, de otro, se indica que se debe reparar íntegramente y limitar la compensación. Una interpretación global e integradora debe llevarnos a entender que se refiere a que se pueda limitar la cobertura del daño por el agente responsable, haciéndose cargo del resto los fondos de compensación que contempla.

- Decidir si conviene crear un fondo general para todos los robots autónomos inteligentes o crear un fondo individual para cada categoría de robot, así como la elección entre un canon único al introducir el robot en el mercado o pagos periódicos durante la vida del robot.

- Crear un número de matrícula individual que figure en un registro específico de la UE que asegure la asociación entre el robot y el fondo del que depende y que permita que cualquier persona que interactúe con el robot esté al corriente de la

naturaleza del fondo, los límites de su responsabilidad en caso de daños materiales, los nombres y las funciones de los participantes y otros datos pertinentes.

- Crear una personalidad jurídica específica para los robots, de forma que como mínimo los robots autónomos más complejos puedan ser considerados personas electrónicas responsables de reparar los daños que puedan causar, y posiblemente aplicar la personalidad electrónica a aquellos supuestos en los que los robots tomen decisiones autónomas inteligentes o interactúen con terceros de forma independiente.

En cualquier caso, el Parlamento Europeo concluyó en la Resolución precitada que cualquier decisión de naturaleza política sobre las normas de responsabilidad civil aplicables a robots e inteligencia artificial debería adoptarse tras un previo proyecto de investigación y su desarrollo a escala europea especializado en robótica y neurociencia, al objeto de obtener una evaluación de los riesgos y consecuencias asociadas a nivel científico.

Tras esta Resolución, la inteligencia artificial pasó a protagonizar la estrategia y agenda política europea, especialmente a partir de la *Comunicación de la Comisión, de 25 de abril de 2018* bajo el título *inteligencia artificial para Europa*[754], a la que siguieron sendas comunicaciones posteriores, en particular, las comunicaciones de la *Comisión de 7 de diciembre de 2018* bajo el título *Plan coordinado sobre la inteligencia artificial*[755] y la *de 8 de abril de 2019,* titulada *Generar confianza en la inteligencia artificial centrada en el ser humano*[756].

Por el camino, en la Resolución del Parlamento Europeo, de 12 de septiembre de 2018**,** sobre los sistemas armamentísticos autónomos[757], la autoridad legislativa europea pidió una posición común sobre los sistemas armamentísticos autónomos letales que garanticen

[754] Comunicación de la Comisión, de 25 de abril de 2018, titulada «Inteligencia artificial para Europa» (COM(2018)0237)
[755] Comunicación de la Comisión, de 7 de diciembre de 2018, titulada «Plan coordinado sobre la inteligencia artificial» (COM(2018)0795)
[756] Comunicación de la Comisión, de 8 de abril de 2019, titulada «Generar confianza en la inteligencia artificial centrada en el ser humano» (COM (2019)0168)
[757] Resolución del Parlamento Europeo, de 12 de septiembre de 2018, sobre los sistemas armamentísticos autónomos. DO C 433 de 23.12.2019, p. 86.

un control humano significativo de las funciones esenciales de los sistemas armamentísticos, incluso durante su despliegue.

Mediante la Resolución del Parlamento Europeo, de 12 de febrero de 2019, sobre una política industrial global europea en materia de inteligencia artificial y robótica[758], el órgano europeo instó a los Estados miembros a que se centraran en la reconversión de los trabajadores de las industrias más afectadas por la automatización de las tareas, que identificaran sus riesgos y establecieran las estrategias.

Entre otros aspectos, destacó que el uso malintencionado o negligente de la inteligencia artificial podría constituir una amenaza para la seguridad digital, la seguridad física y la seguridad pública, y que la ciberseguridad era un aspecto importante de la inteligencia artificial, especialmente teniendo en cuenta los retos en materia de transparencia de sistemas alto nivel.

También destacó como la inteligencia artificial puede ser al mismo tiempo una amenaza para la ciberseguridad y una herramienta para luchar contra los ciberataques.

Asimismo, consideró que las acciones y aplicaciones de inteligencia artificial deben respetar los principios éticos y el Derecho pertinente a escala nacional, de la Unión e internacional y significó que deben establecerse normas éticas para garantizar el desarrollo de una inteligencia artificial centrada en el ser humano, la rendición de cuentas y la transparencia de los sistemas algorítmicos de toma de decisiones, así como unas normas claras en materia de responsabilidad y equidad.

Por último, pidió la creación de una carta ética de buenas prácticas para la inteligencia artificial y la robótica que deben seguir las empresas y expertos.

Posteriormente, el precitado *Grupo de expertos de alto nivel sobre inteligencia artificial* tuvo una prolífica actividad durante 2019 mediante la emisión de sendos informes el 8 de abril de 2019, uno bajo el título *Directrices éticas para una IA fiable* y, el otro, *Una definición de la inteligencia artificial: Principales capacidades y disciplinas científicas.*

[758] Resolución, de 12 de febrero de 2019, sobre una política industrial global europea en materia de inteligencia artificial y robótica. Textos Aprobados, P8_TA (2019) 0081

Posteriormente, emitió el informe titulado *Recomendaciones estratégicas y de inversión para una IA fiable* de 26 de junio de 2019.

Con posterioridad, el *Grupo de expertos sobre responsabilidad y nuevas tecnologías* emitió un informe titulado *Responsabilidad civil sobre inteligencia artificial y otras tecnologías digitales emergentes -Liability for artificial intelligence and other emerging digital technologies-*, de fecha 21.11.2019.

Asimismo, debo destacar otros estudios e informes como el *Estudio de la Dirección General de Políticas Interiores del Parlamento Europeo, de octubre de 2016, para la Comisión de Asuntos Jurídicos, sobre normas de Derecho civil europeo en materia de robótica*[759] así como el *documento informativo de STOA del Servicio de Estudios del Parlamento Europeo, de junio de 2016, sobre reflexiones éticas y jurídicas en materia de robótica*[760].

Durante 2020, la actividad legislativa y ejecutiva de los órganos europeos ha sido especialmente prolífica en esta materia. En febrero se aprobó la Resolución del Parlamento Europeo, de 12 de febrero de 2020, sobre los procesos automatizados de toma de decisiones: Garantizar la protección de los consumidores y la libre circulación de bienes y servicios[761].

Una semana después, en particular, el 19 de febrero de 2020, se presentó el *Informe de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo, de 19 de febrero de 2020, sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica*[762],

[759] https://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU(2016)571379_EN.pdf

[760] https://www.europarl.europa.eu/RegData/etudes/STUD/2016/563501/EPRS_STU(2016)563501(ANN)_EN.pdf

[761] Textos Aprobados, P9_TA(2020) 0032.

[762] Informe de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo, de 19 de febrero de 2020, sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica (COM (2020)0064)

al que le siguió el conocido como *Libro blanco sobre la inteligencia artificial*[763] *de la Comisión Europea.*

Y en el último trimestre de 2020, se aprobaron las tres resoluciones del Parlamento Europeo con recomendaciones a la Comisión, sobre el futuro marco europeo de la inteligencia artificial, en particular, la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial[764], la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas[765] y, por último, la Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial[766], lo que significa que estas materias pretenden regularse desde y para la UE, mediante un instrumento general de directa aplicación.

Asimismo, la UE había mostrado previamente su preocupación en materia de armonización de disposiciones legales, reglamentarias y administrativas en materia de protección de consumidores, usuarios, seguridad de los productos y responsabilidad por daños causados por productos defectuosos, y la necesidad de la actualización de los marcos existentes, conforme referí al abordar estos aspectos.

Durante el cierre de la primera versión de esta investigación, se publicó la Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[767], la cual se focaliza en los sistemas de inteligencia artificial de riesgo inadmisible y los de alto riesgo, no abordando directamente los aspectos de responsabilidad salvo algunas cuestiones específicas

[763] Libro Blanco de la Comisión, de 19 de febrero de 2020, sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza (COM (2020)0065)
[764] 2020/2014(INL)
[765] 2020/2012(INL)
[766] 2020/2015(INI)
[767] COM (2021) 206 final 2021/0106 (COD)

analizadas en el capítulo IV de esta investigación, siendo estos aspectos objeto de la propuesta previa que será objeto de análisis en el siguiente apartado.

8.3. Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre un régimen de responsabilidad civil en materia de inteligencia artificial.

8.3.1 Cuestiones generales

La Resolución del Parlamento Europeo, de 20 de octubre de 2020[768], estableció las recomendaciones del Parlamento destinadas a la Comisión para el establecimiento de un régimen legal de responsabilidad civil en materia de inteligencia artificial en la UE, incorporando la Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial.

Las recomendaciones que incorpora se sustentan en un sistema dual de responsabilidad, objetiva para los operadores de sistemas de inteligencia artificial de alto riesgo, y subjetiva para otros, aunque no pretende "sustituir" sino "integrar y complementar" el régimen actual de responsabilidad por productos defectuosos.

La Propuesta de Reglamento que acompaña, justifica en sus consideraciones previas la atribución de una responsabilidad objetiva al operador de sistemas de inteligencia artificial de alto riesgo en base, principalmente, a que los mismos controlan un riesgo asociado a los mismos, equiparable al del propietario de un automóvil, y que aquél suele ser el primer punto de contacto en muchos casos para la persona afectada ante la complejidad y conectividad de estos sistemas.

La responsabilidad deberá corresponder al operador, y al fabricante cuando el control del riesgo recaiga en aquél.

[768] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))

La Resolución es fruto de la insuficiencia anteriormente analizada de los marcos actuales de responsabilidad para responder a todas las situaciones que la inteligencia artificial puede plantear ante su incesante desarrollo y aplicación, y la inquietud de la UE por garantizar la armonización y necesario equilibrio entre el desarrollo y la innovación europea en inteligencia artificial y la necesaria confianza y seguridad de todos agentes relacionados con la misma, desde fabricantes a ciudadanos.

8.3.2. Objetivos

Conforme recoge en sus consideraciones, la progresiva introducción de los sistemas de inteligencia artificial en la sociedad, lugar de trabajo y economía es una de las cuestiones más importantes de la agenda política europea en la actualidad.

De hecho, durante los últimos años el legislador europeo ha evidenciado su firme convicción en que los sistemas de inteligencia artificial pueden y deben mejorar la vida humana en todas las esferas, desde la personal, a la laboral y la empresarial, así como servir para acometer los retos mundiales como la emergencia climática, la asistencia sanitaria, la nutrición o la logística.

De inicio, la Resolución refleja en sus "Considerandos" la conveniencia de asociar al concepto jurídico de inteligencia artificial al término "*toma de decisiones automatizadas*" para evitar su ambigüedad y distanciarse del atributo "*autonomía*", conforme he analizado entre otros apartados, lo que refleja un enfoque sencillo y concreto de la realidad que pretende contemplar, alejándose de conceptos más complejos e indeterminados, y apartándose inicialmente de una inteligencia artificial más avanzada o "fuerte", focalizándose en la "débil", que es la que actualmente se está desarrollando e implementando principalmente. La nueva propuesta reguladora europea de 21 abril de 2021 se aleja todavía más del concepto "autonomía".

Sin embargo, debo significar que estas consideraciones sobre sus objetivos chocan frontalmente con la Propuesta de Reglamento que se acompaña a la Resolución, en especial, en relación con lo previsto en su artículo 2 en relación con su ámbito de

aplicación, y en su artículo 3 continente de las definiciones de "sistema de inteligencia artificial" y "autónomo", que posteriormente abordaré con mayor profundidad.

De manera sintética, en primer lugar, respecto del ámbito de aplicación, el artículo precitado establece la aplicación del Reglamento propuesto en los casos en que una actividad física o virtual, un dispositivo o un proceso gobernado por un sistema de inteligencia artificial haya causado daños. Es decir, está haciendo referencia a una inteligencia artificial que "gobierna" no que gestione, aplique o constituya el medio o instrumento, de modo que parece estar refiriéndose a una inteligencia artificial más avanzada que la considerada "débil". Pero, además, en segundo lugar, relacionando su ámbito de aplicación con las definiciones precitadas contenidas en el artículo 3, parece que está confirmando y haciendo referencia a una inteligencia artificial más avanzada que la "débil" , lo que podría llevar a la conclusión inicial de que la propuesta no resultaría de aplicación a otros sistemas distintos, en particular, a lo que hoy es la realidad extendida de la inteligencia artificial y que es la que se está desarrollando y aplicando en la actualidad, es decir, una inteligencia artificial débil. Estas consideraciones han sido igualmente analizadas y significadas por algunos autores como Atienza Navarro[769]. No era ni es éste el propósito del Parlamento Europeo.

Como luego expondré, el Reglamento propuesto define "sistema de inteligencia artificial" como todo sistema basado en programas informáticos o incorporado en dispositivos físicos que muestra un comportamiento que simula la inteligencia, entre otras cosas, mediante la recopilación y el tratamiento de datos, el análisis y la interpretación de su entorno y la actuación, con cierto grado de autonomía, para lograr objetivos específicos. Es decir, incorpora la autonomía como rasgo inherente a la misma.

Sin embargo, cuando la define, lo hace considerando que se considerará autónomo a los efectos de esta propuesta "todo sistema de inteligencia artificial que funciona interpretando determinados datos de entrada y utilizando un conjunto de instrucciones predeterminadas, sin limitarse a ellas, a pesar de que el comportamiento del sistema esté

[769] ATIENZA NAVARRO, Mª. L. (2020). "Una aproximación a la Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial de 27 de abril de 2020". *Revista Aranzadi de derecho y nuevas tecnologías*. ISSN 1696-0351, Nº. 54. 2020. Pp. 27-47.

limitado y orientado a cumplir el objetivo que se le haya asignado y otras decisiones pertinentes de diseño tomadas por su desarrollador". Es decir, un sistema que tiene la capacidad de actuar no limitada a las instrucciones predeterminadas, lo que situaría al mismo más próximo a una inteligencia más avanzada y superior a la "débil". De este modo, cabría considerar de manera consecuente, que la propuesta regulatoria no tendría como objeto la inteligencia artificial actual (débil) y su aplicación se limitaría a una más avanzada y próxima a la "fuerte".

No es éste el objetivo de la propuesta ni de su articulado, sino todo lo contrario, como he expuesto y significan autores como la precitada Atienza Navarro, en la medida que pretende regular un marco de responsabilidad civil, cualquiera que sea el sistema de inteligencia artificial involucrado, objetivo para sistemas de alto riesgo y subjetivo para el resto. En consecuencia, todo ello requeriría una revisión de las definiciones incorporadas en el artículo 3 del Reglamento propuesto para su adecuación y congruencia con el espíritu, finalidad y contenido de la norma.

La Resolución objeto de análisis recoge igualmente los retos jurídicos que plantean los sistemas de inteligencia artificial para el marco actual de responsabilidad civil.

En primer lugar, destaca especialmente el riesgo jurídico de su opacidad, en la medida que podría comportar enormes dificultades e incluso la imposibilidad de identificar a la persona o entidad que tiene el control sobre el riesgo asociado al sistema o el algoritmo, entrada o datos que provocan el funcionamiento lesivo, lo que de manera consecuente podría conllevar la dificultad de identificar la relación entre el daño o perjuicio y el comportamiento que lo causó, lo que en la práctica, a su vez, podría tener como resultado que la persona afectada no fuera resarcida, quebrantando así el principio y finalidad esencial sobre el que se sustenta el derecho de daños.

En segundo lugar, también significa otros retos jurídicos de los sistemas de inteligencia artificial, como el creciente aumento de la autonomía de estos sistemas que usan técnicas de aprendizaje autónomo y profundo, su vulnerabilidad ante ciberataques, su dependencia de datos externos o su asociación y conexión con otros sistemas, estén o no dotados de inteligencia artificial.

Y, en tercer lugar, destaca también el reto social que plantean los sistemas dotados de inteligencia artificial, como es su aceptación por sus usuarios y la ciudadanía en general.

El Parlamento Europeo consideró en esta Resolución que, para poder abordar todos estos retos se debe disponer de normas éticas sólidas y procedimientos de indemnización consistentes y justos, de manera que una persona que sufra un daño o perjuicio causado por un sistema dotado de inteligencia artificial tenga garantizado, en primer lugar, el mismo nivel de protección que tendría de no intervenir un sistema de este tipo, en segundo lugar, la cobertura por un seguro adecuado y, en tercer lugar, un procedimiento claramente definido para obtener el resarcimiento.

El Parlamento Europeo consideró igualmente que estos marcos, éticos y jurídicos, contribuirán a la necesaria seguridad jurídica para todas las partes implicadas, condición que se considera esencial, tanto para la innovación y desarrollo de sistemas basados en inteligencia artificial por parte de *startups*, microempresas y PYMEs, como para su aplicación y uso, creando un ecosistema en equilibrio que garantice dicha seguridad.

Respecto del instrumento legislativo propuesto, el Parlamento Europeo ha venido igualmente evidenciando la necesidad de crear un marco jurídico uniforme y horizontal en materia de inteligencia artificial para, de un lado, aprovechar de manera eficaz todas sus ventajas y, de otro, impedir usos inadecuados y evitar la fragmentación normativa en el seno de UE, conforme refleja la Propuesta de Reglamento que integra la Resolución en su Considerando 2º.

En este sentido, el Parlamento estimó necesario disponer de una legislación uniforme basada en principios comunes, horizontal y preparada para el futuro para garantizar la seguridad jurídica, la uniformidad en la UE y proteger de manera eficaz los valores europeos y derechos de los ciudadanos. Ello no debe ser un obstáculo para el desarrollo de una reglamentación específica a nivel sectorial para sus distintas aplicaciones que, el legislador europeo, considera preferible, especialmente para poder acometer sus especificidades ante su diversidad.

Además de ello, el Parlamento Europeo motivó en sus consideraciones el establecimiento de un marco común y armonizado como una necesidad para el marco único digital, para

mantener la soberanía digital e impulsar la innovación digital en Europa, especialmente ante los flujos de datos internacionales.

En definitiva, la Resolución propuesta pretende proponer un marco jurídico en materia de responsabilidad civil eficaz y orientado al futuro. Y éste sería su objetivo inmediato, dado que el mediato sería proporcionar la confianza en la seguridad, fiabilidad y coherencia en productos, servicios y tecnologías asociadas con la finalidad de alcanzar un equilibrio entre la protección -eficaz y equitativa- de las potenciales víctimas de daños o perjuicios derivados de su uso, y la inversión, innovación y desarrollo de nuevas tecnologías, productos y servicios, es decir, proporcionar seguridad jurídica a todas las partes implicadas, desde el fabricante, el operador, la persona afectada o cualquier tercero.

En este contexto, el Parlamento Europeo consideró necesario crear este marco para proteger adecuadamente a la sociedad frente a los daños o perjuicios que puedan provocar los sistemas dotados de inteligencia artificial, en la medida que, como he analizado, los sistemas actuales no dan una adecuada cobertura a los distintos supuestos que plantea su incesante desarrollo, despliegue y aplicación.

El Parlamento Europeo apostó por la inclusión del resarcimiento de los daños inmateriales, considerando que deben tenerse en cuenta tanto el daño material como el daño inmaterial, sobre la base, entre otros documentos, de su *Informe de 19 de febrero de 2020 sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica*, y se instó a la Comisión Europea a que analizara en profundidad las tradiciones jurídicas de todos los Estados miembros, así como las disposiciones legislativas vigentes que concedan una indemnización por daños inmateriales, a fin de evaluar si la inclusión de los daños inmateriales en la futura Propuesta de Reglamento es jurídicamente sólida y necesaria desde el punto de vista de la persona afectada, en la medida que el Parlamento Europeo considera que, conforme a la información actualmente disponible, debe incluirse el daño inmaterial significativo si la persona afectada sufre una pérdida económica apreciable, es decir, una pérdida económica comprobable.

Asimismo, otro aspecto especialmente relevante de la Resolución objeto de análisis, es la posibilidad de crear un *sandbox* regulatorio, y ello ante la imposibilidad de poder ofrecer

una solución única adecuada para todo el conjunto de riesgos de distinta naturaleza, la diversidad de sistemas dotados de inteligencia artificial y sus diversas aplicaciones. Este instrumento también se haya contemplado en el marco de la nueva Propuesta de Reglamento regulador de la inteligencia artificial de 21 de abril de 2021.

La Resolución apostó por la creación de estos espacios limitados de experimentación o pruebas para testear iniciativas innovadoras bajo supervisión estatal, dotados de la conveniente flexibilidad y necesaria certidumbre regulatoria en un espacio seguro y de máxima garantía para los usuarios, que permita elaborar soluciones adecuadas y testadas para situaciones y sectores específicos.

Por último, enlazando con el posicionamiento estratégico de la UE en materia de inteligencia artificial al que hacía referencia en el capítulo I de esta investigación, el Parlamento Europeo se reafirmó en esta Resolución en su convicción de "que la carrera mundial de la IA ya está en marcha y que la Unión debe asumir el liderazgo en ella explotando su potencial científico y tecnológico".

Cierto todo ello como expuse, pero es una carrera ya iniciada hace algún tiempo y que está siendo liderada en la actualidad, de un lado y a nivel de naciones, por las grandes potencias como China y EEUU y, de otro y a nivel empresarial, por las grandes tecnológicas, en la que la UE debe posicionarse, consciente de su situación y potencial científico y tecnológico actual y futuro, en base a objetivos estratégicos alcanzables y realizables y congruente con sus valores y rasgos diferenciadores.

8.3.3. Personalidad jurídica de los sistemas de inteligencia artificial

El Parlamento Europeo abordó de nuevo y consideró en su Resolución que no es necesario atribuir personalidad jurídica a los sistemas de inteligencia artificial, aspecto desechado en el seno de la UE en la actualidad como se ha abordado en anteriores apartados de esta investigación, y ello al considerar que los daños o perjuicios causados por dichos sistemas son el resultado de las acciones previas de las personas que los construyeron, desplegaron o interfirieron en ellos.

8.3.4. Sujetos responsables

La pluralidad de agentes involucrados en el desarrollo, despliegue, uso o aplicación de los sistemas de inteligencia artificial y la opacidad, la conectividad y la autonomía de los mismos podrían dificultar o incluso imposibilitar en la práctica la trazabilidad de acciones específicas de los sistemas de inteligencia artificial causantes de un daño o perjuicio hasta una intervención humana específica o decisiones de diseño.

Todo ello supone un reto y un obstáculo que el Parlamento Europeo propuso resolver aplicando teorías y conceptos de responsabilidad civil ampliamente aceptadas, haciendo responsables a las diferentes personas de toda la "cadena de valor", o quizás más acertadamente en mi opinión, de todo el "ciclo de vida del sistema" -de modo que no sólo se pueda considerar el ciclo desde su creación hasta su distribución, sino su uso posterior, mantenimiento y monitorización-, y que puedan tener el control sobre el riesgo asociado al mismo en cada momento.

8.3.5. Sistemas de responsabilidad preexistentes

El Parlamento Europeo, como he anticipado anteriormente, consideró en su Resolución que no era necesaria una revisión completa de los regímenes de responsabilidad civil que funcionen adecuadamente, ya sean de la UE o nacionales, si bien, ante los nuevos retos que plantean los sistemas de inteligencia artificial para su eficacia -opacidad, complejidad, vulnerabilidad, multitud de agentes, etc.-, consideró necesaria una adaptación coordinada de los preexistentes para evitar que una persona afectada por un daño o perjuicio no sea resarcida, es decir, que se garantice el derecho de resarcimiento efectivo de las personas que sufran un daño.

De este modo, el Parlamento Europeo apostó firmemente en su Resolución por conservar los sistemas actuales, considerando que tanto el futuro Reglamento como la Directiva vigente sobre responsabilidad por los daños causados por productos defectuosos constituirán los dos pilares esenciales de un marco común de responsabilidad civil para

los sistemas de inteligencia artificial que, en cualquier caso, requerirán la necesaria y estrecha coordinación entre todos los agentes políticos de la UE y sus Estados miembros.

De nuevo, el Parlamento vuelve a poner en valor la eficacia de la Directiva sobre responsabilidad por los daños causados por productos defectuosos durante los últimos treinta (30) años, como ya lo hizo la Comisión Europea en su *Informe al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo sobre la aplicación de esta* Directiva, de fecha 07.05.2018[770], y del mismo modo, considera y reitera que debe ser revisada para adaptarla al mundo digital y abordar los nuevos retos que plantea para seguir garantizando protección y seguridad jurídica para empresas y usuarios.

No obstante, el Parlamento fue más allá e instó expresamente a la Comisión para que evalúe si esta Directiva debería transformarse en un reglamento, así como a aclarar el concepto de "producto", si debemos considerar incluido en su ámbito de aplicación los contenidos o servicios digitales y adaptar otros conceptos como "daño", "defecto" y "productor", que ya eran cuestiones a revisar significadas por la Comisión en el informe citado en los apartados precedentes y cuya necesidad de actualización se ha evidenciado en el análisis de estos aspectos que he llevado a cabo a lo largo de esta investigación.

De otro modo, se perdería la oportunidad de solucionar un contexto de inseguridad jurídica, especialmente en aspectos particulares como la consideración de "productor" a efectos de esta Propuesta de Reglamento, en la medida que el mismo se remite a la definición de este concepto recogida en la Directiva, con lo que todo ello comporta a efectos de imputación de la responsabilidad regulada en la Propuesta de Reglamento objeto de análisis.

Según las consideraciones del Parlamento Europeo en esta nueva Resolución, el concepto de "productor" debe incluir a fabricantes, desarrolladores, programadores, prestadores de servicios y operadores finales.

[770] COM/2018/246 final. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52018DC0246

En sus consideraciones, el Parlamento Europeo determinó igualmente que cualquier actualización del marco de responsabilidad establecido por la Directiva precitada debe acompañarse de la actualización de la Directiva 2001/95/CE del Parlamento Europeo y del Consejo, de 3 de diciembre de 2001, relativa a la seguridad general de los productos[771], a fin de garantizar que los sistemas de inteligencia artificial incorporen los principios de seguridad y protección desde el diseño, esto es, la "*Security by design*", en congruencia con su exigencia ética y también jurídica en la Propuesta de Reglamento incorporada en la coetánea Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas.

Del mismo modo, el Parlamento Europeo pidió en sus consideraciones que la Comisión considere la posibilidad de revertir las normas que rigen la carga de la prueba para los daños ocasionados por las tecnologías digitales emergentes en casos claramente definidos y tras una adecuada evaluación, si bien, no especificó a que tecnologías digitales emergentes se refiere, en la medida que, como he referido en anteriores apartados, la inteligencia artificial, entre otras, no es precisamente una tecnología emergente aunque sí de despliegue y aplicación exponencial en la actualidad, en todo tipo de sectores y esferas de la vida personal, administrativa y empresarial.

El Parlamento Europeo también reflejó la importancia de legislar de manera "estratégica" de modo que se permita la inclusión de futuros avances tecnológicos, lo que me parece un acierto ante el objeto de los futuros marcos, si bien, destacando entre los mismos la apuesta por el *software* libre o programas de código abierto, entiendo que por la transparencia e independencia que aporta, entre otros aspectos. Asimismo, considero que el Parlamento Europeo se refiere al despliegue y aplicación de esta modalidad de programación en lo sucesivo más que un avance futuro, en la medida que ya empieza a ser una realidad cada vez más extendida en la actualidad.

Las consideraciones del legislador europeo que preceden al articulado de la Propuesta de Reglamento aclaran que norma utilizar ante la coexistencia de la Directiva sobre responsabilidad por los daños causados por productos defectuosos con el Reglamento

[771] DO L 011 de 15.1.2002, p. 4.

propuesto en materia de responsabilidad civil derivada de sistemas de inteligencia artificial.

En este sentido, el Parlamento determina que la Directiva precitada deberá seguir utilizándose en relación con las reclamaciones por responsabilidad civil contra el productor de un sistema de inteligencia artificial defectuoso, cuando el sistema de inteligencia artificial cumpla los requisitos para ser considerado un producto con arreglo a dicha Directiva, resultando pues de preferente aplicación en tales supuestos, por lo que el Reglamento se aplicaría en el resto de situaciones, en particular, cuando sea considerado un servicio o se dirija al operador.

Esto exigirá reflexionar sobre la necesidad de revisar también la TRLGDCU española que regula el régimen de responsabilidad establecido en España como resultado de la transposición de la Directiva citada europea, y que contempla no sólo la responsabilidad derivada de productos sino de servicios.

De inicio, de aprobarse el Reglamento propuesto y de regular específicamente la aplicación de su régimen de responsabilidad a los sistemas de inteligencia artificial, dada la naturaleza especial de la norma y su rango superior, considero que resolvería cualquier conflicto de leyes que pudiera cuestionarse que, en mi opinión, tampoco se produciría en atención a la especialidad de la norma y la generalidad del TRLGDCU que no contempla específicamente la inteligencia artificial dentro de sus preceptos.

En este sentido, como he manifestado a analizar estas cuestiones en anteriores apartados, la responsabilidad contemplada tanto en dicha Directiva como en la TRLGDCU se focaliza inicialmente en la figura del consumidor, por lo que la condición amplia de perjudicado le permitiría utilizar el régimen de responsabilidad recogido en el Reglamento propuesto, objetiva o subjetiva en función del contexto, para exigir la responsabilidad por los daños sufridos, sin perjuicio de que la situación particular y su valoración le pudiera permitir utilizar igualmente el régimen contemplado en aquéllas en la medida que le beneficie y el amparo de la jurisprudencia a la que aludí en su análisis.

No obstante, considero que esta recomendación del Parlamento Europeo debería revisarse en el futuro, en la medida que la *Directiva sobre responsabilidad por los daños causados*

por productos defectuosos debería revisarse, actualizarse y adecuarse previamente, conforme analicé en los apartados precedentes.

Por último, el Parlamento Europeo consideró que los marcos reguladores vigentes de los Estados miembros de responsabilidad subjetiva ofrecen, en la mayoría de los casos, un nivel de protección suficiente a las personas que sufran daños o perjuicios causados por un tercero que pueda interferir en los sistemas como un "pirata informático" o a las personas que sufren menoscabo al patrimonio por ese tercero, considerando que esta interferencia, por regla general, constituye una acción de responsabilidad civil subjetiva.

Como analizaré en posteriores apartados, estas conductas no sólo pueden ser constitutivas de responsabilidad civil subjetiva, sino que, en atención a la legislación nacional, especialmente en España, podrían ser constitutivas de delito, en función del contexto y elementos concurrentes.

No obstante, como igualmente referiré al abordar esta cuestión más adelante, debido a las características inherentes del ciberespacio en el que vivimos y nos interrelacionamos, existe una enorme dificultad de depurar las distintas responsabilidades derivadas de actos de intrusión, entre otras, el anonimato, la dificultad de identificación y localización, ubicuidad, perseguibilidad o cooperación judicial, lo que en la práctica puede conllevar la dificultad práctica de depuración de las responsabilidades concurrentes y no alcanzar el objetivo principal del derecho de daños, cualquiera que su origen, que no es otro que el resarcimiento efectivo de la víctima.

La solución que plantea la Propuesta de Reglamento sobre estos aspectos, muy discutible a mi juicio, como abordaré con mayor detalle al analizar el mismo, es complementar las legislaciones nacionales imputando la responsabilidad al operador cuando el tercero no sea rastreable o sea insolvente, lo que sin duda debería ser un incentivo adicional para mejorar la seguridad por parte de la industria de la inteligencia artificial.

8.4. Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial.

La Resolución del Parlamento Europeo incorpora una Propuesta de Reglamento sobre responsabilidad civil derivada del funcionamiento de los sistemas de inteligencia artificial.

8.4.1. Objeto y alcance

El objeto de la Propuesta de Reglamento es establecer un marco normativo en relación con las reclamaciones de responsabilidad civil de las personas físicas y jurídicas, cualquiera que sea su condición -salvo que sea el "operador", como luego expondré-, contra los operadores de sistemas dotados de inteligencia artificial, conforme así lo recoge expresamente su artículo 1.

De inicio, ello comporta superar las limitaciones y dificultades asociadas a la Directiva europea de productos defectuosos y a la normativa de transposición española.

8.4.2. Conceptos jurídicos

La Propuesta de Reglamento define jurídicamente distintos conceptos que, como he referido en anteriores apartados, no se hallaban consensuados hasta la fecha.

a) Sistema de inteligencia artificial

En primer lugar y aunque ya fue objeto de análisis en el capítulo I de esta investigación, se define "sistema de inteligencia de artificial" como aquel sistema basado en programas informáticos o incorporado en dispositivos físicos que muestra un comportamiento que simula la inteligencia, entre otras cosas, mediante la recopilación y el tratamiento de datos, el análisis y la interpretación de su entorno y la actuación, con cierto grado de autonomía, para lograr objetivos específicos.

De nuevo, esta definición se focaliza en sistemas que simulan la inteligencia, no que se hallen dotados de la misma, lo que significa el apartamiento definitivo y concreción del concepto para evitar equívocos sobre la realidad que representa y pretende regular.

El Reglamento propuesto pretende regular exclusivamente la inteligencia artificial y, en especial la "débil", si bien, la relación de la definición de "sistema de inteligencia artificial" con la de "autonomía" parece redirigirse siguiendo su tenor literal hacia sistemas más avanzados que pueden operar al margen o más allá de instrucciones predeterminadas, lo que podría llevar a considerar que su aplicación se circunscribiría a una inteligencia más avanzada que la débil y más próxima a la fuerte. Como he referido no es el objetivo de la propuesta normativa, pero debe motivar, a mi juicio, la revisión de las definiciones para adecuarlas y dotarlas de mayor precisión y claridad.

b) Autonomía

Del mismo modo, define el concepto de autonomía asociado a los sistemas de inteligencia artificial, en particular, como todo sistema de inteligencia artificial que funciona interpretando determinados datos de entrada y utilizando un conjunto de instrucciones predeterminadas, sin limitarse a ellas, a pesar de que el comportamiento del sistema esté limitado y orientado a cumplir el objetivo que se le haya asignado y otras decisiones pertinentes de diseño tomadas por su desarrollador.

Es decir, conforme se deduce de su tenor literal, la autonomía incorpora la automatización y su limitación y orientación al cumplimiento del objetivo asignado y decisiones incorporadas en su diseño, pero adiciona como factor diferencial la no limitación del sistema a las instrucciones predeterminadas para cumplir el objetivo o adoptar las decisiones adecuadas conforme a su diseño, lo que le confiere un cierto margen de libertad para seguir instrucciones no predeterminadas en su diseño, aunque limitadas por el objetivo definido y decisiones permitidas en su concepción, es decir una libertad y cierta autonomía restringida y sometida a control desde su concepción -inicialmente humano, aunque podría serlo por parte de otro sistema inteligente conforme a la definición-, es decir, no se trata de una autonomía plena sino restringida en su diseño, lo que supone la exigencia jurídica del control humano, como principio y norma ética esencial.

De nuevo, supone una concreción necesaria del término, que se aparta de un concepto de autonomía plena o total sino sujeta a controles y decisiones adoptadas en su diseño, lo que, a mi juicio, facilita la identificación de la mejor posición para el control de determinados riesgos y consiguiente depuración de responsabilidades por parte de los distintos agentes intervinientes en el ciclo de vida del sistema inteligente.

No obstante, incorpora en la definición la posibilidad de que se limite en su funcionamiento a las instrucciones predeterminadas, lo que, como he referido, deberá ser revisado de cara a su futura regulación, en la medida que, tal y como ha sido redactado, parece referirse a una inteligencia artificial más avanzada que la "débil" -la que actualmente está siendo desarrollada y aplicada de forma masiva-, y más próxima a la "fuerte", conforme he expuesto y analizado anteriormente.

Como he expuesto en su análisis, la Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[772], define un sistema de inteligencia artificial de manera muy distinta, alejándose del concepto "autonomía" y focalizándose en el *software* desarrollado bajo técnicas específicas capaz de generar resultados como predicciones, recomendaciones o decisiones con repercusión en el entorno con el que interacciona, partiendo de un conjunto de objetivos definidos previamente por el ser humano. Es decir, apuesta por una definición estrictamente funcional.

c) Sistemas de inteligencia artificial de alto riesgo

La Propuesta de Reglamento también define en su artículo 3.c) que sistemas de inteligencia artificial deben considerarse de "alto riesgo" a los efectos de esta propuesta, en base al potencial significativo de los mismos para causar daños o perjuicios a una o más personas, físicas o jurídicas, alejándose del concepto de alto riesgo que encontramos, entre otras, en la posterior Propuesta de Reglamento sobre inteligencia artificial de 21 de abril de 2021, conforme analicé en el capítulo anterior de esta investigación.

[772] COM (2021) 206 final 2021/0106 (COD)

En particular, lo define como el potencial significativo en un sistema de inteligencia artificial que funciona de forma autónoma para causar daños o perjuicios a una o más personas "de manera aleatoria y que excede lo que cabe esperar razonablemente".

En este sentido, dicha calificación parece hacerla exclusiva de sistemas que funcionan de forma autónoma y no "con cierta autonomía", en los términos en los que se define este atributo en la Propuesta de Reglamento, si bien, su categorización se haya relacionada con distintos factores, el sector donde opere y la naturaleza de las actividades realizadas.

En consecuencia, interpretando su tenor literal, no todo sistema de inteligencia artificial que, conforme a su definición, pueda tener cierto grado de autonomía, deberá calificarse como tal, especialmente si la autonomía es muy limitada y restringida, especialmente si opera en sectores y actividades de riesgo medio o bajo, en cuyo caso, podría quedar fuera de dicha consideración conforme a una interpretación literal del concepto y en base a las aclaraciones posteriores que el propio precepto analizado incorpora. Sin embargo, conforme a la definición de autonomía, podrían funcionar al margen de las instrucciones predefinidas, con el riesgo potencial que ello comportaría de ser éste el propósito del legislador al definirla del modo analizado anteriormente.

Asimismo, la definición plantea algunos conceptos inicialmente indeterminados como, por ejemplo, que debemos entender por "potencial significativo" y de una forma "que exceda lo razonablemente esperable".

Las consideraciones del Parlamento Europeo que preceden al texto normativo propuesto recogen el criterio del mismo para considerar cuando un sistema de inteligencia artificial debe considerarse como "de alto riesgo", considerando que, debe calificarse como tal, cuando su funcionamiento autónomo conlleva un potencial significativo de causar daño a una o más personas, de forma aleatoria y yendo más allá de lo que cabe esperar razonablemente. Y para calificarlo como tal, también deberá tenerse en cuenta el sector en el quepa esperar que surjan riesgos importantes y la naturaleza de las actividades realizadas.

Sobre lo primero, el propio artículo 3 precitado da las claves para llegar a un posicionamiento en base a criterios de gestión de riesgos, esto es, dicho potencial

"significativo" y su magnitud dependerán de la relación entre la gravedad del posible daño o perjuicio (impacto), del grado de autonomía de la toma de decisiones (autonomía), de la probabilidad de que el riesgo se materialice (probabilidad) y del modo y contexto (contexto) en que se utilice el sistema de inteligencia artificial.

No obstante, todo ello dependerá pues del análisis de riesgos o evaluación de impacto específica que se lleve a cabo del sistema, lo que a su vez dependerá de quién y en base a qué procedimientos realiza el mismo, dado que pueden utilizarse metodologías y parámetros muy distintos de evaluación, especialmente de la probabilidad, el contexto e el impacto, en los que además deberá considerarse la deseable objetividad e independencia de la persona o entidad que efectúa dicha evaluación, de modo que pueda ser cuestionable, especialmente en caso de análisis y evaluaciones internas en el seno de una empresa.

Durante mi trayectoria profesional he dirigido y participado en múltiples análisis de riesgos y evaluaciones de impacto de sistemas de inteligencia artificial, especialmente en materia de conformidad regulatoria y protección de datos, y he podido comprobar que, aun aplicando criterios y procedimientos basados en estándares internacionales de referencia, marcos normativos, procedimientos corporativos internos uniformes y las mejores prácticas a nivel internacional, no siempre los resultados de la evaluación de la probabilidad e impacto son exactamente los mismos, ni tan siquiera en todos los miembros de un mismo equipo sin perjuicio de que posteriormente se debatan y consensuen, especialmente por el inevitable cariz del criterio profesional aplicado, basado en la propia cualificación y experiencia de cada persona, lo que obliga a aplicar mecanismos de consenso en un equipo.

En este sentido, la creación de estándares de evaluación, los análisis y evaluaciones por parte de terceros, la predeterminación de las bases y criterios para llevar a cabo las mismas y su calificación contribuirían notablemente a evitar posibles heterogeneidades o discrepancias significativas.

Y sobre lo segundo, se queda en un concepto jurídico indeterminado. ¿Qué debemos entender por "lo razonablemente esperable"? Estamos ante una tecnología no nueva pero de desarrollo y aplicación disruptiva reciente y exponencialmente creciente -

especialmente por el aumento incesante de las capacidades de computación y disponibilidad de datos-, y en todo tipo de sectores y ámbitos de nuestra vida, con ausencia de marcos éticos uniformes, consensuados en gran medida a nivel internacional pero no vinculantes de inicio, así como con ausencia de marcos jurídicos específicos, por lo que empezamos a tener una cierta experiencia, pero quizás no la suficiente madurez para sentar unas bases sobre lo que debería entenderse como "razonablemente esperable".

Es más, ¿quiere ello decir que debemos esperar como algo innato e inevitable que los sistemas de inteligencia artificial de alto riesgo causen daños o perjuicios a las personas? ¿Debemos esperar el daño y, consecuentemente, generar un temor razonable frente a estos sistemas? Conforme expuse anteriormente, el Considerando 2 del Reglamento propuesto considera los riesgos como inherentes a cualquier nueva tecnología.

La introducción de esta expresión en estos términos no me parece lo más adecuado desde un punto de vista de certidumbre y seguridad jurídica, dado que nos lleva la indefinición para la calificación de un sistema de alto riesgo, con lo que ello significa conforme al texto actual del Reglamento propuesto, lo que de inicio exigiría acudir a la interpretación y valoración por terceros, ya sea autoridades y, en su caso, por jueces y tribunales, en los procedimientos que puedan derivarse en relación con los mismos.

Además, supone volver a un cierto temor frente a la inteligencia artificial que se consideraba ya olvidado, puede chocar con uno de los objetivos de estos marcos como es la seguridad en la misma, y que puede comportar un freno para la necesaria confianza por parte de la ciudadanía y operadores y, en consecuencia, afectar a su despliegue, explotación, competitividad e innovación.

Pero lo cierto, como he comentado anteriormente al abordar otros aspectos durante esta investigación, aspectos como el error y la falibilidad de un sistema, incluso no del mismo sino simplemente de su conectividad, debe ser esperada, al igual que los daños consecuentes, en su caso.

No obstante, el Parlamento Europeo recomienda en sus consideraciones previas al texto normativo propuesto que los sistemas de inteligencia artificial de alto riesgo se enumeren de forma exhaustiva en un anexo del Reglamento propuesto, si bien, ante el rápido

desarrollo tecnológico y los conocimientos técnicos necesarios, conforme recoge la propuesta, correspondería a la Comisión Europea revisar dicho anexo sin demoras injustificadas y, a más tardar, cada seis meses y, en caso de resultar necesario, modificarlo mediante un acto delegado.

La calificación automática en base a este criterio objetivo contrarrestaría la subjetividad interpretativa de aspectos como el comentado relativo a "lo razonablemente esperable".

No obstante, esta definición considero que deberá ser revisada, en la medida que la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021,[773], define de manera taxativa en su artículo 6 los sistemas de inteligencia artificial de alto riesgo alejándose de la definición contenida en esta propuesta objeto de análisis.

d) Sujetos intervinientes

La Propuesta de Reglamento, en particular, su artículo 3.I., define jurídicamente los distintos sujetos involucrados en el marco de la responsabilidad civil por los daños causados por sistemas de inteligencia artificial.

d.1) Persona afectada

En primer lugar, define a la *persona afectada*, como aquella que sufre daños o perjuicios por una actividad física o virtual, un dispositivo o un proceso gobernado por un sistema de inteligencia artificial, y que no sea su operador, cualquiera que sea su condición.

De inicio, la Propuesta de Reglamento deja fuera del concepto de "persona afectada" a su operador en la medida que se podría producir una supuesta identidad entre el causante y el perjudicado, si bien, no necesariamente puede disociarse ambas condiciones como luego analizaré más adelante, en la medida que el operador también podría ser víctima de daños causados por el sistema derivadas de una responsabilidad compartida con otros

[773] COM (2021) 206 final 2021/0106 (COD)

sujetos que también tendrían un cierto grado de control sobre el riesgo asociado, cuya responsabilidad sería exigible incluso, inicialmente, conforme a la Directiva sobre productos defectuosos.

En el caso de que un sistema de inteligencia artificial pudiera causar daños o perjuicios a su operador, incluso en base a su supuesta "autonomía" limitada, no encajaría dentro del concepto de *persona afectada* según la Propuesta de Reglamento, sin perjuicio de que pudiese llegar a ser víctima efectiva del mismo.

Pensemos a modo de ejemplo, el sistema de inteligencia artificial diseñado por una compañía para la gestión de la ciberseguridad en sus sistemas y redes complejas propiedad y/o gestionadas por la misma. En caso de que el sistema de inteligencia artificial causara daños o perjuicios por sus acciones u omisiones a toda la red y a todos sus usuarios, el operador no sería considerado persona afectada.

d.2) Operador

En segundo lugar, la Propuesta de Reglamento define quién debe ser considerado *operador* y, en particular, operador inicial y final.

En particular, conforme al mismo, se consideraría *operador inicial* a toda persona física o jurídica que defina, de forma continuada, las características de la tecnología, y que proporcione datos y un servicio de apoyo final de base esencial y, por tanto, que ejerce también grado de control sobre un riesgo asociado a la operación y el funcionamiento del sistema de inteligencia artificial.

Por otra parte, se consideraría *operador final* a toda persona física o jurídica que ejerza un grado de control sobre un riesgo asociado a la operación y el funcionamiento del sistema de inteligencia artificial y se beneficie de su funcionamiento.

La Propuesta de Reglamento considera operador a los efectos del mismo, tanto al final como al inicial, siempre que la responsabilidad civil de este último no esté ya cubierta

por la Directiva 85/374/CEE[774], es decir, que sea responsable en calidad de productor/fabricante en los supuestos en los que el contexto pueda incardinarse dentro de este marco de responsabilidad.

Con ello, la propuesta pretende constituirse en un sistema coordinado y complementario al existente en materia de responsabilidad derivada de productos defectuosos, al objeto de dar cobertura y respuesta a los distintos supuestos que podrían plantearse en la práctica.

En relación con estas definiciones, el concepto de "control" es determinante para la calificación del sujeto y consecuente responsabilidad asociada.

En este sentido, la Propuesta de Reglamento define ese "control" como "toda acción de un operador que influya en el funcionamiento de un sistema de inteligencia artificial y, por consiguiente, la medida en que el operador expone a terceros a los potenciales riesgos asociados a la operación y funcionamiento del sistema de inteligencia artificial".

En consecuencia, la acción del operador que influya en el funcionamiento del sistema comporta un grado de exposición a los terceros a los riesgos asociados a su operación y funcionamiento, por lo que deberá responder de los mismos en dicha medida.

Es decir, la responsabilidad imputable al operador se asocia al grado de riesgo por la operación y funcionamiento de los sistemas al que se ven expuestas las personas por parte de las acciones de aquél, desde la definición hasta su puesta en operación y funcionamiento.

Las acciones influyentes del operador en los sistemas de inteligencia artificial pueden afectar a su funcionamiento en cualquier fase, y pueden afectar tanto a la entrada y salida de los datos, definición de características, instrucciones, procesos, procesamiento, resultados, así como cambiar las funciones o procesos específicos dentro de propio sistema.

[774] Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos.

Conforme recoge su definición, el grado en que esta afectación de su funcionamiento esté causado por la acción del operador -impacto o determinación- depende del grado de influencia que éste tenga sobre el riesgo relacionado con dicho funcionamiento, es decir, si el control del riesgo depende del operador y se produce una afectación del funcionamiento, la responsabilidad derivada de ello corresponderá a éste sin perjuicio de que corresponsabilidad de otros operadores o responsabilidad concurrente con otros terceros.

De este modo, la Propuesta de Reglamento considera que en las situaciones en que exista más de un operador, lo que ocurrirá frecuentemente, todos ellos deben ser responsables civiles solidarios, esto es, responsabilidad solidaria, aunque dispongan posteriormente del derecho a reclamar en vía de regreso unos de otros de manera proporcional.

Como analicé en los apartados anteriores, ésta constituía una de las posibles soluciones proteccionistas para la depuración de responsabilidades civiles en materia de daños derivados de la inteligencia artificial, ante la pluralidad de sujetos intervinientes y su complejidad.

Los porcentajes de responsabilidad vendrán determinados por los respectivos niveles de control que tengan los operadores sobre el riesgo relacionado con la operación y el funcionamiento del sistema de inteligencia artificial, para lo que la propuesta destaca en sus consideraciones previas, la necesidad de mejorar la trazabilidad de los productos con la finalidad de identificar mejor a los que intervienen en las distintas fases.

No obstante, esta definición más compleja es muy diferente al concepto "operador" que recoge la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021,[775] objeto de análisis en el capítulo anterior. Según la misma, se considerará *operador* a los efectos del Reglamento propuesto, el proveedor, el usuario, el representante autorizado, el importador y el distribuidor, conceptos que a su vez define en su artículo 3.

[775] COM (2021) 206 final 2021/0106 (COD)

d.3) Productor

En tercer lugar, también aborda el concepto "productor", remitiéndose a la precitada Directiva 85/374/CEE de productos defectuosos[776].

La Directiva precitada considera "productor" a la persona que fabrica un producto acabado, que produce una materia prima o que fabrica una parte integrante, y toda aquella persona que se presente como productor poniendo su nombre, marca o cualquier otro signo distintivo en el producto.

Conforme a dicha definición, se consideraría "productor" al fabricante del sistema de inteligencia artificial, si bien, esta definición plantearía distintas cuestiones en la práctica ante las distintas tipologías de sistemas de inteligencia artificial y la naturaleza intangible del *software*, como principal elemento de un sistema de inteligencia artificial, el cual puede estar compuesto por distintos elementos, como *hardware, software*, algoritmos, datos y otros elementos de distinta naturaleza.

En consecuencia, a mi juicio, deberemos interpretar este concepto de manera amplia, considerando productor tanto a la empresa que reúne todos los elementos físicos y digitales para su creación, como al propio desarrollador o programador en el caso de sistemas de inteligencia artificial basados en *software* y algoritmos que pueden operar desde cualquier máquina o servidor gestionado por el mismo o terceros.

Asimismo, debo destacar que la Directiva precitada considera igualmente "productor" al importador, en particular, a toda persona que importe un producto en la UE con vistas a su venta, alquiler, arrendamiento financiero o cualquier otra forma de distribución en el marco de su actividad comercial a los efectos dicha Directiva y, en consecuencia, también a los efectos de esta Propuesta de Reglamento objeto de análisis, al que además la Directiva le imputa la misma responsabilidad que al productor.

[776] Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados miembros en materia de responsabilidad por los daños causados por productos defectuosos

En este sentido, hay que tener en cuenta lo dispuesto en el Anexo III de la Decisión del Consejo y de la Comisión, 13 diciembre 1993, relativa a la celebración del Acuerdo sobre el Espacio Económico Europeo entre las Comunidades Europeas y sus Estados miembros, por una parte, y la República de Austria, la República de Finlandia, la República de Islandia, el Principado de Liechtenstein, el Reino de Noruega, el Reino de Suecia y la Confederación Suiza, por otra parte (DOCEL 3 enero 1994), el cual establece del mismo y a los efectos del Acuerdo que:

> "i) Sin perjuicio de la responsabilidad del productor, cualquier persona que importe un producto al EEE con vistas a su venta, alquiler, arrendamiento financiero o cualquier otra forma de distribución en el marco de su actividad comercial será considerada como productor del mismo y tendrá la misma responsabilidad que el productor.
>
> ii) La anterior disposición se aplicará también con respecto a las importaciones de un Estado de la AELC a la Comunidad, o de la Comunidad a un Estado de la AELC, o de un Estado de la AELC a otro Estado de la AELC. A partir de la fecha de entrada en vigor para cualquier Estado de las CE o de la AELC del Convenio de Lugano relativo a la competencia judicial y a la ejecución de resoluciones judiciales en materia civil y mercantil, de 16 de septiembre de 1988, la primera frase del presente inciso no será de aplicación entre los Estados que hayan ratificado dicho Convenio en la medida en que, en virtud de dicha ratificación, una resolución judicial en favor de la persona perjudicada sea ejecutable contra el productor o el importador, según éstos se definen en el inciso i).
>
> iii) Suiza y Liechtenstein podrán dispensarse mutuamente de la responsabilidad del importador".

En este sentido, el Parlamento Europeo considera y matiza a quién hace referencia el concepto de "productor", que debe incluir a fabricantes, desarrolladores, programadores, prestadores de servicios y operadores finales, conforme recoge en sus considerandos iniciales (Considerando 8) de la Propuesta de Reglamento.

En relación con todo ello, el Parlamento Europeo insta de nuevo a la Comisión Europea, entre otras cosas, a que revise y adapte el concepto de "productor" regulado en la Directiva precitada, conforme se ha analizado anteriormente, y que ya se incluía en el *Informe de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo sobre la aplicación de esta Directiva, de fecha 07.05.2018*[777].

Por último, dicha Directiva establece otra particularidad adicional a considerar, relativa a que, si el productor no pudiera ser identificado, cada suministrador del producto será considerado como su productor, salvo que informe al perjudicado de la identidad del productor o de la persona que le suministró el producto. Y lo mismo establece en el caso de los productos importados, si en éstos no estuviera indicado el nombre del importador, aun indicándose el nombre del productor.

La Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021 analizada en el capítulo anterior, se focaliza principalmente en la figura del *proveedor*, atribuyendo dicha condición a la persona física o jurídica, autoridad pública, agencia u otro organismo que desarrolla un sistema de inteligencia artificial o que tiene un sistema de inteligencia artificial desarrollado con el fin de comercializarlo o ponerlo en servicio con su propio nombre o marca comercial, ya sea de pago o de forma gratuita. Asimismo, aborda y regula con minuciosidad no sólo los requisitos y obligaciones de éste en dicha condición sino las obligaciones del fabricante de determinados productos cuando los sistemas de inteligencia artificial de alto riesgo se integren en los mismos y se comercialicen o pongan en servicio conjuntamente con dichos productos y bajo el nombre del fabricante.

Conforme a esta nueva Propuesta de Reglamento, en estos supuestos y exclusivamente en éstos, el fabricante del producto asumirá la responsabilidad de la conformidad del sistema de inteligencia artificial con el Reglamento propuesto y, en lo que respecta al sistema de inteligencia artificial que integre cualquiera de sus productos, tendrá las mismas obligaciones impuestas por el Reglamento propuesto al proveedor.

[777] COM/2018/246 final. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52018DC0246

d.5) Daño o perjuicio

La Propuesta de Reglamento no diferencia entre ambos conceptos, considerando como tal cualquier impacto adverso que afecte a la vida, la salud, la integridad física de una persona física, los bienes de una persona física o jurídica o bien que produce daños morales significativos que resultan en una pérdida económica comprobable.

La definición relaciona los bienes jurídicos protegidos que incluyen la vida, la salud, la integridad física o bienes de personas físicas o jurídicas, pero de nuevo, no se recogen expresamente los derechos, especialmente los fundamentales, dentro de dicha relación.

No obstante, se incluyen expresamente los daños morales que comporten una pérdida económica comprobable y que, como expuse en su análisis, el sistema de responsabilidad por daños derivados de productos defectuosos vigente en España, incluye parcialmente, al excluir expresamente los daños morales no derivados de la muerte.

8.4.3. Ámbito de aplicación

La Propuesta de Reglamento regula en su artículo 2 su ámbito de aplicación.

Conforme establece el citado precepto, la futura norma se aplicaría en el territorio de la UE y a los casos en los que una actividad física o virtual, un dispositivo o un proceso gobernado por un sistema de inteligencia artificial cause daños o perjuicios a la vida, la salud, la integridad física de una persona física, los bienes de una persona física o jurídica, o daños morales "considerables" que den lugar a una pérdida económica comprobable.

En primer lugar, el precepto identifica la acción y al agente causante del daño y lo define como la actividad física o virtual (lógica), dispositivo o proceso gobernado por un sistema de inteligencia artificial, esto es, un agente sin personalidad jurídica que "gobierna" todo ello, conforme refería al analizar la personalidad jurídica de los sistemas inteligentes avanzados.

De su tenor literal se desprende que se incluyen tanto actividades llevadas a cabo por sistemas "gobernados" por sistemas inteligencia artificial en el mundo físico como en el mundo virtual o digital.

Ello difiere del concepto de "robótica" incorporada por el Parlamento Europeo en la Propuesta de Reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, que incorpora la Resolución del Parlamento Europeo, de 20 de octubre de 2020,[778]en particular en su artículo 4.I.c), en el que, como puse de relieve al efectuar su análisis, se echaban en falta las acciones "lógicas" de los sistemas robóticos en el mundo digital o virtual, y que deberían haber sido expresamente incluidas a mi juicio, ya se trate de sistemas de esta naturaleza integrados en dispositivos físicos dotados de movilidad como en sistemas robóticos basados exclusivamente en *software*.

Esta inclusión pone de relieve la irrelevancia de la dimensión donde se produzca la acción causante del daño a efectos de imputación de la responsabilidad.

Del mismo modo, se incluyen como agentes causantes del año los dispositivos o los procesos gobernados por sistemas de inteligencia artificial.

Quizás, en puridad y para una mayor claridad al introducir conceptos jurídicos indeterminados relacionados con la tecnología, como por ejemplo "gobierno", considero que hubiera sido oportuno reflexionar previamente sobre la opción de hacer referencia al agente causante en general, como "sistemas de información dotados de inteligencia artificial" o meramente "sistemas de inteligencia artificial" en sí mismos considerados, que incluirían *hardware, software*, algoritmos, procesos y datos, entre otros elementos, así como las acciones, procesos, dispositivos o productos dotados de estos sistemas, sin necesidad de matizar y aludir específicamente a que se traten de acciones, dispositivos o

[778] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))

procesos gobernados por la misma, dado que se incluirían también otros aspectos como decisiones o inacciones determinadas por la misma.

De otro modo, se requiere un requisito esencial, a mi juicio absolutamente coherente, para la aplicación del futuro marco, esto es, el gobierno de la actividad, dispositivo o proceso por un sistema de inteligencia artificial que haya causado daños.

El mero hecho de integrar un sistema de inteligencia artificial no debe conllevar la aplicación de este marco de responsabilidad específica. A modo de ejemplo, un coche autónomo podría provocar daños, pero no tener su origen en el funcionamiento del sistema de inteligencia artificial del que estuviere dotado, sino en otros agentes como el pasajero o usuario.

No obstante, del análisis de este primer aspecto me surgen, entre otras, múltiples cuestiones sobre las que reflexionar: ¿Qué debemos entender por "gobernado por un sistema de inteligencia artificial"? ¿Debemos entender que una actividad, dispositivo o proceso se entiende gobernado por un sistema de inteligencia artificial por el mero hecho de que los mismos integren dicho sistema con independencia del grado de "autonomía" que posea? ¿O ese "cierto grado de autonomía" se asocia a aquél en que el sistema decide y dirige plenamente la acción, dispositivo o proceso? ¿Y si solamente la determina, pero no la ejecuta, siendo llevada a cabo por un tercero? ¿Y al margen del ser humano y sin su participación?

Desde mi punto de vista, sin entrar en mayores análisis terminológicos, "gobernar" es dirigir y "gestionar" es realizar las tareas de administración, organización y funcionamiento encomendadas.

La inteligencia artificial debería estar sometida en todo momento al control y supervisión humana, así como a controles y medidas de seguridad predefinidas de modo que su autonomía nunca debería ser plena sino en el mejor de los casos restringida, por lo que sus capacidades y prerrogativas deberían estar más relacionadas con la "gestión" que con el "gobierno" de acciones, dispositivos o procesos, sin perjuicio de que en determinados contextos pueda tener este atributo.

Se trata de una cuestión interpretativa que quizás podría ya resolverse optando por hacer referencia a actividades, dispositivos o procesos "gestionados mediante sistemas de inteligencia artificial". De este modo, considero que la utilización del concepto "gestión" sería más adecuado que "gobierno", si se pretende hacer referencia a la administración, organización y funcionamiento de estas actividades, dispositivos o procesos por estos sistemas inteligentes dado que, en otro caso, estaríamos hablando de sistemas de inteligencia más avanzada o incluso la denominada "fuerte", que verdaderamente estuvieran en condiciones de ejercer su "gobierno". No debe confundirse gobernar con gestionar.

Ello además sería congruente con el concepto de sistema de inteligencia artificial que recoge la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, que omite cualquier referencia a su autonomía.

Por otra parte, la definición de autonomía asociada al concepto de inteligencia artificial que incorpora el texto propuesto, como comenté, parece acotar la misma a sistemas inteligentes más avanzados y con una inteligencia artificial superior a la considerada "débil", al tener la capacidad de funcionar sin limitarse a las instrucciones predeterminadas, lo que se alejaría del objeto de la propuesta, que no es otro que regular los marcos de responsabilidad civil de los daños causados por sistemas inteligentes cualquiera que sea su grado de autonomía, de la tipología de inteligencia de la que se hallen dotados, sea débil o fuerte, o de su nivel de riesgo, si bien esto último será determinante para la aplicación de los distintos regímenes previstos en la misma.

Una vez analizado este primer aspecto, en segundo lugar, como exponía en la introducción del capítulo II, el precepto recoge algunos de los bienes jurídico protegidos, pero no todos.

Las consideraciones previas del Parlamento Europeo que preceden al texto propuesto hacen referencia a los "derechos importantes jurídicamente protegidos", como el derecho a la vida, la salud, la integridad física o la propiedad.

Considero que hubiera sido muy oportuno incluir los derechos de las personas, personalísimos o no, incluidos los fundamentales, sin perjuicio de valorar si también deberían incluirse otros como la libre competencia, la innovación y la competitividad empresarial por parte de empresas que pudieran verse afectadas por determinados sistemas dotados de inteligencia artificial, a escala nacional o internacional.

En tercer lugar, la regulación de su ámbito de aplicación limita la autonomía de la voluntad contractual, dado que se considera nulo cualquier acuerdo celebrado entre un operador de un sistema de inteligencia artificial y una persona física o jurídica que sufra un daño o perjuicio como consecuencia de un sistema de inteligencia artificial que eluda o limite los derechos y obligaciones establecidos en la Propuesta de Reglamento, con independencia de que se hubiere celebrado antes o después de haberse causado el daño o perjuicio.

En consecuencia, la Propuesta de Reglamento, conforme regula su artículo 2.2, considera los derechos y obligaciones regulados en el mismo como irrenunciables y de carácter imperativo, no dispositivo, previendo las consecuencias derivadas de su incumplimiento en particular, la sanción de nulidad del acuerdo entre el operador y la persona afectada.

En cuarto lugar, la Propuesta de Reglamento recoge también en sus consideraciones preliminares que las normas reguladoras de la responsabilidad civil que afecten al operador deberían cubrir todas las operaciones de los sistemas de inteligencia artificial, "independientemente de dónde se lleve a cabo la operación y de que ésta sea física o virtual", sin embargo, el texto normativo propuesto circunscribe el ámbito de aplicación al territorio de la UE, pero sin regular expresamente su aplicación a las acciones o procesos que desarrollen en el seno de la UE pero puedan afectar a personas ubicadas fuera de la misma, o a acciones o procesos realizados fuera de la UE pero con afectación a personas ubicadas en la misma.

En este sentido, sería deseable una mayor concreción del Reglamento propuesto respecto de su aplicación en supuestos de responsabilidad con elementos transnacionales para aportar seguridad jurídica y evitar la interpretación.

Por último, en quinto lugar, la Propuesta de Reglamento contempla su compatibilidad y complemento del resto de sistemas de responsabilidad, regulando un sistema de responsabilidad por daños para su resarcimiento compatible con cualquier otra demanda en materia de responsabilidad civil contractual entre el operador (del sistema de inteligencia artificial) y la persona (física o jurídica) que haya sufrido un daño o perjuicio a causa del sistema de inteligencia artificial, esto es, derivada de las relaciones contractuales, así como de la normativa sobre responsabilidad por daños causados por productos defectuosos, la protección de los consumidores, la lucha contra la discriminación o la protección laboral o del medio ambiente, que se pueda presentar con el operador conforme al Derecho de la Unión o nacional.

Sin duda, un mecanismo absolutamente proteccionista y garante para la persona que sufra un daño o perjuicio que podrá disponer, en función del contexto, de distintas acciones frente al operador, conforme permite el actual marco de responsabilidad en España.

La propuesta justifica además en sus considerandos la necesidad de establecer marcos normativos de responsabilidad distintos para sistemas y riesgos asociados diferentes, lo que de nuevo constituye un argumento en relación con mi posicionamiento respecto de la insuficiencia de los marcos vigentes de responsabilidad a nivel español y de la UE para resolver los distintos supuestos de responsabilidad que pueden plantearse en relación con la inteligencia artificial, especialmente ante la distinta tipología de sistemas, capacidades, características y riesgos asociados.

En definitiva, la inteligencia artificial plantea importantes retos para los regímenes de responsabilidad civil existentes como he comentado anteriormente y analizado con profundidad, y en este sentido la propuesta considera razonable y apuesta por establecer un régimen común de responsabilidad objetiva para los sistemas de inteligencia artificial considerados de alto riesgo, bajo un enfoque basado en los riesgos y en sus distintos niveles, bajo criterios y conceptos jurídicos definidos, especialmente “alto riesgo”, con el objetivo de ofrecer seguridad jurídica.

De este modo, en congruencia con las posibles soluciones que analizaba para superar la falta de adecuación y carencias que suponen los regímenes generales y, en especial, el régimen de responsabilidad por productos defectuosos regulado en el ordenamiento

jurídico español, la Propuesta de Reglamento aborda la responsabilidad de estos sistemas específicos desde un enfoque basado en el riesgo, considerando que un sistema de inteligencia artificial que conlleve un alto riesgo inherente y que actúe de manera autónoma, potencialmente pone en peligro en mucha mayor medida al público en general. En consecuencia, considera que debe establecerse un régimen común de responsabilidad objetiva para este de sistemas de alto riesgo, basados en el mismo.

8.4.4. Responsabilidad objetiva

La Propuesta de Reglamento establece y regula un sistema de responsabilidad objetiva para los operadores de "sistemas de inteligencia artificial de alto riesgo", estableciendo un sistema de responsabilidad subjetiva para operadores de sistemas no calificables como tal, lo que comporta un marco de menor protección para las personas afectadas por estos últimos.

La propuesta ha sido bien acogida por parte de la doctrina, como Platero Alcón[779], si bien, otros autores consideran la misma mejorable en distintos aspectos, entre otros, Ortiz Fernández[780].

La Propuesta de Reglamento establece en su artículo 4.1 que los operadores de sistemas de inteligencia artificial de alto riesgo -sea o no el fabricante o productor del mismo- serán objetivamente responsables de cualquier daño o perjuicio causado por una actividad física o virtual, un dispositivo o un proceso gobernado por dicho sistema de inteligencia artificial.

De nuevo, se hace referencia al concepto jurídico indeterminado de "gobierno", por lo que me remito a las consideraciones realizadas anteriormente al respecto.

---

[779] PLATERO ALCÓN, A. (2021). "Breves notas sobre el régimen de responsabilidad civil derivado de los sistemas de inteligencia artificial: especial referencia al algoritmo de recomendaciones de Netflix". En *Ius et Scientia*. Vol. 7. Nº 1. Universidad de Sevilla. 2021. P. 141.

[780] ORTIZ FERNÁNDEZ, M. (2020). "Reflexiones acerca de la responsabilidad civil derivada del uso de la inteligencia artificial: los 'principios' de la Unión Europea". *Revista de Direito da ULP*. Vol. 14. Nº. 1. 2020. P. 60.

En primer lugar, como he comentado con anterioridad, la Propuesta de Reglamento justifica la atribución de una responsabilidad objetiva al operador de este tipo de sistemas de inteligencia artificial en base, principalmente, a que los mismos controlan un riesgo asociado a los mismos equiparable al del propietario de un automóvil (Considerando 10), así como a la mayor facilidad de identificación de éste en muchos casos para la persona afectada, ante la complejidad y conectividad de estos sistemas.

Conforme he comentado igualmente, la propuesta considera que un sistema de inteligencia artificial presenta un alto riesgo cuando su funcionamiento autónomo conlleva un potencial significativo de causar daño a una o más personas, de forma aleatoria y yendo más allá de lo que cabe esperar razonablemente. Sobre esto último, me remito a mis consideraciones anteriores y a la nueva concepción que ha introducido sobre este tipo de sistemas la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial *-Artificial Intelligence Act-*, de 21 de abril de 2021, que exigirá la necesaria revisión y armonización de los criterios.

Asimismo, la propuesta establece en sus consideraciones que, para determinar si un sistema de inteligencia artificial es de alto riesgo, debe igualmente considerarse el sector en el que cabe esperar que aparezcan riesgos importantes y la naturaleza de las actividades realizadas. Y ese potencial dañino dependerá de la relación entre la gravedad del posible daño, la probabilidad de que el riesgo cause un daño o un perjuicio y el modo en que se utilice el sistema.

En este sentido, el Reglamento propuesto apuesta en el artículo 4.2. por el establecimiento de una relación inicial de sistemas de inteligencia artificial considerados de alto riesgo y de los sectores críticos en los que se utilizan, que se incorporaría al mismo como anexo y que deberá ser revisada, actualizada y modificada por la Comisión Europea. Otra cuestión distinta a considerar será la agilidad con la que se efectúe dicha revisión y actualización por parte de ésta.

La inclusión de un sistema de inteligencia artificial en esta relación comportará su calificación automática como de alto riesgo “per se” y, en consecuencia, la aplicación del

sistema de responsabilidad objetiva previsto en el texto propuesto por los daños que pueda causar el mismo.

En mi opinión, es un acierto este enfoque a nivel de técnica legislativa, al objeto de permitir la adaptación del marco jurídico al contexto, evolución tecnológica y necesidades en cada momento, si bien, como he manifestado al analizar el resto de propuestas europeas, la calificación de alto riesgo no sólo debería derivarse de su tipificación como tal por la Comisión, sino que desde el momento en que sea calificado como tal tras una evaluación de riesgos en atención a sector, ámbito y contexto donde opere y niveles de probabilidad e impacto, debería ser considerado como tal y hallarse sujeto a los requisitos y obligaciones establecidos para los mismos en los futuros marcos reguladores.

Con este propósito, el artículo 13 de la Propuesta de Reglamento, de un lado, otorga los poderes necesarios a la Comisión para adoptar actos delegados para dicha actualización y, de otro, delimita los mismos temporalmente por un período de cinco (5) años a partir de la fecha de aplicación del futuro Reglamento. Además, estos poderes serán revocables en cualquier momento por el Parlamento Europeo o por el Consejo, sin afectación de los actos delegados que ya estén en vigor.

Tal y como ha sido concebida la propuesta, como he indicado, se trataría de un *numerus clausus*, inicial y revisable, de sistemas y sectores a los que resultaría de aplicación este régimen de responsabilidad, conforme aclaran las consideraciones del Parlamento Europeo previas al texto normativo propuesto.

En consecuencia, los sistemas de inteligencia artificial que ocasionen un daño o perjuicio que no estén incluidos en el anexo precitado del futuro Reglamento quedarán sujetos a los marcos de responsabilidad que les resulten de aplicación, incluyendo el de responsabilidad subjetiva previsto en el mismo, aunque, conforme destaca el Parlamento Europeo en sus consideraciones previas, la persona afectada podrá acogerse a una presunción de culpa del operador en este marco de responsabilidad, quien debe poder quedar eximido de la misma demostrando que observó el deber de diligencia.

No obstante, considero que este aspecto debe ser objeto de mayor reflexión, en particular si dicha relación anexa debiera ser considerada enunciativa y no limitativa, dado que bajo

el espíritu y finalidad proteccionista de la norma, el nuevo régimen de responsabilidad debería ser exigible a cualquier sistema de inteligencia artificial de alto riesgo que encaje en la definición dada a este concepto por el Reglamento propuesto, con independencia de que se halle incluido o no a efectos de publicidad, transparencia y seguridad jurídica en el listado inicial anexo al mismo, y sin perjuicio de su posterior inclusión formal por la Comisión a través de un acto delegado, es decir, que dicho listado no sea un *numerus clausus* aunque revisable, conforme al espíritu y finalidad de la norma.

En este sentido me remito a las consideraciones efectuadas en el capítulo III en relación con la Propuesta de Reglamento del Parlamento Europeo y del Consejo, sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas.

La seguridad jurídica para algunas de las partes interesadas se garantiza de mejor manera a través de esta técnica legislativa, pero perjudica a la persona afectada que puede ver desamparada por este nuevo marco jurídico y régimen de responsabilidad específico, hasta su expresa y formal inclusión por la Comisión, con independencia de su encaje efectivo en la definición de "alto riesgo" y de lo que determinen de manera irrefutable los análisis de riesgos realizados que confirmen dicha clasificación.

De este modo, se estaría excluyendo la aplicación del futuro Reglamento a estos sistemas o sectores no incluidos en la enumeración inicial incorporada al mismo, hasta su inclusión posterior en la lista, mediante acto delegado de la Comisión para su modificación, que no entraría en vigor hasta seis meses después de su adopción.

No obstante, en relación con esta cuestión, el Parlamento Europeo incorpora en sus consideraciones previas al texto propuesto que, los sistemas que no hayan sido evaluados por la Comisión y que, en consecuencia, no hayan sido clasificados de alto riesgo ni incluidos en la lista anexa al futuro Reglamento, deben quedar sujetos excepcionalmente a una responsabilidad objetiva si han causado incidentes reiterados que den lugar a un daño o un perjuicio grave. Esta consideración, sin embargo, no tiene su reflejo en el texto propuesto, por lo que debería revisarse el mismo.

El Parlamento Europeo indica que, en estos casos la Comisión debe evaluar, sin demora indebida, la necesidad de revisar dicho anexo para añadir el sistema de inteligencia artificial en cuestión a la lista anexa, si bien, y ésta es una de las cuestiones más importantes que aporta la propuesta a mi juicio, el Parlamento Europeo manifiesta en sus consideraciones que, si tras esa evaluación, la Comisión decide incluir dicho sistema de inteligencia artificial en la lista, dicha inclusión debe tener efectos retroactivos a partir del momento del primer incidente probado y causado por dicho sistema que hubiere causado un daño o perjuicio "grave".

Sin embargo, el texto normativo propuesto, de nuevo, no incorpora ninguna previsión en este sentido que refleje las consideraciones efectuadas, por lo que en su posterior tramitación debería reflexionarse sobre la conveniencia de incorporar esta previsión, sin perjuicio de las consideraciones que pudiésemos efectuar sobre la retroactividad de actos legislativos con carácter obligacional, punitivo o sancionador.

Siguiendo con el análisis de esta propuesta, con el objetivo de abordar estas cuestiones, el Parlamento Europeo considera que la Comisión debe colaborar con un comité permanente a crear, similar, como ejemplifica en sus consideraciones previas, al Comité Permanente sobre Precursores o al Comité Técnico sobre Vehículos de Motor, que incluya a expertos y a las partes interesadas, considerando que la composición equilibrada del *Grupo de expertos de alto nivel sobre la inteligencia artificial* podría servir como ejemplo para la formación del grupo de partes interesadas, adicionando expertos en ética, antropólogos, sociólogos y especialistas en salud mental, así como a expertos consultivos para asesoramiento y soporte permanente.

Como expuse al analizar el mismo, la Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, mantiene la fórmula de generar un listado de sistemas de inteligencia artificial de alto riesgo "per se", a modo de *numerus clausus*, pero a revisar y actualizar de manera continua por la Comisión Europea mediante actos delegados.

En conclusión, dejando a un lado estas cuestiones previas, el nuevo régimen que establece el Reglamento propuesto *objetiviza* de manera absoluta la responsabilidad del operador de sistemas de inteligencia artificial de alto riesgo, de modo que éste no podrá eludir su

responsabilidad civil alegando que actuó con la diligencia debida o que el daño o perjuicio fue causado por una actividad, un dispositivo o un proceso autónomo gobernados por su sistema de inteligencia artificial.

De un lado, desaparece el elemento de culpabilidad para la imputación de la responsabilidad, respondiendo directa y objetivamente el operador de cualquier daño o perjuicio causado por el sistema de inteligencia artificial, siendo irrelevante la diligencia con la que pudiere haber actuado. La responsabilidad se deriva para el operador con independencia de la concurrencia de culpa por su parte.

Y, de otro, la Propuesta de Reglamento se aparta definitivamente de la denominada "personalidad electrónica" que podría permitir imputar responsabilidad al propio sistema, aunque debilitando el régimen de responsabilidad y su objetivo final de resarcimiento efectivo a la persona afectada.

No obstante, conforme establece la Propuesta de Reglamento en su artículo 3, el operador no será responsable si el daño o perjuicio ha sido provocado por un caso de fuerza mayor, concepto que no define y que deberá ser integrado por el resto del ordenamiento jurídico común, conforme analicé en anteriores apartados. Tampoco en el caso de responsabilidad exclusiva de la víctima.

La responsabilidad objetiva del operador, conforme recoge la propuesta en sus consideraciones previas, se justifica por el hecho de que éste controla un riesgo asociado al sistema de inteligencia artificial, comparable, como he referido anteriormente, al del propietario de un automóvil, siendo además la parte más visible para la persona afectada ante la complejidad y conectividad de un sistema de inteligencia artificial.

Del mismo modo, este régimen de responsabilidad objetiva del operador se ha concebido con el objetivo de cubrir todas las operaciones de los sistemas de inteligencia artificial, independientemente de dónde se lleva a cabo la operación y de que está sea física o virtual, tal y como las consideraciones preliminares del Reglamento propuesto recogen expresamente y ha sido expuesto anteriormente.

No obstante, en dichas consideraciones se destaca que las operaciones llevadas a cabo por estos sistemas en espacios públicos y que exponen a muchas personas a un riesgo, constituyen supuestos que requieren una consideración mucho profunda que, sin embargo, no tiene su posterior reflejo en su articulado. Un ejemplo podrían ser los sistemas de reconocimiento facial en espacios públicos que puedan permitir el acceso o no a determinados servicios o instalaciones.

Quizás, en estos y otros supuestos, deberíamos hablar directamente de sistemas de alto riesgo por el volumen de personas expuestas al riesgo y más si se hallan implicados datos personales de carácter especial. No obstante, el texto propuesto deberá integrarse por la Comisión Europea con ese listado de sistemas de inteligencia artificial considerados de alto riesgo por sectores y la posible inclusión de los precitados.

En estos supuestos, se pueden dar situaciones en las que las personas afectadas puedan constituirse en víctimas potenciales de daños o perjuicios sin ni tan siquiera ser conscientes de las operaciones llevadas a cabo y de tener mermados o no disponer de sus derechos para reclamar contra el operador, entre otras vías, a través de la responsabilidad contractual o extracontractual, lo que exigiría ejercer sus pretensiones por responsabilidad subjetiva, con las correlativas dificultades para demostrar la culpa del operador y obtener un resarcimiento efectivo.

Mis consideraciones y preocupaciones en este sentido se han visto refrendadas por lo previsto en la Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, que prohíbe en su artículo 5 algunos de estos sistemas, por ejemplo, los sistemas de identificación biométrica remota en tiempo real en espacios de acceso público con el propósito de hacer cumplir la ley, salvo en supuestos excepcionados, o que los regula como sistemas de inteligencia artificial de alto riesgo en su artículo 6.

Por último, ante el ritmo de desarrollo de sistemas y tecnologías basadas en la inteligencia artificial, el crecimiento exponencial de sus aplicaciones y la necesaria protección de los usuarios, la Propuesta de Reglamento objeto de análisis destaca la necesidad de adoptar un enfoque acelerado y, en mi opinión, dinámico, adaptativo y no estático, que permita analizar de este modo y con agilidad los potenciales riesgos de los nuevos sistemas

dotados de inteligencia artificial que surjan en el mercado, y que simplifique todos los procedimientos relacionados.

Con este propósito, propone que la evaluación por parte de la Comisión para determinar si estamos ante un sistema de inteligencia artificial de alto riesgo, se inicie al mismo tiempo que la evaluación de la seguridad de los productos, de modo que se evite que un sistema de inteligencia artificial de alto riesgo esté ya aprobado para su comercialización pero no se halle clasificado como del alto riesgo, y pueda ser puesto en funcionamiento sin las coberturas y garantías previstas en la Propuesta de Reglamento, en especial, el seguro obligatorio. Estos aspectos deberán ser objeto de armonización con la Propuesta de Reglamento regulador de la inteligencia artificial de 21 de abril de 2021.

Otra de las cuestiones que considero especialmente relevantes en materia de responsabilidad objetiva, es que el Parlamento Europeo considera que la Comisión deberá evaluar cómo la actividad investigadora podría acceder a los datos recogidos, registrados o almacenados en los sistemas de inteligencia artificial de alto riesgo y hacer uso de ellos para recabar evidencias probatorias en caso de daño o perjuicio causado por dicho sistema, y como podría mejorarse la trazabilidad y auditabilidad de dichos datos teniendo en cuenta los derechos fundamentales y el derecho a la intimidad.

A mi juicio, la exigencia de su explicabilidad, transparencia, información, *accountability* y responsabilidad recogida tanto en la Propuesta de Reglamento del Parlamento Europeo y del Consejo de 2020, sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, de 20 de octubre de 2020, como en la posterior Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[781], en caso de aprobación, vinculantes, son las que proporcionarán una solución a la cuestión y determinarán su exigencia jurídica, con las consiguientes consecuencias sancionadoras derivadas de su incumplimiento.

[781] COM (2021) 206 final 2021/0106 (COD)

8.4.5. Seguro obligatorio de responsabilidad civil para sistemas de alto riesgo

Una de las cuestiones más relevantes y significativas de la Propuesta de Reglamento es la exigencia de un seguro obligatorio de responsabilidad civil tanto al operador final como inicial de un sistema de inteligencia artificial de alto riesgo, conforme establece en su artículo 4.4., que ya se recogía en la Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica, analizada en anteriores apartados.

Se trataría de un seguro de responsabilidad civil obligatorio que garantice las operaciones y servicios de estos sistemas dentro del marco de una adecuada gestión de riesgos orientada al resarcimiento efectivo del perjudicado, mediante el traslado del riesgo a un tercero, esto es, la aseguradora, en caso de que el riesgo se materialice, y ello con independencia de que hubiera podido ser o no evitado por quién tenía que gestionarlo y era responsable del mismo y de su control.

El Parlamento Europeo motiva esta medida en garantizar la confiabilidad del público en general en estos sistemas ante la necesidad de explotar y potenciar las ventajas de los sistemas de inteligencia artificial y gestionar adecuadamente los riesgos asociados a los mismos.

Los argumentos del Parlamento Europeo sobre esta exigencia se recogen en sus consideraciones previas al texto propuesto, y con alusión a la Directiva 2009/103/CE del Parlamento Europeo y del Consejo, de 16 de septiembre de 2009, relativa al seguro de la responsabilidad civil que resulta de la circulación de vehículos automóviles, y con el objetivo de controlar la obligación de asegurar esta responsabilidad[782].

De este modo, se considera necesario que todos los operadores de sistemas de inteligencia artificial de alto riesgo relacionados en el anexo del futuro Reglamento deben ser titulares de un seguro de responsabilidad civil que cubra los importes y con el alcance de la indemnización establecidos en la Propuesta de Reglamento.

[782] DO L 263 de 7.10.2009

Asimismo, el Parlamento Europeo considera que la inexistencia de estos seguros puede ser un obstáculo para la investigación y la innovación y destaca la dificultad del sector de los seguros para elaborar productos nuevos o adaptados para los sistemas de inteligencia artificial, especialmente ante la incertidumbre relacionada con los riesgos y la incertidumbre sobre su futura evolución.

Es por ello que propone, de un lado, intervenir normativamente para no dejar el desarrollo de este tipo de seguros en manos del mercado, lo que podría comportar la creación de seguros uniformes con primas desproporcionadas y elevadas, lo que colocaría al operador ante la disyuntiva de optar por la opción más barata en lugar de optar por la de mayor calidad y cobertura, y lo que a su vez y al final repercutiría en la investigación y en la innovación. Y de otro, abrir un diálogo entre la Comisión Europea y el sector de los seguros, para colaborar, estudiar los modelos y crear pólizas de seguros que proporcionen una cobertura adecuada a un precio asequible.

Por lo que se refiere al operador final, se requiere un seguro de responsabilidad civil como garantía de los riesgos asociados al funcionamiento y operación de dichos sistemas, mientras que, respecto del operador inicial, se requiere como seguro de responsabilidad empresarial (en su caso) o de responsabilidad civil como garantía de sus servicios.

No obstante, el Reglamento propuesto prevé expresamente que, si cualquier régimen vigente de responsabilidad que exija un seguro obligatorio al operador final o inicial o fondos voluntarios existentes de seguros de empresas, tanto de la UE como de cualquier Estado miembro, ya dan cobertura y garantizan el funcionamiento del sistema o el servicio prestado, la obligación de suscribir la modalidad de seguro regulada en la Propuesta de Reglamento se consideraría cumplida si el seguro obligatorio existente o los fondos voluntarios existentes de seguros de empresas cubren los importes y alcance de la indemnización previstos en éste.

Por lo que se refiere a sus características, la Propuesta de Reglamento no detalla en exceso las mismas, si bien, establece que debe ser “adecuado en relación con los importes y el alcance de la indemnización previstos en los artículos 5 y 6 del mismo”, de modo que su cobertura deberá alcanzar el importe y alcance que serán objeto de análisis en el siguiente apartado.

8.4.6. Cuantía y alcance de la indemnización

La Propuesta de Reglamento establece en su artículo 5 el límite máximo de las indemnizaciones que corresponderán a las personas afectadas, tanto por daños materiales como inmateriales.

Como contemplaba entre las distintas opciones a valorar en relación con los futuros marcos reguladores de la responsabilidad en el ámbito de la inteligencia artificial, se trata de una medida dirigida, de un lado, a garantizar que no se desincentive la innovación y la iniciativa empresarial para el desarrollo y aplicación de la inteligencia artificial por parte de los operadores -especialmente ante la cuantía de los daños y número de personas que podrían verse afectadas-, y de otro, a garantizar la existencia en el mercado de los seguros requeridos que proporcionen las coberturas necesarias y garanticen unas coberturas mínimas a las personas afectadas.

De este modo se pretende crear un ecosistema equilibrado donde se garantice y salvaguarde la protección de todos los intereses en juego, con el consiguiente sacrificio para todas las partes que lo integran, pero que proporcione confianza y seguridad en su desarrollo, despliegue, aplicación y uso.

De un lado, la propuesta fija un importe máximo de indemnización de dos millones de euros (2.000.000€) en caso de fallecimiento o de daños causados a la salud o a la integridad física de una persona afectada como resultado o del funcionamiento de un sistema de inteligencia artificial de alto riesgo.

De otro, establece un importe máximo de un millón de euros (1.000.000€) en caso de daños morales significativos que resulten en una pérdida económica comprobable o en daños a bienes, también cuando distintos bienes propiedad de una persona afectada resulten dañados como resultado de un único funcionamiento de un único sistema de inteligencia artificial de alto riesgo.

No obstante, el Parlamento Europeo insta a la Comisión en las consideraciones previas al texto normativo propuesto para que ésta revalúe y adecúe los umbrales relativos a los daños y perjuicios, así como para que analice en profundidad las tradiciones jurídicas de

todos los Estados miembros y sus legislaciones nacionales vigentes que contemplan indemnizaciones por daños inmateriales, con la finalidad de evaluar si la inclusión de éstos en actos legislativos específicos sobre inteligencia artificial es necesaria y si contradice el marco jurídico vigente de la UE o puede afectar al Derecho nacional de los Estados miembros.

La Propuesta de Reglamento, como he referido anteriormente, contempla la concurrencia de la responsabilidad objetiva regulada en el mismo con la responsabilidad contractual del operador, estableciendo que si la persona afectada dispone de un derecho a reclamar por responsabilidad contractual contra el operador, no se le abonaría ninguna indemnización conforme al futuro Reglamento si el importe total de los perjuicios materiales o el daño moral es de un valor inferior al importe que finalmente se establezca en el texto final que, por el momento y en ésta propuesta, se fija en quinientos euros (500€).

La Propuesta de Reglamento establece en su artículo 5.2. una limitación de la indemnización y su reducción proporcional, cuando la indemnización combinada por daños materiales y morales a abonar a varias personas afectadas por el mismo funcionamiento de un mismo sistema de inteligencia artificial de alto riesgo supere los importes totales máximos precitados.

La Propuesta de Reglamento también define los conceptos que serán tomados en consideración para la cuantificación de la indemnización.

En caso de daños físicos seguidos de la muerte de la persona afectada, la indemnización incluirá los costes del tratamiento médico recibido por la persona afectada antes de su muerte, así como de los perjuicios económicos sufridos antes del fallecimiento como consecuencia del cese o la reducción de la capacidad de generar ingresos o el aumento de sus necesidades mientras durase el daño antes del fallecimiento. Asimismo, el responsable deberá reembolsar los gastos funerarios de la persona afectada a la parte responsable de sufragar los mismos

Si la persona afectada mantenía una relación con un tercero y tenía la obligación jurídica de asistirle en el momento del incidente que causó el daño que condujo a su muerte, el

responsable deberá indemnizar al tercero mediante el pago de una pensión alimenticia proporcional a la que la persona afectada se habría visto obligada a pagar, durante un período equivalente a la esperanza de vida media de una persona de su edad y teniendo cuenta su estado general. Y el responsable también indemnizará al tercero si, en el momento del incidente que provocó la muerte, el tercero había sido concebido, pero todavía no había nacido.

En caso de daños para la salud o para la integridad física de la persona afectada, la indemnización incluirá el reembolso de los gastos del tratamiento médico seguido, así como el pago del perjuicio económico sufrido por la persona afectada como consecuencia de la suspensión temporal, la reducción o el cese definitivo de su capacidad de generar ingresos o del aumento consiguiente de sus necesidades, acreditado mediante certificado médico.

8.4.7. Prevalencia del régimen previsto en el Reglamento

La Propuesta de Reglamento regula su prevalencia sobre cualquier régimen nacional de responsabilidad civil en caso de que resultara una clasificación distinta de la responsabilidad objetiva en relación con los sistemas de inteligencia artificial, conforme al mismo.

No obstante, como ha sido analizado anteriormente, el Parlamento Europeo determina en las consideraciones incorporadas a la propuesta objeto de análisis, que la Directiva sobre responsabilidad por los daños causados por productos defectuosos deberá seguir utilizándose en relación con las reclamaciones por responsabilidad civil contra el productor de un sistema de inteligencia artificial defectuoso, cuando dicho sistema cumpla los requisitos para ser considerado un producto conforme a la Directiva precitada, siendo ésta la ley aplicable.

8.4.8. Plazo de prescripción de acciones

La Propuesta de Reglamento regula unos plazos especiales de prescripción de acciones en su artículo 7.

Las demandas por responsabilidad civil relativas a daños a la vida, la salud o la integridad física estarán sujetas a un plazo de prescripción especial de treinta años a partir de la fecha en que se produjo el daño.

Las demandas por responsabilidad civil relativas a perjuicios materiales o daños morales considerables que resulten en una pérdida económica comprobable estarán sujetas a distintos plazos de prescripción especial, siendo aplicable el que venza antes.

Los plazos son los siguientes: a) 10 años a partir de la fecha en que se produjo el menoscabo a los bienes o la pérdida económica comprobable resultante del daño moral significativo, respectivamente o; b) 30 años a partir de la fecha en que tuvo lugar la operación del sistema de inteligencia artificial de alto riesgo que causó posteriormente el menoscabo a los bienes o el daño moral.

Los plazos de prescripción podrán suspenderse o interrumpirse de conformidad con la normativa nacional que lo regule, conforme establece el artículo 7.3 de la Propuesta de Reglamento.

8.4.9. Responsabilidad subjetiva

La Propuesta de Reglamento regula en su artículo 8, el régimen de responsabilidad subjetiva para otros sistemas de inteligencia artificial distintos a los de alto riesgo.

Como he comentado anteriormente, el régimen de responsabilidad objetiva regulado en esta Propuesta de Reglamento se circunscribe a sistemas de inteligencia artificial de alto riesgo, de modo que, los operadores de sistemas de inteligencia artificial que no sean calificables como de alto riesgo conforme al mismo y que, en consecuencia, no figuren en el anexo del mismo, estarán sujetos a responsabilidad subjetiva respecto de cualquier daño o perjuicio causado por una actividad física o virtual, un dispositivo o un proceso "gobernados" por los sistemas de inteligencia artificial operados por aquellos.

De este modo, la propuesta regula un régimen de responsabilidad por culpa orientada al operador, como elemento esencial de la misma, pero con inversión de la carga de la prueba.

De manera consecuente, un operador no será responsable del daño o perjuicio causado por una actividad física o virtual, un dispositivo o un proceso gobernados por el sistema de inteligencia artificial, si puede demostrar que no tuvo culpa en el daño o perjuicio causado, en particular y conforme establece la Propuesta de Reglamento, en base a cualquiera de los siguientes motivos tasados, de modo que no cabría eludir su responsabilidad en otros supuestos:

- El sistema de inteligencia artificial se activó sin su conocimiento a pesar de haberse adoptado todas las medidas razonables y necesarias para evitar dicha activación fuera del control del operador, o

- Concurrió la diligencia debida en todo el proceso a través de la realización de las siguientes acciones: La selección de un sistema de inteligencia artificial adecuado para las tareas y las capacidades pertinentes, la correcta puesta en funcionamiento del sistema de inteligencia artificial, el control de las actividades y el mantenimiento de la fiabilidad operativa mediante la instalación periódica de todas las actualizaciones disponibles.

En relación con las circunstancias expresamente previstas en la Propuesta de Reglamento que eximirían de responsabilidad al operador en caso de daño o perjuicio, considero que merecen una reflexión más profunda y no comparto plenamente algunas de ellas.

La primera de estas circunstancias eximentes hace referencia a la activación del sistema sin conocimiento del operador a pesar de haberse adoptado todas las medidas razonables y necesarias para evitar dicha activación fuera del control del operador.

Si la activación se produce por un tercero sin conocimiento ni consentimiento del operador que adoptó todas las medidas, considero que debe constituir una causa de exención de responsabilidad, aunque no plena como analizaré más adelante.

Sin embargo, si la activación se produjo automáticamente por el sistema a modo de *reboot* o reinicio automatizado del sistema, intencionado o no, no considero que pueda constituir una causa de exención de responsabilidad para determinados perfiles de operadores, dado que dicha activación nunca se debería producir técnicamente si se han adoptado "todas las medidas razonables y necesarias para evitar dicha activación fuera del control del operador". Y todo ello en congruencia además con lo previsto en el párrafo segundo del artículo 8.2 de la Propuesta de Reglamento.

En mi opinión, la responsabilidad derivada de esta activación automática o autónoma no debería recaer en el operador cuando éste no es ni el fabricante o desarrollador, pero en el caso de que lo sea, que se sea un operador inicial que defina las características de la tecnología, considero que debe responder precisamente por no haber adoptado las medidas adecuadas, sin perjuicio de que pudiera resultar de aplicación preferente, en función del contexto la Directiva 85/374/CEE.

La segunda hace referencia a la diligencia debida en todo el proceso que parece circunscribirse a tres acciones concretas que, conforme a su tenor literal, deberían concurrir para apreciar dicha "diligencia debida".

- La selección de un sistema de inteligencia artificial adecuado para las tareas y las capacidades pertinentes.

- La correcta puesta en funcionamiento del sistema de inteligencia artificial.

- El control de las actividades y el mantenimiento de la fiabilidad operativa mediante la instalación periódica de todas las actualizaciones disponibles.

De nuevo, no comparto la limitación de la diligencia debida a estas acciones específicas, dado que, en función del contexto, dicha diligencia debida puede requerir otras muchas otras cosas, entre otras, la exigencia de análisis de riesgos y/o evaluaciones de impacto previa sobre el sistema a seleccionar o seleccionado, de declaración o certificación al desarrollador/fabricante sobre determinados aspectos, de nivel de servicio asociado (en su caso), esto es ANS o SLA (*Service Level Agreement*), de cumplimiento legal, de seguridad o de privacidad. No obstante, todas estas cuestiones podrían entenderse subsumidas en el concepto de la adecuación del sistema seleccionado.

Del mismo modo que se contempla respecto de la responsabilidad objetiva en el artículo 4.3 de la Propuesta de Reglamento, se establece que los operadores no podrán eludir su responsabilidad civil alegando que el daño o perjuicio fue causado por una actividad, un dispositivo o un proceso autónomos gobernados por su sistema de inteligencia artificial.

Sin embargo y en congruencia con el régimen de responsabilidad subjetiva, a diferencia de lo que establece el precitado artículo 4.3 en materia de responsabilidad objetiva, los operadores de estos sistemas si podrán eludir su responsabilidad civil alegando que actuaron con la diligencia debida conforme contempla expresamente el artículo 8.2.b) de la Propuesta de Reglamento.

Los operadores tampoco serán responsables si el daño o perjuicio ha sido provocado por un caso de fuerza mayor, cuestión que ha sido objeto de comentario anteriormente, como en los casos de culpa exclusiva de la víctima o de un tercero.

Respecto de la exoneración de responsabilidad por culpa de la víctima, se halla contemplada en el artículo 10, el cual incluye no sólo la negligencia de la propia víctima sino de las personas de cuya actuación deba responde, como luego expondré con mayor profundidad, en el próximo apartado.

Respecto de la exoneración de responsabilidad en el caso de culpa de un tercero, el artículo 8.3 de la Propuesta de Reglamento introduce una importante ampliación de la responsabilidad civil del operador a modo de *garante* por hechos ajenos, en la medida que, cuando el daño o perjuicio haya sido causado por un tercero que haya interferido en el sistema de inteligencia artificial mediante una modificación de su funcionamiento o

sus efectos, el operador será igualmente responsable del pago de la indemnización del daño o perjuicio causado en el caso de que dicho tercero esté ilocalizable o sea insolvente.

Esta previsión no se contempla en el régimen de responsabilidad objetiva en el caso de sistemas de alto riesgo, en la medida que el operador responde directa y objetivamente de cualquier daño o perjuicio causado por el sistema de inteligencia artificial, siendo irrelevante la diligencia con la que pudiere haber actuado, derivándose la misma con independencia de la concurrencia de culpa por su parte.

Esta ampliación de la responsabilidad civil del operador bajo un criterio proteccionista hacia la persona afectada y siguiendo la finalidad esencial del derecho de daños, esto es, el resarcimiento efectivo de la víctima, difiere radicalmente del régimen de responsabilidad derivado de las conductas ilícitas con posible relevancia penal.

De un lado, las conductas relacionadas en el precitado artículo de la Propuesta de Reglamento, pueden situarse en un contexto en el que se produzca la intrusión y la posible interceptación ilegítima de comunicaciones entre sistemas de información de este tipo, esto es, las dos modalidades del delito de *hacking* previstas en los artículos 197 bis.1 y 197 bis.2 del *Código Penal* español, que serán objeto de análisis en el capítulo VI de esta investigación.

De otro, la alteración de datos, programas y documentos, entre otras acciones, puede ser constitutiva de un delito de *cracking* previsto en los artículos 266, siguientes y concordantes del *Código Penal* español, Y, por último, incluso podrían conllevar la comisión de un delito de obstaculización o interrupción de un sistema informático previsto en el artículo 264 bis del *Código Penal* español).

Todas estas conductas delictivas serán analizadas con más profundidad en el capítulo VI de esta investigación.

De este modo, a pesar de que sea el propio tercero (*cracker*) el responsable y causante del daño o perjuicio, la responsabilidad penal recaerá en el mismo, si bien, la civil, conforme al marco de la Propuesta de Reglamento, recaería en el operador conforme al mismo cuando aquél no pueda ser localizable o sea insolvente, circunstancia que ocurre

frecuentemente en la práctica, dados los perfiles, contexto y rasgos característicos de estas acciones, en su mayoría, susceptibles de calificación como *ciberdelitos*.

La experiencia profesional con estas conductas, su investigación y posterior enjuiciamiento, evidencia la dificultad de depurar las distintas responsabilidades, especialmente por los rasgos propios de estas acciones en el denominado *ciberespacio*, en especial la anonimato, la dificultad de identificación y localización, ubicuidad, perseguibilidad, cooperación judicial, simultaneidad, alto impacto en múltiples jurisdicciones, sistemas y personas, o la propia falta de solvencia de la primera línea dentro de las organizaciones y estructuras delictivas, cada vez más complejas[783].

Esta atribución de responsabilidad con origen en las acciones llevadas a cabo por un tercero comporta un importante riesgo legal adicional para el operador de este tipo de sistemas, que deberá considerar evaluar y gestionar adecuadamente, entiendo que trasladándoselo igualmente a un tercero, esto es, a las compañías de seguro. Por lo que se deberá prestar especial atención a los seguros que se ofrezcan al mercado, al objeto de verificar que este riesgo se cubre adecuadamente a través de las mismas.

Este riesgo puede ser un incentivo para la mejora de la seguridad de estos sistemas con la consiguiente inversión, de un lado, en seguridad y, de otro, en la contratación de seguros con las coberturas necesarias, que exigirán la existencia de una seguridad adecuada por el operador conforme el estado de la tecnología y tipología de sistemas.

De manera consecuente, podría constituir un aspecto más a considerar para la inversión, innovación, desarrollo y despliegue de estos sistemas, aunque muy conveniente a mi juicio por los motivos expuestos desde un punto de vista proteccionista. Todo ello también podría comportar mayor inversión y costes, lo que podría conllevar la elevación del precio de los sistemas, sin perjuicio de que puede constituir un factor diferenciador, un valor y una ventaja competitiva que determine su adquisición.

[783] MUÑOZ VELA, J. M. (2020). Conclusiones finales en la Conferencia impartida el 18.12.2020 en la Facultad de Derecho de Valencia, organizada por su Departamento de Derecho Penal, bajo el título "*Ciberdelincuencia y Ciberseguridad. Delitos intrusivos, hacking, cracking y otros*".

Por último, la Propuesta de Reglamento impone, en el marco de la responsabilidad subjetiva, un deber de colaboración del productor de un sistema de inteligencia artificial con operadores y personas afectadas, en la medida que deberá cooperar y facilitarles la información que soliciten y que sea adecuada y proporcional a la "relevancia" de la demanda, para posibilitar la depuración de responsabilidades. No obstante, este deber de colaboración no se contempla para la responsabilidad objetiva en base a la atribución directa de la responsabilidad al operador.

Por lo que se refiere a la cuantía y alcance de la indemnización a solicitar por responsabilidad civil subjetiva y los plazos de prescripción de las acciones, la Propuesta de Reglamento se remite en su artículo 9 a la legislación del Estado miembro en el que se haya producido el daño o perjuicio, lo que supone una aclaración de la ley aplicable desde una construcción legislativa proteccionista de la persona afectada, similar a los marcos regulativos de protección del consumidor y daños por productos defectuosos.

8.04.10. Cuestiones comunes

a) Responsabilidad solidaria

La Propuesta de Reglamento establece la responsabilidad solidaria en los casos en que haya más de un operador de un sistema de inteligencia artificial, conforme regula su artículo 11. En estos supuestos, todos los operadores serán responsables solidarios, aunque dispondrán de la opción de utilizar la vía de regreso entre los mismos de forma proporcional, conforme abordo a continuación.

b) Concurrencia de culpas y exención de responsabilidad

La Propuesta de Reglamento prevé en su artículo 10.1 que, si la actuación de una persona afectada o una persona de la que ésta sea responsable causó o contribuyó a causar el daño o perjuicio, el alcance de la responsabilidad civil del operador, conforme al régimen establecido en el mismo, se reducirá en consecuencia.

Si la persona afectada o una persona de la que ésta sea responsable es la única a la que se puede imputar el daño o perjuicio causado, en estos supuestos el operador no será responsable, por ser responsable exclusivo la propia persona afectada.

En relación con la prueba de la negligencia concurrente de la persona afectada -y entiendo que también de la persona de la que ésta sea responsable-, el operador estará legitimado para utilizar los datos generados por el sistema de inteligencia artificial para su acreditación de conformidad con lo previsto en el artículo 10.2 de la Propuesta de Reglamento, lo que constituye una habilitación o base legal principal para el acceso legítimo y tratamiento de dicha información por parte del operador y sin consentimiento del interesado titular de la misma y, por ende, de conformidad con lo previsto en el Reglamento General de Protección de Datos (RGPD)[784] y en las demás normas en materia de protección de datos.

Del mismo modo, el precitado artículo habilita legalmente para que la persona afectada también pueda tener acceso y usar legítimamente esos datos con "fines probatorios o aclaratorios" en la demanda de responsabilidad civil, que podrían ser propios o de terceros, lo que supondría la base legal principal para ello en cumplimiento de lo previsto en el precitado Reglamento General de Protección de Datos (RGPD).

c) Prevalencia del régimen de responsabilidad de la Propuesta de Reglamento.

Como se ha referido anteriormente, la Propuesta de Reglamento prevé que, en aquellos casos en los que el operador final del sistema de inteligencia artificial sea también el productor del mismo, el futuro Reglamento sería de prevalente aplicación sobre la Directiva sobre responsabilidad por los daños causados por productos defectuosos[785], así como, por ende, sobre la legislación nacional de transposición de ésta.

[784] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento General de Protección de Datos o RGPD/GDPR).

[785] Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos.

Y lo mismo, en los que casos en los que solo exista un operador y éste sea el productor del sistema. Sin embargo, en los casos en los que el operador inicial ostente también la condición de productor conforme a dicha Directiva, se aplicará la misma en lugar del futuro Reglamento.

En consecuencia, en caso de daño o perjuicio derivado de un sistema de inteligencia artificial, será determinante calificar adecuadamente a los sujetos intervinientes y su condición respecto del sistema en cuestión dado que, en el caso de coincidir la condición de productor y operador inicial en la misma persona, no se aplicaría el régimen de responsabilidad previsto en este Reglamento sino el previsto en la Directiva sobre responsabilidad por los daños causados por productos defectuosos, que podría repercutir negativamente en la persona afectada, en atención al análisis anteriormente efectuado.

d) Cláusulas contractuales de exención y limitación de la responsabilidad

Los pedimientos del Parlamento Europeo a la Comisión incorporados en sus consideraciones previas, incluyen que la Comisión "evalúe la necesidad de disposiciones jurídicas a escala de la Unión en materia de contratos para evitar la inclusión de cláusulas contractuales de exención de la responsabilidad, también en las relaciones entre empresas y entre empresas y la administración".

Las cláusulas contractuales de exención de responsabilidad en contextos regulados por marcos imperativos y no dispositivos, deberán seguir siendo nulas al amparo del Derecho de la UE y de los ordenamientos jurídicos internos de los Estados miembros, especialmente cuando se vea afectado un consumidor, si bien, el Parlamento Europeo considera necesario revisar los marcos normativos actuales y evaluar la necesidad de nuevas disposiciones normativas en materia contractual para evitar la inclusión de cláusulas contractuales de exención de la responsabilidad, extendiendo esta protección legal a las relaciones entre empresas y entre empresas y Administraciones públicas.

El objetivo es que las mismas no se vean afectadas por cláusulas contractuales que puedan eximir de responsabilidad a los operadores, en caso de inexistencia de marcos regulativos imperativos que prohíban estas cláusulas y su efectividad en este contexto de sujetos y relaciones. En este sentido, considero necesario que esta evaluación alcance tanto a los

marcos regulativos que permitan la exención plena de responsabilidad, así como aquellos que permitan su limitación, en aras de la necesaria armonización.

e) Acciones de regreso

La Propuesta de Reglamento también prevé la acción de regreso del operador, estableciendo que únicamente será viable conforme a lo previsto en la Propuesta de Reglamento, artículo 12, en caso de que la persona afectada haya percibido la totalidad de la indemnización a la que ésta tenga derecho a percibir conforme a las cuantías y alcance indemnizatorios previstos en el mismo.

Si un operador fuese considerado responsable solidario junto con otros operadores y hubiese indemnizado íntegramente a la persona afectada, aquél podrá recuperar parte de la indemnización de los otros operadores en proporción a su responsabilidad. En este sentido, los porcentajes de responsabilidad se determinarán conforme a los respectivos niveles de control por parte de los distintos operadores sobre el riesgo asociado a la operación y funcionamiento del sistema de inteligencia artificial.

En relación con todo ello, las consideraciones previas al texto propuesto recogen la necesidad de mejorar la trazabilidad de los productos con el fin de identificar mejor a los agentes que intervienen en las distintas fases, si bien, no tiene su correlativo reflejo en el texto normativo propuesto, lo que considero de vital importancia al objeto de facilitar la identificación de los responsables y aplicación consecuente de este régimen de responsabilidad, para obtener el íntegro y efectivo resarcimiento de los daños o perjuicios sufridos por la persona afectada y derechos de los operadores.

En este sentido, la Propuesta de Reglamento establece expresamente en su artículo 12.2.II que, si un operador, solidariamente responsable, indemnizara a la persona afectada y solicitara un ajuste de los anticipos a los demás operadores responsables, el operador podrá subrogarse en el crédito de la persona afectada frente a los demás operadores responsables, salvo que dicha subrogación puede resultar perjudicial para la demanda inicial.

Del mismo modo, conforme a su artículo 12.3 del Reglamento propuesto, si no puede obtenerse de un operador responsable solidariamente la cuantía correspondiente a su porcentaje de responsabilidad, el déficit será asumido por los demás operadores.

Si un operador de un sistema de inteligencia artificial defectuoso indemnizara íntegramente a la persona afectada conforme a lo previsto en la Propuesta de Reglamento, sea por responsabilidad objetiva o subjetiva, según establece su artículo 12.3, éste se hallará facultado igualmente para ejercitar una acción de resarcimiento frente al productor del sistema de inteligencia artificial defectuoso de conformidad con la Directiva 85/374/CEE y las disposiciones nacionales en materia de responsabilidad por los daños causados por productos defectuosos.

Y, por último, si fuese el asegurador del operador el que indemnice a la persona afectada, será el propio asegurador del operador el que se subrogará en el crédito que la persona afectada tenga frente a cualquier otra persona por responsabilidad civil por el mismo daño, hasta el importe con el que el asegurador haya indemnizado a la persona afectada.

8.04.11. Actualización del Reglamento

La Propuesta de Reglamento prevé la actualización permanente del mismo y su revisión periódica trienal, lo que reitero me parece una técnica legislativa acertada y necesaria, al objeto de permitir la adaptación del marco jurídico vigente al contexto y necesidad en cada momento, es decir un mantenimiento evolutivo del marco regulador, especialmente ante el desarrollo exponencial de la inteligencia artificial y su evolución.

Conforme prevé el artículo 14 de la Propuesta de Reglamento, cada tres años, la Comisión deberá presentar al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo un informe detallado en el que se revise el Reglamento conforme al desarrollo experimentado por la inteligencia artificial durante el período transcurrido, al que deberá acompañar, en su caso, propuestas legislativas dirigidas a cubrir las necesidades de regulación que se identifiquen en el mismo.

Para su elaboración, la Comisión solicitará la información pertinente a los Estados miembros sobre jurisprudencia, acuerdos judiciales y estadísticas sobre accidentes relacionados con sistemas de inteligencia artificial, entre otros aspectos, número de accidentes, daños sufridos, aplicaciones de inteligencia artificial involucradas o indemnizaciones abonadas por las compañías de seguros, así como una evaluación de la cantidad de demandas presentadas por las personas afectadas, ya sea de forma individual o colectiva, y de los plazos dentro de los cuales se tramitaron esas demandas en los tribunales, lo que sin duda contribuirá a identificar el contexto real y actualizado de la responsabilidad de los sistemas de inteligencia artificial en la UE, que permita identificar las necesidades y carencias del régimen establecido por el futuro Reglamento, de modo que sirva de base para su actualización y adaptación continua al contexto real en cada momento, conforme a la evolución y estado de la tecnología.

8.04.12. Consideraciones finales sobre la propuesta

Durante el análisis de los distintos aspectos contemplados en el Reglamento propuesto he expuesto mis consideraciones y opiniones sobre los mismos, no obstante, a modo de conclusión, reiterar que la Propuesta de Reglamento en materia de responsabilidad civil se sustenta en un enfoque basado en el riesgo, diferenciando dos regímenes de responsabilidad distintos en función de ello, objetiva para los denominados sistemas de inteligencia artificial de alto riesgo, así catalogados formalmente, y subjetiva para el resto.

El régimen de responsabilidad objetiva se inspira en la Directiva 2009/103/CE del Parlamento Europeo y del Consejo, relativa al seguro de responsabilidad civil de los automóviles[786], en especial para el establecimiento de un seguro obligatorio de responsabilidad civil y sus límites indemnizatorios.

[786] Directiva 2009/103/CE del Parlamento Europeo y del Consejo, de 16 de septiembre de 2009 , relativa al seguro de la responsabilidad civil que resulta de la circulación de vehículos automóviles, así como al control de la obligación de asegurar esta responsabilidad (Texto pertinente a efectos del EEE). OJ L 263, 7.10.2009. P. 11–31.

La responsabilidad será imputable siempre a la persona que haya creado, mantenga, controle o intervenga en el sistema inteligente, en la medida que estos sistemas carecen de personalidad jurídica y de la capacidad para ser titular de derechos y obligaciones.

Por último, el Reglamento propuesto supondría solventar la cuestión relativa a la aplicación de la Directiva 85/374/CEE al *software* que no se halle incorporado a un bien tangible, conforme expuse en su análisis en los apartados precedentes.

9. La responsabilidad en sectores específicos

La utilización de la inteligencia artificial en distintos ámbitos y sectores puede hallarse regulada tanto a nivel de *hard law*, como de *soft law* en instrumentos específicos y sectoriales.

A continuación, en congruencia con el enfoque horizontal de esta investigación en relación con la inteligencia artificial, no sectorial, me limitaré a realizar algunas consideraciones generales respecto de uso en el ámbito sanitario, de la conducción autónoma, de las aeronaves no tripuladas (drones), de la ciberseguridad u otros, sin posibilidad de poder realizar una inmersión vertical en sistemas concretos dentro de cada sector, ámbito y aplicación, dado su objeto y alcance limitados.

9.1. Inteligencia artificial en el ámbito sanitario

La aplicación de la inteligencia artificial en el ámbito sanitario es creciente a un ritmo exponencial, desde la pre-diagnosis hasta el tratamiento médico quirúrgico, evidenciando la necesidad de gestionar adecuadamente sus retos y riesgos, especialmente ante los bienes y derecho en juego, esto es la salud y la vida humana.

El estudio *Adverse Events in Robotic Surgery. A Retrospective Study of 14 years of FDA Data*[787], analizó los datos publicados por la Agencia de Alimentos y Medicamentos de EE.UU. -*FDA*- entre el año 2.000 y 2.013, identificando al menos 144 fallecimientos a causa de los errores cometidos por los sistemas DA VINCI -sistema de cirugía robótica- y ZEUS -sistema robótico de operaciones laparoscópicas-, incluyendo errores de los sistemas, así como problemas e interrupción de las transmisiones de video e imagen (7,4%), funcionamiento involuntario o pérdida de control de los aparatos (10,1%), descargas eléctricas y quemaduras (10,5%) o caída de piezas rotas o quemadas dentro del paciente (14,7%), entre otras.

La responsabilidad por daños derivados de estos productos en España sería exigible a través de los regímenes analizados previamente en materia de responsabilidad por productos defectuosos previsto en la TRLGDCU y los generales dimanantes del Código Civil español en materia de responsabilidad contractual y extracontractual, así como en el régimen administrativo regulado de la responsabilidad patrimonial de la Administraciones públicas, en especial, en la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas y en la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público.

Adicionalmente, deben considerarse los marcos reguladores de estos productos sanitarios, en particular, el Real Decreto 1591/2009, de 16 de octubre, por el que se regula los productos sanitarios, el Real Decreto Legislativo 1/2015, de 24 de julio, que aprueba el texto refundido de la Ley de garantías y uso racional de los medicamentos y productos sanitarios, y el Reglamento (UE) 2017/745, de 5 de abril de 2017, sobre productos sanitarios[788]. Todo ello en relación con el precitado Real Decreto 1801/2003, de 26 de diciembre, sobre seguridad general de los productos, que transpone la Directiva 2001/95/CE.

---

[787] ALEMZADEH, H; RAMAN, J.; LEVESON, N.; KALBARCZYK, Z E IYER, R. K. (2016). "Adverse Events in Robotic Surgery. A Retrospective Study of 14 years of FDA Data". PLoS ONE. Vol. 11. Nº 4. 20.04.2016. Recuperado de: https://pubmed.ncbi.nlm.nih.gov/27097160/. Consultado el 23.03.2021.

[788] Reglamento (UE) 2017/745 del Parlamento Europeo y del Consejo, de 5 de abril de 2017, sobre los productos sanitarios, por el que se modifican la Directiva 2001/83/CE, el Reglamento (CE) n.° 178/2002 y el Reglamento (CE) n.° 1223/2009 y por el que se derogan las Directivas 90/385/CEE y 93/42/CEE del Consejo. OJ L 117, 5.5.2017. Pp. 1-175

El objeto y alcance limitados de esta investigación me impiden abordar los contextos y supuestos específicos que se pueden plantear en cada uso en el ámbito sanitario, si bien, con remisión al análisis efectuado en los apartados anteriores, inicialmente, cuando se produzcan daños por defecto en el producto o la prestación de un servicio podría aplicarse la TRLGDCU, o el Código Civil español cuando concurra negligencia imputable a un médico o al hospital, debiendo considerar su titularidad pública a los efectos de la posible responsabilidad patrimonial de la Administración pública.

La mayoría de supuestos en los que resulte de aplicación, considero que resultará preferible la aplicación del régimen previsto en el TRLGDCU frente a las reglas generales de responsabilidad del Código Civil por las ventajas que puede suponer para el perjudicado, en la medida que supone una responsabilidad por riesgo *cuasi objetiva* aunque no absoluta, en la medida que el fabricante podrá exonerarse de responsabilidad en determinados contextos previstos en la norma precitada.

La implantación de dispositivos tecnológicos sanitarios en el cuerpo humano que puedan estar dotados, ser gestionados u operar con sistemas inteligentes comportan un especial riesgo. El marco de responsabilidad indicado exigiría la concurrencia de un defecto, considerando como tal, tanto cuando el mismo no ofrezca el estándar o nivel de seguridad que legítimamente cabe esperar o que ofrezcan los demás ejemplares de la misma serie, e incluiría también como tal la falta de comprobaciones necesarias para excluir estos riesgos para el ser humano[789].

Los defectos en este tipo de dispositivos pueden tener su origen en la fabricación, en su diseño y concepción como en la información asociada a su uso e interacción.

La cuestión es si el fabricante podría exonerarse de responsabilidad esgrimiendo como causa el denominado "riesgo de desarrollo" que analicé con profundidad en los apartados precedentes. Según lo dispuesto en el artículo 140.3 del TRLGDC, no se podrá invocar

[789] MOLINA MIRANDA, A. Y JUBERÍAS SÁNCHEZ, A. (2017). "Producto sanitario defectuoso". En Juberías Sánchez, A. (Coor.). *Medicamentos, productos sanitarios y protección del consumidor.* Editorial Reus. Madrid 2017. Pp 155-157.

esta causa de exoneración por parte de los sujetos responsables en los casos de medicamentos, alimentos o productos alimentarios destinados al consumo humano.

Una interpretación literal de la norma nos llevaría a concluir la imposibilidad de impedir la alegación de esta causa por los fabricantes de este tipo de dispositivos al no encajar en las categorías de productos indicadas, si bien, parte de la doctrina considera que esta protección debe considerarse extensiva a los mismos por considerarlos productos asimilados a los previstos expresamente en el precepto indicado.

Los argumentos a favor de esta interpretación se basan principalmente en la equiparación y similitudes entre el medicamento y el dispositivo electrónico sanitario, dado que ambos se implantan en el cuerpo humano, principal razón por la que se consideran potencialmente más peligrosos y presentan mayores riesgos de causar daños por su utilización. En este sentido, citar a Camacho Clavijo[790], Gil Saldaña[791] y Parra Lucán[792]. No obstante, otra parte de la doctrina se opone a ello, como los precitados Molina Miranda y Juberías Sánchez[793].

Asimismo, Camacho Gil adiciona a los mismos la finalidad tuitiva de la exclusión y la interpretación amplia del concepto de producto farmacéutico que igualmente han mantenido autores como Bercovitz[794].

En consecuencia, en base al análisis efectuado de la posible aplicación de esta cláusula de exoneración, se evidencia de nuevo la insuficiencia y la necesidad de revisión de los marcos actuales de responsabilidad para adecuarlos a la nueva realidad que supone el

[790] CAMACHO CLAVIJO, S. (2017). "La subjetividad "cyborg". En NAVAS NAVARRO, S. (Dir.). *Inteligencia artificial, tecnología, derecho*. Editorial Tirant lo Blanch. Valencia. Pp 255-256.
[791] GIL SALDAÑA, M. (2008). *El producto sanitario defectuoso en Derecho español*. Editorial Atelier. Barcelona. 2008. P. 144
[792] PARRA LUCÁN, Mª.A. (2011). *La protección del consumidor frente a los daños. Responsabilidad civil del fabricante y del prestador de servicios*. Op. cit. P. 184
[793] MOLINA MIRANDA, A. Y JUBERÍAS SÁNCHEZ, A. (2017). "Producto sanitario defectuoso". En Juberías Sánchez, A. (Coor.). *Medicamentos, productos sanitarios y protección del consumidor*. Op.cit. Pp 155 y ss.
[794] BERCOVITZ RODRÍGUEZ-CANO, R. (1992). "Artículo 28". En Bercovitz Rodríguez-Cano, R; Salas Hernández, J. y Bercovitz Rodríguez-Cano. A. *Comentarios a la Ley General para la defensa de los consumidores y usuarios*. Civitas. Madrid. 1992. P. 179.

despliegue y aplicación de la inteligencia artificial en los distintos ámbitos y sectores, como he ido exponiendo a lo largo de este capítulo.

El uso de la tecnología médica puede ser un referente para los futuros marcos reguladores de la inteligencia artificial, en la medida que ya prevén la sujeción a permisos y autorizaciones previos a su utilización, cumplimiento de determinados deberes de cuidado y vigilancia durante su puesta en marcha y su funcionamiento, monitorización, seguridad, mantenimiento y actualización.

Estos marcos obligacionales son relevantes para determinar el sujeto obligado, así como el que tiene el control y supervisión sobre el sistema para imputar la responsabilidad.

9.2. Conducción autónoma

La aplicación de la inteligencia artificial en el sector de la automoción está permitiendo el desarrollo de la conducción automatizada, más que "autónoma", si bien, los apasionantes y complejos retos y riesgos que genera en el ámbito ético, jurídico y de seguridad y, especialmente en materia de responsabilidad, está dificultando su despliegue, conforme expuse en el capítulo II de esta investigación.

Los accidentes relacionados con este tipo de vehículos han motivado la reflexión y ralentización de su despliegue, algunos motivados por la no detección de otros vehículos u obstáculos, confusión con el entorno, elección de acción inadecuada o exceso de velocidad.

El 18 de marzo de 2018 un vehículo automatizado en pruebas (en entorno real) operado por Uber Technologies, Inc, atropelló sobre las 22'00 horas a un peatón que cruzaba la calzada andando con su bicicleta, fuera del paso de peatones en Temple, Arizona[795].

[795] *Informe de accidentes de carretera, NTSB/HAR-19/03 PB2019-101402*. Junta Nacional de Seguridad del Transporte, Colisión entre un vehículo controlado por un sistema de conducción automatizada de desarrollo y un peatón Tempe, Arizona 18 de marzo de 2018.

En el momento del accidente, el vehículo disponía de un conductor humano de apoyo que, según el informe de la *Junta Nacional del Transporte* competente parecía haberse distraído visualmente durante el trayecto por el uso de su móvil personal.

Recientemente se publicó por Associated Press[796] que el conductor de apoyo ha sido acusado de homicidio por negligencia cuyo juicio se resolverá en fechas próximas. La acusación pública no se dirigió hacia Uber, y ésta resolvió con celeridad las reclamaciones civiles con la familia de la víctima. Se publicó por los medios de comunicación internacionales como el primer atropello mortal protagonizado por un coche autónomo[797].

Los daños se produjeron mientras se entrenaba al sistema para conducir el vehículo que lo integraba para llevar a cabo una conducción automatizada y, supuestamente, autónoma.

Previamente, en 2016 se publicó por los medios de comunicación la primera muerte conocida de un pasajero de un vehículo automatizado de Tesla, al colisionar con un camión en Florida mientras aquél veía una película[798].

Álvarez Olalla[799] recoge distintos ejemplos sobre los que aborda las cuestiones relativas a la responsabilidad asociada a los mismos.

Estos contextos evidencian las complejidades en materia de responsabilidad que genera la automatización y la inteligencia artificial, y tanto respecto de la responsabilidad civil, contractual o extracontractual, como penal, donde en un solo contexto se interrelacionan en tiempo real *hardware* (vehículo, sensores, …), *software*, redes, datos previos y propios del contexto y personas.

---

[796] BILLEAUD, J. Y KRISHER, T. (2020). "Se acusa al conductor de respaldo en el accidente autónomo mortal de Uber en Arizona" *Associated Press*. Publicado el 16 de septiembre de 2020. Recuperado de: https://apnews.com/article/homicide-arizona-transportation-archive-phoenix-fdd1574ac6a3c418d4f2b569b797dc16. Consultado el 28.12.2021.

[797] JIMÉNEZ, R. (2018). "Primer atropello mortal de un coche sin conductor". Publicado en *El País*, el 20.03.2018. Recuperado de: https://elpais.com/tecnologia/2018/03/19/actualidad/1521479089_032894.html. Consultado el 12.02.2021.

[798] JIMÉNEZ, R. (2018). "Primer atropello mortal de un coche sin conductor". Op. Cit.

[799] ÁLVAREZ OLALLA, P. (2019). "Responsabilidad Civil en la circulación de vehículos autónomos", en Monterrosso Casado, E. (Dir.), *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch 2019. Pp. 147-148.

La aplicación de la "inteligencia artificial" en este ámbito y sector sí ha tenido una cierta prolífica actividad regulatoria, conforme expuse en el capítulo IV.

De inicio, en materia de responsabilidad por daños causados por la circulación de vehículos "autónomos" confluyen varios marcos normativos, especialmente, el relativo a productos defectuosos y el de responsabilidad civil derivada de la circulación. Desde una valoración jurídica del contexto, en atención al principio de especialidad de las leyes y proteccionista hacia las personas afectadas, el régimen de responsabilidad automovilística debería ser de preferente aplicación, en la medida que sea aplicable, además, de naturaleza indiscutiblemente objetiva.

En estos supuestos, las cuestiones más complejas serán determinar la causa del daño -fallo en sensores, error en el sistema inteligente, decisión de la inteligencia artificial que gobierne o gestione el vehículo, acto doloso o negligencia del conductor o actuaciones de terceros- y el sujeto responsable, si el operador o usuario del vehículo autónomo -conductor, semiconductor o no- o, en cualquier caso, el propietario, sin perjuicio de los contextos en los que, además, deba responder el fabricante por defectos en el producto, ya sea vía acción principal o de repetición, o terceros.

En consecuencia, estos regímenes podrían resultar compatibles en función del contexto y complementarse para depurar todas las responsabilidades concurrentes, máxime cuando conforme a determinados ordenamientos jurídicos, pudiese atribuirse responsabilidad penal por delitos objetivos, que abordaré sucintamente en el próximo capítulo.

La cuestión sobre la inclusión de los accidentes en los que se vean involucrados este tipo de vehículos en la Ley sobre responsabilidad civil y seguro en la circulación de vehículos a motor[800] y otras normas complementarias, como el Real Decreto 1507/2008, de 12 de septiembre, por el que se aprueba el Reglamento del seguro obligatorio de responsabilidad civil en la circulación de vehículos a motor[801], ha generado cierta discusión doctrinal.

[800] RDLeg. 8/2004 de 29 Oct. (Texto refundido de la Ley sobre responsabilidad civil y seguro en la circulación de vehículos a motor). BOE 5.11.2004

[801] Real Decreto 1507/2008, de 12 de septiembre, por el que se aprueba el Reglamento del seguro obligatorio de responsabilidad civil en la circulación de vehículos a motor. BOE 13.09.2008

Álvarez Olalla[802] no ve ninguna dificultad en su inclusión al igual que García Cantero[803], que no cree ni necesario ni conveniente crear una nueva categoría de responsabilidad civil. Sin embargo, otros autores, como Iturmendi Morales[804], consideran que requeriría una modificación en su regulación, a los que me uno en mi posicionamiento al respecto.

En mi opinión, partiendo del concepto de vehículo a motor del que parten ambas normas, no considero que haya inconveniente alguno en incluir en el mismo los vehículos automatizados o, supuestamente, "autónomos". Sin embargo, el concepto de "hecho de circulación" es más restringido y, a juicio de autores como García Teruel[805], se exigiría una conducción por una persona.

A mi juicio, el artículo 2.1 del Reglamento del seguro obligatorio de responsabilidad civil en la circulación de vehículos a motor, aprobado mediante el precitado Real Decreto 1507/2008, de 12 de septiembre, no recoge expresamente esta exigencia: "A los efectos de la responsabilidad civil en la circulación de vehículos a motor y de la cobertura del seguro obligatorio regulado en este Reglamento, se entiende por "hechos de la circulación" los derivados del riesgo creado por la conducción de los vehículos a motor a que se refiere el artículo anterior, tanto por garajes y aparcamientos, como por vías o terrenos públicos y privados aptos para la circulación, urbanos o interurbanos, así como por vías o terrenos que sin tener tal aptitud sean de uso común".

Sin embargo, la definición de conductor que efectúa la Ley sobre responsabilidad civil y seguro en la circulación de vehículos a motor, si lo asocia a una persona, estableciendo que el conductor es la persona que maneja el mecanismo de dirección o va al mando de

[802] ÁLVAREZ OLALLA, P. (2019). "Responsabilidad Civil en la circulación de vehículos autónomos", en Monterrosso Casado, E. (Dir.), *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento.* Tirant lo Blanch 2019. P. 164.
[803] GARCÍA CANTERO, G. (2021). "Responsabilidad civil por accidentes originados en la circulación de vehículos sin conductor", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Tomo II. Ed. Aranzadi-Thomson Reuters. 2021. P. 781.
[804] ITURMENDI MORALES, G. (2017). "Coches autónomos y conectados. El papel de las aseguradoras". *Revista de la Asociación Española de Abogados Especializados en Responsabilidad Civil y Seguro*. Nº 61. 2017. P. 23.
[805] GARCÍA TERUEL, R.M. (2021). "El derecho de daños ante la inteligencia artificial y el machine learning: una aproximación desde las recomendaciones del Parlamento Europeo y del Grupo de Expertos de la Comisión Europea", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Tomo II. Ed. Aranzadi-Thomson Reuters. 2021. P. 1.034.

un vehículo. En vehículos que circulen en función de aprendizaje de la conducción, tendrían la consideración de conductor la persona que está a cargo de los mandos adicionales, lo que podría ya apuntar una de las vías hacia donde podría orientarse la futura revisión de estos marcos normativos para incorporar los vehículos automatizados o "autónomos".

La Ley sobre responsabilidad civil y seguro en la circulación de vehículos a motor prevé que el conductor es el responsable de los daños causados por la conducción, si bien esta responsabilidad puede extenderse no sólo a la aseguradora, en virtud del seguro obligatorio exigido por esta normativa o voluntario suscrito con la misma, sino también al propietario no conductor cuando exista vinculación con el conductor conforme a los artículos 1903 Código Civil o 120.5 del Código Penal español, esto es, dependientes, representantes o personas autorizadas. Asimismo, dicha responsabilidad podría extenderse al propietario no conductor cuando el vehículo no tuviese suscrito el seguro de responsabilidad civil obligatorio, salvo acreditación de su sustracción.

Este marco normativo establece dos regímenes de responsabilidad:

a) Responsabilidad objetiva respecto de daños personales, lo que comporta que el conductor únicamente puede exonerarse de responsabilidad por culpa exclusiva de la víctima o fuerza mayor ajena a la conducción;

b) Responsabilidad subjetiva respecto de daños a bienes, lo que requeriría la negligencia del conductor con inversión de carga de la prueba, correspondiendo a éste probar su diligencia, quedando exonerado de responsabilidad no sólo en caso de fuerza mayor, sino también de caso fortuito. Se trata de un régimen de responsabilidad con ciertas similitudes al acogido por la Propuesta de Reglamento del Parlamento Europeo sobre responsabilidad civil de la inteligencia artificial, de 20 de octubre de 2020.

La susceptible aplicación de estos marcos a los coches automatizados o "autónomos", dejando a un lado la exigencia de conductor comentada, precisa partir precisamente de su clasificación y grado de autonomía, ese atributo del que curiosamente tanto se aparta la

nueva Propuesta de Reglamento, del Parlamento Europeo y del Consejo, para la regulación de la inteligencia artificial, de 21 de abril de 2021.

La clasificación por niveles de autonomía de los vehículos autónomos fue publicada originalmente por la Society of Automotive Engineers -SAE- en el año 2014 como parte de su informe *Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems*[806], en el que describe una hoja de ruta en el que los fabricantes de vehículos se están basando para desarrollar las capacidades de sus vehículos autónomos.

En EE.UU., la National Highway Traffic Safety Administration -NHTSA- había ya creado su propia escala en 2013, si bien, en septiembre de 2016 decidió descartarla y adoptar la escala de SAE precitada como estándar.

Conforme a la misma, los vehículos se clasifican del siguiente modo:

- Nivel 0: Sin automatización de conducción
- Nivel 1: Asistencia al conductor (conducción asistida)
- Nivel 2: Automatización de conducción parcial
- Nivel 3: Automatización de conducción condicional
- Nivel 4: Alta automatización de conducción
- Nivel 5: Automatización de conducción completa

La legislación española debe revisarse, como ya lo han hecho distintos estados en EE.UU. y otros países, para incorporar los vehículos automatizados y, supuestamente "autónomos", en los marcos específicos vigentes en materia de tráfico, circulación y seguro de vehículos a motor.

La conducción plenamente autónoma no está permitida legalmente en la actualidad salvo en determinadas jurisdicciones y con limitaciones, sin perjuicio de las pruebas o ensayos previamente autorizados. En algunos países se permite el uso experimental o regular de

[806] Última versión recuperada de: https://www.sae.org/standards/content/j3016_202104/. Consultado el 28.05.2021.

vehículos con un alto grado o totalmente automatizados que, además, prevén la cobertura de los daños causados, bien mediante seguros o remisión a normas generales[807].

Respecto de los vehículos autónomos de nivel 5, a mi juicio, no se podría aplicar inicialmente el régimen de responsabilidad previsto en la Ley sobre responsabilidad civil y seguro en la circulación de vehículos a motor, en la medida que se trataría de accidentes de circulación donde no intervendría un conductor en sentido estricto, por lo que el perjudicado debería exigir la responsabilidad por daños al propietario o poseedor del vehículo por la vía del régimen de responsabilidad extracontractual previsto en los artículos 1902, siguientes y concordantes del Código Civil español, con las dificultades y falta de adecuación del mismo para estos supuestos, especialmente ante la necesidad de probar la culpa/negligencia del propietario o poseedor -licenciatario, operador o usuario- y la relación de causalidad entre la acción u omisión y el daño. Del mismo modo, podría valorarse esta vía para depurar las responsabilidades, cuando se incurra en ellas, especialmente en caso de vehículos de nivel 3 y 4.

Respecto a la posible imputación de la responsabilidad al operador-usuario en relación con vehículos de nivel 5 plenamente automatizados o "autónomos", podría valorarse esta posibilidad en determinados contextos en los que el operador-usuario tuviera que tomar una decisión. Desde luego la falta de mantenimiento del vehículo y el sistema por el mismo podría comportar dicha responsabilidad en el mismo o el propietario, así como cuando el operador-usuario pueda activar u operar con culpa o de manera negligente el

[807] El artículo 19 del Decreto italiano de 28 de febrero de 2018 sobre las pruebas de vehículos conectados y automatizados en la vía pública -*Modalità attuative e strumenti operativi della sperimentazione su strada delle solu- zioni di Smart Road e di guida connessa e automatica, 18A02619, GU n° 90 de 18 de abril de 2018*- establece que la persona que solicite la autorización para realizar pruebas de vehículos automatizados en la vía pública debe acreditar una cobertura de seguro de responsabilidad civil suficiente. La circular de la Dirección General de Tráfico de España de 13 de noviembre de 2015 (Instrucción 15/V-113) también autoriza las pruebas de los coches automatizados y exige que el seguro de responsabilidad civil cubra los límites del seguro obligatorio para los vehículos de motor. El artículo 7 de la Ley de Tráfico Alemana -*Straßenverkehrsgesetz*- establece la responsabilidad objetiva del poseedor del vehículo. Esta norma se dejó deliberadamente sin modificar cuando la Ley de Tráfico se adaptó a la aparición de los vehículos automatizados. El Decreto francés n° 2018-211, de 28 de marzo de 2018, relativo a la experimentación con vehículos automatizados en la vía pública, se basa en la precitada Ley Badinter, Nº. 85/677, de 5 de julio de 1985. El ejemplo más destacable es la reciente Ley de Vehículos Automatizados y Eléctricos del Reino Unido de 2018 (c 18), Sección 2, que establece que el asegurador es responsable de los daños sufridos por el asegurado o cualquier otra persona en un accidente causado por un vehículo automatizado. Si el vehículo no está asegurado, el responsable es el propietario del vehículo.

sistema de conducción, o cuando el operador no intervenga en contextos de aviso de desactivación o no operación del sistema de conducción autónoma.

Adicionalmente, se podrían dar supuestos en los que no coinciden en la misma persona la condición de operador y de usuario (conductor o no) del vehículo, ni de operador y propietario, es decir, contextos en los que se podría poner a disposición del usuario un servicio de transporte, en el que su operador o proveedor (prestador) debe responder por los daños causados y ello por distintas vías en función del contexto.

Si la persona afectada es el usuario del servicio y se considera consumidor o usuario, se podría valorar la aplicación del régimen de responsabilidad objetiva por servicios defectuosos vigente en España analizado anteriormente, previsto en el TRLGDCU (y no previsto en la Directiva 85/374/CEE), dada la posibilidad de su inclusión tanto en la categoría de vehículos a motor como relativos a medio de transporte, conforme a su artículo 148 del TRLGDCU.

Si la persona afectada es un tercero, por ejemplo, un peatón, a juicio de la precitada García Teruel, no resultaría aplicable inicialmente el régimen de responsabilidad por servicios defectuosos habida cuenta de la condición subjetiva del perjudicado -consumidor o usuario-, sin perjuicio de incardinarse en el régimen general de responsabilidad extracontractual previsto en el Código Civil español.

No obstante, la precitada García Teruel[808] considera que podría aplicarse el régimen de responsabilidad por hechos de la circulación en determinados supuestos y vehículos automatizados de los que sería responsable el conductor, por aplicación del principio *ubi emolumentum ibi onus*, de modo que el responsable debería ser el prestador del servicio u operador, siguiendo la doctrina del Tribunal Supremo español recogida en su Sentencia de 7 de abril de 1997 (RJ 1997, 2743), de modo que se atribuya la misma a quien obtiene el provecho del quebranto sufrido por tercero, "a modo de contrapartida del lucro obtenido con la actividad peligrosa". Este criterio podría ser igualmente extensible a

[808] GARCÍA TERUEL, R.M. (2021). "El derecho de daños ante la inteligencia artificial y el machine learning: una aproximación desde las recomendaciones del Parlamento Europeo y del Grupo de Expertos de la Comisión Europea", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Tomo II. Ed. Aranzadi-Thomson Reuters. 2021. P. 1.045.

determinados supuestos relacionados con los vehículos de niveles inferiores como principio básico de la responsabilidad por riesgo[809], situando la responsabilidad en la órbita del quién se lucra o beneficia de ello, esto es, el fabricante u operador.

De hecho, a mi juicio, precisamente este principio debería ser igualmente considerado en los futuros marcos reguladores de la responsabilidad civil de la inteligencia artificial, sin perjuicio del criterio del control sobre el riesgo, especialmente a la hora atribuir la responsabilidad y su distribución.

Adicionalmente, en estos contextos se precisarían también nuevas previsiones reguladoras en materia de aseguramiento que contemplen un seguro obligatorio, similar al seguro obligatorio de viajeros, a concertar por los operadores de estos servicios.

Por lo que se refiere a los niveles inferiores a esa conducción plenamente automatizada, sí podemos encontrarnos ya supuestos de responsabilidad por hechos de la circulación, conforme abordé en el capítulo IV al analizar su regulación, que podrían ser abordables conforme a los marcos jurídicos especiales precitados en España, siempre que concurra la figura del conductor.

Aun así, respecto de los vehículos clasificados en los niveles 3 y 4 donde la autonomía es parcial o se halla condicionada, en los que se requeriría alguna acción por parte del conductor en determinados contextos, el régimen de responsabilidad previsto en la Ley sobre responsabilidad civil y seguro en la circulación de vehículos a motor se muestra para algunos autores, como la precitada García Teruel, como inadecuado para los distintos supuestos que se pueden plantear.

Efectivamente pueden darse múltiples contextos, en los que la intervención humana en una conducción semi-autónoma o condicionada sea la que cause el daño, en cuyo caso, considero que podría ser aplicable el régimen de responsabilidad especial previsto en la precitada Ley sobre responsabilidad civil y seguro en la circulación de vehículos a motor.

[809] Ver STS, Sala 1ª, de 5 de enero de 2007, Rec. 161/2000: "Los sistemas de responsabilidad objetiva o responsabilidad por riesgos, desde su inicio en la doctrina civilista, se han basado en un principio de equidad jurídica: *ubi emolumentum, ibi onus*".

Sin embargo, conforme a los mismos, si el accidente que causó el daño tiene su causa en una decisión del sistema sin intervención humana, en este contexto de análisis, la responsabilidad debería recaer en el "conductor" conforme a la normativa especial, si bien, respondería de distinta manera en función de la naturaleza de los daños: a) Una responsabilidad objetiva respecto de los daños personales de la que sólo podría exonerarse en supuestos legalmente tasados[810], en particular, en caso de fuerza mayor ajena a un hecho de la circulación -no por caso fortuito-, respondiendo de hechos imprevisibles o inevitables para el mismo; b) Una responsabilidad subjetiva respecto de los daños materiales.

Si fue una decisión del sistema en la que el conductor no tuvo la oportunidad de intervenir, de nuevo se evidencia la falta de adecuación de este marco especial en materia de circulación de vehículos a motor a estos supuestos específicos, de modo que habría que acudir al precitado principio general *ubi emolumentum ibi onus,* para intentar encajar la responsabilidad en la órbita del quién se lucra o beneficia de ello, esto es, el fabricante u operador.

No obstante, en mi opinión y en función del contexto, considero que los perjudicados podrían disponer de la opción de utilizar la vía de responsabilidad por productos defectuosos al fabricante en determinados contextos, analizada en los apartados precedentes, con las dificultades inherentes conforme a su marco actual, especialmente la prueba del defecto y la relación de causalidad -a cuya resolución contribuiría enormemente la exigencia jurídica de los requerimientos éticos de transparencia, explicabilidad y responsabilidad- así como la acotación de las posibles causas de exoneración de responsabilidad del productor.

Y adicionalmente a todo ello, en el mismo contexto analizado en el párrafo anterior, puede darse la circunstancia en la que solo se hayan causado daños propios, personales o materiales. En este caso, el conductor no podría ser resarcido conforme a esta normativa

[810] La precitada García Teruel considera que "de igual forma que el conductor debe responder en la actualidad por un fallo en los frenos…también debe responder cuando la IA falle". En GARCÍA TERUEL, R.M. (2021). "El derecho de daños ante la inteligencia artificial y el machine learning: una aproximación desde las recomendaciones del Parlamento Europeo y del Grupo de Expertos de la Comisión Europea", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Tomo II. Ed. Aranzadi-Thomson Reuters. 2021. P. 1.037.

especial en materia de circulación de vehículos a motor, sin perjuicio de que pueda utilizar la precita vía de la responsabilidad por productos defectuosos, si resultara aplicable.

La conducción efectivamente "autónoma" comportaría que el sistema inteligente sería quien monitorizaría y ejecutaría todas o algunas de las tareas de la conducción, sin necesidad de supervisión o intervención humana activa, de modo que en estos supuestos, no estaríamos ante la presencia de un conductor a los efectos previstos y exigidos por el marco jurídico previsto en el Texto Refundido de Ley sobre Tráfico, Circulación de Vehículos a Motor y Seguridad Vial, aprobada mediante Real Decreto Legislativo 6/2015, de 30 de octubre, y en la Ley sobre responsabilidad civil y seguro en la circulación de vehículos a motor. De este modo, estaríamos más bien, en su caso, ante un operador, como igualmente consideran autores como Castells i Marqués[811], la cual analizó esta cuestión en 2017, remitiéndose como argumento de su postura, a la regulación de los vehículos autónomos entonces en distintos estados en EE.UU. a la que hice una referencia general en el Capítulo IV de esta investigación, que optaron por esta consideración en su regulación específica, como Florida o California.

Conforme estos marcos, se consideraba operador de un vehículo autónomo a la persona que hiciera que la tecnología autónoma del vehículo se activara, independientemente de que la persona estuviera físicamente presente en el vehículo mientras estuviera funcionando en modo autónomo, como es el caso de Florida. Sin embargo, el *California Vehicle Code*[812], revisado en octubre de 2017, considera operador a la persona que esté sentada en el asiento del conductor, o si no hay ninguna persona el mismo, la que haga que la tecnología autónoma se active.

La propia Castells i Marqués citaba a parte de la doctrina, en particular, a Goodrich, que criticaba esta última definición a la hora de imputar responsabilidad y con absoluto acierto en mi opinión, y lo hacía ilustrándolo en un sencillo ejemplo: Si una madre situaba a su hijo de 8 años en un vehículo autónomo para mandarlo al colegio, según esta definición,

[811] CASTELLS I MARQUÉS, M. (2017). "Drones Civiles", en Navas Navarro, S. (Dir.). *Inteligencia artificial, tecnología, derecho*. Valencia, Tirant lo Blanch, 2017. Pp. 112 y 113.
[812] S 38750 (4).

el operador sería el menor y su madre quedaría liberada de responsabilidad en caso de colisión.

Sin embargo, el *Florida Status* de 2021[813], establece un importante cambio en esta materia. En primer lugar, un vehículo completamente autónomo puede operar en Florida sin importar si un operador humano está físicamente presente en el vehículo. Y, en segundo lugar, que, salvo que el contexto requiera lo contrario, cuando esté activado, el sistema de conducción automatizado es el que se considerará el operador de un vehículo autónomo, independientemente de si una persona está físicamente presente en el vehículo mientras el vehículo está funcionando con el sistema de conducción automatizado activado. Es decir, se consideraría operador al propio sistema automatizado.

Una posible interpretación para la aplicación de los marcos especiales en España en materia de responsabilidad sobre el tráfico y circulación a los vehículos automatizados o "autónomos" y su posible adecuación a los supuestos que plantea la inteligencia artificial, la responsabilidad debería recaer en el propietario del vehículo en todo caso -sin perjuicio de la posible responsabilidad del usuario cuando sea conductor del mismo en los casos que no sea una conducción plenamente autónoma o cuando existan alteraciones del *software* o falta de actualización del sistema-, conforme recoge la nueva normativa en Reino Unido citada en el anterior capítulo, cuando no tenga asegurado el mismo.

El régimen especial actual no exige culpa o negligencia del conductor o propietario para atribuirle responsabilidad por daños corporales, bastando que la circulación del vehículo sea la causa, salvo supuestos de culpa exclusiva de la víctima o fuerza mayor.

Las cuestiones adicionales que se plantean son, entre otras, si debería extenderse el régimen de responsabilidad objetiva a los daños materiales o si el seguro obligatorio debería complementarse con un seguro de daños que indemnizara al propio conductor por los daños sufridos por el mismo. Esta última cuestión, ha sido igualmente contemplada en la nueva normativa en Reino Unido citada.

[813] S 316.85 (3) (a).

En definitiva, conforme a esta interpretación de estos marcos especiales, en caso de daños derivados de un accidente causado por decisiones automatizadas del vehículo (no plenamente automatizado) no imputables a fallos de *software*, debería imputarse la responsabilidad objetiva al propietario y, en su caso, al usuario en función de los contextos precitados, no siendo imputable inicialmente al fabricante, dado que se trataría de una responsabilidad derivada de los riesgos de la circulación.

Si los daños se derivan de accidentes causados por un defecto de *software* que motiva una decisión inadecuada en el contexto que provoca el accidente y los daños, podrán concurrir y debería valorarse, de un lado, la responsabilidad del fabricante del vehículo o del *software* (solidaria) y sus compañías aseguradoras con la responsabilidad objetiva o subjetiva del conductor o propietario y el seguro obligatorio, en función de la autonomía del vehículo.

Por otra parte, en caso de daños causados por un accidente imputable a un tercero, por ejemplo, cuando un vehículo se vea obligado a tomar una decisión motivada por una actuación culposa o dolosa de otro usuario de la vía, por ejemplo, un peatón que se cruce por la misma, un individuo que altera las velocidades máximas indicadas en las señales ubicadas en la vía o un *hacker* que interfiere el sistema de inteligencia artificial del vehículo. En estos casos, se abriría la vía de repetición contra el responsable, en especial, por parte de la compañía aseguradora.

En cualquier caso, algunas de las referencias legislativas precitadas en EE.UU. y UK evidencian los diferentes posicionamientos respecto de la responsabilidad. En mi opinión, en el caso de vehículo plenamente "autónomos", la responsabilidad debería orientarse hacia el operador o productor del vehículo o el sistema inteligente que lo gobierna.

En el caso de productores, dejando a un lado el perfil de importador y proveedor, debemos considerar que, de nuevo, podrán concurrir varios perfiles, el fabricante del vehículo (que inicialmente puede no estar automatizado), el fabricante del sistema inteligente (hardware como sensores, software, algoritmos, datos y conectividad), el programador del software y el integrador, conversor o *automator* que lo transforma en un vehículo automatizado. Estos perfiles pueden concurrir en una sola persona física o jurídica, como así podrá ocurrir en el caso del fabricante del vehículo, el fabricante del sistema y el que automatiza

el vehículo e integra el sistema inteligente. Adicionalmente, intervendrán otros agentes, como el entrenador del sistema.

La responsabilidad de los mismos debería construirse conforme a los marcos actuales bajo el amparo del TRLGDCU por intervenir todos ellos en el "proceso de fabricación", sin perjuicio del derecho de repetición en función de su participación en la causación del daño. Sin embargo, conforme analicé en los apartados precedentes, este régimen especial de responsabilidad comporta dificultades para su adecuación a los distintos contextos y para su aplicación, especialmente ante la necesidad de probar el daño y la relación de causalidad por la víctima, a lo que contribuirán los futuros marcos jurídicos que prevean determinadas obligaciones de registro de eventos y de transparencia como, por ejemplo, la Propuesta de Reglamento europeo sobre inteligencia artificial de 21 de abril de 2021, y como ya recogen algunos marcos específicos en materia de vehículos autónomos en otros países como, por ejemplo, el precitado *California Vehicle Code*[814].

En función del mismo, si el defecto o daño es imputable al sistema automatizado posteriormente integrado en el vehículo fabricado que inicialmente no tenía esta condición y fue posteriormente convertido en autónomo, la responsabilidad debería recaer en el fabricante del sistema y no en el del fabricante del vehículo.

Estas iniciativas en otros países evidencian, de nuevo, la necesidad de revisar la normativa española.

La víctima, como he referido, podrá ser el conductor, el pasajero/usuario o un tercero, y los defectos en los que estos agentes deberán responder, podrían agruparse en defectos de fabricación (del vehículo o del *hardware* asociado al sistema inteligente, como por ejemplo, sensores), defectos en la concepción y diseño (incluyendo errores del *software* o interacción deficiente entre conductor y sistema de conducción), así como defectos de información al usuario/conductor (por ejemplo, advertencias de no intervención, no priorización de su integridad personal o limitaciones del sistema en entornos específicos).

[814] S 38750

Otra cuestión apasionante que abordé en el capítulo III de esta investigación en relación con los aspectos éticos es el relativo a la disyuntiva del sistema inteligente que "gobierne" o "gestione" el vehículo en determinados contextos, ilustrado en el denominado "dilema del tranvía" y otros citados.

En caso de tener que decidir entre evitar atropellar a uno o varios peatones, o maniobrar y colisionar el vehículo con uno o varios ocupantes contra un muro ¿quién sería el responsable de los daños? ¿El diseñador, el fabricante, el propietario? Conforme a los marcos especiales analizados debería ser el propietario, pero ¿y si se trata de un solo peatón con 90 años de edad y los ocupantes se trata de una familia con los progenitores y dos niños de 1 y 3 meses de edad, y el vehículo toma la decisión de esquivar al peatón y estrellar el vehículo contra el muro falleciendo uno de los ocupantes y con importantes secuelas el resto? ¿Quién decide quién asumirá la materialización del riesgo? ¿Quién decide quién vive y quién muere?

Y, por último, en caso de daños derivados de un accidente causado por culpa única y exclusiva de la víctima podría aplicarse el Ley sobre responsabilidad civil y seguro en la circulación de vehículos a motor, con exoneración de responsabilidad para el resto de sujetos. Del mismo modo, se aplicaría dicha norma en caso de concurrencia de culpa de la víctima.

9.3. Aeronaves no tripuladas

De inicio, las aeronaves no tripuladas o *UAS*, conocidas comúnmente como "drones", no tienen por qué integrar inteligencia artificial, si bien, cada vez más se está incorporando a las mismas para permitir no sólo su automatismo en su funcionamiento, sino su autonomía.

El concepto "dron" no es el más adecuada desde un punto de vista técnico, aunque es el término por el que comúnmente se conocen este tipo de aeronaves no tripuladas con

control remoto, y ha pasado a adoptarse en las distintas iniciativas legislativas, como la *Uniform Tort Law Relating to Drones Act*[815] de 2019 en EEUU.

A nivel de la UE, el marco jurídico básico se integraba por el Reglamento (UE) 2018/1139 del Parlamento Europeo y del Consejo[816], de 4 de julio de 2018, sobre normas comunes en el ámbito de la aviación civil y por el que se crea una Agencia de la Unión Europea para la Seguridad Aérea y por el que se modifican los Reglamentos (CE) n.° 2111/2005, (CE) n.° 1008/2008, (UE) n.° 996/2010, (CE) n.° 376/2014 y las Directivas 2014/30/UE y 2014/53/UE del Parlamento Europeo y del Consejo y se derogan los Reglamentos (CE) n.° 552/2004 y (CE) n.° 216/2008 del Parlamento Europeo y del Consejo y el Reglamento (CEE) n.° 3922/91 del Consejo, el Reglamento de Ejecución (UE), 2019/947 de la Comisión, de 24 de mayo de 2019[817], relativo a las normas y procedimientos aplicables a las utilización de aeronaves no tripuladas, el Reglamento Delegado (UE) 2019/945 de la Comisión, de 12 de marzo de 2019[818], sobre los sistemas de aeronaves no tripuladas y los operadores de terceros países de sistemas de aeronaves no tripuladas.

Al cierre de esta investigación se ha aprobado por la Comisión Europea el conjunto de normas que regularán el denominado U-Space[819], aplicables a partir del 26 de enero de 2023: El Reglamento de Ejecución (UE) 2021/664 de la Comisión, de 22 de abril de 2021, sobre un marco regulador para el U-Space[820], el Reglamento de Ejecución (UE) 2021/665 de la Comisión, de 22 de abril de 2021, por el que se modifica el Reglamento de Ejecución (UE) 2017/373, en lo que se refiere a los requisitos para los proveedores de servicios de gestión del tránsito aéreo/navegación aérea y otras funciones de la red de gestión del tránsito aéreo en el espacio aéreo U-Space designado en espacio aéreo controlado[821], y el Reglamento de Ejecución (UE) 2021/666 de la Comisión, de 22 de abril de 2021, por el

[815] KURTZ, P.M.; GLASER, M. Y HEVERLY, R. (2019). "Memorandum Uniform Tort Law relating to Drones Act". *Uniform Law Commission*, 2019. P. 1.
[816] OJ L 212, 22.8.2018. Pp. 1-122
[817] OJ L 152, 11.6.2019. Pp. 45-71
[818] OJ L 152, 11.6.2019. Pp. 1-40
[819] Definido en el artículo 2 del *Reglamento de Ejecución (UE) 2021/664 de la Comisión, de 22 de abril de 2021, sobre un marco regulador para el U-Space*, define *espacio aéreo U-Space* como la "zona geográfica de UAS designada por los Estados miembros, en la que solo se permite que se lleven a cabo operaciones de UAS con el apoyo de servicios de U-Space". A su vez define *Servicio de U-Space* como el "servicio basado en servicios digitales y automatización de funciones diseñados para facilitar un acceso protegido, eficiente y seguro al espacio aéreo U-Space para un gran número de UAS".
[820] OJ L 139, 23.4.2021. Pp. 161-183
[821] OJ L 139, 23.4.2021. Pp. 184-186

que se modifica el Reglamento (UE) 923/2012, en lo que se refiere a los requisitos para la aviación tripulada que opera en el espacio aéreo U-Space[822].

Este conjunto de normas constituye la base para una regulación armonizada en el seno de la UE de la seguridad aérea que contemple la misma desde la doble dimensión de la misma, esto es, *safety* (eliminación de riesgos no intencionados y daños a terceros) y la *security* (prevención de interferencias ilícitas. La protección de la privacidad y el medioambiente constituyen objetivos adicionales de la mismas, por lo que igualmente resultan de aplicación los marcos generales en estas materias.

A nivel nacional (España), se regulan en el Real Decreto 1036/2017, de 15 de diciembre, por el que se regula la utilización civil de las aeronaves pilotadas por control remoto, y se modifican el Real Decreto 552/2014, de 27 de junio, por el que se desarrolla el Reglamento del aire y disposiciones operativas comunes para los servicios y procedimientos de navegación aérea y el Real Decreto 57/2002, de 18 de enero, por el que se aprueba el Reglamento de Circulación Aérea[823], así como en el reciente Real Decreto 1088/2020, de 9 de diciembre, por el que se completa el régimen aplicable a la notificación de sucesos de la aviación civil y se modifica el Real Decreto 1036/2017, de 15 de diciembre, por el que se regula la utilización civil de las aeronaves pilotadas por control remoto, y se modifican el Real Decreto 552/2014, de 27 de junio, por el que se desarrolla el Reglamento del aire y disposiciones operativas comunes para los servicios y procedimientos de navegación aérea y el Real Decreto 57/2002, de 18 de enero, por el que se aprueba el Reglamento de Circulación Aérea[824].

El Real Decreto 1036/2017, de 15 de diciembre, regula en sus artículos, entre otras responsabilidades y obligaciones, las de información del fabricante (Artículo 16.1), las de mantenimiento y conservación del operador (Artículo 16.2.) y las de aseguramiento del operador que cubra la responsabilidad civil frente a terceros por los daños que puedan ocasionarse durante y por causa de la ejecución de las operaciones aéreas especializadas o vuelos experimentales (Artículo 26.I.c)).

[822] OJ L 139, 23.4.2021. Pp. 187-188
[823] BOE 29.12.2017
[824] BOE 10.12.2020

Las responsabilidades derivadas de daños causados por estas aeronaves se rigen por los marcos especiales precitados, así como los generales que resulten de aplicación al contexto.

Los contextos de daños pueden ser muy variados, especialmente significativos los causados a terceros en superficie o derivados de la colisión con naves tripuladas.

En este contexto, las normas especiales de referencia serían la Ley 48/1960, de 21 de julio, de Navegación Aérea[825] -especialmente su artículo 120-, el *Convenio sobre daños causados a terceros en la superficie por aeronaves extranjeras*[826] -especialmente sus artículos 1.1 y 11-, firmado en Roma, el 7 de octubre de 1952, y el Convenio sobre indemnización por daños causados a terceros por aeronaves[827], firmado en Montreal, el 2 de mayo de 2009. No obstante, también debe considerarse el Convenio sobre la responsabilidad internacional por daños causados por objetos espaciales, hecho en Londres, Moscú y Washington, el 29 de marzo de 1972[828].

Este marco parte de un régimen de responsabilidad por daños objetiva, basada en el riesgo y limitada en sus indemnizaciones en función del contexto.

De inicio, los daños causados por una aeronave durante el vuelo a terceros deben ser indemnizados conforme a los artículos 1, 3, 11 y 19 del Convenio sobre daños causados a terceros en la superficie por aeronaves extranjeras y 119 de la Ley de Navegación Aérea precitada.

No obstante, en caso de tratarse de aeronaves no tripuladas se pueden plantear múltiples contextos, como la pérdida de control, la caída de la misma o algunos de sus componentes, colisión con aeronaves tripuladas, expulsión o rociado de material expulsado por la misma o pérdida de la carga, con todo lo pueden comportar adicionalmente en función de la naturaleza de la misma como, por ejemplo, medicamentos, mercancías peligrosas, etc.

[825] BOE 12.08.1960
[826] BOE 17.05.1961
[827] BOE 07.05.1961
[828] BOE 02.05.1980

Algunos de estos supuestos podrían tener cabida en el marco jurídico vigente, otros no, lo que obligaría acudir al régimen común.

No se aplicarían límites indemnizatorios en el caso de utilización de la aeronave sin consentimiento del operador, por ejemplo, en el caso de la sustracción o toma de control por un tercero, conforme a los artículos 122 de la precitada Ley de Navegación Aérea española y 12.2 del precitado Convenio sobre daños causados a terceros en la superficie por aeronaves extranjeras, ni tampoco cuando el operador o las personas bajo su control (piloto remoto u observador) actúan de forma dolosa o por medio de culpa grave - entendida como conducta indebida deliberada u omisión manifiesta, importante y grave del deber de diligencia ante un riesgo evidente y una falta profesional grave a la hora de prestar la diligencia debida en ese contexto, conforme al artículo 16 del Reglamento 376/2014 precitado.

En el caso de actuaciones realizadas por empleados del operador, aunque justifiquen que obraron con diligencia, la responsabilidad del operador sería limitada y objetiva, sin perjuicio de responder ilimitadamente si se probara un elemento culpabilístico por incumplimiento de sus obligaciones o la de sus agentes.

En la práctica, las acciones para reclamar la indemnización por daños por accidentes aéreos suelen sustentarse conjuntamente en los artículos 121, siguientes y concordantes de la Ley de Navegación Aérea, y en los artículos 1902 y 1903 del Código Civil español, conforme recoge expresamente el Tribunal Supremo en sus Sentencias de 3 de mayo (RJ 2738) y 10 de junio (RJ 4868) de 1.988, entre otras razones, en la medida que el régimen común se considera menos restrictivo y más favorable para la víctima, en la medida que el grado de culpa exigible es menor, aplicándose la teoría del riesgo como criterio de imputación y una responsabilidad *cuasi objetiva* desplazándose la prueba al demandado[829], se amplían los posibles sujetos responsables más allá del operador y su personal dependiente y el plazo de prescripción para el ejercicio de la acción es mayor (1 año).

[829] Conforme con la interpretación del Tribunal Supremo en Sentencias de 23 de octubre de 2012 (RJ9725) 201218.03. 2016 (RJ983) y 18.11.2016,

De este modo, se podría ejercitar una acción directa de responsabilidad extracontractual vía 1902 del Código Civil español contra el piloto remoto, el observador, la organización, la persona encargada del mantenimiento o el gestor responsable de la aeronave no tripulada[830] e incluir en la demanda al operador por culpa "in vigilando" conforme al artículo 1903.4 del Código Civil español. Asimismo, cabría interponer la acción del artículo 1903.4 precitado únicamente frente al operador, sin perjuicio que la aseguradora integre la relación jurídico procesal, incluso siendo directamente demandado en virtud del artículo 76 de la Ley de Contrato de Seguro.

El operador podría responder en exclusiva o solidariamente con otros posibles agentes intervinientes conforme a los artículos 120 a 122 de la Ley de Navegación Aérea y 1902 y 1903 del Código Civil español. Entre otros supuestos de responsabilidad en exclusiva, podríamos situar determinadas acciones como no atender las comunicaciones del fabricante sobre la actualización de los sistemas para su correcto funcionamiento o no seguir el manual de mantenimiento.

Especial significación merece la interferencia, toma de control y utilización de drones con fines maliciosos mediante ciberataques o terrorismo, lo que de antemano ya puede ser constitutivos de varios delitos conforme analizaré en el capítulo VI y que, tras su comisión, puede ser utilizado para llevar a cabo actos de terrorismo como para la comisión de otros delitos contra las personas y cosas.

Asimismo, estos actos pueden suponer el incumplimiento de determinadas obligaciones y garantías de seguridad previstas en los marcos reguladores precitados con relevancia a efectos de responsabilidad, en la medida que las mismas recaen en el operador, el cual debe adoptar las medidas de protección contra interferencias ilegales y accesos no autorizados, y evitarlas en el uso del espacio radioeléctrico.

Los actos maliciosos de interferencia, toma de control o utilización de la máquina, dispositivo, vehículo o aeronave con la finalidad de cometer actos ilícitos o delictivos, pueden producirse respecto de cualquiera de estos elementos gobernados por un sistema de inteligencia artificial como coches, autobuses, taxis de tráfico rodado o aéreo, robots,

[830] Sentencia Audiencia Provincial de Baleares 2007, 2 5, JUR 289587.

máquinas industriales, sistemas de gestión financiera e inversión, exoesqueletos o prótesis médicas, y máxime cuando se utilizan redes y tecnologías de comunicación entre el sistema y el dispositivo, donde puede estar una de sus vulnerabilidades a explotar por un tercero.

De nuevo se evidencia la importancia y necesidad de la ciberseguridad en relación con la aplicación y uso de la inteligencia artificial.

En caso de drones, como cualquier otro sistema inteligente o máquina, robot o dispositivo dotado del mismo, cuyo uso se halle permitido a menores de edad, responderían los padres o tutores conforme a lo dispuestos en los apartados 2 y 3 del artículo 1903 del Código Civil español, o aquellos directamente en virtud del artículo 1902, si tenían capacidad de discernimiento.

No obstante, el breve análisis expuesto hace referencia a la responsabilidad de aeronaves no tripuladas conforme a los marcos comunes y especiales citados, pero no a aquélla derivada de aeronaves no tripuladas y gobernadas por un sistema inteligente, que podrían ir del mero automatismo a la autonomía parcial o plena.

Las aeronaves no tripuladas plenamente "autónomas" y sus operaciones, inicialmente quedarían excluidas del sistema de aviación civil, conforme a la *Circular 328/AN/190 de la Organización de la Aviación Civil Internacional* -OACI- en relación con el Convenio sobre Aviación Civil Internacional de 1944 (Convenio de Chicago) y el *Manual de 2015 de la propia OACI*.

Conforme ya analicé al abordar los regímenes generales de responsabilidad y reflexionar sobre esta posibilidad, Castells i Marqués[831] plantea la posibilidad de equiparar los drones autónomos con capacidad de aprendizaje con los animales para la aplicación analógica del régimen de responsabilidad previsto en el artículo 1905 del Código Civil español, fijando la responsabilidad en la órbita de su poseedor (sea o no el propietario) o que se sirva de él, más que en el fabricante, en base al riesgo que tanto los drones como los

[831] CASTELLS I MARQUÉS, M. (2017). "Drones Civiles", en Navas Navarro, S. (Dir.). *Inteligencia artificial, tecnología, derecho*. Valencia, Tirant lo Blanch, 2017. Pp. 98 y 99.

animales comportan, aun adoptando precauciones, así como la influencia del usuario a través de sus interacciones o no correcciones de determinas conductas que puede determinar su comportamiento. Como la misma autora destaca, la aplicación de esta responsabilidad dependerá en todo caso del contexto, especialmente si concurre una causa de exoneración (fuerza mayor o culpa de la víctima) o de concurrir la responsabilidad del fabricante.

De nuevo, en estos supuestos se evidencia la insuficiencia y falta de adecuación de los marcos actuales especiales para resolver todos los supuestos que se pueden plantear en la práctica donde se hallen involucrados sistemas inteligentes con cierta autonomía, sin personalidad jurídica ni capacidad de responder de sus actos, y la necesidad de revisar, adecuar y complementar dichos marcos, conforme expuse en los apartados precedentes, al objeto de definir marcos de responsabilidad objetiva basada en el riesgo por los daños causados por sistemas inteligentes imputable a su operador o, en su caso a otros agentes involucrados que pueda intervenir desde su diseño, fabricación, entrenamiento, dotación, operación y uso. Esos marcos deberían contemplar aspectos complementarios, como los seguros obligatorios de responsabilidad civil o identificación electrónica de los mismos, al igual que contempla la legislación[832] que regula de manera adicional la tenencia de animales potencialmente peligrosos.

9.4. Ciberseguridad

Los sistemas de inteligencia artificial pueden ser utilizados para la gestión de sistemas y redes informáticas, así como para gestionar la ciberseguridad de las mismas o instalaciones, como expuse en el capítulo II.

Los posibles escenarios de responsabilidad que pueden suscitarse en este contexto son innumerables y resultaría inabordable realizar un análisis vertical en el seno del objeto y alcance limitados de esta investigación.

[832] Ley 50/1999, de 23 de diciembre, sobre el Régimen Jurídico de la Tenencia de Animales Potencialmente Peligrosos). BOE 307 24.12.1999.

No obstante, significar que pueden producirse decisiones, actuaciones y omisiones por parte de esto sistemas en estos contextos cuyos resultados pueden causar no sólo enormes daños a una organización, sino incluso su suspensión de actividades o afectar a su continuidad y supervivencia, lo que en el contexto de proveedores de servicios esenciales puede significar dejar a un país sin conexión a Internet, a una ciudad sin control de tráfico, a una región sin energía eléctrica, sin combustible o sin abastecimiento de otros productos básicos.

Estos no son supuestos de laboratorio, ya ha ocurrido, como he expuesto a lo largo de esta investigación y han sido objeto de los encabezamientos y primeras páginas de los distintos medios de comunicación.

Los supuestos que pueden plantearse son múltiples, desde errores de programación o defectos funcionales, errores de diseño y configuración que generan vulnerabilidades, actos maliciosos que provoquen el error o que aprovechen dichas vulnerabilidades en el diseño del sistema o su programación -por ejemplo, puertas traseras o *backdoors*-, generación o transmisión de software malicioso no detectado (virus o troyanos) voluntaria o involuntariamente a través del sistemas inteligentes ya sea a nodos de la red, usuarios específicos o terceros, ataques de denegación de servicio simple o distribuidos (DDos), secuestro o sustracción de información (*ransomware*), etc.

Todos estos actos pueden generar daños y la consecuente responsabilidad para su reparación, si bien, en función del contexto será de naturaleza contractual o extracontractual, con posibilidad de utilizar los marcos específicos de responsabilidad sobre productos defectuosos dado su carácter *cuasi objetivo* y no necesidad de culpa, cuando resulten de aplicación.

En este ámbito, nos podemos encontrar con múltiples agentes: El proveedor del sistema inteligente para la gestión de redes o sistemas o para la gestión de su ciberseguridad, que podrá ser el propio fabricante o un tercero, por ejemplo, comercializadores, distribuidores o importador, el operador si lleva servicios asociados de monitorización continua, el implantador o un tercero.

En función del origen del daño deberán dirimirse, imputarse y, en su caso, distribuirse las responsabilidades.

Si el sistema presenta errores de programación o defectos funcionales que afecten a la ciberseguridad, inicialmente deberá responder el fabricante/productor, sin perjuicio de las responsabilidades de naturaleza penal y/o civil del agente que pueda explotar los mismos y cause daños.

Si las vulnerabilidades explotadas por terceros (*hackers, crackers,* etc.) tienen su origen en un incumplimiento de las medidas de seguridad requeridas al sistema o de los acuerdos de nivel de servicio -ANS o *SLA* según sus siglas en inglés- la responsabilidad contractual debería corresponder inicialmente al proveedor/fabricante, sin perjuicio de nuevo de las responsabilidades de naturaleza penal y/o civil del agente que pueda explotar los mismos y cause daños.

Si las vulnerabilidades y daños tienen su origen en factores como la inadecuada configuración del sistema, diseño de su arquitectura, integración e implantación, la responsabilidad corresponderá inicialmente al implantador que acometió estos servicios, que podría ser el fabricante, el proveedor o un tercero prestador de este servicio específico, de nuevo, sin perjuicio de las responsabilidades de naturaleza penal y/o civil del agente que pueda explotar los mismos y cause daños.

Si las vulnerabilidades y daños tienen su origen en el entrenamiento inadecuado o insuficiente del sistema, la responsabilidad debería recaer en quién acometió el entrenamiento del mismo.

Además de los escenarios descritos, se podrían dar otros como la utilización de los propios sistemas inteligentes por un *cracker* para realizar ataques a terceros. En este caso, parece que inicialmente deberíamos descartar la posibilidad de atribuir la responsabilidad al titular o usuario del sistema, correspondiendo la penal y, en su caso, la civil derivada del delito al *ciberatacante*, si bien, a mi juicio, sería exigible la misma también a los sujetos precitados, si la vulnerabilidad explotada por aquel tiene su origen en los mismos y, por tanto, responde a algo que debió ser considerado conforme al estado de la técnica y, por tanto, previsible.

No obstante, como analicé anteriormente, la Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial[833], prevé supuestos en los que el operador respondería en caso de no identificación de los terceros causantes del daño.

El artículo 8.3 de la Propuesta de Reglamento citada introduce una importante ampliación de la responsabilidad civil del operador a modo de *garante* por hechos ajenos, en la medida que, cuando el daño o perjuicio haya sido causado por un tercero que haya interferido en el sistema de inteligencia artificial mediante una modificación de su funcionamiento o sus efectos, sin especificar que interfiera de buena o mala fe, el operador será igualmente responsable del pago de la indemnización del daño o perjuicio causado, en el caso de que dicho tercero esté ilocalizable o sea insolvente

En cualquier caso, el titular o usuario del sistema inteligente deberá responder, contractual o extracontractualmente por los daños causados por el mismo derivados de su uso que no tengan su origen en terceros, máxime si no adoptó las medidas que fueran exigibles al mismo generales o sectoriales, en especial, en virtud de marcos como el Reglamento General de Protección de Datos (RGPD), la Ley Orgánica de Protección de Datos y Garantía de Derechos Digitales (LOPDGDD), la Directiva NIS o el Real Decreto-ley 12/2018, de 7 de septiembre, de Seguridad de las Redes y Sistemas de Información, así como más específicos, como por ejemplo en el ámbito bancario o de las telecomunicaciones.

A esta perspectiva general debemos adicionar el papel de las distintas compañías aseguradoras de los respectivos agentes y sus coberturas -responsabilidad civil, error informático ciberataques, etc.-, así como, en caso de explotación de vulnerabilidades por un tercero, consciente o inconscientemente, el de los agentes que llevan a cabo estas acciones de explotación, que serían los causantes inmediatos del daño y responsables, en su caso, tanto a nivel penal y/o civil de los daños provocados.

[833] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial.

9.5. Otros ámbitos

Conforme he expuesto al inicio de este capítulo, el enfoque horizontal y el objeto y alcance limitados de esta propuesta me impiden abordar con mayor detalle ámbitos, sectores, aplicaciones y sistemas concretos como, por ejemplo, protección de datos, gestión contable, fiscal o laboral, etc, sin perjuicio de los aspectos generales que efectuaré en el capítulo VI de esta investigación.

10. Consideraciones finales

Conforme al marco jurídico vigente en la UE y en España, los sistemas de inteligencia artificial y los robots, máquinas u otros productos dotados de la misma, dispongan o no de cierta o total autonomía, carecen de personalidad jurídica, no pueden ser titulares de derechos y obligaciones y no pueden ser considerados responsables civiles de los daños o perjuicios causados por los mismos.

Las argumentaciones en este sentido han sido expuestas al analizar la postura al respecto tanto del Parlamento Europeo como del Comité Económico y Social (CESE), y a la vista del precitado marco regulador propuesto.

De hecho, considero que no deberíamos pues hablar de responsabilidad de la inteligencia artificial sino de responsabilidad de los sistemas de inteligencia artificial o mejor, de la responsabilidad derivada de su funcionamiento y uso.

Las propuestas normativas en la UE en materia de responsabilidad civil de la inteligencia artificial, en particular, la Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial[834], que se incorporó como recomendación en la Resolución del Parlamento Europeo de 20 de octubre de 2020, cerraría inicialmente cualquier interpretación actual al respecto, si bien, considero que la evolución tecnológica y de los

[834] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial.

sistemas de inteligencia artificial en particular, especialmente por lo que se refiere a su autonomía, capacidades de aprendizaje automático y profundo, y el consecuente alejamiento de la supervisión y control humano que se pretende instaurar como principio y norma irrenunciable en la UE -inicialmente respecto de cualquier sistema inteligente, aunque a la vista de la última Propuesta de Reglamento de 21 de abril de 2021, sólo respecto de sistemas de alto riesgo-, posiblemente requerirán revisar este posicionamiento actual en el futuro.

Mi opinión sobre estos aspectos, manifestada al abordar las distintas cuestiones, posicionamientos y propuestas, es coincidente con esta postura, en la medida que mientras exista la imperativa supervisión y control humano respecto de cualquier sistema inteligente o máquina, robot, vehículo, dispositivo o producto que lo integre, cualquiera que sea el grado de autonomía, existirá una necesaria dependencia, por lo que, en función del contexto, la responsabilidad debería focalizase de inicio en la/s persona/s que esté/n detrás del mismo y que disponga/n del control sobre el riesgo materializado que produzca el daño o el perjuicio a la/s persona/as afectada/s, sin perjuicio de las responsabilidades exclusivas o concurrentes que puedan producirse.

Hasta la aprobación de un nuevo marco regulador en materia de responsabilidad civil de los sistemas de inteligencia artificial deberemos aplicar los marcos vigentes en materia de responsabilidad por productos defectuosos regulado por la Directiva 85/374/CEE y, en particular en España, en la TRLGDCU, que transpone la misma pero ampliando su ámbito de aplicación, así como los generales de responsabilidad civil contractual y extracontractual prevista en el ordenamiento jurídico español en los artículos 1902, siguientes y concordantes del Código Civil, sin perjuicio de la aplicación de los regímenes de responsabilidad especiales en determinados sectores y ámbitos, especialmente en el ámbito de la circulación de vehículos a motor, cuando resulten de aplicación.

No obstante, como he analizado y expuesto al abordar los mismos, estos marcos precisan su revisión y adecuación, presentando distintas carencias que, a mi juicio, impiden su aplicación a la totalidad de contextos que puede plantear la inteligencia artificial, especialmente en atención a sus características y capacidades y, en particular en relación con su grado de autonomía, autoaprendizaje, complejidad, opacidad, impredecibilidad e intervención de múltiples sujetos en la producción del resultado lesivo, lo que podría

afectar al derecho a un resarcimiento efectivo por parte de las personas afectadas, que es el fin último de los marcos reguladores de la responsabilidad civil por daños.

En relación con esta necesidad de adecuación me permito significar algunas de las recomendaciones efectuadas por el Parlamento Europeo en las distintas Resoluciones analizadas en esta investigación y las recogidas en el precitado informe *Liability for Artificial Intelligence and other emerging digital technologies*, elaborado por el *Grupo de expertos en responsabilidad y nuevas tecnologías* de la Comisión Europea, especialmente respecto a la necesidad objetivación de la responsabilidad del operador de una tecnología basada en inteligencia artificial -construida desde un enfoque de riesgos y ante la pluralidad de sujetos intervinientes que pueden considerarse corresponsables dentro de su ciclo de vida-, la necesidad de facilitar la prueba cuando sea requerida y exista dificultad de la misma, el aseguramiento obligatorio y la posible creación de fondos de compensación.

Las orientaciones que estableció el precitado Grupo de Expertos en el informe mencionado para los futuros marcos reguladores de la responsabilidad, considero que constituyen una buena base de referencia a considerar, que incluyen la responsabilidad objetiva del operador de sistemas que impliquen riesgo de daño significativo para terceros, la responsabilidad según el grado de control sobre la tecnología y sus riesgos, el deber del usuario de actuar con diligencia al seleccionar, operar, controlar y mantener adecuadamente el sistema, siendo responsable por incumplimiento de estas obligaciones si causa daños a terceros concurriendo su culpa, la responsabilidad del usuario, la responsabilidad de los fabricantes por productos por los defectos que causen, incluso si el defecto fuese causado por cambios introducidos en el producto una vez comercializado, el seguro de responsabilidad civil obligatorio, la inversión de la carga de la prueba (especialmente ante dificultades o costes desproporcionados para establecer el nivel de seguridad pertinente o para demostrar que no se ha cumplido dicho nivel de seguridad), el registro de actividad de los dispositivos (logs) con inversión de la carga de la prueba ante su inexistencia, la solidaridad en caso de concurrencia de múltiples agentes, la responsabilidad por pérdida de datos del usuario que debe ser indemnizable y la no necesidad de otorgar personalidad jurídica a los sistemas autónomos en la medida que el daño que causa puede y debe ser atribuible a personas.

El sistema de responsabilidad extracontractual previsto en el Código Civil español, el de responsabilidad del fabricante por producto defectuoso o el de responsabilidad por daños derivados de la circulación de vehículos a motor vigentes en España, no dan respuesta a la totalidad de situaciones y de daños causados por estos sistemas en los distintos contextos que se pueden plantear, por lo que los nuevos marcos jurídicos deben actualizar, adaptar y complementar a aquéllos.

Los requisitos de aplicabilidad, prueba y posibles causas de exoneración comportan dificultades para garantizar un derecho al resarcimiento efectivo de la persona afectada o *restitutio in integrum* en el contexto de sistemas inteligentes.

La legislación española vigente no contempla específicamente la responsabilidad del proveedor de un servicio que usa inteligencia artificial y del operador del mismo, si bien, conforme he analizado, podrían incardinarse en el marco de la responsabilidad contractual o extracontractual previsto en el Código Civil español, así como en el de la responsabilidad por productos defectuosos previsto en el TRLGDCU, al contemplar expresamente la posible aplicación de éste no sólo a productos sino también a servicios, a diferencia de la Directiva 85/374/CEE. Tampoco se prevé en nuestro ordenamiento jurídico la responsabilidad o grado de diligencia del usuario, la obligatoriedad de contratar un seguro obligatorio para la fabricación o tenencia de máquinas o robots dotados de inteligencia artificial o la creación de fondos de compensación.

La revisión y reformulación de los sistemas clásicos de responsabilidad civil en base este marco, no sólo es deseable sino obligatoria en mi opinión, como la de diversos autores que han analizado esta cuestión a los que he citado a lo largo de mi análisis, como la precitada Núñez Zorrilla o Martínez Mercadal[835].

Los marcos actuales de responsabilidad sólo contemplan los supuestos en los que es posible atribuir los daños de un sistema de inteligencia artificial al agente humano relacionado con el mismo, ya sea el fabricante, el operador, el propietario o el usuario,

[835] MARTÍNEZ MERCADAL, J.J. (2018). "Vehículos autónomos y derecho de daños. La estructura clásica de la responsabilidad civil frente al avance de la inteligencia artificial". *Revista de la Facultad de Ciencias Económicas UNNE*, nº 20. Universidad de la República de Montevideo Uruguay 2018. Pp. 57-67. Disponible en https://revistas.unne.edu.ar/index.php/rfce/article/view/3267. Consultado el 31.01.2021.

según el contexto, en la medida que dicho agente debería haber previsto y evitado la conducta lesiva.

La Propuesta de Reglamento sobre responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial constituye un paso firme hacia adelante por parte del Parlamento Europeo, elaborado bajo un enfoque evolutivo, dinámico y global, que pretender ir adaptándose a la realidad de la tecnología y sus riesgos.

El sistema objetivo de responsabilidad que prevé para sistemas de inteligencia artificial de alto riesgo me parece adecuado, máxime al verse complementado con la limitación de las indemnizaciones y la exigencia de seguros obligatorios con el objetivo de proteger los intereses de todas las partes interesadas en su desarrollo, despliegue y aplicación, tanto de las empresas para no desincentivar el desarrollo y la innovación segura, como de la sociedad en general de modo que, de un lado, pueda acceder a los sistemas y soluciones tecnológicas más vanguardistas de manera segura y a precios asumibles y no exclusivos de una minoría y, de otro, garantizar su derecho al resarcimiento en caso de daños provocados por dichos sistemas.

Sin embargo, la tipificación inicial de los sistemas de inteligencia artificial considerados del alto riesgo en un listado cerrado, aunque revisable y actualizable, debe revisarse en la medida que pueden quedar fuera determinados sistemas que por los riesgos que comportan, tanto por sus características, capacidades, aplicaciones y usos deberían estar de inicio en dicho listado.

La no inclusión de los mismos determinaría la aplicación de un sistema de responsabilidad subjetiva a los daños causados por dichos sistemas, no siendo exigible con carácter retroactivo en caso de inclusión posterior en el listado, salvo que el texto propuesto se modifique, conforme a sus considerandos previos, que sí lo preveía.

Además de todo ello, la coordinación pretendida con otros sistemas de responsabilidad, en particular, con la responsabilidad derivada de productos defectuosos, podría conllevar la aplicación preferente de éstos en el caso de imputarla al productor y, en consecuencia, privar a la persona afectada de las ventajas del sistema previsto en el Reglamento

propuesto y permitir que el fabricante pueda resultar beneficiado de aquél, en función del contexto y circunstancias.

El texto propuesto, en caso de que adopte como base para la construcción de un marco específico de responsabilidad civil derivada de la inteligencia artificial, deberá ser objeto de un análisis más profundo en relación con aspectos como los comentados, incluida la definición y alcance del concepto "autonomía", sin perjuicio de la necesidad de una mayor coordinación con la propuesta coetánea del Parlamento Europeo simultánea en materia de ética y obligaciones jurídicas relacionadas de la inteligencia artificial así como con la nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021.

Asimismo, en relación con el ámbito de aplicación y causas de exoneración de responsabilidad en los marcos vigentes en España y la UE en materia de productos defectuosos, considero que sería deseable una urgente revisión de la Directiva 85/374/CEE, en primer lugar, para considerar la inclusión o no de los servicios y su ampliación, como prevé el TRLGDCU, en segundo lugar, valorar la eliminación como causa de exoneración de responsabilidad la inexistencia del defecto en el momento de la puesta en circulación en casos de sistemas inteligentes más avanzados, especialmente los dotados de capacidad de autoaprendizaje y, en tercer lugar, valorar cuanto menos la posibilidad de regular una presunción *iuris tantum* de defecto cuando un sistema inteligente cause un daño, con posibilidad de exoneración cuando el defecto o daño del producto no le es imputable.

Por último, considero que los marcos reguladores de la responsabilidad de la inteligencia artificial deberán construirse desde el enfoque de riesgos sobre el que se está sustentando la UE en sus informes y propuestas, y ello comporta, en congruencia y en el ámbito de la responsabilidad, lo siguiente: a) Asumir que el riesgo forma parte de cualquier actividad de negocio, de cualquier sistema de información o de cualquier tecnología que utilicemos aisladamente, integrada con otras e interconectada, y por supuesto, de la aplicación y uso de sistemas inteligentes; b) Identificar cual es la capacidad de riesgo que se puede asumir en base al impacto; c) Identificar la cantidad de riesgo que estamos dispuestos a aceptar para garantizar conseguir los objetivos pretendidos -*risk apetite*-, entre otros, una inteligencia artificial segura y confiable y un resarcimiento íntegro y efectivo de los daños

que pueda causar la misma; d) Gestionar adecuadamente las incertidumbres que genera la inteligencia artificial por sus propias características, capacidad y distintos sujetos involucrados en su ciclo de vida, mediante el establecimiento de algunas de las posibles soluciones comentadas en los apartados precedentes, esto es, inversión de la carga de la prueba, responsabilidad solidaria, distribución de la responsabilidad o incluso valorar un régimen combinado de responsabilidad individual y de responsabilidad colectiva, de modo que todos y cada uno de los miembros integrantes del colectivo de riesgo puedan ser considerados responsables por parte de la persona afectada, sin perjuicio de la definición de normas que exijan cumplir o determinen la asignación de la responsabilidad dentro de los colectivos de riesgo.

En este sentido, me permito significar los argumentos en este sentido de Argyri Panezi[836], en relación con una futura responsabilidad colectiva en el ámbito de la inteligencia artificial, no prevista en el ordenamiento jurídico vigente. Según esta autora, utiliza el símil de la denominada "responsabilidad por cuota de mercado" aplicada en ocasiones en casos de catástrofes sanitarias o ecológicas, comparando una fuga de petróleo con una fuga de datos, en la medida que ambos activos son valiosos, necesitan salvaguardas de seguridad y protección para hacer frente a sus vulnerabilidades en su almacenamiento y transporte y para evitar brechas y fugas.

Sin duda, la posibilidad de "consorcios de riesgo de inteligencia artificial" propuestos por esta autora podría ser una opción muy interesante a explorar y debatir y he querido hacer eco de la misma, con sus ventajas y sus inconvenientes, especialmente por su posible rechazo por el mercado y de determinados agentes para entrar en grupos de industrias que les comportaría asumir cargas de responsabilidad colectiva, sin perjuicio de que ello pueda beneficiar y hacer más sólida la industria de la inteligencia artificial.

En cualquier caso, en congruencia con los distintos enfoque políticos y jurídicos sobre los riesgos asociados a la innovación, lo deseable es una actuación global desde un enfoque igualmente global, en la línea iniciada por la UE, al objeto de no establecer marcos de responsabilidad que se limiten a intentar resolver de la manera más adecuada

[836] PANEZI, A. (2021). "IA: un enfoque ecosistémico para gestionar el riesgo y la incertidumbre", en García Mexia, P. y Pérez Bes, F. (Eds); Panezi. A. (Coord). (2021). *Artificial Intelligence and the Law*. Ed. La Ley (Wolters Kluwer). 2021. Edición electrónica (SMARTECA).

las incertidumbres asociadas la inteligencia artificial, que al final son instrumentos reactivos para restaurar las consecuencias de la materialización de un riesgo con impacto, sino marcos reguladores de carácter preventivo que definan requisitos y obligaciones éticas y jurídicas esenciales y que exijan una adecuada gestión de los riesgos asociados a los sistemas inteligentes, desde principios básicos como el de precaución, considerando los riesgos inherentes y residuales en relación con la capacidad de riesgo aceptable definida, al control y supervisión humana durante todo el ciclo de vida.

Capítulo VI

Responsabilidad penal y otras

1. Introducción

La realidad siempre va por delante del Derecho. Los mitos y la ciencia ficción van muy por delante de ambos, si bien, la historia del ser humano ha evidenciado una y otra vez su grandeza al hacer real lo irreal, posible lo imposible.

La consciencia y la inteligencia como atributos de la mente humana que nos permiten darnos cuenta de nuestra propia existencia, de la del resto del mundo y de las cosas que pasan, que nos permiten procesar la información en el interior de nuestro cerebro para obtener un resultado inteligible, que nos permite aprender, entender, razonar, tomar decisiones y formarnos una idea determinada de la realidad, parecen exclusivas de los seres humanos pero conforme he anticipado en anteriores capítulos, ¿podrían dejar de serlo en el futuro?, ¿podrían crearse artificialmente?, ¿podrían atribuirse a entes con mayor capacidad que la del ser humano?, ¿podría ser diseñada, construida y atribuida a un sistema o a una máquina?

Hoy, esta posibilidad sólo existe en la ciencia-ficción.

Como he analizado en capítulos anteriores, para muchos es imposible, para otros improbable, para otros meramente lejana y para algunos pocos posible, como igualmente comentaré a lo largo de este capítulo.

Los avances tecnológicos y la innovación en su aplicación nos sorprenden permanentemente y crean nuevas realidades para las que el Derecho no tiene soluciones previstas, dado que los marcos reguladores vigentes fueron creados en otros contextos.

La computación cuántica no es todavía una realidad, pero siguiendo un criterio científico, antes o después lo será. La cuestión es si debemos empezar a pensar ya, “quién, bajo qué principios y normas y cómo deberá utilizarse esta tecnología”. La posesión de la misma

en manos privadas podría llegar a romper una red *Blockchain* o cualquier otra red cifrada en minutos o seguramente en segundos. ¿Somos conscientes del poder que supone esta tecnología?

La inteligencia artificial supone retos múltiples, diversos y algunos similares, y más ante su interacción con otras tecnologías, las crecientes capacidades de computación y el acceso y tratamiento masivo de datos. ¿Quién tendrá el control y bajo qué principios y normas deberá utilizarse?

Como he expuesto a lo largo de esta investigación, los distintos gobiernos están definiendo sus estrategias de inteligencia artificial y están intentando definir en la actualidad su marco ético y jurídico, sin embargo, su avance, aumento de capacidades, perfeccionamiento incesante y aplicación masiva hace que los retos y riesgos que plantea aumenten progresivamente si no se aborda desde un enfoque adecuado a nivel legislativo, de manera flexible, evolutiva y adaptativa.

La realidad de esta tecnología dentro de diez años será muy diferente a la actual.

En este sentido, ¿qué ocurrirá si realmente se logran construir sistemas de inteligencia artificial con capacidad de decisión similar a la humana, con hipotética "plena autonomía", con capacidad de autoaprendizaje automatizado y profundo, con independencia absoluta respecto de quién los diseño o construyó, y con relativa impredecibilidad en sus actuaciones para adaptarse a cualquier contexto?, ¿Se permitirá la creación de sistemas autónomos no sujetos a la supervisión y control humano efectivo? ¿Y si se consiguiera construir una conciencia artificial similar a la humana?

Nos encontramos ante retos apasionantes desde un punto de vista científico, ético y jurídico que deberán ser abordados en los próximos años de manera global y que obligará a establecer marcos que regulen su diseño, su uso y su explotación ante los riesgos que puede suponer para personas, empresas o gobiernos.

Una de las cuestiones centrales sobre las que se está debatiendo para la construcción de los marcos presentes y futuros de la inteligencia artificial es la posibilidad de atribuirle una personalidad jurídica propia a los sistemas inteligentes más avanzados, distinta a la

de las personas físicas y jurídicas, pero asimilada en algunos aspectos, especialmente respecto de determinadas prerrogativas y obligaciones.

Durante el tratamiento de los distintos aspectos abordados con anterioridad, he manifestado mi opinión personal conforme al estado actual de la tecnología y la postura del legislador europeo al respecto, sin perjuicio de que sea algo a reflexionar en el futuro conforme evolucione ésta.

En mi opinión y como he reflejado durante el tratamiento de otros aspectos, considero que algunos de esos sistemas podrían llegar a ser considerados agentes o entes sin personalidad jurídica, pero con determinadas prerrogativas y obligaciones acotadas y restringidas, de las que deberán responder las personas, físicas o jurídicas que estén detrás de su control real o potencial.

La responsabilidad civil derivada de los defectos, usos, acciones y omisiones por parte de estos sistemas ha sido analizada en los apartados precedentes, reflexionando sobre cómo aplicar el derecho vigente a los mismos, sus dificultades y limitaciones, reflexionando sobre las propuestas regulatorias actuales en la UE y adicionando mi opinión personal sobre posibles soluciones a explorar en relación con la legislación vigente y con los futuros marcos reguladores.

La posible responsabilidad penal derivada de los usos, acciones y omisiones de estos sistemas será igualmente objeto de análisis en los posteriores apartados.

2. Inteligencia artificial y autoría

El análisis actual sobre cómo aplicar el derecho vigente a los sistemas de inteligencia artificial -actuales y futuros-, supone un importante reto y esfuerzo interpretativo para los juristas, dado que los marcos jurídicos vigentes no fueron pensados para sistemas autónomos con plena capacidad de autoaprendizaje -automatizado y profundo-, en constante evolución y creciente despliegue en todo tipo de ámbitos, aplicaciones y

sectores, y constituye una preocupación constante para los expertos y para el propio legislador[837], como he puesto de relieve anteriormente al analizar distintos aspectos.

Durante la introducción al capítulo V de esta investigación abordé la distinción general y relación entre el denominado ilícito civil del ilícito penal, y las responsabilidades dimanantes de los mismos.

Sin embargo, abordar ahora la responsabilidad penal exige partir necesariamente del elemento de culpabilidad para su atribución, esto es, no hay pena sin dolo o imprudencia -*nullum crimen sine culpa*-, consecuente al derecho fundamental a la presunción de inocencia consagrado en la Constitución.

No cabe inicialmente la responsabilidad objetiva en el ámbito penal en el Derecho occidental, lo que nos lleva inevitablemente a la cuestión relativa a si un sistema de inteligencia artificial puede tener moral y consciencia o no, cuya carencia, de inicio, podría imposibilitar su imputabilidad desde un punto de vista penal.

Esta cuestión ha sido abordada por autores como Domínguez Peco[838], Sánchez del Campo[839] o Mengotti[840] que reflejan la imposibilidad de imputar responsabilidad penal a un robot o sistema dotado de inteligencia artificial avanzada, sin perjuicio de que pueda ser un instrumento o medio para la comisión de un delito.

De conformidad con lo previsto en el artículo 5 del Código Penal español[841], no puede haber pena sin dolo o imprudencia.

Dicho precepto enlaza con lo dispuesto en sus artículos 27 y 28 del mismo que establecen que serán responsables criminalmente de los delitos, los autores y los cómplices, y se considerarán autores de un delito a quienes realicen el hecho por sí solos, conjuntamente

---

[837] ASARO, P. (2007). "Robots and responsibility from a legal perspective". En *Proceedings of the IEEE*. IEEE 2007. Recuperado de https://peterasaro.org/writing/ASARO%20Legal%20Perspective.pdf. Consultado el 11.01.2021

[838] DOMÍNGUEZ PECO, E.M. (2018). "Los Robots en el Derecho Penal", en Barrio Andrés, M. (Director). *Derecho de los Robots*. Wolkers Kluwer, Madrid 2018.

[839] SÁNCHEZ DEL CAMPO, A. (2016). *Reflexiones de un replicante legal. Los retos jurídicos de la robótica y las tecnologías disruptivas*. Editorial Aranzadi, Navarra 2016.

[840] NIETO MENGOTTI, J.P. (2016). *El Derecho Penal frente a los robots*. FIDE Papers, Madrid, 2016.

[841] Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal.

o por medio de otro del que se sirven como instrumento. Asimismo, conforme a dicho precepto, se considerarán también autores los que induzcan directamente a otro u otros a ejecutarlo, así como los que cooperen a su ejecución con un acto sin el cual no se habría efectuado.

El Código Penal español también contempla en su artículo 31 la responsabilidad personal del administrador de hecho o de derecho de una persona jurídica, o en nombre o representación legal o voluntaria de otro, aunque no concurran en él las condiciones, cualidades o relaciones que la correspondiente figura de delito requiera para poder ser sujeto activo del mismo, si tales circunstancias se dan en la entidad o persona en cuyo nombre o representación obre.

Esto podría invitar a una primera reflexión sobre la responsabilidad por actos de terceros como en la vía civil, de concurrir los requisitos exigidos para ello, si bien, la ausencia de consciencia, moralidad, personalidad jurídica y capacidad de obrar en un sistema de inteligencia artificial o en el robot, máquina o producto que la integre y, en consecuencia, de dolo o culpa, imposibilita de inicio su atribución.

Del mismo modo, la responsabilidad penal de las personas jurídicas[842] regulada en el artículo 31.bis del Código Penal español -exigible de forma restrictiva en los supuestos y delitos expresamente previstos en el mismo como *numerus clausus* (aunque un listado extensísimo, a diferencia del resto de países que han regulado también esta responsabilidad)-, tampoco resultaría aplicable a los sistemas de inteligencia artificial, ni tan siquiera a los más avanzados o "fuertes", y ello al carecer de personalidad jurídica y no disponer de la condición de "personas jurídicas" por el ordenamiento jurídico, en la medida que este régimen específico de responsabilidad penal está prevista para los delitos cometidos, en nombre o por cuenta de personas jurídicas y en su beneficio directo o

[842] A destacar las reflexiones de González Cussac sobre el fundamento de la responsabilidad penal de las personas jurídicas, el cual, tras analizar distintas posiciones doctrinales al respecto de dicho de dicho fundamento, concluye afirmando que el art. 31 bis del Código Penal español no precisa que las personas jurídicas cometan el delito. Lo que hace es establecer "un sistema en el cual, bajo ciertos requisitos (presupuestos comunes y hechos de conexión), traslada también la responsabilidad penal a la persona jurídica desde la original conducta delictiva cometida por determinadas personas físicas". GONZÁLEZ CUSSAC, J.L. (2021) "Sobre el fundamento de la responsabilidad penal de las personas jurídicas", en GALÁN MUÑOZ, A. Y MENDOZA CALDERÓN, S. (Coord.). *Derecho penal y política criminal en tiempos convulsos: libro homenaje a la Profa. Dra. Maria Isabel Martínez González*. Tirant lo Blanch. 2020. Pp. 109-120.

indirecto, por sus representantes legales o por aquéllos que, actuando individualmente o como integrantes de un órgano de la persona jurídica, estén autorizados para tomar decisiones en nombre de la persona jurídica u ostenten facultades de organización y control dentro de la misma.

Idéntica conclusión debemos alcanzar respecto de los delitos cometidos, en el ejercicio de actividades sociales y por cuenta y en beneficio directo o indirecto de las personas jurídicas, por quienes, estando sometidos a la autoridad de las personas físicas mencionadas en el párrafo anterior, hayan podido realizar los hechos por haber incumplido gravemente por aquéllos los deberes de supervisión, vigilancia y control de su actividad atendidas las concretas circunstancias del caso.

No obstante, algunos autores se han planteado la posibilidad de aplicar a los casos de utilización de robots o sistemas de inteligencia artificial avanzados la teoría de la autoría mediata, sustentándose en la "teoría de la autoría" establecida por el Tribunal Supremo español[843], que plantea que pueda haber autoría mediata, por ejemplo, en casos de incapacidad para la culpabilidad.

La base de la autoría mediata implica que el autor penalmente responsable de un delito no es, necesaria, exclusiva y excluyentemente, el que realiza la acción material que describe el tipo penal, sino que también puede serlo quién consigue que un tercero realice la acción tipificada.

De este modo, en la autoría mediata, el ejecutante de la acción la llevaría a cabo sin plena consciencia de su relevancia penal o sin pretender llevar a cabo la misma, concurriendo engaño, violencia o intimidación entre el autor material y el mediado.

Distintos autores como Nieto Mengotti[844] han reflexionado sobre dicha posibilidad, aunque este último considera que no es directamente aplicable por cuanto que, los robots

[843] Sentencia 507/2019, Tribunal Supremo, Sala de lo Penal, Sección 1, Rec. 1473/2018 de 25 de octubre de 2019; Sentencia Penal Nº 1111/2010, Tribunal Supremo, Sala de lo Penal, Sección 1, Rec 1626/2010 de 17 de Diciembre de 2010

[844] NIETO MENGOTTI, J.P. (2016). *El Derecho Penal frente a los robots*. FIDE Papers, Madrid 2016

o sistemas de inteligencia artificial más avanzada, son tratados en la actualidad como objetos muebles cualquiera que sea su grado de autonomía.

Sin embargo, según el mismo, no debemos rechazar de plano esta construcción teórica de cara al futuro, en función del avance y realidad de los sistemas de inteligencia artificial en los próximos años, aspecto con el que coincido plenamente.

De hecho, expertos de inteligencia artificial de la *National Crime Agency* británica y del *Europol European Cybercrime Center* llevan anunciado durante los últimos años el alarmante incremento de los delitos en los que intervienen y robots, y algunas predicciones estiman que para el año 2040 los robots "cometerán" más crímenes que los humanos[845]. Conforme reflejó la Europol en su informe *Exploring tomorrow's organised crime*[846] de 2015, los sistemas inteligentes son, y lo serán cada vez en mayor medida, utilizados en el marco de la delincuencia tradicional.

A mi juicio, la futura pretensión de la aplicación de esta autoría exigiría una profunda revisión de la misma, tal y como ha sido construida por la jurisprudencia, así como una revisión de las instituciones básicas del Derecho Penal y el marco actual previsto en el Código Penal español.

Según Domínguez Peco, anteriormente citada, aún el supuesto de que en el futuro se le pudiese reconocer a estos sistemas alguna capacidad subjetiva como base para poder plantear su posible responsabilidad penal, la teoría de la autoría mediata, tal y como está concebida en la actualidad, no resolvería los problemas futuros.

Conforme a las bases de esta teoría, se pretende imputar el delito al "sujeto que domina la acción más allá del ejecutor, que carece de esa responsabilidad", por lo que, de poder aplicarse en el futuro a sistemas dotados de inteligencia artificial con autonomía, la responsabilidad sería del usuario, no del robot o sistema, considerando que éste no

[845] DUNCAN, J. (2016). "Robots and computers will commit more crime than humans by 2040, expert warns". Publicado en *Mailonline*. 08 de septiembre de 2016. Recuperado de: https://www.dailymail.co.uk/news/article-3780314/Robots-computers-commit-crime-humans-2040-expert-warns.html. Consultado el 15.12.2020.

[846] Recuperado de: https://www.europol.europa.eu/sites/default/files/documents/Europol_OrgCrimeReport_web-final.pdf. Consultado el 17.12.2020.

actuaría con voluntad propia sino como mero instrumento, por lo que se aplicaría la teoría del autor material tradicional.

Al igual que expuse al abordar los retos y marcos de responsabilidad civil, en materia de responsabilidad penal, en el futuro se podrían plantear supuestos en los que el usuario podría esgrimir su argumento absolutorio en que la decisión o acción fue llevada a cabo por el sistema plenamente autónomo y con capacidad de autoaprendizaje, por "decisión" del mismo -más que por "voluntad consciente"-, y al margen de la del usuario.

Entre los posibles contrargumentos a esta postura, siguiendo el razonamiento de la precitada Domínguez Peco, se hallaría la propia aplicación de la teoría de la autoría mediata, que considero, como he referido, inaplicable en la actualidad.

En definitiva, un hipotético sistema de inteligencia artificial autónomo, actualmente carecería de consciencia, lo que comporta su inimputabilidad penal.

En base a lo expuesto, en mi opinión, actualmente no resulta posible la aplicación de esta teoría en base a la consideración actual de los sistemas de inteligencia artificial más avanzada o incluso "fuerte" o los robots, máquinas o productos dotados de la misma, como objeto y no sujeto de Derecho, junto con los demás aspectos precitados. Esta cuestión relativa a la posibilidad de atribuir responsabilidad penal a la inteligencia artificial ha sido también analizada por autores como De la Cuesta Aguado[847], Azuaje Pirela[848] o Posada Maya[849].

No puede considerarse concurrente ni el error, ni el engaño, ni la coacción ni la intimidación en la relación entre usuario y el sistema de dotado de inteligencia artificial que integran esta responsabilidad.

---

[847] DE LA CUESTA AGUADO, P.M. (2020). "Inteligencia artificial y responsabilidad penal". *Revista penal México*. ISSN 2007-4700. Nº. 16-17, 2019-2020. Pp. 51-62.

[848] NAVARRO-DOLMESTCH, R. Y VIDAL-TAMAYO I. (2020) "Sobre la justificación de aplicar el derecho penal a las entidades de inteligencia artificial". En AZUAJE PIRELA, M. Y CONTRERAS P. *Inteligencia artificial y Derecho: Desafíos y perspectivas*. Tirant lo Blanch. 2020. Pp. 261-278.

[849] POSADA MAYA, R. (2019). "La responsabilidad penal de los agentes de inteligencia artificial: entre la ficción y una realidad que se aproxima". En PORTILLA CONTRERAS, G. Y VELÁSQUEZ F. (Dir.); POMARES CINTAS, E. Y FUENTES OSORIO, J.L. (Coord.). *Un juez para la democracia. Libro homenaje a Perfecto Andrés Ibáñez*. Dykinson 2019. Pp. 561-581.

En estos supuestos habría una acción material y un dolo directo por parte del usuario que sería el único autor y sobre el mismo debe recaer la imputación en base a la comisión directa y personal de la conducta típica.

El robot o sistema dotado de inteligencia artificial avanzada sería herramienta del delito y, consiguientemente susceptible de ser decomisado como tal, conforme al artículo 127 del Código Penal español.

Sobre la responsabilidad penal de los sistemas de inteligencia artificial, merece significarse el trabajo de Gabriel Hallevy, referenciado tanto por Kingston[850] como por Martínez Rey y Pazos Sierra[851]. Según el precitado autor, la responsabilidad penal requiere tanto una acción como una intención mental, esto es, el *actus reus* y el *mens rea*. Conforme a ello, plantea tres posibles escenarios.

Hallevy habla del perpetrador por intermedio de terceros, aplicable cuando una ofensa o daño ha sido cometida por una persona mentalmente deficiente o un animal, al que se considera inimputable penalmente, en consecuencia, no culpable, pudiendo corresponder la misma a quien instruyó a la persona o animal. El autor podría ser el programador o el usuario que da instrucciones al sistema.

Del mismo modo, habla de los delitos de consecuencia natural o probable, en los que la responsabilidad podría recaer en quién pudiera haber previsto que el sistema se utilizara en la forma en la que se hizo y si era una consecuencia probable, debiendo valorar la atribución de la misma al programador o proveedor, y menos probable al usuario, salvo que las instrucciones que acompañen al producto o servicio detallen las limitaciones del sistema y las posibles consecuencias de un mal uso con un detalle inusual.

Y, por último, plantea la responsabilidad directa que requeriría una acción como una intención, planteando como ejemplo circular excediendo la velocidad permitida por parte

[850] KINGSTON, J. (2016). "Artificial Intelligence and Legal Liability" (2016). En Bramer M. y Petridis M. (Eds) *Investigación y desarrollo en sistemas inteligentes XXXIII*. SGAI 2016. Springer, Cham. https://doi.org/10.1007/978-3-319-47175-4_20.

[851] MARTÍNEZ REY, M.A. Y PAZOS SIERRA, J. (2019). "La inteligencia artificial y el derecho: Pasado, presente y futuro", en MONTERROSSO CASADO, E. (Dir.), *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch 2019. Pp. 561-562.

de un vehículo sin conductor, en el que la ley podría asignar responsabilidad penal al sistema de inteligencia artificial que dirigía la acción en ese momento, con lo que estoy absolutamente en desacuerdo por lo expuesto anteriormente. En estos supuestos el conductor podría no ser el responsable. Según estos escenarios, el precitado Kingston considera que podría serlo el programador, o si del diseñador del programa experto que aporto los conocimientos o el directivo que lo designó.

Como conclusión, siguiendo al precitado Kingston, la responsabilidad legal de los sistemas de inteligencia artificial dependería de al menos tres factores:

a) Si la inteligencia artificial es un producto o un servicio, en este escenario no estaría bien definido en los marcos vigentes y genera distintas opiniones en diferentes sentidos.

b) Si la acción es constitutiva de delito, que intención requeriría. Este autor considera poco probable que los programas de inteligencia artificial puedan infringir marcos legales que exijan que se sepa que se está cometiendo un acto delictivo, sin perjuicio de que sea muy probable que estén infringiendo leyes para las que "un hombre razonable" hubiere sabido que una actuación de aquel tipo podía dar lugar a un delito, y es casi seguro que pueda contravenir los delitos de responsabilidad objetiva.

c) Si las limitaciones de los sistemas de inteligencia artificial se comunican al adquirente. En la medida que los sistemas de inteligencia artificial tienen limitaciones tanto generales como específicas, los conflictos en esta materia deberían enfocarse desde las advertencias informadas sobre dichas limitaciones.

Adicionalmente, deberá abordarse la cuestión sobre quién debe ser considerado responsable, lo que dependerá, siguiendo a Hallevy, del modelo que se aplique, esto es, autor por otro, consecuencia natural probable o responsabilidad directa.

De nuevo, la reflexión a realizar para el futuro sobre esta relevante conclusión es si podrán llegar a ser considerados estos sistemas autores de un delito o simplemente un medio o instrumento idóneo para su comisión en el *ciberespacio* de manera masiva, desde

cualquier lugar del mundo y refugiándose en el anonimato, la dificultad de su persecución y, en ocasiones su atipicidad.

En conclusión, conforme al ordenamiento jurídico vigente en España no es posible imputar responsabilidad penal a un sistema de inteligencia artificial, ya sea más avanzado o "fuerte", sin perjuicio de su utilización como medio o instrumento para su comisión.

3. La inteligencia artificial como instrumento para la comisión de delitos

De manera congruente con lo expuesto hasta este momento, conforme he analizado en el apartado anterior, los sistemas de inteligencia artificial podrán ser utilizados como instrumento para la comisión de un delito por parte de un sujeto, sobre el que recaerá la imputación de la responsabilidad penal y sus consecuencias, en base al dolo o imprudencia concurrente.

No obstante, con el aumento de la supuesta autonomía, capacidad de aprendizaje e impredecibilidad de los sistemas de inteligencia artificial, se producirán situaciones que podrían exigir en el futuro una revisión de las teorías del delito actuales.

La decisión o acción de un sistema de inteligencia artificial podrá venir determinada por acciones llevadas a cabo por el ser humano o por el propio sistema, esto es, por su diseño y programación, por las instrucciones predefinidas, su interacción con el entorno y por su capacidad de autoaprendizaje principalmente, pero también podrá verse afectada por errores de programación, defectos funcionales, automatismos preconcebidos, restricciones no dadas, inconsciencia, entrenamiento y formación posterior, órdenes del usuario, engaño por parte de éste (por ejemplo presentando como reales datos de entrada o situaciones falsas), factores medioambientales, caso fortuito o fuerza mayor.

La conducta del operador o usuario humano que utiliza un sistema de inteligencia artificial como instrumento para la comisión de un delito podrá sustentarse en un dolo directo, dolo eventual o en la denominada culpa con previsión.

El dolo directo exige dos elementos sustanciales: De un lado, el elemento intencional, volitivo o voluntario, es decir, lo que el sujeto quiere hacer, y de otro, el elemento cognitivo, cognoscitivo o intelectual, es decir, el saber o conocer que esa acción es ilícita.

Como he referido anteriormente, no cabría apreciar la concurrencia de este dolo en el propio sistema de inteligencia artificial en la actualidad, ya sea de naturaleza "fuerte" o dotado de una supuesta autonomía plena, ante la ausencia real de consciencia y voluntad asociada.

La decisión o acción autónoma por parte de un sistema de inteligencia artificial o que utilice otros sistemas de esta naturaleza para accionar, partiendo de la ausencia de personalidad de jurídica y de su inimputabilidad, deberá analizarse previamente si podría incardinarse dentro del caso fortuito o fuerza mayor como hipotética causa eximente de responsabilidad penal, identificando si el usuario tenía un control de la acción y si la acción era previsible o podría enmarcarse dentro de la impredecibilidad. En base de todo ello deberá determinarse la posible concurrencia de dolo eventual o, en su caso, de imprudencia punible.

El dolo eventual comportaría que el sujeto, aun sabiendo el resultado y el daño que puede provocar una determinada acción, continúa la misma y no descarta el resultado que puede llegar a ocurrir. Es decir, en el dolo eventual, el sujeto ve el resultado solo como probable, pero lo acepta para el caso de que se produzca.

La doctrina jurisprudencial del Tribunal Supremo español[852] estima que concurre dolo eventual "en quien, conociendo que genera un peligro concreto jurídicamente desaprobado, actúa voluntariamente, no obstante, y realiza la conducta que somete a la víctima a un riesgo de producción altamente probable, que el agente no tiene la seguridad de poder controlar, por lo que, sin perseguir directamente la causación del resultado comprende que existe un elevado índice de probabilidad de que su acción lo produzca". Es decir, que el conocimiento de la posibilidad de que se produzca el resultado y la conciencia del alto grado de probabilidad de que realmente se produzca caracteriza la

[852] Sentencias del Tribunal Supremo nº 11/2017 de 19 de enero de 2017 (Rec. 10371/2016) y nº 301/2011 de 31 de marzo de 2011 (Rec. 1414/2010).

figura del dolo eventual desde el prisma de la doctrina de la probabilidad o representación, frente a la *teoría del consentimiento* que centra en el elemento volitivo -asentimiento, consentimiento, aceptación, conformidad o, en definitiva, "querer" el resultado- el signo de distinción respecto la culpa consciente.

Según el Alto Tribunal, "afirmando que la aceptación del resultado existe cuando el autor ha preferido la ejecución de la acción peligrosa a la evitación de sus posibles consecuencias no se rompe, en realidad, con la teoría del consentimiento, tratándose en el fondo de una cuestión probatoria: el dolo requiere, en cualquier caso, conocimiento y voluntad, pero la voluntad se infiere del hecho de que, conociendo el agente el peligro generado por su acción y la elevada probabilidad de causación de un resultado, decida voluntariamente actuar, de lo que cabe deducir que acepta o asume el resultado que acaba derivándose de su voluntaria decisión".

Según Bullemore y Mackinnon[853] *"*la diferencia radical entre el dolo directo y el eventual se encuentra en que en el primero la verificación del tipo objetivo es la meta de la voluntad del sujeto y su conducta es el medio escogido para tal verificación, en tanto que en el dolo eventual la verificación del tipo penal es una consecuencia previsible de la propia conducta del agente, pero que no es perseguida ex profeso por el agente, de tal modo que su conducta no es un medio escogido para arribar a tal verificación, sino el medio para alcanzar otros fines, siendo la verificación del tipo objetivo solo una consecuencia colateral previsible y ante la cual el sujeto activo manifiesta indiferencia para el caso de producirse".

En definitiva, el dolo es eventual cuando, de un lado, el sujeto se representa el resultado como relativamente probable y, de otro, incluye esa probabilidad de que se materialice el resultado en la voluntad realizadora[854].

---

[853] BULLEMORE V.R. Y MACKINNON, J. (2004). "Fin y Función del Derecho Penal y de la pena: las teorías de la pena". en *Anales Facultad de Derecho Universidad de Chile*, N.1. 2004. Pp. 5-6.

[854] Sobre el dolo, ante la ausencia de definición en el Código Penal español, y sus clases, destacar a Vives Antón, citado por González Cussac en Compendio de Derecho Penal, Parte General, 4ª edición, de 2014, referenciado a continuación. Según el mismo, la doctrina española tradicionalmente el dolo ha sido entendido como un proceso psicológico. Frente a ello se alzan nuevas concepciones de naturaleza normativa. Dentro de esta orientación, Vives Antón propone su entendimiento en términos estrictamente normativos, como compromiso. Así, para averiguar si existió una intención concreta, tendremos que examinar las reglas sociales y jurídicas que definen su acción (por ejemplo, matar) y ponerlas en relación

Sin embargo, la denominada "culpa con previsión" se sustenta sobre un elemento volitivo, en la medida que el sujeto es consciente del resultado como probable, pero no lo acepta, espera que no se produzca, es decir, una falta de "diligencia intencional".

La posible casuística es inabordable en el seno de esta investigación -de objeto y alcance limitados-, y puede ser casi infinita en atención a las características de cada sistema de inteligencia artificial, aplicación, sector y contexto de uso, si bien, la posibilidad de que el sistema tenga capacidad de aprendizaje y autonomía con posibilidad de tomar decisiones o realizar acciones con objetivos predefinidos, pero sin necesidad de circunscribirse a las instrucciones inicialmente dadas, y de manera "supuestamente" impredecible, podría permitir valorar la posible concurrencia de caso fortuito o fuerza mayor. Idéntica valoración se podría llevar a cabo cuando es el propio sistema el que pudiera haberse reprogramado de manera ajena y fuera del control del responsable.

Las mayores dificultades se plantearán, de un lado, en determinar la concurrencia de dolo eventual o culpa con previsión, identificando si el sujeto conocía el riesgo de estas circunstancias en el momento de iniciar su uso y como actuó una vez consciente de la desviación de la conducta supuestamente esperada por el mismo y, de otro, en la prueba.

Y, además, dicha casuística exigirá valorar la posible existencia o concurrencia de responsabilidades por parte de otros sujetos ajenos inicialmente a la conducta delictiva,

---

con las competencias del autor (las técnicas que domina). Solo de esta forma podremos determinar lo que efectivamente sabía, esto es, lo que podía ser capaz de entender (o sea, si dominaba una técnica). Ha de tenerse en cuenta que el dolo no puede demostrarse entrando en la mente del autor y viendo su intención. Únicamente es posible juzgarlo por sus manifestaciones externas y de éstas sí podemos averiguar los conocimientos del autor, las técnicas que dominaba, lo que podía y no podía prever o calcular y, entonces ya podremos entender sus intenciones expresadas en la acción. El entendimiento normativo del dolo se proyecta en el elemento volitivo, de modo que le querer del autor no se identifica con sus deseos, sino que reside en la acción misma. Si por tanto la voluntad se expresa en el mismo actuar del sujeto, ya no se puede explicar como un proceso natural (psicológico), sino en términos normativos de compromiso con la acción, y, en consecuencia, con un compromiso con la lesión del bien jurídico protegido (STC 68/1998). A continuación distingue dolo directo (de primer grado o intención y de segundo grado o de consecuencias necesarias) de dolo eventual que es "cuando el autor se representa como probables las consecuencias de su comportamiento y, no obstante, decide actuar asumiéndolas" que en ocasiones es muy complicado distinguirlo de la imprudencia.

Por cuanto a la imprudencia y su estructura, se define negativamente con relación al dolo y sus elementos son: a) ausencia de intención, se dice que el que así actúa lo hace "sin compromiso con el resultado típico", b) infracción de un deber específico de cuidado subjetivo, de un proceder diligente conforme a la experiencia, las normas socio-culturales y la normativa vigente (previsibilidad del riesgo y del resultado) y c) evitabilidad de producción del resultado: análisis de que hubiera pasado si el sujeto hubiese obrado conforme al deber de cuidado. ORTS BERENGUER, E. Y GONZÁLEZ CUSSAC, J.L. (2014). *Compendio de Derecho Penal. Parte General*. 4ª edición. Tirant lo Blanch 2014. P. 306.

como el diseñador, el desarrollador, el fabricante o propietario, que se abordará en el siguiente apartado.

El caso fortuito o la fuerza mayor no plantean específicas consideraciones respecto de su aplicación como factores externos que no hubieran podido preverse o que previstos fueran inevitables, es decir situaciones imprevisibles o inevitables en las que podría no imputarse su causa a la acción u omisión del sujeto.

Como expuse al analizar los regímenes generales de responsabilidad civil, el caso fortuito se identifica con un suceso que no pudo preverse, pero de haberse previsto se hubiera podido evitar, mientras que la fuerza mayor es un suceso inevitable, se hubiera o no previsto.

El Tribunal Supremo español[855] considera que en el derecho moderno y, naturalmente, en el español, para que el caso fortuito o la fuerza mayor tengan virtud de exoneración, se requiere la simultánea presencia de dos presupuestos: Hechos imprevisibles o, aun previstos, inevitables, y la inimputabilidad al deudor respecto al accidente fortuito y sus causas. Dentro de lo supuestos de fuerza mayor podría considerarse la participación de la víctima en determinados supuestos[856].

En mi opinión, las mayores dificultades se plantearán a la hora de probar su concurrencia, que requerirá pruebas periciales a través de expertos cualificados que puedan determinar la existencia o no de dichas circunstancias causantes de la acción, lo que resulta complejo especialmente ante las propias características y capacidades de las que pueda estar dotado un sistema, sin perjuicio de que los futuros marcos reguladores de los requisitos éticos y jurídicos de la inteligencia artificial, exijan la información, transparencia, explicabilidad y derecho de acceso, de modo que se facilite la prueba en caso de exigencia de responsabilidad.

[855] Sentencia del Tribunal Supremo 06/04/1987.

[856] Sentencia del Tribunal Supremo 11/07/1990: "Entre la acción u omisión del agente y el resultado dañoso debe existir una relación directa, sin interferencias de otras posibles conductas o eventos ajenos al agente y, más concretamente, no debe interferir ninguna acción negligente por parte de las víctimas. En este sentido, jurisprudencia precedente matiza que la teoría del riesgo no descansa en la mera causación de un evento físico dañoso, ya que si la víctima se interfiere en la cadena causal quedará el agente exonerado de responsabilidad, por tratarse de un suceso imprevisto o inevitable".

En este sentido, destacar el marco de requerimientos y obligaciones que establece la Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[857], si bien, respecto de sistemas de inteligencia artificial considerados de alto riesgo y no respecto del resto.

4. Sujetos responsables

El usuario de un sistema de inteligencia artificial podrá ser considerado autor de determinados delitos mediante la imputación de la responsabilidad penal por dolo directo o imprudencia en función de los elementos concurrentes y contexto.

Sin embargo, ¿los diseñadores, desarrolladores o fabricantes de un sistema de inteligencia artificial podrían ser considerados responsables penalmente de las acciones supuestamente autónomas de los mismos, especialmente llevadas a cabo al margen o en contra de las instrucciones del usuario?

El objeto limitado de esta investigación y la inagotable casuística que se puede producir en la práctica me impide abordar con detalle estas cuestiones, si bien, analizaré algunos de los escenarios generales que nos podemos encontrar.

Nos podemos hallar, entre otros contextos, ante sistemas de inteligencia artificial concebidos en su diseño y programación para tomar decisiones o realizar acciones autónomas fuera del control del usuario y con capacidad de autoaprendizaje, sistemas de inteligencia artificial que adicionan a todo ello su capacidad de reprogramación autónoma al margen de la inicial, sistemas de inteligencia artificial con capacidad para tomar decisiones o realizar acciones autónomas, con capacidad de aprendizaje autónomo y sometidos a cierta supervisión y control del usuario, o sistemas de inteligencia artificial con capacidad para tomar decisiones o realizar acciones autónomas sometidos a la

[857] COM (2021) 206 final 2021/0106 (COD)

supervisión y control del usuario y con capacidad de aprendizaje también bajo el control de éste.

El grado de impredecibilidad es distinto en cada supuesto y su concurrencia o consciencia debería afectar directamente a la determinación de la responsabilidad penal.

¿Qué ocurre si el sistema toma una decisión o realiza una acción al margen de la programación prevista por el usuario, de la descripción técnica, de sus instrucciones, de su entrenamiento, de su control, o en contra de las instrucciones del usuario, determinando la comisión de un delito?

La falta de control sobre el riesgo y dominio sobre el hecho, entre otros factores, puede dificultar o imposibilitar su imputabilidad, sin perjuicio de la concurrencia de ausencia de intencionalidad y la posible imprudencia en el diseño y programación que debía empoderar al usuario o, cuanto menos, asegurar su supervisión y control, ya sea previa, coetánea o posterior a la toma de la decisión o ejecución de la acción.

En este sentido, considero que la información que acompañe al sistema por parte del desarrollador y fabricante, especialmente sobre sus riesgos -conforme exige la nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, aunque respecto de sistemas de alto riesgo- debe tener igualmente relevancia jurídica también en el ámbito penal para alejar a los mismos de una posible responsabilidad penal, en caso de conocimiento efectivo de los riesgos por parte del propietario o usuario.

En cualquier caso, en mi opinión, de conformidad con lo previsto en el artículo 28 del Código Penal, diseñadores, desarrolladores o fabricantes no encajan inicialmente en las categorías previstas en el mismo, en la medida que no realizarían la acción punible ni por sí mismos ni por medio de otros de los que se servirían como instrumento. Tampoco encajarían en la condición de inductores o cooperadores.

Y refiero "inicialmente", en la medida que podría darse la circunstancia de que un sistema de inteligencia artificial pueda diseñarse, compilarse y/o construirse por su diseñador, su

desarrollador y/o su fabricante para cometer un delito, ya sea por sí mismo o a través de un tercero, como podría ser un usuario.

A modo de ejemplo, pensemos en sistemas involucrados en la gestión contable y fiscal de una organización, donde no solamente pueden derivarse responsabilidades penales sino administrativas para el propio desarrollador o comercializador al amparo de los nuevos marcos propuestos en España, como ya existe en otros países europeos, en particular mediante el Proyecto de Ley de medidas de prevención y lucha contra el fraude fiscal, de transposición de la Directiva (UE) 2016/1164, del Consejo, de 12 de julio de 2016, por la que se establecen normas contra las prácticas de elusión fiscal que inciden directamente en el funcionamiento del mercado interior, de modificación de diversas normas tributarias y en materia de regulación del juego, cuya tramitación se inicia en fechas coincidentes con el cierre de esta investigación en el Senado.

La norma propuesta aborda la reforma de Ley General Tributaria española e introduce la prohibición del denominado *software* de doble uso, prohibiendo el desarrollo, la tenencia o el uso de sistemas y aplicaciones informáticas de gestión contable que permitan alterar y falsear la información contable de las empresas.

No se trata de una cuestión nueva, dado que precisamente fue una cuestión que tuve la ocasión de analizar en una conferencia impartida en marzo de 2019[858], donde precisamente analicé todas estas cuestiones desde una visión global de responsabilidad desde un punto de vista administrativo, fiscal, mercantil-civil, administrativo y penal la responsabilidad actual y futura de consultoras, desarrolladoras y comercializadoras de software contable, de facturación o de gestión empresarial.

Prosiguiendo con mi análisis penal, la ciberdelincuencia está utilizando los medios y tecnologías más vanguardistas para llevar a cabo su actividad a escala mundial y la inteligencia artificial no ha quedado al margen de sus instrumentos.

[858] MUÑOZ VELA, J. M. (2019). Conclusiones finales en la Conferencia impartida el 25.03.2019 en la Universidad Politécnica de Valencia, organizada por el Instituto Tecnológico de Informática (ITI), bajo el título "Responsabilidad de consultoras, productoras y comercializadoras de software contable, de facturación o de gestión empresarial: Situación actual y tendencias regulatorias".

La creación de un sistema inteligente que bajo la apariencia de inocuidad pudiere ser utilizada por múltiples o incluso millones de usuarios conectados a una red pública para la comisión de delitos podrían incardinarse en estos supuestos. A modo de ejemplo, pensemos a ataques de denegación de servicio utilizando los sistemas precitados para administrar la conexión y equipo de múltiples usuarios.

En relación con sujetos responsables, el Código Penal español prevé algunos supuestos de autoría especial en su parte general.

El artículo 30.1 del Código Penal introduce un supuesto especial de autoría respecto de los delitos que se cometan utilizando medios o soportes de difusión mecánicos respecto de hechos en los que puedan estar involucrados terceros vinculados al hecho delictivo y relacionados con el dominio de la acción, en el que se establece que no responderán criminalmente ni los cómplices ni quienes los hubieren favorecido personal o realmente. No obstante, no ofrece una solución a la cuestión planteada.

Del mismo modo, la responsabilidad de los responsables de las personas jurídicas y la responsabilidad de éstas como entes con personalidad jurídica, en los supuestos expresamente previsto en el Código Penal español, constituyen otros supuestos especiales de autoría. En todos ellos, la atribución de la responsabilidad penal está directamente relacionada con su dominio de la acción y se hallan vinculados al hecho delictivo. No obstante, no resulta aplicable este marco ante la ausencia de la condición indicada.

Asimismo, el Código Penal español también prevé en su parte especial otros supuestos de autoría especial, considerando penalmente responsables a sujetos que no realizan la acción material pero que crean una situación de riesgo con relevancia penal, produzca o no un resultado lesivo, dejando de cumplir una obligación de diligencia y cuidado, pero tampoco se ajustaría a los sistemas de inteligencia artificial autónomos por lo expuesto anteriormente.

En definitiva, el marco legal vigente en materia penal imposibilita inicialmente la imputación de responsabilidad penal a diseñadores y programadores de sistemas de inteligencia artificial, salvo que exista dolo directo por parte de éstos y dominio de la

acción delictiva conforme a los ejemplos planteados, sin perjuicio de la que corresponda al usuario.

Ello no implica la posibilidad de analizar su posible responsabilidad civil, especialmente en el caso de incumplir los marcos y normas éticas o jurídicas que sean consideradas vinculantes para los mismos y que conformen su diligencia debida en el futuro, sobre lo que ya se hicieron algunas propuestas por el propio Parlamento Europeo, como he expuesto, especialmente desde su Resolución de 16 de febrero de 2017 relativa a la creación un código de conducta ética en el campo de la robótica, integrado por un *Código de Conducta Ética* para los ingenieros en robótica, un *Código Deontológico* para los comités de ética de la investigación, un modelo de licencia para los usuarios -como conjunto de derechos y obligaciones de los mismos- y un modelo de licencia para los diseñadores de robots o sistemas dotados de inteligencia artificial avanzada.

El *Código de Conducta Ética* para los ingenieros en robótica, conforme abordé en el apartado 3 del capítulo III de esta investigación, incorporaba principios como el de beneficencia[859], no perjuicio o maleficencia[860], así como normas de rendición de cuentas, seguridad y reversibilidad, entre otras.

La reversibilidad, tal y como fue concebida en la Resolución precitada, constituiría una condición necesaria para el control pretendido de los sistemas de inteligencia artificial supuestamente autónomos y un concepto fundamental en su programación para que se comporten de manera segura y fiable. Según fue definida en el código de conducta precitado, la reversibilidad indicaría al robot o sistema dotado de inteligencia artificial avanzada qué acciones son reversibles y, en su caso, el modo de invertirlas. De este modo, su programación permitiría al usuario anular las acciones no deseadas.

La licencia para los diseñadores que igualmente integraba el anexo de la precitada Resolución, recogía un conjunto de principios y normas que los diseñadores de un sistema deben considerar, dentro de la denominada *Ethics by design*.

[859] Los robots deben actuar en beneficio del hombre.
[860] Los robots no deberían dañar o perjudicar a las personas.

Entre su principios y normas recogía la necesidad por parte de los diseñadores de no perjudicar, herir, engañar o explorar a los usuarios (vulnerables), de garantizar que un robot funcione de modo conforme con los principios éticos y jurídicos a nivel local, nacional e internacional, garantizar la transparencia, la trazabilidad y la reconstrucción en la toma de decisiones, de analizar la previsibilidad de un sistema humano-robot teniendo en cuenta la incertidumbre en la interpretación y en la acción, así como los posibles fallos de los robots o del hombre, entre otros aspectos.

Recientemente, el Parlamento Europeo aprobó su Resolución de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas[861], que incorpora la Propuesta Reglamento del Parlamento europeo y del Consejo sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas, que fue objeto de análisis en el capítulo III y que establecía el carácter vinculante de un conjunto de principios y normas éticas y obligaciones jurídicas a observar en el diseño, desarrollo, despliegue y uso de sistemas de inteligencia artificial avanzada.

El Reglamento propuesto exige una inteligencia artificial con origen y centrada en el ser humano, sometida a la supervisión y control humano, conforme a los marcos reguladores, segura, resiliente, fiable, transparente, respetuosa con los derechos fundamentales, trazable, reproductible, auditable, que rinda cuentas, responsable socialmente o respetuosa con el medio ambiente. Sin embargo, no recoge expresa y específicamente su reversibilidad, sin perjuicio de pueda formar parte de ese control y supervisión humana en su diseño, despliegue y uso.

Asimismo, comporta otras obligaciones asociadas, como la evaluación de sus riesgos, su evaluación de conformidad o su certificación de conformidad ética. En este sentido, considero incompatible con este certificado la ausencia de reversibilidad que garantice

[861] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL))

esa supervisión y control humano, previo, coetáneo y posterior, es decir, en todo momento durante su diseño, despliegue y uso, esto es, durante todo el ciclo de vida.

De esta propuesta destacar que los requerimientos y exigencias para cualquier sistema de inteligencia artificial fuera de nivel alto o no, incluían aspectos como el respeto a la dignidad, autonomía y seguridad humana o el respeto de los derechos fundamentales, incluyendo el derecho a la protección de datos personales, a diferencia de la nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, que se focaliza en la prohibición de determinados sistemas y en los requisitos y obligaciones para los sistemas de alto riesgo, pero sin una declaración expresa de las normas éticas y jurídicas que deben regir el resto de sistemas, a excepción de las obligaciones de transparencia e información para determinados sistemas.

Del mismo modo, el Parlamento Europeo aprobó en la misma fecha de aquella, su Resolución de 20 de octubre de 2020, sobre un régimen de responsabilidad civil en materia de inteligencia artificial[862], que fue también objeto de análisis en el capítulo anterior y que incluye una Propuesta de Reglamento relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial, en el que igualmente se prevén nuevas obligaciones legales, como por ejemplo, el registro de los sistemas de inteligencia artificial, la suscripción de seguros de responsabilidad civil obligatoria y la creación de fondos de compensación para reparación en caso de daños.

Por último, la reciente Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley de inteligencia artificial) y modificando determinados actos legislativos de la Unión[863], regula todo un conjunto de requerimientos de los sistemas y obligaciones para proveedores y demás sujetos participantes, incluyendo aquéllas que debe incorporarse en su diseño, si bien, respecto de los sistemas considerandos conforme al mismo de alto

[862] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial (2020/2014(INL))
[863] COM (2021) 206 final 2021/0106 (COD)

riesgo, remitiendo el resto al cumplimiento de las obligaciones de transparencia y a la adscripción voluntaria a futuros códigos de conducta.

Los futuros marcos reguladores de la inteligencia artificial en el ámbito civil, actualmente en proceso de análisis, reflexión y tramitación, deberían motivar también una reflexión desde un enfoque penal, no tanto quizás para replantearse la personalidad jurídica e imputabilidad penal de los sistemas de inteligencia artificial más avanzada o "fuerte", dado el estado actual de la tecnología y la previsión actual de su desarrollo en los próximos años, sino más bien una reflexión focalizada en la revisión y, en su caso, posible adición de nuevos tipos penales en atención a los nuevos riesgos que puede conllevar la inteligencia artificial y su interacción con otras tecnologías y medios, especialmente sobre bienes jurídicos de especial relevancia.

A mi juicio, la evaluación, certificación y, en su caso el futuro registro de los robots y sistemas dotados de inteligencia artificial que se está barajando en el ámbito civil pueden convertirse en elementos o aspectos relevantes en el ámbito penal para la concepción de nuevos tipos penales, de modo que su ausencia pueda determinar el nacimiento del riesgo y de una posible responsabilidad penal.

Se trataría de un control previo como destacaba Domínguez Peco[864] en relación con esta cuestión y que se remite a las reflexiones de Pagallo[865] sobre inteligencia artificial y robots "malos", en las que este autor mantiene la conveniencia de realizar pruebas de los robots en laboratorios que simulen la vida humana, siguiendo el modelo implementado por el gobierno japonés.

El propietario del sistema inteligente cuando se comete un delito mediante su uso ¿podría ser responsable penalmente? De inicio, siguiendo los principios generales de Derecho Penal, el propietario de un robot o sistema dotado de inteligencia artificial no podría incurrir en responsabilidad penal, salvo que haya conocido y quiera la acción delictiva,

[864] DOMÍNGUEZ PECO, E. M. (2018). "Los robots en el Derecho Penal", en BARRIO ANDRÉS, M. (Director), *Derecho de los robots*. Wolkers Kluwer. Madrid 2018.
[865] PAGALLO, U. (2017). "AI and bad robots. The criminology of automation", en HOLT, T. (Coord.). *The Routledge Handbook of Technology, Crime and Justice*. Ed. Routledge, Londres, 2017.

contribuyendo a que la misma se lleve a cabo mediante la puesta a disposición del objeto, instrumento o medio.

Ello no obsta a que el mismo pueda ser considerado responsable civil, como he analizado en el capítulo anterior, especialmente conforme a los futuros marcos europeos en materia de responsabilidad civil comentados en el mismo.

No obstante, de mismo modo en que en el ámbito civil y administrativo se creará un marco específico para la inteligencia artificial, especialmente por lo que se refiere a la responsabilidad derivada de los daños causados por ésta, a mi juicio, se debería llevar a cabo una reflexión profunda sobre la necesidad de revisar el marco penal en relación con la misma.

En ese sentido, la responsabilidad civil derivada de un delito se regula en los artículos 116, siguientes y concordantes del Código Penal español, los cuales contemplan múltiples supuestos específicos de responsabilidad, incluyendo a las aseguradoras, si bien, por la propia naturaleza del Derecho Penal, no considero adecuado ni procedente hacer una interpretación analógica o extensiva de dichos preceptos para la inclusión de sujetos no recogidos en los mismos, y los propietarios de un sistema dotado de inteligencia artificial no aparecen reflejados en dichos preceptos.

El artículo 120.5.º del Código Penal español establece que las personas naturales o jurídicas titulares de vehículos susceptibles de crear riesgos para terceros son civilmente responsables por los delitos cometidos en la utilización de aquellos por sus dependientes o representantes o personas autorizadas.

Es precisamente este precepto el que podría permitir considerar hoy civilmente responsable al propietario de un vehículo dotado de un sistema de inteligencia artificial, derivada de los delitos cometidos en la utilización de aquellos por sus dependientes o representantes o personas autorizadas, que podría ser el operador remoto o conductor asistido de un vehículo automatizado.

En estos supuestos, es el operador o usuario del "vehículo" el que podría ser considerado el determinante de la acción delictiva, no la máquina, por el momento inimputable

penalmente por lo referido anteriormente. Sin embargo, si la conducta delictiva es cometida por el propio sistema al margen y sin intervención del dependiente, representante o persona autorizada, estaríamos ante un supuesto de inimputabilidad y, de manera consecuente, no nacería la responsabilidad civil del delito, sin perjuicio de la reclamación de la responsabilidad civil correspondiente que proceda.

En cualquier caso, los sistemas dotados de inteligencia artificial que no pudieran ser considerados "vehículos"[866] a estos efectos, quedan inicialmente fuera de esta relación, lo que debería llevar a una reflexión profunda, como he referido anteriormente, sobre la necesidad de revisar y, en su caso, modificar los marcos penales vigentes y, en especial, el Código Penal español, para la modificación de algunos de sus preceptos al objeto de adaptarlo al contexto y realidad actual.

Por último, considero relevante efectuar una reflexión sobre los supuestos en los que el propietario sea una persona jurídica.

Cuando el propietario sea una persona física, será habitual que el mismo sea igualmente usuario del sistema, pero cuando se trata de una persona jurídica, ésta no será el usuario sino alguno de sus administradores, representantes, gestores, empleados, dependientes o usuarios supuestamente autorizados.

En este contexto resulta de indudable interés el particular régimen de responsabilidad penal de las personas jurídicas incorporado en el Código Penal español, en particular en su artículo 31 bis, en virtud del cual, se considera penalmente responsable de determinados delitos a la persona jurídica cuando se cometan en su nombre o por su cuenta y en su beneficio por sus representantes o personas con capacidad para tomar decisiones, o por quienes puedan realizar la acción delictiva por el incumplimiento, por parte de las personas jurídicas, de sus deberes de vigilancia y control.

El precitado artículo establece que en los supuestos previstos en el Código Penal español, las personas jurídicas serán penalmente responsables de: a) Los delitos cometidos en

[866] Conforme al Diccionario panhispánico del español jurídico, el concepto general "vehículo" haría referencia, tanto al medio de desplazamiento o de transporte de personas o cosas como al aparato apto para circular por las vías o terrenos a que se aplica la normativa de tráfico y seguridad vial.

nombre o por cuenta de las mismas, y en su beneficio directo o indirecto, por sus representantes legales o por aquellos que actuando individualmente o como integrantes de un órgano de la persona jurídica, están autorizados para tomar decisiones en nombre de la persona jurídica u ostentan facultades de organización y control dentro de la misma; b) Los delitos cometidos, en el ejercicio de actividades sociales y por cuenta y en beneficio directo o indirecto de las mismas, por quienes, estando sometidos a la autoridad de las personas físicas mencionadas en el párrafo anterior, han podido realizar los hechos por haberse incumplido gravemente por aquéllos los deberes de supervisión, vigilancia y control de su actividad atendidas las concretas circunstancias del caso.

A diferencia los antecedentes legislativos en otros países, como Reino Unido a través de la denominada *Bribary Act 2010* o Italia a través de su Decreto Legislativo 231/2001, de 8 de junio de 2001, el número de delitos de los que puede ser responsable la persona jurídica en España es amplísimo, aunque sea efectivamente una lista cerrada (*numerus clausus*).

Entre otros, esta lista incluye en España delitos contra la privacidad y datos reservados (Artículos 197 del Código Penal español -CP español-), acceso no autorizado a sistemas de información (197 bis CP español), descubrimiento y revelación de secretos (Artículos 197, 278-280 CP español), delitos contra la propiedad intelectual e industrial (Artículos 270-277 CP español), delito de cohecho (corrupción) entre particulares (Artículo 286 bis CP español), estafa (Artículos 248-251 bis CP español) e insolvencias punibles (Artículos 257-261 bis CP español), daños informáticos (Artículo 264 CP español), obstaculizar o interrumpir el funcionamiento de un sistema informático ajeno sin autorización (264 bis CP español), delitos contra la Hacienda Pública y la Seguridad Social (Artículos 305, 307 CP español), fraude de subvenciones (Artículo 308), delito contable (Artículo 319 CP español), delito contra los derechos de los trabajadores (Artículos 311-318 CP español), delitos contra los recursos naturales y el medio ambiente (Artículo 325 CP español), delito de establecimiento de depósitos o vertederos tóxicos (Artículo 328 CP español) o delitos de contrabando en ciertos supuestos (Ley Orgánica 6/2011 de 30 de junio, por la que se modifica la Ley Orgánica 12/1995, de 12 de diciembre, de represión del contrabando).

No obstante, la persona jurídica se halla obligada a la implantación de un sistema efectivo de prevención de delitos o *compliance program*, lo que le permitiría la exención o atenuación de su responsabilidad penal en determinados supuestos.

De los distintos delitos previstos en la parte especial del Código Penal español en los que se prevé la responsabilidad penal de la persona jurídica, debo destacar algunos en los que será más común que puedan estar "implicados" sistemas de inteligencia artificial, en especial, los delitos contra la privacidad y datos reservados (Artículos 197 del *Código Penal* español), acceso no autorizado a sistemas de información (197 bis *Código Penal* español), descubrimiento y revelación de secretos (Artículos 197, 278-280 *Código Penal* español) o en los artículos 264, 264 bis y 264 ter del Código Penal, en relación con sus artículos 264 quáter y 31 bis, es decir, delitos de *cracking* o daños informáticos, obstaculización o interrupción de un sistema informático (ataques masivos de denegación de servicio DoS o DDoS, o redes *botnet*) y el de facilitación de herramientas informáticas dañinas que se analizarán posteriormente.

Pero podrían también esta involucrados en delitos económicos, contables, fiscales, contra la propiedad intelectual, etc.

En caso de comisión de cualquiera de estos delitos, la persona jurídica podría ser además declarada responsable por no disponer de un sistema efectivo de prevención de delitos, lo que podría conllevar para la misma la imposición de penas pecuniarias, esto es, de multa, o interdictivas que, en función de las circunstancias, podrían consistir o conllevar su disolución, suspensión de sus actividades, clausura de locales y establecimientos, prohibición de realizar en el futuro su actividades, inhabilitación para obtener subvenciones y ayudas públicas, para contratar con el sector público y para gozar de beneficios e incentivos fiscales o de la Seguridad Social, o su intervención judicial, tal y como prevé el artículo 33.7 del Código Penal.

En caso de demostrar la existencia y aplicación efectiva de un modelo de prevención de delitos, como he referido, dicha responsabilidad podría verse atenuada o eximirse de la misma, en cuyo caso la responsabilidad penal se imputaría exclusivamente al sujeto responsable.

En estos supuestos, la persona jurídica podría responder directamente de esta nueva responsabilidad penal prevista para la misma. Del mismo modo, la responsabilidad penal de una persona jurídica comportará su responsabilidad civil de forma solidaria con las personas físicas que fueren condenadas por los mismos hechos, conforme a los artículos 116.3 y 119 del Código Penal español.

5. Inteligencia artificial y *ciberdelitos*.

Conforme he expuesto, la inteligencia artificial puede tener un papel protagonista como medio o instrumento para la comisión de determinados delitos más vinculados a la tecnología, tanto a nivel cualitativo como cuantitativo. A continuación, analizaré sucintamente algunos de los que considero tienen una relación más estrecha con la misma.

5.1. *Hacking* de intrusión informática o intrusismo informático

El denominado *hacking* se encuentra regulado como delito en el artículo 197.bis.1 del Código Penal español, el cual establece que, el que por cualquier medio o procedimiento, vulnerando las medidas de seguridad establecidas para impedirlo, y sin estar debidamente autorizado, acceda o facilite a otro el acceso al conjunto o una parte de un sistema de información o se mantenga en él en contra de la voluntad de quien tenga el legítimo derecho a excluirlo, será castigado con pena de prisión de seis meses a dos años.

La conducta típica conforme al precitado artículo, consiste en acceder o facilitar a otro el acceso al conjunto o una parte de un sistema de información o mantenerse en el mismo, en contra de la voluntad de su propietario o administrador, sin estar autorizado y vulnerando las medidas de seguridad establecidas para impedirlo.

El bien jurídico protegido estaría constituido por la intimidad informática y la seguridad de los sistemas de información.

Conforme está concebido, se trata de un delito de peligro abstracto y de comisión dolosa, de modo que se requiere que el sujeto sea consciente y pretenda acceder, mantener o facilitar el acceso a un sistema.

En ese sentido, conforme he analizado en los anteriores apartados, en la medida que se exige dolo y consciencia, el sistema de inteligencia artificial no sería imputable desde el punto de vista penal, el cual, además, carecería de personalidad jurídica.

Sin embargo, estos sistemas pueden constituir el instrumento idóneo por sus capacidades -especialmente ante el potencial de los mismos para realizar cálculos matemáticos complejos y aumento de su capacidad de procesamiento- para la comisión de este tipo de delitos, pudiendo ser utilizados tanto con finalidades lícitas y legítimas desde un punto de vista ético y legal, como para todo lo contrario, por parte del lado más oscuro del *ciberespacio*.

Y, es más, incluso el propio sistema de inteligencia artificial podría ser el sistema y objeto afectado por esta intrusión, conforme analicé en anteriores capítulos, es decir, sería la "víctima", aunque carente de personalidad jurídica y, en consecuencia, no siendo posible reconocerle dicho atributo.

La intrusión en estos sistemas podría comportar riesgos de un altísimo potencial lesivo en función de sus características, capacidades, contexto, sector y aplicación, así como facilitar la comisión de múltiples delitos o ilícitos civiles de diversa naturaleza, incluso afectar a la integridad física de las personas y a su vida.

La intrusión en un sistema de inteligencia artificial de un vehículo autónomo, de un dron, de una prótesis humana, de un sistema de intervención quirúrgica remota asistida por un sistema inteligente y ejecutada por un robot o de cualquier máquina, robot o producto que integre un sistema inteligente podría comportar la alteración del mismo conforme a sus directrices, instrucciones y finalidades incorporadas en su diseño o adquiridas posteriormente, lo que podría conllevar a convertir dichos sistemas o productos en armas que podrían ser utilizadas contra la vida humana, así como contra los derechos de las personas, cosas e instalaciones. Es decir, el delito de intrusión permitiría la comisión

posterior de otros delitos o ilícitos de cualquier otra naturaleza, ya sea en el orden civil, tributario u otros.

5.2. Interceptación ilegítima de comunicaciones entre sistemas de información

El delito de interceptación ilegítima de comunicaciones entre sistemas de información se halla regulado en el artículo 197 bis.2 del Código Penal español.

La conducta delictiva tipificada consiste la interceptación de transmisiones no públicas de datos informáticos que se produzcan desde, hacia o dentro de un sistema de información -incluidas las emisiones electromagnéticas de los mismos-, mediante la utilización de artificios o instrumentos técnicos y sin estar debidamente autorizado.

Estos instrumentos pueden consistir en herramientas técnicas de escucha, transmisión, grabación o reproducción de sonido o de imagen, e incluiría la interceptación de comunicaciones personales, interceptación entre sistemas informáticos y máquinas, y entre personas y máquinas.

Inicialmente el bien jurídico protegido en este delito sería la intimidad y se trataría de un delito para cuya comisión bastaría el dolo general.

De nuevo, los sistemas de inteligencia artificial se pueden constituir en un medio o instrumento idóneo para la comisión de este tipo acciones.

No obstante, en atención al contexto, podrían ser utilizados con fines lícitos y legítimos en el marco de investigaciones y lucha con la delincuencia y el terrorismo internacional, pero también con fines ilícitos e ilegítimos como el espionaje industrial, la obtención de datos para su uso posterior fraudulento como claves de acceso, órdenes de transferencia, tarjetas de crédito o cuentas bancarias.

Y del mismo modo, los sistemas inteligentes podrían ser la supuesta “víctima” en el caso de que pudiera atribuírsele dicha condición y considerando afectada su intimidad cuando se pudiera producir la interceptación de la comunicación entre dos sistemas de

inteligencia artificial o entre un sistema y una persona. No obstante, como he tratado anteriormente, no resulta posible atribuir esta condición ante la ausencia de personalidad jurídica del sistema.

5.3. Facilitación de herramientas informáticas dañinas para la comisión de los delitos anteriores

La conducta delictiva consistiría en la producción, la adquisición para su uso, la importación o la facilitación a terceros de un programa informático, una contraseña, un código de acceso o datos similares que permitan acceder a la totalidad o a una parte de un sistema de información, sin estar autorizado y con la intención de facilitar la comisión de alguno de los delitos de descubrimiento y revelación de secretos regulados en los apartados 1 y 2 del artículo 197 del Código Penal español -Delitos contra la privacidad- o en el artículo 197 bis del mismo y que han sido analizados anteriormente, esto es, los delitos de *hacking* o intrusismo informático, e interceptación ilegítima de comunicaciones entre sistemas de información.

Se trata de un delito de dolo que exige la intención específica de facilitar la comisión de determinados delitos, y se aplicaría la doctrina de los programas informáticos genéricos o con tecnología de doble uso concebidos inicialmente para fines legítimos, pero con herramientas y capacidades para ser usado como instrumento idóneo para la comisión de este delito, por lo que, en función de la naturaleza y funciones del programa informático, el hecho podría típico, atípico o lícito.

Los sistemas de inteligencia artificial podrían ser objeto o instrumento idóneo para la comisión de este tipo de delitos, mediante la producción o facilitación a terceros de herramientas o claves de acceso a sistemas para la comisión de los delitos precitados.

Se trataría de un delito más focalizado en el fabricante, desarrollador, adquirente, importador o suministrador del *software* (incluido de inteligencia artificial), de contraseñas o de claves, por lo que los sistemas de inteligencia artificial más avanzados que pudieran ostentar dicha condición, en concreto la de “desarrollador” de otros

sistemas, carecerían de responsabilidad penal por los motivos expuestos en los anteriores apartados al analizar la misma, ante su inimputabilidad y ausencia de personalidad jurídica.

5.4. *Cracking*. Daños informáticos y sabotajes

El denominado *cracking* se halla regulado en el artículo 264 del Código Penal español, como delito de daños informáticos y sabotajes.

La conducta delictiva consiste en borrar, dañar, deteriorar, alterar, suprimir o hacer inaccesibles datos informáticos, programas informáticos o documentos electrónicos ajenos, por cualquier medio, sin autorización, de manera grave y cuando el resultado causado sea igualmente grave.

El bien jurídico protegido sería esencialmente la información contenida en el sistema que la aloja y para su comisión no importa el medio utilizado, esto es, virus, gusanos, bombas lógicas, *phishing*, *spywares*, accesos remotos sin conocimiento ni autorización mediante *rootkits*, *ransomware* o *malware*, entre otros.

Se trata de un delito de resultado y doloso, entre cuyos elementos destaca la gravedad del acto y del impacto para determinar su concurrencia, que deberá ser valorado económicamente basándose en los perjuicios derivados del delito y no en el valor del elemento informático afectado. La apreciación de dicha gravedad puede determinar que la acción sea atípica.

De nuevo, nos encontramos ante un delito en el que los sistemas de inteligencia artificial podrían ser el instrumento o medio idóneo para su comisión, así como para la potenciación de su impacto y efectos perjudiciales, tanto a nivel cualitativo, por la naturaleza de los bienes y derechos afectados, como cuantitativo por el número de personas, sistemas, bienes y derechos afectados.

De hecho, los sistemas de inteligencia artificial no sólo podrían facilitar la labor intrusiva de los *hackers*, sino la labor posterior de los mismos y otros considerados *crackers* que

realizarían la nueva acción delictiva una vez dentro del sistema, como borrar, dañar, deteriorar, alterar, suprimir o hacer inaccesible datos, programas y documentos electrónicos de terceros, con las graves consecuencias que todo ello puede significar en función de la naturaleza, características y usos de los sistemas de información implicados, entre lo que podrían igualmente encontrarse sistemas de inteligencia artificial, potenciando la capacidad lesiva de la acciones delictivas en función de las características, aplicación y uso de estos sistemas.

Además, constituiría una herramienta para posibilitar los ataques de *ransomware* o secuestro de información, especialmente mediante su cifrado para solicitar un rescate a la persona o entidad afectada para poder acceder a la información.

De nuevo, en relación con estos delitos, pensemos la mera alteración de los datos o *software* de un coche autónomo, que le permitiera activar el sistema de frenado hallándose en circulación a 120 km/hora, la alteración de un robot asistencial que pudiera agredir al usuario del mismo o sistemas de inteligencia artificial quirúrgicos que puedan verse alterados en plena intervención.

Estas cuestiones fáciles de pensar pero difíciles, por no decir que casi imposibles en la práctica, en función de la calidad de su diseño y desarrollo, evidencian de nuevo la interrelación absoluta entre ética, seguridad y responsabilidad, tanto civil como penal, porque los principios y normas éticas que exigen la seguridad, la fiabilidad, la exactitud, la supervisión y control humano de los sistemas inteligentes, junto con los marcos jurídicos de responsabilidad y ciberseguridad deben impedir el diseño y desarrollo de sistemas de inteligencia artificial, especialmente de alto riesgo, que puedan ser vulnerables ante ciberataques de esta naturaleza sin medidas adecuadas, que puedan permitir todas estas acciones, omisiones y sus consecuencias.

De este modo, la supervisión y control humano, la seguridad, la fiabilidad o la exactitud constituyen principios y normas éticas esenciales que deben ser consideradas en el diseño, despliegue, aplicación y uso de estos sistemas inteligentes y durante todo su ciclo de vida, para lo que el futuro establecimiento de su carácter vinculante a nivel legal garantizará su aplicación, siendo deseable su acompañamiento por un régimen sancionador en caso de incumplimiento, conforme así lo prevé la nueva Propuesta de Reglamento del Parlamento

Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, eso sí, respecto de los sistemas de inteligencia artificial considerados de nivel alto conforme al mismo.

Además, estos principios y normas deberán exigir la integración de mecanismos, medidas y controles, proactivos y preventivos para evitar que acciones de este tipo pudiesen ocurrir, así como medidas y controles detectivos -para identificarlos y bloquearlos antes de que impacten- o reactivos y correctivos para eludir o minimizar su impacto.

La mejor garantía para el cumplimiento de estos principios y normas éticas es el acompañamiento de los mismos con un marco normativo que los convierta en jurídicamente vinculantes para todas las partes implicadas en su diseño, fabricación despliegue, aplicación y uso, por lo que las propuestas regulatorias europeas, objeto de análisis en esta investigación, van en la dirección correcta si finalmente las exigen.

De manera consecuente, los sistemas de inteligencia artificial deberán construirse bajo el principio de *Ethics by design* y, de manera asociada, bajo el principio *Security by design*, que abordé en los capítulos II y III.

La seguridad técnica debe ser integrada en el diseño, despliegue, aplicación y uso de estos sistemas durante todo su ciclo de vida, por lo que debe ser igualmente garantizada y, de nuevo, a mi juicio, la mejor manera de hacerlo es definir marcos de medidas y controles que no sólo queden en el ámbito de buenas prácticas de la industria, sino que se vean acompañados de un marco jurídico que los convierta en obligatorios, con las consecuencias legales pertinentes en caso de incumplimiento.

En este sentido las propuestas europeas en materia de ética y responsabilidad de la inteligencia artificial refuerzan la necesaria existencia de dicha seguridad, pero, en mi opinión, deberán acompañarse de normas técnicas y de seguridad que los complementen en el futuro, en lo que ya está trabajando la *European Union Agency for Network and Information Security* -ENISA por sus siglas en inglés-, conforme expuse al abordar los aspectos de seguridad en el capítulo II.

Asimismo, por el momento, la nueva Propuesta de Reglamento de 21 de abril regula los requisitos y obligaciones generales ciberseguridad, pero para los sistemas de inteligencia artificial considerados de alto riesgo conforme al mismo, exigiendo evaluación y gestión de los riegos y medidas adecuadas conforme a los riesgos identificados y contexto.

Además, conforme ya estamos viendo en la práctica, la seguridad debe además contemplar los elementos, componentes, redes y tecnologías que se integren o con las que interactúe un sistema de inteligencia artificial, que pueden ser el punto más vulnerable junto al factor humano a explotar por los *ciberatacantes*, máxime si el medio es otro sistema de inteligencia artificial.

A modo de simple ejemplo de automatización más que de inteligencia artificial: Una aspiradora "inteligente" dotada con cámara IP conectada a Internet vía WIFI, la cual es *crackeada* y controlada por terceros para visualizar el interior de viviendas y locales.

En conclusión, la mejor garantía de cumplimiento de los marcos éticos y de seguridad es la construcción de marcos normativos sólidos que exijan dicha ética y seguridad, y que integran igualmente el denominado *Compliance by design* o conformidad normativa global en su diseño, fabricación, provisión, despliegue, aplicación y uso.

5.5. Obstaculización o interrupción de un sistema informático

El *Código Penal* español regula en su artículo 264 bis, el delito de obstaculización o interrupción del funcionamiento de un sistema informático.

La conducta delictiva consiste en la obstaculización o interrupción del funcionamiento de un sistema informático ajeno sin estar autorizado y de manera grave. Además, contempla supuestos agravados si los hechos hubieran perjudicado de forma relevante la actividad normal de una empresa, negocio o de una Administración pública, así como en los demás supuestos previstos en el 264 bis.2 del *Código Penal* español. Dentro de este delito podrían calificarse acciones como los ataques masivos de denegación de servicio (DoS o DDoS) o mediante redes *botnet*.

Uno de los aspectos que mayores problemas plantea en su exigencia práctica es la gravedad exigida a la interrupción y obstaculización que quedaría en manos de la interpretación *a posteriori* por parte de los tribunales. En mi opinión, esa gravedad no debe exclusivamente medirse en términos de tiempo sino de otros factores en función del contexto, especialmente el impacto.

A modo de ejemplo, una acción que pudiera provocar la "caída" de una página web o red durante intervalos de minutos podría llegar a ser considerada atípica conforme a los fundamentos jurídicos de algunas resoluciones judiciales. Sin embargo, la gravedad en estos casos no puede medirse exclusivamente en unidades temporales sino considerar aspectos como el impacto. El nivel de gravedad de la caída de una página web durante 10 minutos, ¿sería el mismo para una panadería que utiliza la web para simplemente informar de su ubicación sin comercializar ningún tipo de producto a través de la misma que la web del Corte Inglés o Amazon?

5.6. Delito de facilitación de herramientas informáticas para facilitar la comisión de los delitos anteriores

De la misma manera que expuse respecto de los delitos contra la privacidad, intrusismo informático e interceptación ilegítima de comunicaciones entre sistemas de información, el artículo 264 ter del Código Penal español regula el delito de producción, adquisición para su uso, importación o facilitación a terceros, sin autorización y por cualquier medio, programas -concebidos para ello o adaptados-, contraseñas o claves de acceso con la intención de facilitar la comisión de un delito de *cracking* o de obstaculización o interrupción de un sistema de información, sin autorización .

La conducta se tipifica como delito autónomo de dolo, que podría ser preparatoria para la comisión de uno de los delitos precitados, en particular, el de *cracking*.

De nuevo, la responsabilidad sobre el delito se situaría en la esfera del diseñador, desarrollador, adquirente, importador o facilitador de estos programas, por lo que los sistemas de inteligencia artificial que pudieran situarse en la órbita de quien ostente dicha

condición -especialmente la de desarrollador- carecerían de responsabilidad penal por los motivos expuestos en los anteriores apartados al analizar la misma, dada su inimputabilidad penal.

Los sistemas de inteligencia artificial podrían ser objeto o instrumento idóneo para la comisión de este tipo de delitos, mediante la producción o facilitación a terceros de herramientas o claves de acceso a sistemas para la comisión de los delitos precitados, tanto a personas -físicas o jurídicas- como a otros sistemas.

5.7. Descubrimiento y revelación de secretos

Por último, abordaré los delitos de descubrimiento y relevación de secretos, regulados en el artículo 197 del Código Penal español, en sus distintas modalidades.

El tipo básico se regula en apartado 1 del precitado artículo, en el que se considera delito la acción de apoderarse de los papeles, cartas, mensajes de correo electrónico o cualesquiera otros documentos o efectos personales de otro, interceptar sus telecomunicaciones o utilizar artificios técnicos de escucha, transmisión, grabación o reproducción del sonido o de la imagen, o de cualquier otra señal de comunicación, para descubrir sus secretos o vulnerar su intimidad, sin su consentimiento.

Se trata de un delito de dolo y de finalidad específica, en el que el bien jurídico protegido sería la intimidad y la inviolabilidad de las comunicaciones.

La segunda modalidad prevista en el apartado 2 del precepto indicado consiste en la acción de apoderarse, utilizar o modificar, en perjuicio de tercero, datos reservados de carácter personal o familiar de otro que se hallen registrados en ficheros o soportes informáticos, electrónicos o telemáticos, o en cualquier otro tipo de archivo o registro público o privado, acceder por cualquier medio a los mismos, así como alterarlos o utilizarlos en perjuicio del titular de los datos o de un tercero, sin estar autorizado.

Se trata de un delito doloso llevado a cabo en perjuicio de otro, en el que el bien jurídico protegido lo constituye la protección de datos y la privacidad.

La última modalidad prevista en el apartado 3 del artículo 197 del Código Penal español consiste en la revelación de secretos de origen ilícito, consistente en difundir, revelar o ceder a terceros los datos o hechos descubiertos o las imágenes captadas a las que se refieren las dos modalidades precitadas, si participó en su descubrimiento, castigándose también a quién no participe en su descubrimiento, pero las revele con conocimiento de su origen ilícito.

De nuevo, nos encontramos ante un delito doloso y en perjuicio de otro, cuyo bien jurídico protegido es la protección de datos y la privacidad.

El precepto indicado regula igualmente distintos supuestos agravados en atención a su difusión, por razón del sujeto activo (responsables y encargados ficheros, soportes, archivos y registros), en casos de utilización no autorizada, datos especiales, menores y personas con discapacidad, finalidad lucrativa, pertenencia a organización o grupo criminal o la condición de autoridad o funcionario público del sujeto activo.

La inteligencia artificial podría ser utilizada en relación con todas estas modalidades como instrumento o medio idóneo para su comisión, no resultando posible su imputabilidad ante la exigencia de dolo y ausencia de personalidad jurídica.

5.8. Delitos contra la propiedad intelectual, industrial y otros.

Por último, la inteligencia artificial avanzada puede ser medio o instrumento para la comisión de todo tipo de delitos por su naturaleza, si bien, en algunos de ellos la intervención de robots o sistemas dotados de inteligencia artificial, previsiblemente, será cada vez más habitual, especialmente delitos relacionados con su uso médico, los delitos contra la propiedad intelectual e industrial regulados en los artículos 270 a 277 del Código Penal o los delitos contra el mercado previstos en los artículos 278 a 286.

A modo de ejemplo significar, que el tipo básico regulado en el apartado 1 del artículo 270 *del Código Penal* español, castiga con la pena de prisión de seis meses a cuatro años y multa de doce a veinticuatro meses el que, con ánimo de obtener un beneficio económico directo o indirecto y en perjuicio de tercero, reproduzca, plagie, distribuya,

comunique públicamente o de cualquier otro modo explote económicamente, en todo o en parte, una obra o prestación literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier medio, sin la autorización de los titulares de los correspondientes derechos de propiedad intelectual o de sus cesionarios.

Del mismo modo, los apartados 2 y 3 del artículo 270 del Código Penal español regula los delitos relacionados cometidos por agentes proveedores de servicios de sociedad de la información en determinados contextos, cuando faciliten el acceso a obras protegidas:

> "2. La misma pena se impondrá a quien, en la prestación de servicios de la sociedad de la información, con ánimo de obtener un beneficio económico directo o indirecto, y en perjuicio de tercero, facilite de modo activo y no neutral y sin limitarse a un tratamiento meramente técnico, el acceso o la localización en internet de obras o prestaciones objeto de propiedad intelectual sin la autorización de los titulares de los correspondientes derechos o de sus cesionarios, en particular ofreciendo listados ordenados y clasificados de enlaces a las obras y contenidos referidos anteriormente, aunque dichos enlaces hubieran sido facilitados inicialmente por los destinatarios de sus servicios.
>
> 3. En estos casos, el juez o tribunal ordenará la retirada de las obras o prestaciones objeto de la infracción. Cuando a través de un portal de acceso a internet o servicio de la sociedad de la información, se difundan exclusiva o preponderantemente los contenidos objeto de la propiedad intelectual a que se refieren los apartados anteriores, se ordenará la interrupción de la prestación del mismo, y el juez podrá acordar cualquier medida cautelar que tenga por objeto la protección de los derechos de propiedad intelectual.
>
> Excepcionalmente, cuando exista reiteración de las conductas y cuando resulte una medida proporcionada, eficiente y eficaz, se podrá ordenar el bloqueo del acceso correspondiente".

Por último, el artículo 270.6 tipifica como de delito la fabricación, importación, puesta en circulación o posesión con una finalidad comercial de cualquier medio principalmente

concebido, producido, adaptado o realizado para facilitar la supresión no autorizada o la neutralización de cualquier dispositivo técnico que se haya utilizado para proteger programas de ordenador o cualquiera de las otras obras, interpretaciones o ejecuciones protegidas.

Conforme establece su artículo 272, la extensión de la responsabilidad civil derivada de los delitos tipificados en los artículos 270 y 271 se regirá por las disposiciones de la LPI relativas al cese de la actividad ilícita y a la indemnización de daños y perjuicios, conforme abordaré en el capítulo VII.

5.9. Otros delitos

Los sistemas inteligentes podrían ser utilizados para la comisión de cualquier otro tipo de delito de distinta naturaleza como delitos de terrorismo, fraudes o incluso asesinatos, en los que habrá que identificar la *mens criminis*, es decir, la persona detrás del delito, sea el desarrollador, el fabricante o el usuario.

Del mismo modo, se debe ya plantear una reflexión sobre la posibilidad de conductas criminalmente imprudentes de modo que se pudiera plantear la imputación a un sistema inteligente de la gestión negligente de un riesgo, por ejemplo, el atropello de un peatón que cruzó una vía fuera del paso de peatones por un vehículo autónomo de Uber en 2018. En este tipo de supuestos, de nuevo, la atribución de la responsabilidad civil debería situarse en el marco de fabricante, operador o usuario en función del contexto.

6. Otras responsabilidades

La aplicación y uso de la inteligencia artificial puede comportar otras responsabilidades de distinta naturaleza en función del contexto.

Conforme he analizado en el presente capítulo, los sistemas inteligentes pueden ser utilizados para la comisión de delitos y convertirse en el instrumento o medio idóneo para su comisión.

No obstante, la utilización de este tipo de sistemas en otros ámbitos puede dar lugar a otro tipo de responsabilidades de distinta naturaleza, por ejemplo, en el ámbito de la circulación de los vehículos "autónomos", la navegación aérea o marítima no tripulada o la utilización con finalidades de gestión empresarial, pudiéndose derivar no sólo responsabilidad civil por los daños ocasionados, sino administrativas o incluso penales.

La utilización de sistemas inteligentes para la gestión económico-financiera, contable y fiscal de una organización puede generar responsabilidades en el orden fiscal e incluso con relevancia penal, especialmente en el ámbito de delitos económicos.

En relación con el software de doble uso, en el que tendría igualmente su encaje los sistemas de inteligencia artificial con las finalidades indicadas, conforme he expuesto al abordar algunos aspectos de esta investigación, está siendo regulado por los distintos ordenamientos jurídicos. Francia y Portugal tienen marcos en este sentido, y en España, como expuse anteriormente, se acaba de aprobar por el Congreso de los Diputados, el Proyecto de Ley de medidas de prevención y lucha contra el fraude fiscal, de transposición de la Directiva (UE) 2016/1164, del Consejo, de 12 de julio de 2016, en la que se aborda la reforma de Ley General Tributaria[867] española e introduce la prohibición del denominado software de doble uso, prohibiendo el desarrollo, la tenencia o el uso de sistemas y aplicaciones informáticas de gestión contable que permitan alterar y falsear la información contable de las empresas.

La infracción de estas obligaciones puede dar lugar a importantes responsabilidades administrativas tipificadas y sancionadas legalmente, mediante la inclusión del artículo 201 bis en la Ley General Tributaria española, que prevé como infracción tributaria, la fabricación, producción, comercialización y tenencia de sistemas informáticos que no

[867] Ley 58/2003, de 17 de diciembre, General Tributaria. BOE 18 diciembre 2003

cumplan las especificaciones exigidas por la normativa aplicable. Del mismo modo determinadas conductas podrían incluso ser constitutivas de delito.

La utilización de sistemas inteligentes para gestionar aspectos en el ámbito laboral o social de una organización también puede dar lugar a distintas responsabilidades.

Entre otros ejemplos, de un lado, una incidencia en redes o sistemas, inteligentes o no pero gestionados por éstos, que pueda afectar a los sistemas de seguridad de los trabajadores y que puedan provocar un accidente, puede conllevar la correspondiente responsabilidad de distinta naturaleza y clase, desde la administrativa a la penal, incluyendo el resarcimiento de los daños personales sufridos por los trabajadores afectados. Y todo ello en relación con la Ley 31/1995, de 8 de noviembre, de Prevención de Riesgos Laborales[868]. Entre otros supuestos, pensemos simplemente en una planta industrial robotizada.

De otro, se podría plantear la responsabilidad civil por vulneración de derechos fundamentales del trabajador ante determinados usos de los sistemas inteligentes, al amparo del artículo 4 del Estatuto de los Trabajadores[869] aprobado mediante Real Decreto Legislativo 2/2015, de 23 de octubre, en relación con el artículo 35 de la Constitución Española y de la Ley 36/2011, de 10 de octubre, de la Jurisdicción Social[870].

Del mismo modo, la utilización de estos sistemas puede dar lugar a otras responsabilidades en el ámbito administrativo y regulatorio, especialmente en materia de protección de datos, en lo que constituye uno de sus principales riesgos y retos.

La vulneración del derecho a la privacidad y la protección de datos personales mediante un sistema inteligentes podrá generar no sólo responsabilidades administrativas, sancionables en virtud del Reglamento General de Protección de Datos (RGPD) precitado y, en el caso de España, en la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD), sino que incluso

[868] Ley 31/1995, de 8 de noviembre, de Prevención de Riesgos Laborales. BOE 10.11.1995.
[869] Real Decreto Legislativo 2/2015, de 23 de octubre por el que se aprueba el texto refundido de la Ley del Estatuto de los Trabajadores. BOE 24.20.2015.
[870] Ley 36/2011, de 10 de octubre, reguladora de la jurisdicción social. BOE 11.10.2011.

puede generar en determinados contextos responsabilidad penal, conforme expuse en los apartados precedentes, así como civil.

En caso de vulneración de estos derechos, con independencia de su sanción administrativa, permitirá a la persona afectada reclamar la responsabilidad civil derivada de la misma por los daños y perjuicios causados en sus derechos y libertades.

Y a mayor abundamiento, cuando los hechos cometidos, en atención a sus elementos y contexto, puedan ser calificados delito conforme analicé anteriormente, su autor deberá responder tanto de las responsabilidades penales como de las civiles dentro del procedimiento penal.

El RGPD exige una seguridad proactiva y gestionada, en el diseño y por defecto -*privacy by design & by default*-.

El artículo 82 del Reglamento General de Protección de Datos (RGPD), regula el derecho a indemnización y la responsabilidad derivados de la infracción del mismo.

Conforme al mismo, toda persona que haya sufrido daños y perjuicios materiales o inmateriales como consecuencia de una infracción del presente RGPD tendrá derecho a recibir del responsable o el encargado del tratamiento una indemnización por los daños y perjuicios sufridos.

Del mismo modo, el apartado 2 del precepto citado, establece que cualquier responsable que participe en la operación de tratamiento responderá de los daños y perjuicios causados en caso de que dicha operación no cumpla lo dispuesto por el RGPD y, respecto del encargado del tratamiento, establece que únicamente responderá de los daños y perjuicios causados por el tratamiento cuando no haya cumplido con las obligaciones del RGPD dirigidas específicamente a los encargados o haya actuado al margen o en contra de las instrucciones legales del responsable, lo que constituye una causa de exención de responsabilidad.

En este sentido, la responsabilidad del encargado se limita al incumplimiento por el mismo de sus propias obligaciones conforme al RGPD o la actuación al margen o en contra de las instrucciones legales del responsable.

El sistema de responsabilidad regulado en el RGPD se sustenta sobre la necesidad de resarcimiento efectivo de los daños y perjuicios causados a la persona afectada, y no es suficiente que se haya producido una infracción normativa, de lo que derivará la correspondiente responsabilidad administrativa, sino que es necesario que se haya producido un daño o perjuicio en los bienes y/o derechos de la persona afectada y que se acredite por ésta. De manera consecuente, deberá concurrir igualmente la existencia de un nexo causal entre la conducta infractora, ilegítima o ilícita llevada a cabo por el responsable o el encargado del tratamiento y el daño.

No obstante, en cualquier caso, tanto el responsable como el encargado del tratamiento estarán exentos de responsabilidad si demuestran que no son en modo alguno responsables del hecho que haya causado los daños y perjuicios, invirtiendo la carga de la prueba, de modo que corresponde a los mismos la prueba de esta causa eximente.

Por otra parte, el RGPD regula una previsión específica en su artículo 82.4 especialmente proteccionista para la persona afectada cuando concurran más de un responsable o encargado, o un responsable y un encargado en la misma operación de tratamiento, con el objetivo de garantizar la indemnización efectiva del interesado, estableciendo que, en estos supuestos y cuando sean responsables conforme a lo indicado en los párrafos precedentes de cualquier daño o perjuicio causado por dicho tratamiento, cada responsable o encargado será considerado responsable de todos los daños y perjuicios, esto es, en su totalidad, sin perjuicio del derecho de repetición o regreso. El RGPD apuesta por la solidaridad y la corresponsabilidad en el tratamiento.

De este modo, un responsable o encargado del tratamiento que haya pagado una indemnización total por el daño causado, dicho responsable o encargado tendrá derecho a reclamar a los demás responsables o encargados que hayan participado en esa misma operación de tratamiento la parte de la indemnización correspondiente a su parte de responsabilidad por los daños y perjuicios causados. El artículo 30.1 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales[871] (LOPDGDD), regula de manera similar estos aspectos.

[871] BOE 06.12.2018

De nuevo, la cuestión que se suscita en este contexto es qué responsabilidad podría exigirse en este sentido a los sistemas de inteligencia artificial.

La ausencia de personalidad jurídica y de capacidad para ser titular de derechos y obligaciones, nos lleva a las cuestiones analizadas por mi parte a lo largo del capítulo V, de modo que la responsabilidad deberá situarse en la esfera de la persona física o jurídica que actúe como responsable o encargado del tratamiento llevado a cabo por el sistema inteligente en cualquier fase de su ciclo de vida que provocó la infracción del marco jurídico en materia de privacidad (RGPD) y causó el daño consecuente.

El tratamiento podría llevarse a cabo en fase de su desarrollo y concepción si se trataran datos reales, en su fase de entrenamiento y formación, en la entrada de datos, en su procesamiento, en su autoaprendizaje y en la salida de datos.

De este modo, la figura del responsable, corresponsable o encargado, podría recaer en el desarrollador, fabricante, proveedor, formador, entrenador, operador o usuario, de manera individual o solidaria.

En el caso de que un responsable o encargado del tratamiento no se halle establecido en la UE, pero se lleven a cabo tratamientos de datos relativos a personas situadas en España, la responsabilidad prevista en el RGPD podrá imponer al representante solidariamente con el responsable o encargado del tratamiento, de conformidad con lo previsto en el RGPD.

El RGPD regula un principio general de responsabilidad solidaria de todos los sujetos involucrados en el tratamiento de datos, como he indicado, lo que puede ser una buena referencia para la definición del futuro marco regulador de la responsabilidad civil por los daños causados por la inteligencia artificial. Esta responsabilidad solidaria de responsables, encargados y representantes se recoge expresamente del mismo modo en el artículo 30.2 de la posterior Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales[872] (LOPDGDD).

[872] BOE 06.12.2018

La responsabilidad por daños se depurará, en el ejercicio del derecho fundamental a una tutela judicial efectiva, ante los Juzgados y Tribunales del orden jurisdiccional civil que resulten competentes de conformidad con la Ley 1/2000, de 7 de enero, de Enjuiciamiento Civil[873] en relación con la legislación sustantiva recogida en los artículos 1902, siguientes y concordantes del Código Civil español, sin perjuicio de que, en el caso de que el responsable o el encargado del tratamiento y sistema sea una Administración pública, deba utilizarse de manera previa la vía la administrativa hasta su agotamiento para la reclamación de responsabilidad patrimonial de la misma, en su caso, regulada en los artículos 32 a 35 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público[874]. Una vez agotada la vía administrativa, se deberá iniciar la vía judicial a través del orden jurisdiccional contencioso-administrativo.

Las acciones deberán ejercitarse ante los tribunales del Estado miembro en el que el responsable o encargado tenga su establecimiento. No obstante, al objeto de proteger a los afectados, de conformidad con lo previsto en el artículo 79 del RGPD, estas acciones podrán ejercitarse alternativamente ante los tribunales del Estado miembro en el que el interesado tenga su residencia habitual, a menos que el responsable o el encargado sea una autoridad pública de un Estado miembro que actúe en ejercicio de sus poderes públicos.

Por último, significar que la derogada Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal (LOPD), si contemplaba expresamente el derecho a indemnización en su artículo 19, si bien, la posterior LOPDGDD, no incorpora expresamente el derecho a la indemnización contemplado en el RGPD, entiendo que al hallarse ya regulado por éste y dada la propia naturaleza de este instrumento normativo de carácter general y de directa aplicación en España.

Otra cuestión es la relativa a su ubicación en el RGPD, en la medida que se incluye en el Capítulo VIII, relativo a los "Recursos, responsabilidad y sanciones", conforme significan autores como Amat Llombart[875], en lugar de incorporarlo en el Capítulo III

[873] BOE 08.01.2000
[874] BOE 02.10.2015
[875] AMAT LLOMBART, P. (2020). "La protección de las personas físicas en relación al tratamiento de sus datos personales. Condiciones para el ejercicio de sus derechos y en el marco de la normativa comunitaria

relativo a los "Derechos del interesado", ubicación que parece la más adecuada sistemáticamente, en la medida que el derecho a un indemnización efectiva por parte de la persona afectada constituye el último elemento del sistema global de protección de las personas físicas en relación con el tratamiento de sus datos personales y vulneración de sus derechos y libertades en esta materia.

Similares responsabilidades pueden derivarse ante sistemas inteligentes y uso que vulneren el derecho a la no discriminación.

Por último, merece significarse igualmente la responsabilidad por daños provocados por tratamientos automatizados de datos no personales. El tratamiento masivo de estos datos podría parecer inocuo, si bien, su utilización para la obtención de valoraciones, predicciones, comportamientos y patrones de conducta podría impactar en los bienes y derechos de las personas.

La responsabilidad se orientaría hacia los productores y responsables de estos tratamientos y su marco jurídico se conformaría principalmente en el artículo 13 de la Ley 34/2002, de 11 de julio, de servicios de la sociedad de la información y de comercio electrónico[876] (LSSICE), en relación con los artículos 128 y 129 del TRLGDCU y 1902 y 1903 del Código Civil, sin perjuicio de las responsabilidades contractuales que puedan igualmente derivarse en función el contexto concreto.

La responsabilidad, conforme a los preceptos citados, podría derivarse tanto para las empresas productoras del *Big data* o de los resultados obtenidos del tratamiento masivo y automatizado de datos, como para las empresas que adquieren las bases de datos por perfiles de usuarios para determinados usos, especialmente para dañar o discriminar a determinados usuarios, conforme destaca Vázquez de Castro[877]. En consecuencia, las responsabilidades que podrían derivarse para los prestadores de servicios de la Sociedad de la Información regulados en la LSSICE, sean o no intermediarios, podrían quedar tanto

y española", en Bello Janeiro, D. (Coord.) *Nuevas tecnologías y responsabilidad civil*. Editorial Reus 2020. P. 141.

[876] BOE N. 166. 12.07.2002.

[877] VÁZQUEZ DE CASTRO, E. (2020). "Aproximación a la responsabilidad derivada de los riesgos de la inteligencia artificial en Europa". Op.cit. P. 271.

sujetos a responsabilidades civiles, como penales y administrativas en relación con los contenidos elaborados por los mismos o por cuenta de estos.

Por lo que se refiere a las civiles, responderán subjetivamente por los hechos propios por incumplimiento de la diligencia exigible conforme a lo previsto en el régimen de responsabilidad extracontractual regulado en España en el artículo 1902 del Código Civil. Algunos autores como el citado Vázquez de Castro[878], proponen la aplicación del artículo 1903 del Código Civil, anteriormente comentado, máxime en la medida que es la empresa la que desarrolla una actividad económica de la que obtiene un beneficio.

En conclusión, la aplicación y uso de sistemas inteligentes puede dar lugar a responsabilidades de distinta naturaleza en función del sector, contexto y uso como las que se han expuesto u otras.

7. Consideraciones finales

La reflexión y análisis de la posible responsabilidad penal de los robots y los sistemas dotados de inteligencia artificial es ineludible a mi juicio, en congruencia con las reflexiones y avances realizados en materia de responsabilidad civil.

La atribución de la personalidad jurídica a estos sistemas la considero inviable en estos momentos conforme al estado de la técnica, en base a los motivos expuestos en este capítulo y en los precedentes.

La imputabilidad penal y atribución de un dolo directo a los robots y sistemas dotados de inteligencia artificial es inviable conforme el marco penal vigente.

La posibilidad de construir robots y sistema de inteligencia artificial realmente autónomos y que pudieran tener consciencia la considero inviable conforme al estado de desarrollo

878 VÁZQUEZ DE CASTRO, E. (2020). "Aproximación a la responsabilidad derivada de los riesgos de la inteligencia artificial en Europa". Op.cit. P. 272.

de la tecnología actual, los límites establecidos por los marcos éticos más consensuados a nivel internacional y por las propuestas regulatorias europeas en estudio y tramitación, si bien, como he manifestado al abordar distintos puntos de esta investigación, la grandeza del ser humano nos impide afirmar su imposibilidad.

Si llegara ese día, sería entonces cuando podríamos valorar su posible imputación penal y la posible consideración de la concurrencia de dolo directo o imprudencia, conforme a un nuevo marco previamente definido que regule un nuevo sistema de autoría y de las instituciones clásicas, pasando por el reconocimiento de una personalidad jurídica.

Del mismo modo que he comentado en otros capítulos, son varios los autores que adoptan la responsabilidad de las personas jurídicas como punto de partida para la atribución futura de la responsabilidad civil a robots y sistemas de inteligencia artificial avanzados, y del mismo modo también para la penal como los citados Pagallo y Asaro, si bien, por la propia naturaleza del Derecho Penal, considero que todo ello deberá conllevar una profunda reflexión en lo sucesivo para adaptar el mismo a la realidad actual y futura, revisando su parte general y modificando algunos tipos y creando otros integrados en su parte especial, bajo un principio de conservación e intervención mínima.

La comisión de delitos en los que se verán involucrados sistemas dotados de inteligencia artificial o robots, máquinas u otros productos que integren la misma como medio o instrumento para su comisión, no cesará de crecer de manera paralela a su despliegue y utilización en todo tipo de ámbitos y sectores, de un lado, y a la digitalización de nuestra sociedad y el imparable y exponencial avance tecnológico de otro, y, en mi opinión, será especialmente significativo y exponencial su crecimiento tanto cualitativa como cuantitativamente en relación con los denominados *ciberdelitos* y otros relacionados más directamente relacionados con la tecnología.

La responsabilidad penal de los robots o los sistemas avanzados de inteligencia artificial no debe ser una prioridad en este momento, sí la seguridad jurídica para todas las partes involucradas y el fomento y apoyo de la innovación y el avance tecnológico que proporcione valor al ser humano y sí, también, la aprobación de los marcos éticos y jurídicos en materia de inteligencia artificial y responsabilidad civil relacionada con la misma.

El avance en la definición de los marcos civiles y administrativos pueden servir de base para la reflexión y construcción de la dimensión penal de la inteligencia artificial y las máquinas, robots y otros dispositivos dotados y gobernados por la misma.

De igual forma, la legislación procesal en materia criminal, centenaria en España, debe revisarse y adaptarse a los nuevas realidades y necesidades para dotar de los medios necesarios para la adecuada investigación y persecución de los delitos y su enjuiciamiento, potenciando la cooperación internacional y su enfoque transnacional que supere los límites territoriales que comporta el concepto de ordenamiento jurídico y que no existen en el *ciberespacio,* donde todos -personas, empresas, Administraciones públicas, gobiernos, sistemas y máquinas- vivimos e interactuamos.

El Derecho debe evolucionar en paralelo a como lo hace nuestra sociedad, y el Derecho Penal no es una excepción.

El legislador debe abordar un enfoque dinámico, flexible, adaptativo y evolutivo para dar soluciones eficaces a los problemas que tengamos hoy o mañana, sin perjuicio de su constante revisión y adaptación.

Como he referido al abordar distintos aspectos dentro de esta investigación, considero necesario legislar sobre tendencias, no sobre novedades que pueden convertir una norma específica en obsoleta antes de su aprobación final. En este sentido, debo significar de nuevo que, en tecnología, “el hoy, es ayer”, “el ahora, es antes”.

Capítulo VII

Robots

1. Introducción

¿Qué es un robot? De nuevo, considero que abordar adecuadamente los contenidos de este capítulo exige partir de una definición clara del concepto, dado que se trata de otra realidad de la que todo el mundo ha oído hablar, pero sobre la que quizás, son pocos los que realmente tienen un concepto claro de lo que es.

Incluso, me planteo si deberíamos ya migrar definitivamente a un concepto distinto, el de "máquinas" o "sistemas" en lugar de robots, para alejarnos de un concepto alterado en la percepción social que los asocia más a una tipología en particular, esto es, a humanoides, con formas o características humanas.

De hecho, la nueva Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021, parece construirse sobre la base de estos conceptos.

Desde el inicio de la tecnología, el ser humano no ha cesado de crear máquinas capaces de realizar comportamientos cada vez más complejos, automáticos y progresivamente ha ido incorporando a los mismos la imitación del ser humano y sus procesos mentales.

La literatura científica y jurídica habitualmente significa los mecanismos animados de Arquitas de Tarento y su ave mecánica (siglo IV a.C.), los mecanismos animados de Herón de Alejandría (siglo I) que algunos consideran los pre-robots, el hombre de hierro de Alejandro Magno (siglo XII), la cabeza parlante de Roger Bacon (siglo XIII), el gallo del reloj de la catedral de Estrasburgo (Siglo XIV), la máquina *Ars Generalis Ultima* descrita por el filósofo mallorquín Ramón Llull (Siglo XIV) para analizar y validar/invalidar teorías utilizando la lógica, el león mecánico o el robot humanoide de Leonardo da Vinci (siglo XV), el monje autómata o "hombre de palo" de Juanelo Turriano

(siglo XVI), incluso algunos de los desconocidos inventos del fraile, pero también científico e inventor burgalés, Mariano Díez Tobar (siglo XIX), hasta llegar en la actualidad a robots dotados de apariencia humana e inteligencia artificial como Sophia[879], un robot humanoide desarrollado por Hanson Robotics e inspirado en la actriz Audrey Hepburn, así como a *roboadvisors* de nueva generación, *nanobots* capaces de integrarse en el cuerpo humano, etc.

Todo ello evidencia la constante inquietud del ser humano por ir más allá y la evolución que ha experimentado la robótica, especialmente exponencial durante los últimos años y de manera asociada al avance científico, la capacidad de computación, su asociación con otras tecnologías y su uso en determinados sectores como la biotecnología.

2. Concepto "robot".

No existe una definición universal o única y es un concepto variable en el tiempo y muy evolucionado desde sus orígenes.

Al igual que la inteligencia artificial, no existe un consenso sobre su definición ni al nivel científico, ético ni jurídico, especialmente por la diversidad de clases y características.

La primera definición que el diccionario de la *Real Academia de la Lengua* ofrece de este concepto es "máquina o ingenio electrónico programable que es capaz de manipular objetos y realizar diversas operaciones".

Algunas definiciones matizan esas "operaciones" haciendo alusión que se trataría de "operaciones antes reservadas sólo a las personas", si bien, considero que matizan en

[879] Sophia un robot humanoide que puede mostrar en su cara más de 60 tipos de sentimientos mientras conversa con un ser humano, aunque no disponga de la capacidad de sentir, al que incluso se le ha reconocido la ciudadanía por Arabia Saudí. Procesa datos visuales y utiliza reconocimiento facial y tecnología de reconocimiento de voz, analizando conversaciones y extrayendo datos para mejorar su respuesta en el futuro. Stone, Zara. "*Everything You Need To Know About Sophia, The World's First Robot Citizen*". Publicado en Forbes, 7.11.2017. Disponible en: https://www.forbes.com/sites/zarastone/2017/11/07/everything-you-need-to-know-about-sophia-the-worlds-first-robot-citizen/?sh=7dd2352f46fa. Consultado el 21.02.2021.

exceso la definición general y las operaciones que, en la actualidad, pueden no ser exclusivas del ser humano sino asociadas a animales, plantas, procesos químicos o medioambientales, por ejemplo.

No obstante, me permito significar la tercera definición que incluye el Diccionario de la Real Academia de la Lengua de "robot" como persona que actúa de manera mecánica o sin emociones, que implícitamente está asociando a dicho concepto esa mera actuación mecánica o automática, más que autónoma, y sin emociones propias del ser humano.

A nivel científico-técnico, una definición clásica lo define como una máquina controlada por ordenador y programada para moverse, manipular objetos y realizar trabajos a la vez que interacciona con su entorno.

Según Ryan Calo[880], un robot debe disponer de tres (3) capacidades esenciales: La de recoger datos mediante sensores (sentir), la de procesar los datos (pensar) y la de planificar y cumplir acciones mediante conocimientos e informaciones adquiridas generalmente en función de objetivos prefijados (actuar), y todas ellas sin perjuicio de que pueda contar con otras como la capacidad de autoaprendizaje, autoprogramación o de comunicación con redes, operadores u otros sistemas.

García-Prieto[881] considera que un robot es una máquina, provista de cierta complejidad tanto en sus componentes como en su diseño o en su comportamiento, y que manipula información acerca de su entorno para así interactuar con él.

Este mismo autor, cita al profesor Roth, catedrático de la Escuela de Ingeniería Mecánica de la Universidad de Standford, que significa que "a medida que la máquina se incorpora a las funciones de un humano, solemos llamarlo robot. A medida que nos acostumbramos a esta función y volvemos a interpretar que dicha función no es propia de humanos, volvemos a llamarlo máquina", considerando que la definición es "borrosa y variable".

[880] CALO, R. (2015). "Robotics and the Lessons of Cyberlaw". *California Law Review*. Vol. 103:513. 2015. P. 529.
[881] GARCÍA-PRIETO, J. (2018). "¿Qué es un robot?". En BARRIO ANDRÉS, M. (Dir.). *Derecho de los Robots*. Op.cit. Pp. 26-32.

De todas estas definiciones, debo extraer algunas reflexiones:

- ¿Un robot únicamente debe tratarse de hardware que puede integrar software, es decir, una máquina? ¿O podría consistir exclusivamente en software?

Un robot puede estar basado exclusivamente en *software*, por ejemplo, un *roboadvisor* de nueva generación, sin perjuicio de que opere a través de sistemas de información integrados por *hardware*, *software, datos y* redes. De hecho, algunos estándares como la ISO 8373:2012 "*Robots and robotics devices - Vocabulary*"[882] incluye dentro del concepto de robot las prótesis biónicas, los robots biológicos, los micro y los *nanobots o* robots nanoscópicos.

- ¿Inerte, inanimado? No, un robot puede estar compuesto por secciones o partes vivas u operar interconectado o a través de seres y tejidos vivos, como los precitados nanobots, sensores internos de estimulación/reacción neurológica, válvulas complejas.

- ¿Complejidad equivale a impredecibilidad? En opinión de García-Prieto antes citado sí, aunque no comparto esta equivalencia, en la medida que, en general, todo su comportamiento debería ser predecible conforme a su diseño y componentes, salvo en los casos en los que el robot estuviera dotado de un sistema de inteligencia artificial con cierta "autonomía" para actuar conforme a un contexto concreto, en base a unos objetivos prefijados y con una relativa impredecibilidad dentro de unos límites acotados. Aun así, se discute por algunos autores, como se ha analizado en anteriores capítulos, si realmente deberíamos hablar de automatismo más que de autonomía.

- ¿Dotado de inteligencia artificial? No. El concepto robot no conlleva de forma inherente que el mismo este dotado de inteligencia artificial, para ello deberíamos irnos al concepto de "robot inteligente" de Murphy[883], definido una criatura mecánica que puede funcionar de manera autónoma. Definición que tampoco comparto en la medida que una máquina "gestionada" o "gobernada" por un sistema inteligente, no necesariamente tiene que sea

[882] Recuperado de https://www.iso.org/standard/55890.html
[883] MURPHY, R. (2000). *Introduction to AI Robotics 2e (Intelligent Robotics and Autonomous Agents series)*. Prensa del MIT. 2000. Cambridge EE.UU.

autónoma para ser considerada como tal, sin perjuicio de que ejecute tareas de manera automatizada, y ello en base al necesario control y supervisión humana y a la seguridad.

En este sentido, conforme igualmente signifiqué en su análisis, la nueva Propuesta de Reglamento europeo de inteligencia artificial de 21 de abril de 2021 se aleja del concepto de "autonomía" para regular los sistemas inteligentes como tales o integrados en máquinas o productos.

Santos González[884] define el concepto de robot inteligente como "aquella máquina física que de manera autónoma a través de programas y sensores inteligentes pueda llegar a tomar decisiones basándose en la lógica e inteligencia artificial, prediciendo las necesidades de los humanos y de las situaciones en las que se ven envueltos actuando, alternado e interactuando con el mundo físico, todo ello sin estar sometidos al control continuo de los humanos".

No comparto esta definición por múltiples razones como he argumentado a lo largo de estas investigación y párrafos precedentes, en primer lugar, en la medida que la definición se basa en la existencia de una máquina física, cuando como he expuesto, disponemos de robots basados en *software* y, en segundo lugar, en la medida que no debería crearse un solo robot inteligente que no se halle sometido al control humano. Asimismo, considero que el concepto robot no tiene que llevar asociado necesariamente que el mismo esté dotado de un sistema inteligente que lo gobierne o gestione.

Otros autores como Barrio Andrés[885], consideran que un robot *strictu sensu* podría considerarse como aquel "objeto mecánico que capta el exterior, procesa lo que percibe y, a su vez, actúa positivamente sobre el mundo", es decir, algo próximo a lo que los profesores Pfeifer y Scheier[886] denominaron el paradigma de "sentir-pensar-actuar". En consecuencia, los atributos esenciales de un robot serían sentir, pensar y actuar.

---

[884] SANTOS GONZÁLEZ, M.J. (2017). "Regulación legal de la robótica y la inteligencia artificial: retos de futuro". *Revista Jurídica de la Universidad de León*, Nº 4. 2017. P.31.
[885] BARRIO ANDRÉS, M. (Dir.) (2018). *Derecho de los Robots*. Wolters Kluwer España, S.A. 2018 Madrid. P. 70.
[886] PFEIFER, R. Y SCHEIER, C. (1999). *Understanding Intelligence*. Editorial MIT Press. Cambridge 1999. P. 37.

No obstante, estas últimas definiciones citadas vienen referidas a una clase específica de robot dotado de inteligencia artificial, pero, en modo alguno pueden asociarse estos atributos a todo tipo de robot para su categorización bajo este concepto, los cuales puede basar sus actuaciones en automatismos propios de su diseño y no en autonomía. Procesar no es pensar, siendo esto último un atributo propio del ser humano y no de las máquinas, por el momento.

Una cosa es la robótica y otra muy distinta la inteligencia artificial, y la primera no integra necesariamente la segunda.

Un sistema robótico puede ser absolutamente eficiente resolviendo cosas, pero eso no significa que sea autónomo y, ni mucho menos, que sea inteligente.

De hecho, a mi juicio, debemos partir de una definición simple del concepto "robot" diferenciándolo de las propiedades, capacidades, características o tecnologías adicionales de las que podrá estar dotado o no como, por ejemplo, inteligencia artificial.

Así, podríamos definir robot como sistema o máquina programable que es capaz de manipular objetos y realizar diversas operaciones de manera automática. Y "automática" no debe confundirse con "autónoma", capacidad de la que también podría estar dotado, en caso de integrar un sistema de inteligencia artificial para su "gobierno" o gestión.

Como fundamento de mi opinión a este respecto, me remito a la propia Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica, objeto de análisis previo en esta investigación en su capítulo III, que realmente se halla orientada a la inteligencia artificial que gobierna realmente a la máquina, distinguiendo y exigiendo una definición europea común de "robots autónomos inteligentes" y sus subcategorías teniendo en cuenta sus características.

En consecuencia, podríamos encontrarnos ante robots físicos y virtuales, si bien, éstos últimos deberán necesariamente utilizar como medio para operar otros sistemas de información que integren el *hardware* que los aloja y desde el que operan, así como otros elementos, como conectividad a redes para posibilitar la misma.

De este modo, considero que no debemos abordar sistemáticamente el concepto "robot" como agente electrónico dotado de inteligencia artificial, sino como lo que es, sin perjuicio de su dotación o interacción con otros sistemas y tecnologías. De otro modo se estaría desnaturalizando el concepto.

De manera consecuente, no soy partidario de diferenciar de manera general el denominado *Derecho de los Robots* como categoría diferenciada, dado que considero que actualmente no tiene sustantividad propia que lo requiera, aunque si debiéramos reflexionar sobre la necesidad de un *Derecho de la inteligencia artificial* que englobaría sistemas, máquinas, robots o productos dotados de la misma y/o que puedan ser "gobernados" o gestionados por la misma.

Quizás lo más adecuado para superar esta falta de consenso, sería definir un concepto general de robot del que se deriven tipologías específicas en función a las características, capacidades y contextos para los que ha sido diseñado, adaptadas a la concreta y variada realidad que pretende representar.

Por lo que se refiere a la clasificación de los "robots" es una tarea compleja y no pacífica, por lo que abordarlo en el marco de esta investigación resulta imposible, si bien, al menos de manera general, indicar que, según su cronología, podemos diferenciar entre robots de 1ª, 2ª y 3ª generación.

Según su estructura, podríamos diferenciar en poliarticulados, móviles, androides, zoomórficos e híbridos. Siguiendo a Barrio Andrés, según su complejidad, pueden diferenciarse de tipo A, B, C y D[887].

Según sus componentes, podemos diferenciar entre electromecánicos, microscópicos (*nanorobots* y máquinas moleculares) y *softbots*.

Según su aplicación, podemos diferenciar medioambientales, cirugía, prostética, salud, asistencial, militar, educación, juguetes, entretenimiento o arte.

[887] BARRIO ANDRÉS, M. (Dir.) (2018). *Derecho de los Robots*. Wolters Kluwer España, S.A. 2018 Madrid. P. 40.

La clase más próxima al ser humano serían los robots humanoides, esto es, robots con forma y/o características humanas, que pueden adquirir por su forma o por sus propiedades la capacidad para interactuar con humanos de forma que se asemejan en su comportamiento o en su aspecto al de un humano.

Los robots forman parte ya de los procesos industriales y progresivamente se convertirán en algo cotidiano en nuestros hogares, ciudades y en todos los sectores de actividad, lo que comportará la irrupción en nuestra sociedad de nuevos agentes, corpóreos o incorpóreos, dotados o no de inteligencia artificial, con un mayor o menor grado de automatismo o "autonomía", que desarrollarán todo tipo de tareas, desde movilizar y desplazar a una persona tetrapléjica, cuidar a nuestros mayores, conducir un vehículo de transporte o gestionar nuestros ahorros o finanzas.

Y el que no sea algo cotidiano para la mayoría de la sociedad no significa que no sea ya una realidad. Tenemos coches autónomos "en circulación" o *roboadvisors* de segunda generación que gestionan nuestras finanzas.

3. Retos éticos, jurídicos y de seguridad de los robots inteligentes

Los robots inteligentes, esto es, los dotados de inteligencia artificial, plantean los retos éticos, jurídicos y de seguridad propios de los sistemas de inteligencia artificial analizados en esta investigación, remitiéndome a los capítulos precedentes sobre estos aspectos.

No obstante, adicionan los retos propios asociados a su materialización física y mecánica de funcionamiento, en su caso, en la medida que ejecutan físicamente lo que el sistema o usuario determina.

China, Corea del Sur, Japón, EEUU o Canadá consideran la robótica junto con la inteligencia artificial aspectos clave de sus estrategias industriales, y ya están trabajando en la adopción de marcos y medidas normativas en el ámbito de la robótica y la inteligencia artificial, como he expuesto a lo largo de este investigación y así lo recogía expresamente, entre otras resoluciones y comunicaciones, la propia Resolución del

Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica, a que aludiré más adelante.

Los robots plantean los desafíos éticos y jurídicos propios de los sistemas de inteligencia artificial, así como específicos de los mismos por su estructura, características, capacidades, sector y contexto donde operan y su modo de interaccionar con el ser humano, su entorno y otras tecnologías.

Por lo que se refiere a la dimensión ética de los retos que plantean, la preocupación por la misma es incesante durante las últimas dos décadas, de hecho, la *Conferencia Internacional de Robótica* celebrada en Japón en 2004 hizo pública su *Declaración Mundial de la Robótica* continente de las tres expectativas esenciales sobre el futuro desarrollo de la denominada roboética[888].

Con posterioridad, tanto la *Organización de las Naciones Unidas* (ONU) como la *Organización para la Educación Científica y Cultural* (UNESCO) significaron la necesidad de un código ético relacionado con la robótica en el *Programa de Ética en la Ciencia y la Tecnología -Ethics of Science and Technology Programme-*.

La UE ha requerido la creación de un marco ético para el desarrollo y despliegue de los robots, así como la adopción de normas para regular los distintos tipos de robots, incluyendo los marcos de responsabilidad, y especialmente en materia de vehículos "autónomos".

El sexto Programa Marco de la Comisión Europea financió dentro del programa *Science and Society Work Programme* el proyecto *EthicBots* para promover y coordinar un grupo de expertos multidisciplinar con investigadores de inteligencia artificial, robótica, antropología, moral, filosofía, psicología y ciencia cognitiva, con el objetivo común de identificar y analizar los problemas tecno-éticos relativos a la integración de seres humanos y robots. Asimismo, el *Programa SPARC*[889] de la UE se focalizó en profundizar en la resolución de problemas éticos relacionados con el uso de robots.

[888] Recuperado de: htpp://robots.net/article/1113.html. Consultado el 04.01.2021.
[889] Recuperado de: https://www.eu-robotics.net/sparc/about/index.html. Consultado el 04.01.2021

A pesar de los esfuerzos, el seguimiento e implantación de los códigos éticos es moderado, por lo que los expertos demandan la necesaria armonización internacional de las diferentes iniciativas normativas y éticas, debiendo destacar el proyecto coordinación de robótica para Europa denominado *RockEU*[890], relacionado con el proyecto EURON, financiado por la UE y continuador del precitado proyecto EthicBots, a través del que los mayores actores de la industria pretenden conseguir, entre otros propósitos, una declaración universal relacionada con la ética de la robótica, o "roboética".

La Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica, la cual será abordada de nuevo posteriormente con más detalle por su singularidad inicial respecto de robótica, como ya referí en su análisis parcial en relación con el objeto de capítulos anteriores, incorporó en su anexo una *Carta sobre Robótica* como código de conducta ética en el campo de la robótica, que integra un *Código de Conducta Ética para los ingenieros en robótica*, un *Código Deontológico para los comités de ética de la investigación*, un *Modelo de licencia para los usuarios* -como conjunto de derechos y obligaciones de los mismos- y un *Modelo de licencia para los diseñadores de robots*, consistentes en un conjunto de principios y normas que los diseñadores deberán considerar sobre la base de la denominada *Ethics by design*, la *Security by design* y la *Privacy by design*.

Por lo que se refiere a la dimensión jurídica, las grandes potencias están analizando y abordando los marcos reguladores que deberán abordar los retos jurídicos que suponen los robots y la inteligencia artificial, complementando normas existes y/o creando nuevos marcos regulativos específicos.

Conforme abordé en los capítulos precedentes, la mayoría de países están aprobando nuevas normas sobre el uso de la inteligencia artificial y robótica y, en particular, algunos estados de EE.UU. han aprobado ya algunas leyes para limitar el uso de drones y están trabajando en marcos específicos para los coches autónomos, conforme expuse en el

[890] Recuperado de: http://cordis.europa.eu/project/rcn/111361_en.html. Consultado el 04.01.2021

capítulo IV de esta investigación. Del mismo modo, la UE pretende liderar a nivel internacional la regulación de la inteligencia artificial en distintos aspectos específicos.

En materia de robótica, la preocupación y concienciación sobre la necesidad de definir un marco jurídico es global, pero en especial, en referencia a esa categoría de robots a los que me he referido anteriormente, dotados de cierta "Inteligencia" y "autonomía". Es decir, las principales preocupaciones, discusiones y retos se sitúan en esto último, en la inteligencia artificial que pueda asociarse a una máquina o sistema, lo que ha motivado una iniciativa normativa para la aprobación de un nuevo Reglamento sobre Máquinas al que aludí en el capítulo II.

En este sentido, la UE ha venido liderando distintas propuestas con esta finalidad.

Una de las primeras reflexiones formales sobre robots inteligentes que se llevaron a cabo en el seno de la UE se produjo a partir de una pregunta parlamentaria[891] relativa a los derechos de los robots en 2013, que la Comisión resolvió rechazando la posibilidad de otorgar personalidad a los robots, especialmente al considerar la tecnología no estaba preparada todavía para ofrecer al grado de autonomía necesario para su reconocimiento. No obstante, abrió la puerta al debate y reflexión científica, ética y jurídica sobre si más adelante, en caso de disponer de esta autonomía, podrían ser dotados de dicha personalidad en el ámbito de la UE, dado que en otros países ya se les estaba reconociendo personalidad electrónica.

Por el camino, la Comisión ya había financiado un proyecto bajo el título *Regulating Emerging Robotic Technologies in Europe: Robotics facing Law and Ethics*[892],

[891] Texto de la pregunta: "La ley de robots parece estar materializándose de varias maneras, dado que los abogados están trabajando en la creación de una 'personalidad de robot' y en la asignación de números de seguridad social de robots, el Ministerio de Recuperación Productiva de Francia está redactando un borrador de carta de ética no vinculante. y la Comisión está considerando otorgar a los robots personalidad jurídica. Los robots que reemplazan a las personas son algo que la Comisión está comenzando a introducir con su proyecto Petrobot. Junto con un consorcio de 10 empresas europeas lideradas por la petrolera Shell, busca desarrollar robots que puedan reemplazar a los humanos en "inspecciones de recipientes a presión y tanques de almacenamiento ampliamente utilizados en la industria del petróleo, el gas y la petroquímica". Dice que otorgar estatus legal a robots y sistemas inteligentes es una opción y nada más. 1. ¿Cuál es la realidad de la situación? 2. ¿Cuál es el objetivo? 3. ¿Cuál es el presupuesto para esta área política?" http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+WQ+E-2013-011289+0+DOC+XML+V0//EN. Consultado el 05.01.2021

[892] Recuperado de: http://www.robolaw.eu/. Consultado también: KOOPS BJ Y A. PIRNI (2014) (Eds): "Aspectos éticos y legales de la mejora de las capacidades humanas a través de la robótica", en *Derecho,*

denominado también "Robolaw", con la finalidad de elaborar un informe detallado sobre las cuestiones éticas y jurídicas que plantear los robots, con definición de principios y orientaciones en la materia que sirvieran de guía para los legisladores europeos y nacionales. El informe se publicó en septiembre de 2014.

Entre otros aspectos, el informe precitado apuesta por la regulación tanto mediante *hard law* para garantizar la dignidad, la justicia, la solidaridad, la discapacidad, la no discriminación, los derechos fundamentales, la asistencia sanitaria o los derechos de los consumidores, así como mediante *soft law*, a través de normas y estándares técnicos, códigos de conducta y de buenas prácticas para alcanzar una mayor precisión y flexibilidad.

El informe también planteó la cuestión relativa al reconocimiento de personalidad jurídica diferenciada al robot para realizar ciertas transacciones, para ser considerado responsable o para comparecer en juicio.

Posteriormente al mismo, el Parlamento Europeo aprobó la precitada Resolución de 16 de febrero de 2017[893] con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica. La Resolución concluyó sobre la imperiosa necesidad de regular esta materia y se instó a la Comisión a que presentara una propuesta de Directiva relativa a las normas de la legislación civil en materia de robótica.

De nuevo, reiterar que esta Resolución se orientó especialmente a robots inteligentes y a la inteligencia artificial que gobierna realmente a la máquina, al robot, no a la máquina física o virtual en sí misma, distinguiendo y exigiendo una definición europea común de robots autónomos "inteligentes", incluyendo las definiciones de sus subcategorías teniendo en cuenta sus características, esto es, la capacidad de adquirir autonomía mediante sensores y/o mediante el intercambio de datos con su entorno (interconectividad) y el análisis de dichos datos, la capacidad de aprender a través de la

innovación y tecnología. Número especial. IX, 2/2013; y BERTOLINI A. (2013). "Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules*"*, en *Law, Innovation and Technology* 5 (2), Pp. 214-247. DOI: http://dx.doi.org/10.5235/17579961.5.2.214.

[893] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))

experiencia y la interacción, la forma del soporte físico del robot y la capacidad de adaptar su comportamiento y acciones al entorno.

Asimismo, el *Comité Económico y Social Europeo* (CESE) publicó su Dictamen sobre la inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad[894], de 31 de mayo, de 2017.

El Dictamen agrupó en once áreas los distintos retos que plantea la inteligencia artificial y abordó las cuestiones que mayor controversia están suscitando en la actualidad.

En primer lugar, aunque no expresamente, partió de un enfoque de la inteligencia artificial como un medio o instrumento y no como un fin en sí misma, sujeta a control humano y a su dominio.

En segundo lugar, el Comité Económico y Social Europeo negó cualquier tipo de personalidad jurídica para los robots o la inteligencia artificial por el riesgo moral que ello implicaría y el posible uso indebido que podría conllevar.

En tercer lugar, abordó la responsabilidad civil de los daños causados por un sistema de inteligencia artificial.

El Dictamen también abordó otras cuestiones de indudable repercusión, solicitando la elaboración de un código deontológico para el desarrollo, despliegue y utilización de la inteligencia artificial, abogando por un sistema de normalización en la industria en materia de seguridad, transparencia e inteligibilidad, apoyando la prohibición de armas autónomas y proponiendo una regulación mundial que debe liderar la UE.

Como expuse al abordar los retos y riesgos de la inteligencia artificial, los robots transformarán nuestro modo de vida y forma de trabajar. Impactan y seguirán impactando en el mercado laboral y la industria, suponen y supondrán un enorme impacto y reto para muchas profesiones, y no sólo las más físicas o mecánicas, sino otras más cualificadas,

[894] Dictamen del Comité Económico y Social Europeo sobre la "Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad» (2017/C 288/01). https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52016IE5369&from=ES

como en materia de asesoramiento financiero, diagnóstico médico o cirugía, y que exigirá su reconversión en lo que considero deberá ser un futuro construido sobre la base de la unión hombre-máquina, aprovechando la precisión de la máquina como medio a utilizar por el profesional para realizar mejor sus tareas y cometidos y mejorar al ser humano.

La Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica, resumió entonces muy bien la situación actual, reflejando las preocupaciones y retos que supone pero también algunas de las ventajas que supone la inteligencia artificial para el ser humano: "Considerando que durante los últimos doscientos años las cifras de empleo han aumentado de manera continuada gracias al desarrollo tecnológico; que el desarrollo de la robótica y de la inteligencia artificial tiene potencial para transformar el modo de vida y las formas de trabajo, aumentar los niveles de eficiencia, ahorro y seguridad y mejorar la calidad de los servicios, y que se espera que, a corto y medio plazo, la robótica y la inteligencia artificial traigan consigo eficiencia y ahorro, no solo en la producción y el comercio, sino también en ámbitos como el transporte, la asistencia sanitaria, las operaciones de salvamento, la educación y la agricultura, permitiendo que los seres humanos dejen de exponerse a condiciones peligrosas, como, por ejemplo, las que entraña la limpieza de lugares contaminados con sustancias tóxicas; (…) Considerando que el envejecimiento de la población se debe al aumento de la esperanza de vida propiciado por los avances en las condiciones de vida y en la medicina moderna, y que se trata de uno de los principales retos políticos, sociales y económicos a los que se enfrentan las sociedades europeas del siglo XXI; que en 2025 más de un 20 % de los europeos habrá cumplido los sesenta y cinco años, con un aumento especialmente rápido de la población mayor de ochenta años, lo que dará lugar a un equilibrio radicalmente diferente entre las generaciones dentro de nuestra sociedad, y que redunda en beneficio de la sociedad y de las familias que las personas de edad avanzada se mantengan saludables y activas el mayor tiempo posible*"*.

Una de las principales cuestiones que la robótica plantea, en congruencia con el objeto de esta investigación, es la responsabilidad derivada de las acciones de un robot que causen daños o perjuicios a un tercero, conforme ha sido objeto de análisis en esta investigación.

Si estamos en contexto contractual, se derivarán las responsabilidades propias del mismo, pero en el extracontractual, de nuevo, se diluyen algunas cuestiones, especialmente sobre quién será el sujeto responsable: ¿El fabricante, el diseñador, el programador, el propietario, el fabricante, el proveedor, el operador, el usuario, los entrenadores o la propia víctima de daños?

De nuevo, se hará necesario separar la responsabilidad moral o ética de la jurídica, dado que, en algunos casos, por eficiencia y con el objetivo de proporcionar una adecuada y efectiva protección de la víctima será más adecuado situar la responsabilidad jurídica en la órbita del fabricante o del operador en lugar de la del desarrollador, sin perjuicio de su derecho de repetición. En relación con todo ello, me remito al análisis efectuado en el capítulo V de esta investigación sobre responsabilidad civil de los sistemas de inteligencia artificial.

4. Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica[895]

4.1. Consideraciones iniciales

La Resolución abordó las principales inquietudes del Parlamento Europeo y sus recomendaciones a la Comisión sobre determinados aspectos a considerar en la elaboración del futuro marco europeo, considerando que un enfoque adecuado, eficiente, transparente, coherente y global en la UE beneficiará a sus industria y permitirá que tanto la UE como los Estados miembros conserven el control sobre el marco normativo que se haya de establecer, de modo que no se vean obligados a adoptar o aceptar normas establecidas por terceros países que también están a la vanguardia del desarrollo de la robótica y la inteligencia artificial.

[895] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))

De inicio, me permito significar algunos aspectos recogidos en la misma.

Las consideraciones preliminares se encabezaron con una alusión a la literatura de ciencia-ficción, y no es la única, al señalar que los seres humanos han fantaseado siempre con la posibilidad de construir máquinas inteligentes, sobre todo androides con características humanas, desde *"*el monstruo de Frankestein creado por Mary Shelley al mito clásico de Pigmalión, pasando por el Golem de Praga o el robot de Karel Čapek"[896]".

Y destaco esta cuestión positivamente por las connotaciones negativas y minusvaloraciones que algunos científicos, juristas y filósofos, todavía hoy, dan a obras literarias y cinematográficas -como en su día se le dio a algunos de los diseños y dibujos de Leonardo da Vinci-, que plasmaron realidades entonces sólo existentes en la imaginación de su autor e irrealizables en el momento de su concepción, pero que posteriormente se han convertido en realidad gracias a la grandeza del ser humano y su capacidad de convertir en real algo que, en el momento de su concepción, era el producto de la imaginación de un novelista o un inventor.

La literatura y el cine han supuesto desafíos intelectuales que el ser humano ha abordado y ha resuelto.

En este sentido, la literatura y el cine han reflejado en tantas ocasiones un guión de una realidad posteriormente creada, que considero injusto su rechazo directo y descalificación como factor distorsionador de la percepción actual sobre los riesgos éticos y jurídicos de la inteligencia artificial, cuando precisamente ha servido para anticiparnos de manera muy gráfica la materialización posterior de algunos de esos riesgos.

El Parlamento Europeo justificó su Resolución en la necesidad de que el legislador pondere las consecuencias jurídicas y éticas que comporta una nueva revolución industrial sustentada sobre los robots, sin obstaculizar la innovación.

[896] Escritor en lengua checa al que se la atribuya haber acuñado por primera vez el concepto de "robot". El término aparece por primera vez en su obra de teatro de ciencia ficción R.U.R. -*Rossumovi univerzální roboti*-, escrita en 1.920 y estrenada en 1.921, en referencia a humanos artificiales orgánicos construidos con el fin de aligerar la carga de trabajo del resto de personas, lo que hoy denominaríamos androides.

Asimismo, consideró necesario crear una definición generalmente aceptada de robot y de inteligencia artificial, especialmente ante sus distintas clases y falta de consenso, como he analizado en el anterior apartado.

La Resolución reflejó en sus consideraciones iniciales las dimensiones económicas del mercado de los robots, afirmando que entre 2010 y 2014, las ventas de robots aumentaron un 17 % de media cada año, que en 2014 las ventas registraron el mayor incremento anual observado hasta ahora -a saber, un 29 %-, que los principales motores de este crecimiento fueron los proveedores de componentes de automoción y la industria electrónica y eléctrica y que a lo largo del último decenio se han triplicado las solicitudes anuales de patentes en el sector de la tecnología robótica.

De igual modo, se hizo eco de los aspectos positivos que puede suponer la tecnología y, en particular, los robots y la inteligencia artificial, como el aumento continuado de las cifras de empleo durante los últimos doscientos años, la transformación del modo de vida y las formas de trabajo, aumentando los niveles de eficiencia, ahorro y seguridad y mejora de la calidad de los servicios, y tanto en materia del transporte, como de la asistencia sanitaria, las operaciones de salvamento, la educación y la agricultura, permitiendo que los seres humanos dejen de exponerse a condiciones peligrosas.

También significó una realidad actual innegable, como lo es el envejecimiento de la población[897] gracias al aumento de la esperanza de vida y los retos que supone.

Pero también significó la preocupación ante el desarrollo de máquinas inteligentes y autónomas, con capacidad de ser entrenadas para pensar y tomar decisiones de manera independiente, dado los retos que plantea para garantizar la no discriminación, la privacidad, la intimidad, la dignidad, la autonomía y la autodeterminación del individuo, las garantías procesales, la transparencia y la inteligibilidad de los procesos decisorios, así como para el mercado de trabajo, que exige la necesidad de reflexionar en consecuencia sobre el futuro de la educación, el empleo y las políticas sociales.

[897] Según la Resolución objeto de análisis, en 2025 más de un 20 % de los europeos habrá cumplido los sesenta y cinco años, con un aumento especialmente rápido de la población mayor de ochenta años, lo que dará lugar a un equilibrio radicalmente diferente entre las generaciones dentro de la sociedad.

El Parlamento Europeo consideró que las aplicaciones de inteligencia artificial deben incorporar desde el principio en su desarrollo y comercialización las características de seguridad y éticas, es decir, la *Ethics & security by design* que considero absolutamente necesaria, considerando que los fabricantes deben ser responsables de la calidad de la tecnología que producen.

Y por último, la Resolución concluyó sus consideraciones preliminares destacando que existía la posibilidad de que a largo plazo la inteligencia artificial llegue a superar la capacidad intelectual humana, conforme opinan algunos científicos citados en esta investigación, y que es necesario integrar salvaguardias y la posibilidad de control y verificación por parte de las personas en los procesos de toma de decisiones automatizados y basados en algoritmos, lo que nunca supondría una autonomía plena y garantizaría la supervisión y control humano. Sin duda, una visión de un futuro cada vez más próximo para el que debemos estar preparados.

A continuación, me permito exponer los principales aspectos que abordó la Resolución por su relevancia, muchos ya anticipados al abordar otros aspectos de esta investigación, que deben servir también de base para los futuros marcos reguladores de los robots y sistemas dotados de inteligencia artificial, si bien, muchos de ellos han sido ya considerados, revisados e integrados en las sendas propuestas regulatorias del Parlamento Europeo que acompañaban a la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas, y a la Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial, así como en la reciente Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021.

La exposición seguirá el mismo orden seguido en su tratamiento por la Resolución objeto de análisis.

De inicio, la Resolución objeto de análisis efectuó una nueva referencia expresa a la literatura de ciencia ficción, considerando que ya las leyes de Asimov[898], a las que hice referencia en capítulos anteriores, "*v*an dirigidas a los diseñadores, fabricantes y operadores de robots, incluidos los que disponen de autonomía y capacidad de autoaprendizaje integradas".

De nuevo la literatura se adelantó a la realidad, fijó una ficción deseable, definió las leyes básicas que debían regir la relación entre seres humanos y máquinas, y el ser humano se ha encargado de convertirla en realidad. Pero, además, aquella estableció las normas básicas que debían regirla entonces y ahora, y el órgano legislativo europeo las ha tomado como referencia y pretende otorgarles eficacia vinculante para diseñadores, fabricantes y operadores, de ahí mis consideraciones anteriores sobre la literatura y el cine.

En relación con todo ello, el Parlamento Europeo consideró necesario establecer un marco normativo en materia de responsabilidad, transparencia y rendición de cuentas que refleje los valores humanistas europeos y universales, estableciendo principios éticos que sirvan de marco para el desarrollo, programación y utilización de robots y de inteligencia artificial e incorporándolos a la normativa y códigos de conducta de la UE.

Con este propósito, la Resolución adjuntó una Carta sobre robótica elaborada con la asistencia de la *Unidad de Prospectiva Científica (STOA)* de la *DG European Parliament Research Service*, en la que se propone un *Código de conducta ética para los ingenieros en robótica*, un *Código deontológico para los comités de ética de la investigación*, una *Licencia para los diseñadores* y una *Licencia para los usuarios*. En posteriores apartados abordaré los mismos.

El Parlamento significó la necesidad de adoptar una actitud gradual, pragmática y prudente para abordar el futuro marco normativo, protegiendo todos los intereses en juego y sin perjudicar a innovación.

[898] 1ª. Un robot no hará daño a un ser humano ni permitirá que, por inacción, este sufra daño; 2ª Un robot obedecerá las órdenes que reciba de un ser humano, a no ser que las órdenes entren en conflicto con la primera ley; 3ª Un robot protegerá su propia existencia en la medida en que dicha protección no entre en conflicto con las leyes primera y segunda (véase Isaac Asimov, Círculo vicioso (Runaround), 1943); y 0.ª Un robot no hará daño a la humanidad ni permitirá que, por inacción, esta sufra daño.

4.2. Autonomía y responsabilidad jurídica

La Resolución definió el concepto "autonomía" relacionada con un robot como la capacidad de tomar decisiones y aplicarlas en el mundo exterior, con independencia de todo control o influencia externos y considera que, cuanto más autónomos sean los robots, más difícil será considerarlos simples instrumentos en manos de otros agentes, como el fabricante, el operador, el propietario, o el usuario.

La toma de decisiones y su aplicación al margen de cualquier control obviamente parece evidenciar una inteligencia artificial no sujeta al necesario control y supervisión humana que proclama dicha Resolución y las posteriores, lo que obviamente exigiría matizar dicha definición de inicio y, además, parece partir de un concepto de inteligencia artificial específico y superior a la débil, que igualmente estaría en conflicto con el propio objeto y contenido de la Resolución.

En relación con este concepto de autonomía planteó dos cuestiones principales: a) La naturaleza jurídica de los robots y la posibilidad de incardinarlos en alguna de las categorías existentes o si debe crearse una nueva categoría por sus características jurídicas; b) Si la normativa general sobre responsabilidad es suficiente o si se requerirían normas y principios específicos que aclaren la responsabilidad jurídica de los distintos agentes y su responsabilidad por los actos y omisiones de los robots cuya causa no pueda atribuirse a un agente humano concreto.

Estos últimos aspectos fueron expuestos y analizados durante en el capítulo V de esta investigación.

El avance incesante en la *biorobótica*, biónica, *neurorobótica* o robótica humanoide dificultan cada vez más el establecimiento de vinculación y relación causal con su actividad y decisiones. En este contexto, los futuros análisis sobre la responsabilidad deberán considerar algunos aspectos clave como una hipotética autonomía plena, el conocimiento, la voluntad y el control sobre la decisión o acto realizado.

Sin embargo, la nueva propuesta europea reguladora de la inteligencia artificial de 21 de abril de 2021, se aparta radicalmente del concepto autonomía para la definición y

regulación de los sistemas de inteligencia artificial, en particular la denominada inteligencia artificial "débil", como analicé anteriormente.

4.3. Principios generales relativos al desarrollo de la robótica para uso civil

4.3.1. Definiciones comunes

El Parlamento instó a la Comisión para que propusiera definiciones europeas comunes de sistema *ciberfísico*, sistema autónomo, robot autónomo inteligente y sus distintas subcategorías, tomando en consideración las características que debe tener un robot inteligente, como son: a) Capacidad de adquirir autonomía mediante sensores y/o mediante el intercambio de datos con su entorno (interconectividad) y el intercambio y análisis de dichos datos; b) Capacidad de autoaprendizaje a partir de la experiencia y la interacción (criterio facultativo); c) Soporte físico mínimo; d) Capacidad de adaptar su comportamiento y acciones al entorno; e) Inexistencia de vida en sentido biológico.

4.3.2. Creación de un registro de robots

El Parlamento Europeo consideró necesario la creación de un sistema global de registro de robots "avanzados" dentro del mercado interior de la Unión para determinadas categorías instando a la Comisión que establezca criterios para la clasificación de los robots que tendrían que registrarse en el mismo. Asimismo, instó a la Comisión a analizar la conveniencia de que la gestión del sistema de registro y de las inscripciones se atribuya a una agencia de la UE para la robótica y la inteligencia artificial.

El Parlamento Europeo significó la tecnología robótica debe orientarse a complementar las capacidades humanas y no a sustituirlas, que el ser humano debe tener en todo momento el control sobre el robot o sistema de inteligencia artificial, y que debe prestarse especial atención a los vínculos entre humanos y robots, especialmente en caso de grupos vulnerables. La Resolución se focaliza en el control y supervisión humana, principio ético esencial que ha formado parte de todas las propuestas posteriores y que, sin embargo, sólo

se exige a los sistemas de alto riesgo en la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial -*Artificial Intelligence Act*-, de 21 de abril de 2021.

4.3.3. Marco armonizado y reconocimiento mutuo en el uso transfronterizo de robots

La realización de ensayos, la certificación y la autorización de comercialización deberían exigirse solo en un Estado miembro según la Resolución.

4.3.4. Necesidad de apoyo a las empresas del sector de la robótica

El Parlamento Europeo determinó la necesidad de apoyar a las PYMEs y *startups* del sector de la robótica que creen nuevos segmentos de mercado en este ámbito o que utilicen robots en sus actividades, si bien, no recogió propuestas específicas para ello.

4.3.5. Investigación e innovación

El Parlamento Europeo consideró esencial que la UE y lo Estados miembros sigan estando en la vanguardia de la investigación en robótica e inteligencia artificial con ayuda de la financiación pública, e instó a que se fomenten los programas de investigación y un aumento del apoyo financiero a proyectos de investigación en materia de robótica y TIC y que apliquen en sus políticas de investigación los *principios de ciencia abierta e innovación ética responsable*. También destacó la necesidad de destinar recursos suficientes a la búsqueda de soluciones a los retos sociales, éticos, jurídicos y económicos que plantea.

Asimismo, instó a la Comisión y a los Estados miembros para que aunaran sus esfuerzos para supervisar la transición de estas tecnologías de la investigación a la comercialización

y uso en el mercado, y para garantizar que esta transición se produzca de forma adecuada tras las oportunas evaluaciones de su seguridad conforme al principio de precaución.

El Parlamento también significó la necesidad una infraestructura digital que ofrezca una conectividad ubicua para el despliegue de todas estas tecnologías, instando a la Comisión a que establezca un marco que satisfaga las necesidades de conectividad y que garantice que el acceso a la banda ancha y a las redes 5G conformidad al principio de neutralidad de la red.

Y por último, destacó la necesidad de garantizar la interoperabilidad entre los sistemas, los dispositivos y los servicios en nube construida sobre la seguridad y la privacidad en el diseño S*ecurity and privacy by design*- como algo imprescindible para proporcionar mayor flexibilidad y autonomía de los robots y la inteligencia artificial, instando a la Comisión a que promoviera un entorno abierto (normas, licencias, plataformas y transparencia) que eviten la restricción a sistemas privados que puedan limitar la interoperabilidad.

4.4. Principios éticos

Los principios éticos fueron analizados en el apartado 4º del capítulo III de esta investigación, a los que me remito.

4.5. Agencia europea para la robótica y la inteligencia artificial

El Parlamento Europeo instó ya a la Comisión para que estudie la posibilidad de crear una agencia europea para la robótica y la inteligencia artificial que proporcionara los conocimientos técnicos, éticos y normativos necesarios para apoyar la labor de los actores públicos pertinentes, tanto a nivel de la Unión como a nivel de los Estados miembros.

4.6. Derechos de propiedad intelectual, flujo de datos, seguridad y privacidad

El Parlamento Europeo instó a la Comisión para que apoyara un enfoque horizontal y de neutralidad tecnológica para la propiedad intelectual en los distintos sectores en que se pueda utilizar la robótica y que la legislación civil en el sector de la robótica se ajuste al RGPD. Igualmente destacó que el libre flujo de datos es fundamental para la economía digital y para el desarrollo en el sector de la robótica y la inteligencia artificial, cuestiones que abordaré en el capítulo VIII.

Como ya anticipé al tratar los aspectos de seguridad en anteriores capítulos, el Parlamento Europeo destacó que la utilización adecuada de la robótica y la inteligencia artificial requiere un alto grado de seguridad de estos sistemas.

En este sentido, significó la necesidad de garantizar la protección de redes de robots y sistemas de inteligencia artificial interconectados para evitar posibles quiebras de seguridad, así como la comunicación entre humanos y robots y sistemas de inteligencia artificial, que debe llevarse a cabo bajo elevados niveles de seguridad y de protección de la privacidad.

El Parlamento significó la responsabilidad de los diseñadores de robots y sistemas de inteligencia artificial de desarrollar productos que sean seguros, fiables y que cumplan su función, e insta a la Comisión Europea y a los Estados miembros a que apoyen e incentiven el desarrollo de la tecnología necesaria con este propósito, en especial, la seguridad desde el diseño -*Security by design*-.

4.7. Normalización, interoperabilidad, seguridad y protección

El Parlamento Europeo destacó la necesidad de definir normas y garantizar la interoperabilidad para la competencia futura en el ámbito de la inteligencia artificial y las tecnologías robóticas, instando a la Comisión que continúe trabajando por la armonización internacional de las normas técnicas a fin de fomentar la innovación, evitar la fragmentación del mercado interior y garantizar un elevado nivel de seguridad de los

productos y protección de los consumidores, también, en su caso, mediante normas mínimas de seguridad adecuadas para el entorno de trabajo.

En sus recomendaciones, significó la conveniencia de disponer de acceso al código fuente, a los datos de entrada y a los detalles de construcción del sistema cuando sea necesario, en el marco de la investigación de los accidentes como los daños causados por robots inteligentes, así como con el objetivo de velar por su funcionamiento, disponibilidad, fiabilidad, seguridad y protección continua.

Este acceso supone un conflicto con otros derechos como los de propiedad intelectual y los secretos empresariales que pueden ver limitados en virtud de la transparencia, explicabilidad, responsabilidad y *accountability* de la inteligencia artificial.

Asimismo, destacó la importancia de la licitud de la ingeniería inversa y las normas abiertas para maximizar el valor de la innovación y garantizar que los robots puedan comunicarse entre sí y se muestra favorable en la creación de comités técnicos especiales, como el ISO/TC 299 Robotics, dedicados exclusivamente a la elaboración de normas sobre robótica.

Por último, en esta materia destacó la necesidad de los ensayos de robots en situaciones reales es esencial para determinar y evaluar los riesgos que puedan entrañar, así como para su desarrollo tecnológico más allá de la mera fase experimental en el laboratorio, instado a la Comisión a establecer criterios uniformes para su realización.

4.8. Medios de transporte autónomos

El Parlamento Europeo abordó sus consideraciones sobre drones y vehículos autónomos y automatizados, especialmente ante la afectación que estos últimos supondrán en materia de responsabilidad civil, seguridad vial, medio ambiente (por ejemplo, eficiencia energética, uso de tecnologías renovables y fuentes de energía), tratamientos de datos (por ejemplo, acceso a los datos, protección de los datos personales y la intimidad, intercambio de datos), infraestructuras TIC (por ejemplo, alta densidad de comunicaciones eficientes y fiables) y también en materia empleo (por ejemplo, creación y pérdida de puestos de

trabajo, formación de los conductores de vehículos pesados para el uso de vehículos automatizados).

4.9. Robots asistenciales, robots médicos, rehabilitación e intervenciones en el cuerpo humano

El Parlamento Europeo trasladó también a la Comisión sus consideraciones y recomendaciones sobre esta tipología.

El Parlamento destacó que la utilización de estas tecnologías era creciente, especialmente en materia de prevención, asistencia, seguimiento, estimulación y compañía en personas de edad avanzada o con determinado grado de dependencia y necesidades especiales, advirtiendo que la sustitución del necesario contacto humano por robots podría deshumanizar la prestación de estos cuidados, si bien, por otra parte, podría mejorar la atención y la rehabilitación en colaboración con el personal médico, más focalizado en el diagnóstico y en la planificación de tratamientos, en mi opinión, más personalizados.

No obstante, significó que, a pesar del potencial de la robótica para mejorar la movilidad e integración de estas personas, seguirá siendo necesaria la intervención humana, en especial, los cuidadores, dado que son una fuente de interacción social imposible de sustituir en su integridad en estos contextos de uso.

Respecto de los robots de uso médico, destacó, la importancia de una adecuada educación, formación y preparación de todos los profesionales de la salud, incluidos médicos como auxiliares sanitarios, para garantizar "el nivel más elevado posible de competencia profesional y proteger y salvaguardar la salud de los pacientes". El Parlamento Europeo significó especialmente la necesidad de definir los requisitos profesionales que debe cumplir un cirujano para poder operar y estar autorizado a utilizar robots quirúrgicos, en cuyas intervenciones deberá estar garantizada la supervisión humana, incluidos cuidados y ejecución. Del mismo modo que respecto de servicios asistenciales, destacó la importancia de la formación de los profesionales y significó la creciente tendencia al autodiagnóstico mediante el uso de robots móviles, lo que no debería afectar a la relación

entre médico y paciente, sino que debería tener una finalidad de apoyo o asistencia al médico para el diagnóstico y/o el tratamiento de los pacientes, con la finalidad reducir el riesgo de error humano y aumentar la calidad y la esperanza de vida.

En este sentido he manifestado mi opinión a lo largo de esta investigación, sobre la necesidad de simbiosis entre hombre-máquina y, en especial, en el ámbito médico, donde los sistemas inteligentes podrían tener encomendado un prediagnóstico y, en su caso, prescripción de tratamiento, conforme a un protocolo preestablecido, todo ello a validar por el profesional médico supervisor, en base a su conocimiento especializado sobre el área médica y sobre la tecnología utilizada. Efectivamente, como destacaba el Parlamento en su Resolución, considero que esto contribuiría a diagnósticos y tratamientos basados en experiencias y conocimiento acumulado y compartido de miles de médicos y millones de casos, y su experiencia profesional acumulada durante decenas de años a lo largo de toda su trayectoria profesional, de modo que podría aportar, entre otros aspectos, agilidad y precisión en los mismos, detección de patrones distintos y reducción de error humano.

El Parlamento también abordó algunas de las bondades de la utilización de robots en medicina avanzada, en especial, la optimización de recursos y la reducción de gastos sanitarios. Del mismo modo, pidió a la Comisión que la seguridad sea garantizada en los procedimientos utilizados para ensayar nuevos dispositivos robóticos médicos, especialmente en el caso de dispositivos implantados en el cuerpo humano. Asimismo, destacó el potencial de la robótica en el ámbito de la rehabilitación de órganos dañados y el restablecimiento de funciones corporales reducidas, si bien, significando las cuestiones éticas que ello comporta y la necesidad urgente de comités de ética sobre robótica en hospitales y otras instituciones sanitarias, para lo que deben desarrollarse directrices para ayudar al establecimiento y funcionamiento de dichos comités por la Comisión y los Estados miembros.

Respecto de aplicaciones médicas vitales de robots, como las prótesis robóticas, consideró que debe garantizarse el acceso continuo y sostenible al mantenimiento, la mejora y, en particular, las actualizaciones de *software* que subsanen fallos y vulnerabilidades, así como crear entidades de confianza independientes que dispongan de los medios necesarios para proporcionar a las personas que lleven dispositivos médicos vitales y avanzados los servicios que precisen, como mantenimiento, reparaciones,

mejoras o actualizaciones, sugiriendo que los fabricantes proporcionen instrucciones de diseño global, incluido el código fuente, a estas entidades de confianza independientes.

En materia de seguridad, destacó los riesgos de manipulación, borrado o desconexión de los sistemas *ciberfísicos* integrados en el cuerpo humano que pueden poner en peligro la salud y vida humanas, debiendo de disponer de la protección adecuada.

Y, por último, significó la necesidad de garantizar la accesibilidad a estas innovaciones tecnológicas por toda la sociedad en plano de igualdad.

4.10. Educación, empleo y medio ambiente

El Parlamento Europeo también reflejó sus consideraciones y recomendaciones sobre estos aspectos en la Resolución, destacando las previsiones de la Comisión sobre la escasez de profesionales en el sector de las TIC y la necesidad del desarrollo de las competencias digitales en todos los grupos de edad, especialmente ante un nuevo entorno con robots.

Asimismo, significó el enorme potencial de la robótica en el ámbito laboral para mejorar la seguridad en los entornos de trabajo mediante la ejecución de las tareas más peligrosas para el ser humano, así como la necesidad de normas que regulen las interacciones entre el ser humano y los robots, con el objetivo de garantizar la salud, la seguridad y el respeto de los derechos fundamentales en el lugar de trabajo.

En relación con el medio ambiente, instó a que el desarrollo de la robótica y de la inteligencia artificial se realice de modo que se limite el impacto en el medio ambiente.

4.11. Responsabilidad civil

El Parlamento Europeo consideró que la responsabilidad civil por los daños y perjuicios causados por robots era una cuestión fundamental que debe analizarse y abordarse a

escala de la Unión, con la finalidad de garantizar el mismo grado de transparencia, coherencia y seguridad jurídica en toda la UE, en beneficio de ciudadanos, consumidores y empresas.

El posicionamiento del Parlamento Europeo en esta Resolución fue abordado al analizar el capítulo V, remitiéndome a mis consideraciones allí expuestas, sin perjuicio de su posterior revisión y postulado en su Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial, que fue igualmente analizada con profundidad en el mismo.

4.12. Aspectos internacionales

El Parlamento consideró que no era necesario modificar de forma sustantiva y con carácter de urgencia las actuales normas de Derecho internacional privado en materia de accidentes de tráfico aplicables en la UE para adaptarlas al desarrollo de los vehículos autónomos, si bien considera que la simplificación del actual sistema dual con el que se determina la legislación aplicable[899] mejoraría la seguridad jurídica y limitaría las posibilidades de búsqueda del foro más favorable.

Sin embargo, si consideró necesario modificar algunos acuerdos internacionales, como el Convenio de Viena sobre la circulación vial, de 8 de noviembre de 1968, y el Convenio de la Haya sobre la ley aplicable en materia de accidentes de circulación por carretera.

Asimismo, el Parlamento animó a la comunidad internacional a cooperar para estudiar los desafíos sociales, éticos y jurídicos para, posteriormente, establecer normas reglamentarias bajos el patrocinio de las Naciones Unidas.

Y, por último, el Parlamento Europeo destacó que las restricciones y condiciones establecidas en el Reglamento (CE) n.º 428/2009 del Parlamento Europeo y del Consejo

[899] Reglamento (CE) n.º 864/2007 del Parlamento Europeo y del Consejo y el Convenio de La Haya de 4 de mayo de 1971 sobre la ley aplicable en materia de accidentes de circulación por carretera.

sobre el comercio de los productos de doble uso deberían extenderse a las aplicaciones de la robótica, esto es, en referencia a productos, programas informáticos y tecnología que puedan utilizarse para aplicaciones tanto civiles como militares o que puedan contribuir a la proliferación de armas de destrucción masiva.

4.13. Carta sobre Robótica

Una de las principales aportaciones de la Resolución analizada es la *Carta sobre Robótica*[900] que incorporó como anexo a la misma y que recogió los principios que debería ser considerado para la elaboración de futuras propuestas legislativas. La misma fue analizada en el apartado 3.2. del capítulo III de esta investigación.

La Carta sobre Robótica se acompañó de un *Código de conducta ética para los ingenieros en robótica*, de un *Código deontológico para los comités de ética de la investigación* y de sendos modelos de licencia para diseñadores y para usuarios. Estos documentos fueron igualmente analizados en el capítulo precitado de esta investigación.

4.14. Aspectos finales

El Parlamento Europeo instó a la Comisión a que presentara una propuesta de Directiva relativa a las normas de legislación civil en materia de robótica, en base a los artículos 225 y 114 del TFUE[901], lo que no se produjo, si bien, sucedieron las Resoluciones y Propuestas de Reglamento objeto de análisis en los capítulos precedentes en materia de regulación de inteligencia artificial, ética y responsabilidad civil.

[900] https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_ES.html#title2
[901] https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:12012E/TXT&from=ES

5. Personalidad jurídica de los robots

La posibilidad de que un robot o un sistema dotado de inteligencia artificial pueda ser titular de derechos y obligaciones es una cuestión controvertida a nivel internacional, como he expuesto en el capítulo V, y suscitó un importante debate en el seno de la UE que, a mi juicio y conforme expuse al analizar distintos aspectos donde he tenido ocasión de abordar esta cuestión, ha sido resuelto por el momento, al quedar al margen de las nuevas propuestas reguladoras de la UE en materia de inteligencia artificial.

El Parlamento Europeo ha dejado clara su postura actual, negando esta posibilidad. No obstante, considero que es una cuestión que necesariamente será objeto de revisión en el futuro en función del grado de autonomía y capacidades de las que realmente pueda llegar a dotarse a un robot o sistema de inteligencia artificial.

En este sentido, sin perjuicio de que la cuestión haya sido ya abordada a lo largo de esta investigación en materia de responsabilidad ética y jurídica de los sistemas de inteligencia artificial, considero necesario abordar y profundizar más sobre esta cuestión en relación con los robots que puedan estar dotados de este tipo de inteligencia.

La persona no es creada por el Derecho, sino que es prexistente al mismo. La personalidad se adquiere por naturaleza desde el momento del nacimiento con vida, de conformidad con lo previsto en artículo 30 del Código Civil español. Pero también se puede adquirir dicha condición por atribución, como sucede con las personas jurídicas, conforme establecen los artículos 35 a 39 del Código Civil precitado.

El concepto de persona jurídica se regula en el artículo 35 del Código Civil español, comprendiendo en el mismo corporaciones, fundaciones y asociaciones a las que la ley conceda personalidad propia, independientemente de la de cada uno de los asociados.

La personalidad otorga capacidad jurídica, concebida como la aptitud para ser titular de derechos y obligaciones.

Un sistema dotado de inteligencia artificial más avanzada, más autónoma o incluso "fuerte" o un robot, máquina o producto dotado de la misma, no es nada de estas dos cosas, ni persona física ni persona jurídica, siendo únicamente concebible como objeto,

bien o servicio en el tráfico, sin perjuicio de que, en mi opinión, conforme he expuesto en otros apartados de esta investigación y profundizaré más adelante, podrían ser considerados *entes sin personalidad jurídica* a los efectos de que pudiera considerarse en el futuro atribuir a los mismos facultades, funciones y obligaciones específicas asociadas.

La personalidad a conferir a una inteligencia artificial sí debe ser creada por el ordenamiento jurídico, otorgándole las capacidades y libertades que éste defina.

En este contexto es en el que se ha planteado la posibilidad o necesidad de crear un tipo de personalidad jurídica por atribución a los sistemas de inteligencia artificial en función de sus características y capacidades, especialmente autoaprendizaje, autonomía, libertad y relativa impredecibilidad que, realmente y en mi opinión, jamás sería plena, entre otras razones, por razones de seguridad y en virtud del necesario control y la supervisión humana impuesto por la ética y, espero que en breve, por los nuevos marcos jurídicos reguladores de la inteligencia artificial.

Esta atribución de "personalidad electrónica" podría permitir hipotéticamente atribuirles la responsabilidad por los daños causados, así como la disposición de determinados fondos y propiedades, liberando en cierto modo de responsabilidades a fabricantes y usuarios, tal y como también refrendan autores como Cerka, Grigiene o Sirbikyté[902].

Abundando sobre estas reflexiones iniciales, la persona física tiene una individualidad innata resultante de su propia naturaleza. Una posible personalidad jurídica a un ente o cosa sería una individualidad creada artificialmente por el Derecho para una realidad igualmente artificial.

Los rasgos y atributos que definen una persona incluyen conciencia, voluntad, moralidad, capacidad de sentir y proyectar afecto o emociones, y libertad. En base a estos atributos, el Derecho reconoce a la persona un conjunto de derechos de la personalidad para que

[902] GARCÍA TERUEL, R.M. (2021). "El derecho de daños ante la inteligencia artificial y el machine learning: una aproximación desde las recomendaciones del Parlamento Europeo y del Grupo de Expertos de la Comisión Europea", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Tomo II. Ed. Aranzadi-Thomson Reuters. 2021. P. 1.020.

pueda desarrollar sus facultades con normalidad y permitirle actuar libremente conforme a su estatus personal[903].

Un robot o sistema dotado de inteligencia artificial más avanzada o "fuerte" no podrá ser considerado auténtica persona, al carecer de los atributos propios de la misma, conforme al estado actual de la tecnología.

El robot no es libre, ni debería nunca serlo conforme a los principios y normas éticas básicas más elementales y al ordenamiento jurídico vigente. Y ni es consciente de sus actos ni de sus consecuencias, sus acciones están predeterminadas[904] en mayor o menor medida, por lo que considero inviable pretender atribuirle una responsabilidad ante la inexistencia de toda capacidad de discernimiento para entender el alcance de sus actos.

No obstante, podríamos reflexionar por analogía sobre los regímenes precitados de responsabilidad en relación con menores de edad, personas con discapacidad o propietarios/poseedores de animales y, en especial, de los considerados peligrosos, conforme analicé en el capítulo V de esta investigación.

Conforme significan algunos autores como Núñez Zorilla o Rogel Vide[905], por muy compleja y avanzada que sea la inteligencia de un robot, no dejan de ser máquinas o cosas que deberían poder ser desconectadas en cualquier momento.

Dejando a un lado mi opinión actual sobre esta cuestión, que he expuesto a lo largo de esta investigación, quizás, a corto plazo, considero que deberíamos reflexionar con mayor profundidad, más que en el reconocimiento de personalidad electrónica, en la posible creación de un estatus jurídico específico de los robots conforme a sus características y dentro de la categoría de "cosa", "servicio" o "ente" con el objetivo de proteger

[903] MONTÉS PENADÉS, V.L. (2011). "El significado institucional y técnico de la idea de persona", en Blasco GASCO, F. (Coord.). *Derecho Civil. Parte General. Derecho de la Persona*. Tirant Lo Blanch, Valencia, 2011. Pp. 153, 154 y 156.
[904] O'SULLIVAN, S. ET AL. (2019). "Legal, regulatory and ethical framework for development of standards in artificial intelligence (AI) and autonomous robotic surgery". *Int J Med Robotics Comput Assist Surg*, Nº 15. 2019. P. 7
[905] ROGEL VIDE, C. (2018). "Robots y personas". *Revista General de Legislación y Jurisprudencia*, Nº 1. Editorial Reus. Madrid 2018. Pp. 87 y 88.

determinados intereses de la sociedad o de las personas, o si sería atribuible una capacidad limitada o restringida, nunca plena, así como una capacidad representada, no directa.

El ordenamiento jurídico ha reconocido personalidad jurídica no sólo a la persona humana individualmente considerada, sino también a entidades, organizaciones o agrupaciones de personas a las que se les reconoce la facultad de relacionarse con otras personas como sujeto independiente a quienes las conformen, y con capacidad para ser titular de derechos y obligaciones. Y dicho reconocimiento se otorga en base a un sentido y utilidad para la sociedad.

El Derecho otorga a las personas jurídicas una individualidad artificialmente creada por el mismo que les confiere una personalidad regulada y limitada por la Ley, creando un "ser" ficticio o lo que autores como Núñez Zorrilla[906] denominan "cosa personificada".

De este modo, las entidades dotadas de personalidad jurídica, como las sociedades mercantiles, son titulares de derechos y responden por ilícitos civiles, administrativos, fiscales o laborales, e incluso también por ilícitos penales en España, tal y como expuse en el anterior capítulo y de conformidad con lo dispuesto en el artículo 31 bis del *Código Penal* español.

La personalidad así otorgada no comporta el mismo estatus que las personas naturales, es diferente, y tiene una finalidad instrumental bajo un marco de actuación, capacidades, prerrogativas y responsabilidades reguladas por el ordenamiento jurídico con determinados objetivos, que determinan sus condiciones de uso y sus límites.

El debate suscitado en el seno de la UE en relación con la posibilidad de otorgar personalidad jurídica a determinados sistemas inteligentes fue planteado como posible opción o medio para la atribución de la responsabilidad derivada de sus actuaciones frente a la sociedad.

En este contexto, se debatió la posibilidad y conveniencia de otorgar en el futuro personalidad jurídica específica a robots y sistemas dotados de inteligencia artificial más

[906] NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Op. Cit. P. 57.

avanzada o "fuerte", con un sentido y un objetivo principal, esto es, el de proteger los intereses de la sociedad, y sometida en todo momento al beneficio e interés del ser humano.

En este supuesto, en mi opinión, se trataría de una personalidad específica creada "a la carta" por el ser humano sobre una cosa "singular" o "especial", concibiendo la tecnología como lo que debe ser, esto es, un medio al servicio de la humanidad para conseguir sus fines, que debería integrarse con obligaciones más que con derechos, como objeto, instrumento, servicio y medio para satisfacer necesidades, solucionar problemas y mejorar la vida del ser humano.

Por ello, el foco de estas reflexiones, a mi juicio y en caso de pretender construir esta personalidad, podría orientarse a otorgar una personalidad jurídica "instrumental" o "mediata" a entes no humanos con sólo obligaciones, como medio instrumento, cosa o servicio, y sin otorgamiento de derechos propios del ser humano y otros seres vivos, y para servir a determinados intereses recocidos y protegidos por el ordenamiento jurídico.

Además, considero que la restricción de esa personalidad a la titularidad de obligaciones sería además un enfoque prudente, preventivo, razonable, ético y de seguridad ante la posibilidad de que esta inteligencia pueda superar a la humana en el futuro y que, cuantos más atributos se le confieran, mayor puede ser capacidad de influencia y manipulación en el ser humano. En este sentido, son varios los autores que coinciden con esta tesis como la precitada Núñez Zorrilla[907].

No obstante, en mi opinión, tal y como he manifestado en diversas ocasiones a lo largo de esta investigación, soy más partidario de que un futuro estatus de los robots dotados de inteligencia artificial más avanzada o "fuerte" se sitúe en el marco de entes sin personalidad jurídica, con determinadas prerrogativas conferidas por el ordenamiento jurídico. Y ello en la medida de una inexistencia de autonomía plena que igualmente impide el otorgamiento de una personalidad jurídica y capacidad de obrar consecuente a una libertad y voluntad inexistentes ante su indisponibilidad.

[907] NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Op. Cit. P. 63.

En este sentido, considero que debe superarse definitivamente la cuestión de su naturaleza jurídica y si los robots encajan en alguna de las categorías jurídicas existentes en la actualidad, como personas físicas, jurídicas, animales u objetos conforme se ha expuesto en esta investigación, o si debería crearse una nueva categoría específica dadas sus particulares capacidades y características, conforme considero más adecuado y sobre lo que igualmente se han pronunciado autores como Santos González[908].

No obstante, ¿sería justo no reconocer derechos a un sistema de inteligencia artificial o robot, máquina o producto dotado de la misma? La ficción del cine ya nos ha mostrado incluso reivindicaciones de robots "inteligentes" que demandan igualdad de derechos, como por ejemplo en la película "*El hombre bicentenario*", donde una inteligencia artificial prefiere morir si así le reconocen su supuesta capacidad para sentir, soñar o actuar como un ser humano.

El debate en el ámbito ético quedó abierto, pero migró también al ámbito jurídico.

El Tribunal Popular[909] del Distrito de Shenzhen Nanshan (China) si reconoció determinados derechos en una disruptiva sentencia de enero de 2020, en particular, consideró que un artículo escrito por un algoritmo de inteligencia artificial desarrollado por la compañía Tencent debía recibir la misma protección a nivel de derechos de autor o "copyright" que los textos creados por seres humanos, lo que no significa, que se le esté reconociendo al sistema o algoritmo la condición de "autor".

El Convenio de Berna para la Protección de las Obras Literarias y Artísticas, el Tratado de la OMPI sobre Derecho de Autor, así como, a nivel nacional, el artículo 5 de la Ley de Propiedad Intelectual española, únicamente reconocen la protección de las creaciones

[908] SANTOS GONZÁLEZ, M.J. (2017). "Regulación legal de la robótica y la inteligencia artificial: Retos del futuro". *Revista Jurídica de la Universidad de León*. Nº 4. 2017. P. 39.

[909] El Tribunal condenó a Shanghai Yingxun Technology Company al pago de una multa de 1.500 yuanes, a causa de la infracción de derechos de autor y por las pérdidas económicas ocasionadas a Tencent, por considerar que el artículo cumplía con los requisitos de análisis, estructura lógica y originalidad que protegen los derechos de autor. https://www.europapress.es/portaltic/sector/noticia-china-protege-copyright-articulo-escrito-inteligencia-artificial-20200113141813.html

de la mente humana, no así las formuladas por algoritmos artificiales, como analizaré con más profundidad en el próximo capítulo de esta investigación.

La Oficina de Derechos de Autor de Estados Unidos -*United States Copyright Office*- también ha adoptado un enfoque en este sentido, aplicando desde 1.973 una "política de autoría humana", no protegiendo las obras que no hayan sido creadas por un autor humano. Y todo ello sin perjuicio de los derechos que puedan recaer en las personas jurídicas relacionadas con los autores humanos.

En cualquier caso, la atribución de derechos de cualquier clase debería conllevar su ejercicio bajo unos valores morales de los que carecen estos sistemas.

En mi opinión, en cualquier caso, el debate sobre su posible capacidad jurídica *ad hoc* debería focalizarse exclusivamente en sus obligaciones y, en su caso, en las facultades o funciones específicas y limitadas que se le puedan atribuir, previamente asignadas legalmente, más que en el reconocimiento de derechos.

Siguiendo la tesis planteada, el propietario, operador o usuario del sistema es quien debería ostentar los derechos sobre las actuaciones, conductas y operaciones llevadas a cabo por el sistema. Y sus obligaciones serían todas aquellas derivadas de sus conductas, acciones u omisiones en el marco de las capacidades, funciones y facultades asignadas en su diseño o adquiridas por el sistema.

La responsabilidad por daños, bajo este enfoque constructivo sobre una hipotética personalidad jurídica, podría recaer en el sistema y subsidiaria o solidariamente en el propietario o usuario, salvo que el daño sea imputable a la víctima o al propio propietario, operador o usuario por culpa o dolo. Todo ello sin perjuicio del necesario traslado del riesgo a un tercero mediante un seguro obligatorio que garantice el derecho de resarcimiento de la víctima. La otra opción que considero más adecuada y acorde a los regímenes vigentes de responsabilidad sería hacerla recaer en el propietario, operador o usuario por las decisiones y actos de su ente auxiliar. Sin perjuicio, obviamente, de los daños que puedan ser exigibles al productor, a través del régimen de responsabilidad por productos defectuosos.

Respecto de los posibles argumentos sobre el reconocimiento de personalidad a los mismos, si el fundamento pues de la atribución de la personalidad a las personas jurídicas se basa en la utilidad que comportan al ser humano para la consecución de sus fines, idéntica razón podría concurrir en la inteligencia artificial más avanzada o "fuerte", conforme exponen distintos autores como González Granado[910], por lo que según los mismos nada impediría reconocer una personalidad jurídica específica y distinta a un ente artificial inteligente si comporta beneficios para el ser humano.

El sistema o ente debería disponer de un estatus jurídico específico creado por el ordenamiento jurídico que establezca sus obligaciones y responsabilidades y, en su caso, cualquier facultad o derecho asociado. Se trataría de un estatus propio acorde a sus circunstancias[911].

Pero, dejando a un lado mi postura y argumentos relacionados, considero necesario una reflexión previa y obligada, ¿es necesario atribuir personalidad jurídica a un objeto o ente para asociar el cumplimiento de una obligación? Ya he expuesto supuestos en el ordenamiento jurídico donde vemos que no sería necesaria esta fórmula.

Por el camino podrían plantearse fórmulas intermedias, como una personalidad jurídica electrónica "limitada" a sistemas dotados de inteligencia artificial avanzada o "fuerte" y a los robots, máquinas y otros productos dotados de la misma, que podría restringir el alcance de las consecuencias jurídicas de sus actos y minimizaría los riesgos que una capacidad plena podría suponer. Asimismo, una capacidad representada podría evitar, con los mecanismos adecuados, los riesgos de capacidad y solvencia para atender las responsabilidades derivadas de su funcionamiento y el resarcimiento de los daños o perjuicios causados derivados de sus acciones para la persona afectada.

No obstante, el objetivo no debería ser, en ningún caso, dotar a los robots y sistemas dotados de inteligencia artificial de derechos propios del ser humano y coincido

---

[910] GONZÁLEZ GRANADO, J. (2016). *Derecho y Robots en la Unión Europea: Hacia una persona electrónica*. Taller de derechos, 27 de junio de 2016. Disponible en: https://tallerdederechos.com/derecho-y-robots-en-la-union-europea-hacia-una-persona-electronica/ Consultado el 30 de diciembre de 2020.

[911] ERCILLA GARCÍA, J. (2018). *Normas de Derecho Civil y Robótica. Robots Inteligentes, Personalidad Jurídica, Responsabilidad Civil y Regulación*. Editorial Thomson Reuters Aranzadi. Navarra 2018. Pp. 17 y 18.

plenamente con autores como Sánchez del Campo[912], que considera que el problema no es tanto el reconocimiento de la personalidad jurídica sino la capacidad de obrar, es decir, la determinación de las acciones que un sujeto de derecho puede realizar, lo que además, a mi juicio, comporta una complejidad añadida en el caso de sistemas dotados de inteligencia artificial o robots, máquinas u otros productos dotados de la misma, como lo es la titularidad, gobierno y acciones cuando aprenden progresivamente, interaccionan con su entorno y pueden determinar su capacidad y sus actos, adicionado impredecibilidad.

Si el concepto de personalidad jurídica se considera coincidente con el de capacidad jurídica como la aptitud para ser titular de derechos subjetivos y deberes jurídicos[913], la capacidad de obrar se refiere a la aptitud para ejercer esos derechos de los que se es titular en virtud de la capacidad jurídica atribuida.

En consecuencia, se podría llegar a reconocer personalidad jurídica o no, pero desde luego el hecho de disponer de la misma no debe conllevar la existencia de capacidad de obrar plena, que puede ser complementada o apoyada.

La capacidad de obrar de las personas físicas se regula en el ordenamiento jurídico español en su Código Civil, que contempla distintos grados en función del estado civil, edad, nacionalidad o vecindad civil. De hecho, contempla distinto grados.

La capacidad plena, sería la que correspondería a toda persona mayor de edad (En España 18 años y que en otros países es distinta), si bien, para determinados supuestos y actos jurídicos no es suficiente la mayoría de edad para disponer de plena capacidad para su realización, sino que se requeriría para la validez y plena eficacia del acto o negocio jurídico algo más, por ejemplo, en la adopción, que se requeriría que el adoptante tenga al menos la edad de 25 años.

[912] SÁNCHEZ DEL CAMPO, A. (2017). "Europa quiere regular a los robots*"*. *Diario La Ley*, Nº 4, Ed. Wolters Kluwer. 28.02.2017.
[913] LASARTE, C. (2016). *Compendio de derecho de la persona y del patrimonio*. Editorial Dykinson, Madrid 2016. P. 4.

Además de la capacidad plena, el ordenamiento jurídico también venía contemplando la incapacidad hasta la reciente Ley 8/2021, de 2 de junio, por la que se reforma la legislación civil y procesal para el apoyo a las personas con discapacidad en el ejercicio de su capacidad jurídica[914], en vigor desde el 3 de septiembre de 2021, lo que suponía la existencia de personalidad y capacidad jurídica, aunque la persona careciera de la aptitud para el ejercicio de derechos subjetivos y deberes jurídicos, por lo que precisaba de un representante legal, que ejercitaría los derechos y deberes de los cuales si era titular.

La reforma planteada pretende adecuar el ordenamiento jurídico español a la Convención internacional sobre los derechos de las personas con discapacidad, hecha en Nueva York, de 13 de diciembre de 2006, en cuyo artículo 12 establece que las personas con discapacidad tienen capacidad jurídica en igualdad de condiciones que las demás en todos los aspectos de su vida, de modo que los Estados deben adoptar las medidas de apoyo que puedan necesitar en el ejercicio de su capacidad jurídica que no debe verse mermada.

Ello supone un cambio de un sistema basado en la sustitución en la toma de decisiones que afectan a personas con discapacidad por otro basado en el respeto de la voluntad y las preferencias de la persona, titular del derecho a la toma de sus propias decisiones y que debe ejercer su capacidad jurídica con los apoyos necesarios. Se elimina en el contexto de la discapacidad la tutela y la patria potestad prorrogada o rehabilitada. Y en relación con la curatela, pasa a concebirse como una institución de naturaleza asistencial más que representativa, salvo supuestos excepcionales.

El ordenamiento jurídico también contempla la capacidad restringida, en la que ésta requiere de un complemento de la misma para determinados y específicos actos o negocios jurídicos, denominado, bien consentimiento, bien autorización o asentimiento. Sería el caso de menores emancipados y, hasta esta reciente reforma, de pródigos, es decir, personas con capacidad restringida requiriendo un complemento de capacidad para determinados actos y negocios jurídicos.

Y del mismo modo, se contemplaba hasta esta reforma a las personas incapacitadas parcialmente, es decir, personas que padecían enfermedades o deficiencias persistentes

[914] BOE 03.06.2021

de carácter físico o psíquico que impedían a la persona gobernarse por sí misma. En este supuesto la sentencia que declarase la incapacidad parcial de la persona, debía fijar taxativamente todos los actos y negocios jurídicos para los cuales precisaría de complemento de capacidad por tutor o curador.

Este contexto y marco podría ser otro de los puntos de partida desde donde considero que se debería reflexionar y valorar el posible otorgamiento futuro de una capacidad de obrar restringida a un sistema dotado de inteligencia artificial, por lo que en mi opinión, en caso de que se pudiera otorgar personalidad jurídica electrónica, ello no debe conllevar asociadamente ni derechos ni la plena capacidad de obrar, sino en cualquier caso, una capacidad restringida y limitada a un conjunto acotado y público de actos y negocios jurídicos autorizados de origen, en función de su naturaleza, características y aplicación, inicialmente con independencia de la capacidad de aprendizaje del sistema, y que constituirá una capacidad que deberá ser controlada y supervisada durante todo el ciclo de vida del sistema, así como complementada por el ser humano responsable del sistema -propietario, operador o usuario (licenciatario)-.

Y, es más, con el ánimo de minimizar riesgos, aquellos actos y negocios jurídicos inicialmente autorizados, en la medida que su naturaleza y características, y ejecución lo permitan, deberían requerir la autorización previa o validación posterior humana.

Dejando a un lado estas reflexiones, la personalidad por representación apuntada anteriormente, similar a la que se reconoce a "entes" o entidades mercantiles, como una sociedad limitada o anónima, es otra de las opciones que podrían evaluarse en el futuro "sin necesidad de reinventar la rueda", aunque no la comparto en la actualidad, de modo que los robots y los sistemas dotados de inteligencia artificial tendrían algunos derechos y facultades, incluyendo la posibilidad de celebrar un contrato o ser parte en un juicio, pero también responsabilidades legales.

La construcción de esta ficción legal permitiría la atribución de una *seudopersonalidad* jurídica que posibilitaría imputar directamente alguna responsabilidad jurídica a los sistemas de inteligencia artificial avanzados, aunque parece una posibilidad más remota que real en la actualidad, conforme recogen distintos autores, entre otros, el precitado Gómez-Riesco.

Otra de las opciones para algunos expertos y autores como Asaro[915], sería la construcción de una personalidad jurídica inspirada en las personas jurídicas, según los cuales sería también viable, si bien, detrás de la misma, está siempre una o más personas que conforman la voluntad de la persona jurídica y a la/s que puede llegar a incidir las decisiones y acciones de la persona jurídica, cuando lo que se pretende es dotar a los sistemas de inteligencia artificial de personalidad propia basada en el supuesto carácter autónomo de sus actos y decisiones adoptados en función de sus instrucciones de programación, autoaprendizaje y contexto determinado, lo que es igualmente una ficción a mi juicio, ante la irrenunciable supervisión y control por el ser humano y seguimiento de las instrucciones predefinidas por el mismo.

Según estos mismos autores, nada impediría que se constituyera una sociedad cuyo principal activo sea un robot o sistema inteligente, si bien, en mi opinión, no sería ni socio ni administrador, no tendría voz ni voto, no tomaría decisiones ni realizaría acciones en nombre o representación de la sociedad -salvo y suponiendo que fuese legalmente apoderado legítima y legalmente para ello-, y desde luego no dejaría de ser precisamente un objeto y activo contable de la sociedad, no un sujeto de derechos y obligaciones.

Otra de las razones por las que no comparto esta tesis se basa en que, si el objetivo principal de dicha atribución es la determinación de la responsabilidad y el íntegro resarcimiento de la persona afectada, al final siempre estaría en última instancia una persona física detrás y además no estaría garantizada ni la solvencia ni el resarcimiento, dado que inicialmente el sistema de inteligencia artificial carecería de bienes con los que responder salvo que respondiera con su propio valor si fuera incautado, embargado o vendido para resarcir a la persona afectada con el precio obtenido, o se estableciesen mecanismos compensatorios a través de seguros obligatorios y de los denominados fondos de compensación, mediante aportaciones de fabricante y diseñadores como incluso de adquirentes, operadores y/o usuarios.

[915] ASARO, P. (2007). "Robots and Responsability from a Legal Perspective", en *Proceedings of the IEE conference of robotics and automation*; *Workshop on roboethics*, Roma, 2007. Así como ROSALES DE SALAMANCA, F. *¿Puede un robot ser sujeto de derecho?*. Recuperado de https://notariofranciscorosales.com/puede-robot-sujeto-derecho/. Consultado el 05.02.2021.

En este sentido, el precitado Dictamen del Comité Económico y Social Europeo -CESE- de 31.05.2017[916] se opone a cualquier tipo de estatuto jurídico o personalidad jurídica a los robots o sistemas dotados de inteligencia artificial en la medida que la responsabilidad dejaría de recaer en el autor y pasaría a la máquina.

La razón principal para su posicionamiento es el riesgo moral inaceptable que ello conllevaría: "La legislación en materia de responsabilidad tiene un efecto correctivo y preventivo que podría desaparecer en cuanto el riesgo de responsabilidad civil dejase de recaer sobre el autor por haberse transferido al robot (o sistema de IA). Además, una forma jurídica así sería susceptible de uso y aplicación indebidos". Asimismo, dictamina que la comparación con la responsabilidad limitada de las sociedades no es válida, puesto que el responsable en última instancia es siempre una persona física.

En relación con este último aspecto, García Sanchez[917] afirma que es cierto que en el caso de las personas jurídicas sus actuaciones y decisiones se llevan a efecto por personas físicas, pero es evidente que el término "persona" es empleado tanto en la referencia a la "persona jurídica" como, en su caso, respecto de la "persona electrónica" para poner de relieve su configuración como un ente dotado de derechos y obligaciones, con el fin de posibilitar una respuesta normativa en torno a las situaciones que su relación e interacción con el entorno generen.

En cualquier caso, reitero que la pretensión inicial del posible reconocimiento de una personalidad electrónica no debe ser otorgar derechos humanos a los robots y sistemas dotados de inteligencia artificial o convertirlos en seres a imagen y semejanza del ser humano, sino asegurar la supervisión humana, la depuración de responsabilidades y el resarcimiento efectivo de las personas afectadas por sus acciones u omisiones, sin olvidar la satisfacción de necesidades, la resolución de problemas y la mejora de la vida del ser humano.

[916] Dictamen del Comité Económico y Social Europeo sobre la "Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad». (2017/C 288/01)

[917] GARCÍA SÁNCHEZ, M. D. (2020). "Inteligencia artificial y oportunidad de creación de una personalidad electrónica". *Ius et Scientiam.* Vol. 6, Nº 2. ISSN 2444-8478. Editorial Universidad de Sevilla 2020. P. 91.

Además de todo ello, en caso de valorar el reconocimiento de derechos y dicha capacidad de obrar restringida, se podrían plantear problemas procesales para reconocer la legitimación pasiva a los sistemas de inteligencia artificial que, de inicio, deberían comprender y estar dotados de la capacidad para articular su defensa y que requeriría una modificación de la legislación procesal.

Autores como Brozek y Jakubiec[918] afirman además que la percepción psicológica de la ciudadanía en relación a la acción humana y la responsabilidad no está capacitada o preparada para considerar a las máquinas autónomas como autores de sus acciones. El derecho y la responsabilidad legal deben ser inteligibles para las personas que son sujetos de derechos y obligaciones. Estos autores defienden que el estatus de agentes legales no puede ser admitido para las máquinas autónomas. Sería posible desde un punto de vista puramente técnico, pero que en la práctica permanecería como un mero "derecho de libro" que no llegaría a materializarse en un "derecho en la práctica".

Nevejans, por su parte, destaca los peligros de derribar las fronteras entre el ser humano y las máquinas, afirmando que el reconocimiento a las mismas de una personalidad electrónica implicaría derribar los límites entre el hombre y la máquina, desdibujando la línea entre lo vivo y lo inerte y "pondría en tela de juicio los fundamentos humanistas de Europa"[919].

Otros destacados expertos en el ámbito del derecho de daños como Wagner[920] cuestionan igualmente esta posibilidad.

Autores como Bryson, Diamantis, Mihailis y Grant[921] consideran que, si bien sería posible declarar a una máquina como una persona legal, consideran que conferir esta

[918] BROZEK, B. Y JAKUBIEC, M. (2017). "On the legal responsibility of autonomous machines", en *Artificial Intelligence and Law*. Vol. 25, nº 3, 2017. DOI 10.1007/s10506-017-9207-8. P. 303.
[919] NEVEJANS, N. (2016). "*Study for Jury Committee European Civil Law Rules in Robotics*". Policy Department C: Citizens' Rights and Constitutional Affairs. European Parliament. Bruselas 2016. P. 16. Disponible en: http://www.europarl.europa.eu/committees/fr/supporting-analyses-search.html
[920] WAGNER, G. (2019). "Robot Liability", en LOHSSE, S., SCHULZE R. Y STAUDENMEYER, D. (Eds.), *Liability for Artificial Intelligence and the Internet of Things*. Nomos. Baden-Baden 2019. Pp. 27-62.
[921] BRYSON, JOANNA J, DIAMANTIS, MIHAILIS E., GRANT, THOMAS D. (2017). "Of, for, and by the people: the legal lacuna of synthetic persons". *Artif Intell Law*, 8 de septiembre de 2017. P. 289.

personalidad legal a los robots resultaría moralmente innecesario y legalmente problemático.

Nieva Fenoll[922] niega también de forma categórica la posibilidad de creación de una personalidad electrónica en materia de atribución de responsabilidad. Considera que un robot no sería responsable, dado que no es una persona, por lo que la responsabilidad recae en el fabricante o el programador del robot.

Sin embargo, otros expertos y autores continúan valorando esta posibilidad en la actualidad. Ercilla García[923] considera que, ante la evolución de los sistemas inteligentes en cuanto a la autonomía y voluntad, la personalidad jurídica de los robots inteligentes más complejos sería factible.

Bardield considera que, si la finalidad es proteger a las víctimas y evitar las dificultades de determinación del sujeto responsable y prueba, sería preferible crear un régimen de responsabilidad específico *ad hoc* amplio que pueda incluir cualquier robot inteligente y que prevea los posibles sujetos responsables en función del tipo de inteligencia artificial, similar al régimen de responsabilidad previsto en el sector de la construcción y, en especial, en la Ley de la Ordenación de la Edificación[924].

Para el precitado Barrio Andrés[925] no habría obstáculo jurídico en que el Derecho pueda reconocer la personalidad jurídica electrónica de robos avanzados en cuanto que tendrían la aptitud de ser titulares de relaciones jurídicas con sus correspondientes derechos y obligaciones. Cabría pues "un cierto reconocimiento jurídico de su subjetividad, fundamentalmente en derechos de naturaleza patrimonial", no respecto de derechos constitucionales o de la personalidad propios del ser humano.

Según el mismo, el reconocimiento de una "personalidad jurídica electrónica" es técnicamente viable para atribuir derechos y obligaciones a robots y sistemas dotados de

[922] NIEVA FENOLL, J. (2020). "Conferencia: inteligencia artificial y Proceso Penal". Universitas Fundación, 22 de julio de 2020. Disponible en: https://www.youtube.com/watch?v=5BrCNVTPp0o
[923] ERCILLA GARCÍA, J. (2018). "Aproximación a una personalidad jurídica específica para los robots". *Revista Aranzadi de Derecho y Nuevas Tecnologías*. Nº 47. 2018. Pp. 1-38.
[924] Ley 38/1999, de 5 noviembre, de Ordenación de la Edificación. BOE 06.11.1999.
[925] BARRIO ANDRÉS, M. (2018) (Dir.). *Derecho de los Robots*. Wolters Kluwer, Madrid 2018.

inteligencia artificial más avanzada, sin que por el momento se precise crear una nueva categoría de sujeto jurídico compuesto principalmente de *software* y a medio camino entre persona y cosa.

Asimismo, Quintero Olivares[926] considera la posibilidad de que existan normas que partan de la (ficticia) personalidad del robot y que permitan reconocerle responsabilidad, en la medida que tales máquinas pueden tener obligaciones y algún tipo de personalidad legal.

Por su parte, García Sánchez[927] considera igualmente esta posibilidad. Para la misma no resulta incompatible la creación de esta personalidad jurídica con la exigencia de un responsable humano por los daños que ocasione un sistema en la medida que "en definitiva, el término 'personalidad' no deja de ser un eufemismo, como en el caso de las personas jurídicas". Esta autora considera que reconocer a los sistemas más avanzados una entidad propia a nivel jurídico no tiene por qué ir necesariamente asociado a la autonomía de su responsabilidad, por lo que esta puede ser distribuida entre el fabricante, programador o usuario, sin que ello impida hablar de una categorización propia. En este sentido, entiendo que la autora considera la concurrencia de una personalidad jurídica que llevaría asociada una capacidad de obrar inexistente o limitada, y una responsabilidad por hecho ajeno asociada al fabricante, programador o usuario.

Según Giovanni Sartor[928] podría atribuirse estados cognitivos a los sistemas inteligentes por su comportamiento, que se puede entender mejor desde una postura intencional, considerado que estos estados cognitivos tendrían relevancia jurídica cuando se les delegue la toma de decisiones basadas en su propio conocimiento, sin intervención humana.

---

926 QUINTERO OLIVARES, G. (2017). "La robótica ante el Derecho Penal: El vacío de respuesta jurídica a las desviaciones incontroladas". *Revista Electrónica de Estudios Penales y de la Seguridad*. Nº. 1, 2017. ISSN:2531-1565. P. 9.

927 GARCÍA SÁNCHEZ, M.D. (2020). "Inteligencia artificial y oportunidad de creación de una personalidad electrónica". *Ius et Scientiam: Revista Electrónica de Derecho y Ciencia*. Vol. 6, Nº 2. ISSN 2444-8478. Editorial Universidad de Sevilla 2020.

928 SARTOR, G. (2009). "Cognitive automata and the law: electronic contracting and the intentionality of software agents". *Artificial intelligence and law*. Vol. 17.4. 2009. Pp. 253 y ss.

Las reflexiones realizadas se complican enormemente cuando barajamos la opción de encontramos ante un futuro sistema de inteligencia artificial que pueda tomar decisiones plenamente autónomas -supuestamente- , con capacidad de autoaprendizaje y bajo una supuesta libertad, y más si pudiera llegar a desarrollar una inteligencia artificial con consciencia o autorreflexión, con capacidad para calificar la justicia o injusticia, maldad o bondad de su actos o sentir, lo que constituyen cualidades humanas que difícilmente podrán atribuirse a un sistema de inteligencia artificial en base a la tecnología actual, sin perjuicio de que su programación pueda integrar parámetros éticos y morales que le permitan emitir juicios de valor en contextos concretos que determinen sus decisiones y acciones.

Para algunos autores como el precitado Ebers[929], esta evolución comporta que el comportamiento de la máquina se determina cada vez menos desde la programación y depende cada vez más de su interacción con el entorno, los procesos de autoaprendizaje y nuevas formas de comportamiento que difícilmente se pueden controlar durante su funcionamiento.

Para otros, como el precitado Díaz Alabart[930] o Badillo Arias[931], las decisiones autónomas de los robots derivarán siempre de su creación y programación, y no de una capacidad volitiva que nunca van a tener.

A mi juicio, como he manifestado durante el tratamiento de otros aspectos a lo largo de esta investigación, la necesaria supervisión y control humano, la consecuente ausencia de una autonomía y libertad real -y más en la medida que la inteligencia artificial opera bajo instrucciones, objetivos, parámetros y restricciones programadas-, así como la ausencia de conciencia real impiden actualmente considerarla sujeto de derecho, aunque sí objeto del mismo, como objeto o servicio. En la misma línea, destacar a Díaz-Limón anteriormente citado.

---

[929] EBERS, M. (2016). "La utilización de agentes electrónicos inteligentes en el tráfico jurídico: ¿Necesitamos reglas especiales en el Derecho de la responsabilidad civil?", en *InDret: Revista para el Análisis del Derecho*. Nº 3, 2016, Pp. 8-9.

[930] DÍAZ ALABART, S. (2018). *Robots y responsabilidad civil*. Editorial Reus, Madrid 2018. P. 75.

[931] BADILLO ARIAS, J.A. "Responsabilidad civil y aseguramiento obligatorio de los robots", en MONTERROSSO CASADO, E. (Dir.). *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch 2019. P. 41.

El sistema podría tener capacidad para decidir y ordenar su propia conducta conforme a su propio razonamiento lógico e incluso al margen de las instrucciones programadas, pero dicha supuesta autonomía e independencia siempre tendrá que ser relativa por razones éticas, jurídicas y de seguridad, por lo que siempre deberá restringirse y nunca tendrá absoluta libertad. Nunca deberían crearse inteligencias artificiales plenamente autónomas y libres.

Si la voluntad es la capacidad para decidir con libertad lo que se desea, un sistema o ente de este tipo estaría sujeto a capacidades restringidas, directrices e instrucciones, objetivos, mecanismos y controles de seguridad internos y externos (que incluirían suspensión, apagado o reversión), de modo que realmente, aunque dispusiera de supuesta autonomía, capacidad de autoaprendizaje y cierta impredecibilidad para sus creadores, carecería de verdadera libertad, no sería un ente libre e independiente, sino que seguiría siendo un objeto, servicio o, en todo caso, un ente sin personalidad jurídica, plenamente dependiente y con capacidad restringida y, en consecuencia, no se le puede imputar una actuación consciente o intencionada.

Además, para existir voluntad, la libertad de opción debería ejecutarse con conciencia, de la que carecen este tipo de sistemas, que no sienten, que no tienen emociones, que no distinguen entre lo justo o injusto, lo bueno y lo moral. Y, por último, como objeto pertenecerá o será operado o usado por una persona física o jurídica que es quien podrá permitir o decidir que siga funcionando y usándose.

Nunca se le podría imputar ni culpa ni dolo en la comisión de sus actos dañinos.

No obstante, profundizando en este análisis razonado, la cuestión relativa a la conciencia no es pacífica en la comunidad científica, como ya puse de manifiesto al abordar las cuestiones éticas en el capítulo III de esta investigación.

De un lado, distintos expertos consideran que la conciencia es una capacidad exclusiva del comportamiento humano y no puede ser copiada por las máquinas. Otros consideran

que el desarrollo tecnológico previsiblemente permitirá dotar de conciencia a las mismas[932].

La *Comisión de Asuntos Jurídicos del Parlamento Europeo*, en su proyecto de informe de 2016 con recomendaciones a la *Comisión de Derecho Civil sobre Normas de Robótica*, ya consideró la posibilidad de "crear una condición jurídica específica para los robots, de modo que al menos los robots autónomos más sofisticados puedan establecerse como personas electrónicas con derechos y obligaciones específicos, incluido el de reparar los daños que puedan causar, y aplicar la personalidad electrónica a los casos en que los robots tomen decisiones autónomas inteligentes o interactúen de otro modo con terceros de forma independiente."

No obstante, con posterioridad, el Parlamento Europeo en la precitada Resolución de 16 de febrero de 2017[933] sobre normas de Derecho civil sobre robótica, consideró que, conforme al actual marco jurídico, los robots y los sistemas inteligentes no pueden ser considerados responsables de los actos u omisiones que causan daños a terceros. No obstante, abrió la vía a su reflexión y debate.

Posteriormente, el Comité Económico y Social Europeo en su Dictamen de 31.05.2017[934], negó cualquier tipo de personalidad jurídica para los robots o la inteligencia artificial por el riesgo moral que ello implicaría y el posible uso indebido que podría conllevar, así por la UNESCO, especialmente en el *Informe de la Comisión Mundial de Ética del Conocimiento Científico y la Tecnología sobre la ética de la robótica* de 14.09.2017[935].

[932] NIETO, M. (2021). "Conciencia e inteligencia artificial, un debate abierto". Disponible en: https://blogthinkbig.com/conciencia-inteligencia-artificial. Consultado el 09.02.2021.
[933] Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL)). Disponible en https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_ES.html
[934] Dictamen del Comité Económico y Social Europeo "Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad".31.05.2017.https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/artificial-intelligence-consequences-artificial-intelligence-digital-single-market-production-consumption-employment-and
[935] https://unesdoc.unesco.org/ark:/48223/pf0000253952

Si en un futuro se le otorgará dicha personalidad jurídica a los robots o sistemas de dotados de inteligencia artificial autónomos, el modelo actual en el que el fabricante, el propietario y/o el usuario serían los responsables, debería transformarse.

Y son muchas más las voces desde todos los ámbitos que se oponen al reconocimiento de la personalidad electrónica de los robots.

La carta pública dirigida a la Comisión[936] por más de dos centenares de expertos europeos sobre esta materia, reflejó su clara oposición al reconocimiento de dicha personalidad. Suscriben la misma expertos del ámbito tecnológico, legal, sanitario y ético.

Los expertos firmantes del documento manifestaron su conformidad sobre la necesidad de que la UE establezca un marco regulador adecuado de derecho civil sobre la robótica y la inteligencia artificial, que garantice, de un lado, un alto nivel de confianza y seguridad de la ciudadanía, así como, de otro, el apoyo a la innovación.

Asimismo, manifestaron la necesidad de que dicho marco aborde la cuestión de la responsabilidad de los robots "autónomos". A juicio de los firmantes, el otorgamiento de una condición jurídica de "persona" electrónica sería "*ideológica y carente de sentido y pragmática*".

Este grupo de expertos mostró su preocupación por lo dispuesto en la Resolución del Parlamento Europeo de 16 de febrero de 2017, sobre las normas de derecho civil de la robótica, en particular en su párrafo 59.f), en el que Parlamento pide a la Comisión que cuando realice la evaluación de impacto del futuro instrumento legislativo, considere lo siguiente: "Crear un estatuto jurídico específico para los robots a largo plazo, de modo que al menos los robots autónomos más sofisticados puedan establecerse como personas electrónicas responsables de reparar los daños que puedan causar, y posiblemente aplicar la personalidad electrónica a los casos en que los robots tomen decisiones autónomas o interactúen de otro modo con terceros de forma independiente."

[936] http://www.robotics-openletter.eu/

Según los mismos, el argumento principal de los defensores de atribuir personalidad jurídica a los robots se basa en la dificultad de determinar la responsabilidad de un robot cuando el daño causado por el mismo es fruto de un proceso de autoaprendizaje, ya que la inteligencia artificial permitiría que dispositivos y máquinas inteligentes tomen decisiones autónomas, aun cuando dichas acciones no estén configuradas en un primer momento.

Los expertos participantes en esta comunicación advierten que la concesión de una personalidad electrónica a los robots podría suponer un reconocimiento de derechos y responsabilidades que entrañaría un impacto económico, legal, social y ético innegable.

Esta cuestión y posibilidad planteada no fue pacífica en su tratamiento en el Parlamento Europeo, que ya motivó una enmienda y 285 miembros del Parlamento votaron a favor de su supresión.

Antes de la votación, Mady Delvaux, vicepresidenta del *Comité de Asuntos Legales del Parlamento Europeo* y principal redactora de la controvertida Resolución, escribió una comunicación a todos los miembros del Parlamento aclarando sus intenciones en la Resolución, ante la complejidad creciente para determinar la responsabilidad en caso de accidente cuando se vean involucrados robots autónomos y autodidactas, capaces de tomar decisiones que no puedan ser trazables hasta un agente humano.

Según el grupo de expertos firmantes, la creación de un estatuto jurídico de "persona electrónica" para los robots "autónomos", "imprevisibles" y "de autoaprendizaje" se justifica por la afirmación incorrecta de que sería imposible probar la responsabilidad por daños.

Desde un enfoque técnico, consideraban que esta afirmación comportaba muchos sesgos basados en “una sobrevaloración de las capacidades reales de incluso los robots más avanzados, una comprensión superficial de la imprevisibilidad y las capacidades de autoaprendizaje y, una percepción de los robots distorsionada por la ciencia-ficción y algunos recientes anuncios sensacionalistas en la prensa”.

Desde un enfoque ético y legal, consideraban que la creación de una personalidad jurídica para un robot es inapropiada, sea cual sea el modelo de estatus legal, y lo fundamentaban en lo siguiente:

a) No cabe considerar que los robots ostenten derechos similares a los seres humanos. Un estatus legal para un robot no puede derivarse del modelo de persona natural, ya que el robot tendría entonces derechos humanos, como el derecho a la dignidad, el derecho a su integridad o el derecho a la ciudadanía, contraviniendo directamente el concepto de Derechos Humanos. Todo ello estaría en contradicción con la Carta de los Derechos Fundamentales de la Unión Europea y el Convenio para la Protección de los Derechos Humanos y las Libertades Fundamentales.

b) La condición jurídica de un robot no puede derivarse del modelo de persona jurídica, ya que implica la existencia de personas físicas detrás de la persona jurídica para representarla y dirigirla. Y este no es el caso de un robot.

c) El estatus legal de un robot no puede derivar del modelo de "Trust" anglosajón (negocio fiduciario), también denominado "*Fiducie und Treuhand*" en Alemania. De hecho, este régimen es extremadamente complejo, requiere competencias muy especializadas, no resolvería la cuestión de la responsabilidad y seguiría implicando la existencia de un ser humano como último recurso -el fideicomisario o fiduciario- responsable de la gestión del robot.

Según la principal impulsora de la carta, la precitada Nathalie Navejans, profesora de Derecho en la Universidad de Artois en Francia, "Al adoptar la personalidad jurídica, vamos a borrar la responsabilidad de los fabricantes". Y en el mismo sentido se manifiesta Noel Sharkey, profesor emérito de inteligencia artificial y robótica en la Universidad de Sheffield y también firmante del documento: "Al buscar la personalidad jurídica para los robots, los fabricantes simplemente intentaban eximirse de la responsabilidad de las acciones de sus máquinas"[937].

[937] DELCKER, J. (2018). "Europe divided over robot 'personhood'". Artículo publicado en el 11.04.2018 en *Politico.eu*. Disponible en: https://www.politico.eu/article/europe-divided-over-robot-ai-artificial-intelligence-personhood/. Consultado el 21.01.2021.

Consideran que su concepción como “personas” a efectos legales, con el objetivo de poder asegurarlos y reclamar responsabilidad por los daños que ocasionen, beneficiaría en mayor medida a los fabricantes que, de esta forma, dispondrían de un subterfugio para eludir su responsabilidad por los daños que aquellos ocasionaren.

El documento suscrito por este grupo de expertos abordó esta cuestión tan controvertida de una manera más superficial, sin profundizar en las distintas opciones que podrían contemplarse en los análisis y reflexiones futuras. De hecho, la Resolución del Parlamento Europeo invita a considerar y analizar estas cuestiones en la futura evaluación de impacto normativo y a reflexionar sobre su estatuto jurídico, pero en modo alguno se planteó o tenía que ser la única solución, conforme recoge la misma.

Por su parte, la precitada Delvaux[938] invitaba a reflexionar sobre la posibilidad de otorgar a los robots una "personalidad electrónica" limitada, comparable a la "personalidad corporativa", un estatus legal que permite a las empresas demandar o ser demandado, al menos en lo que se refiere a la compensación. Esta “personalidad electrónica” no pretendería otorgar a los robots derechos asociados a humanos (derecho a la dignidad o a contraer matrimonio) si no que los pondría a la par que las empresas, que gozan de una reconocida personalidad jurídica. Según la misma, el reconocimiento de la personalidad electrónica podría convertir a cada robot inteligente en una entidad legal, por lo que tendrían que asumir ciertas responsabilidades y obligaciones sociales a concretar.

Desde mi punto de vista, la política legislativa en lo sucesivo debe partir de un enfoque basado en el riesgo y en la flexibilidad y adaptabilidad necesaria de los nuevos marcos a la realidad en cada momento.

Los futuros sistemas dotados de inteligencia artificial más avanzada o “fuerte” con autonomía, capacidad evolutiva y de autoaprendizaje, de interactuar con terceros de forma independiente y relativa impredecibilidad, incluso con *seudoconsciencia* artificial, exigirán marcos jurídicos actualizados y adecuados que, entre otras cosas, permitan determinar la responsabilidad en caso de acciones u omisiones llevadas a cabo por los

[938] Recuperado de: https://www.theeconomyjournal.com/texto-diario/mostrar/595263/ascenso-robots-debe-regulado. Consultado el 23.03.2021.

mismos de manera autónoma y ajena a toda intervención humana, que puedan causar daños o perjuicios.

Conforme al ordenamiento jurídico español y de la UE, no es posible atribuir personalidad jurídica electrónica a un sistema dotado de inteligencia artificial más avanzada o "fuerte" o a los robots, máquinas, dispositivos u otros productos dotados de la misma. No obstante, es un concepto abierto a su reformulación en el futuro en atención al avance de la ciencia y la tecnología y la revisión de los marcos reguladores.

Sobre la creación de una legislación específica sobre robots o una denominada "Lex Robótica" como propone el profesor Ryan Calo[939] y sobre lo que ya anticipé mi opinión, no es una cuestión pacífica en la doctrina, conforme destacan autores como Sánchez-Urán Azaña y Grau Ruiz[940].

De un lado, el enfoque estadounidense propugnado por Calo y otros aboga por un Derecho propio o "Derecho de la Robótica", con lo que parte de la doctrina en España se halla plenamente alienada, como los precitados Barrio Andrés y Santos González en base a sus caracteres distintivos y disruptivos.

Y de otro, el enfoque europeo que se aparta, por el momento, de un cuerpo jurídico específico, y con lo que igualmente se alinea parte de la doctrina en España, como García Mexía[941].

El futuro debate sobre robots dotados de inteligencia artificial no debería ser meramente jurídico sino ético y moral, donde los criterios iniciales para determinar cuándo un "individuo" debería disponer de estatus moral serían, siguiendo a Bostrom y Yudkowsky[942], el de sensibilidad -concebida como la capacidad de un individuo de

[939] CALO, R. (2016). "La robótica y las lecciones del derecho cibernético". *Revista de privacidad y derecho digital*. Nº 2. 2016.
[940] SÁNCHEZ-URÁN, Y. Y GRAU RUIZ, M.A. (2018). "El impacto de la robótica, en especial la robótica inclusiva, en el trabajo: Aspectos jurídicos-laborales y fiscales". Congreso Internacional sobre Innovación Tecnológica y Futuro del Trabajo. *Technological Innovation and the Future of Work: Emerging aspects worldwide*. 5 y 6 de abril 2018. Santiago de Compostela.
[941] GARCÍA MEXÍA, P.L. (2016). "Lex robótica y derecho digital". *Revista de privacidad y derecho digital*. Nº 2. 2016.
[942] BOSTROM, N. Y YUDKOWSKY, E. (2011). "The Ethics of Artificial Intelligence", en RAMSEY, W. Y FRANKISH, K. *Cambridge Handbook of Artificial Intelligence*. Cambridge University Press.2011. Recuperado de: http://www.nickbostrom.com/ethics/artificial-intelligence.pdf. Consultado el 20.02.2021.

reaccionar a fenómenos mediante la capacidad de sufrir o sentir dolor, físico o emocional-, y la sapiencia -entendida como el conjunto de actividades relacionadas con la inteligencia superior, la humana, como la autoconciencia y la capacidad para comprender las consecuencias de sus actos-.

Y adicionalmente a los mismos, han adicionado dos criterios secundarios para ampliar el elenco de individuos susceptibles de disponer de dicho estatus moral: La pertinencia de un individuo a un colectivo que en su generalidad sí tiene estatus moral y la existencia de una relación que vincule al individuo en cuestión con otro individuo que tenga por sí sólo estatus moral[943].

De inicio, desde un enfoque ético, conforme a las reflexiones precedentes, no sería adecuado éticamente negarle el estatus moral a un "sujeto" sólo porque este haya sido diseñado y fabricado artificialmente en base a un principio de "no discriminación", lo que tampoco implica que haya de otorgársele estatus moral a todo sistema de inteligencia artificial, sino llevar a cabo una reflexión y ponderación adecuada de todo que, en caso de otorgársela exigiría un cambio en su tratamiento jurídico reflejado en el derecho positivo. En mi opinión no considero ni tan siquiera viable partir de un principio de no discriminación hombre y máquina en el contexto tecnológico actual.

En cualquier caso y como he referido anteriormente, desde un punto de vista jurídico, la cuestión considero que ha sido zanjada, por el momento, en la medida que esta cuestión ha quedado fuera de la agenda regulatoria de la UE, evidenciado en las recientes propuestas regulatorias en materia de inteligencia artificial del Parlamento Europeo que se integran en su Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas y en su Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial. Del mismo modo, la Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021, estableciendo normas armonizadas sobre inteligencia artificial (Ley

[943] BOSTROM, N. Y YUDKOWSKY, E. "The Ethics of Artificial Intelligence". Op.cit.

de inteligencia artificial) y modificando determinados actos legislativos de la Unión, tampoco recoge esta cuestión.

6. Consideraciones finales

Los conceptos de "Inteligencia artificial" o "robot" deben ser consensuados y acotados para construir el marco que debe regular los mismos y los retos jurídicos que plantean.

Los sistemas de inteligencia artificial pueden hallarse presentes en robots para ir más allá de meros automatismos, en maquinaria o en cualquier otro tipo de producto, dispositivo o de servicio.

Como he puesto de relieve a lo largo de este capítulo, un robot puede estar dotado o no de inteligencia artificial y, en caso de disponer de la misma, puede tratarse de una inteligencia artificial "débil" o "fuerte", con una mayor o menor autonomía, independencia e imprevisibilidad.

Los mayores riesgos y retos actuales los plantean los robots dotados de inteligencia artificial y, dentro de estos, en el futuro, los dotados de una inteligencia artificial más avanzada, autónoma o "fuerte".

En consecuencia, considero que los nuevos marcos reguladores deben focalizarse en la inteligencia artificial, cualquiera que sea el sistema o dispositivo que la contenga, sin perjuicio de que a la hora de regular determinados aspectos especialmente relacionados con sus riesgos y evaluación de su impacto deban considerarse aspectos del sector donde opere o la naturaleza y características del dispositivo o sistema que la contenga, por ejemplo, un robot físico.

Obviamente no es lo mismo una inteligencia artificial avanzada, con capacidad de autoaprendizaje y amplios márgenes de autonomía, libertad e impredecibilidad, que se instale en un robot con forma humana con capacidad de desplazarse, coger, retener, golpear o lanzar objetos, que la misma instalada en una supercomputadora ubicada en un centro de procesamiento de datos (CPD), que gestiona una red de telecomunicaciones o

un hospital. No obstante, la falta de materialización, tangibilidad o identificación física del sistema inteligente no lo hace menos peligroso o que suponga menos retos y riesgos.

Incluso, ante robots dotados de inteligencia artificial más avanzada, a mi juicio, debería ser menos relevante el sector donde puedan operar y más las capacidades de las que esté dotada, ante la posibilidad de actuar al margen de sus objetivos e instrucciones iniciales, por eso mis reflexiones críticas sobre los requisitos y obligaciones de los sistemas de inteligencia artificial considerados de alto riesgo y no del resto, que podrían desarrollar niveles de riesgo muy superiores a los inicialmente evaluados en su diseño, especialmente en relación con su impacto, más que en su probabilidad.

La actual discusión sobre la personalidad jurídica electrónica de los robots dotados de inteligencia artificial debe ser superada y objeto de mayor estudio y reflexión de cara al futuro, con una profunda evaluación de los riesgos asociados.

Si el objetivo es asegurar la responsabilidad en caso de un daño o perjuicio causado por un robot o sistema dotado de inteligencia artificial, la solución no tiene por qué ser exclusivamente el otorgamiento de una personalidad jurídica electrónica, especialmente ante la diversa tipología de sistemas y opciones disponibles analizadas.

En cualquier caso, considero que no se debe prescindir en ningún momento de la seguridad y la supervisión y control humano, de modo que los robots y sistemas dotados inteligencia artificial deben disponer de los mecanismos de seguridad, control y supervisión adecuados por parte del ser humano durante todo su ciclo de vida, lo que incluye mecanismos de suspensión, desactivación y reversión en caso de decisiones, acciones u omisiones que puedan poner en peligro personas, bienes e instalaciones y que deberá estar en manos de personas. Ello impide concebir sistemas inteligentes plenamente autónomos, libres e impredecibles, dejando a un lado su falta de consciencia.

Capítulo VIII

Protección de la inteligencia artificial: Retos, autoría y responsabilidad

1. Introducción

Los sistemas de inteligencia artificial y, en especial, los dotados de aprendizaje autónomo y profundo, permiten la realización de actividades que se consideraban exclusivas de los seres humanos hasta fechas recientes, entre otras, la capacidad de obtener resultados creativos que están suscitando debate y reflexión sobre su condición de obra creativa o invención y, su posible protección por los marcos jurídicos vigentes en materia de propiedad intelectual e industrial.

Los sistemas inteligentes con capacidad para crear contenidos tradicionalmente contemplados y protegidos por la legislación en materia de propiedad intelectual e industrial plantean importantes retos éticos, sociales, políticos y jurídicos.

La capacidad "creadora", más que "creativa", de determinados sistemas de inteligencia artificial es incuestionable a mi juicio, la cual ha sido creada a su vez e integrada en los mismos por el propio ser humano. Otra cuestión será su valor artístico, inventivo, literario, científico o cultural, Es más, en el futuro, esta capacidad podrá ser incluso susceptible de creación por otros sistemas inteligentes.

Estos sistemas, cada vez más complejos y con mayores capacidades, comportan nuevos desafíos para los marcos actuales de propiedad intelectual e industrial, no sólo como objeto de protección por parte de los mismos, como medio o instrumento para su generación o como sistemas de tratamiento de contenidos protegidos por los mismos para su operación o entrenamiento, sino, también, como instrumento para la gestión y protección de los propios derechos de propiedad intelectual e industrial, así como como instrumento para su vulneración, con las responsabilidades derivadas de todo ello en el ámbito civil e incluso penal.

La posibilidad de creación de obras e innovaciones de manera supuestamente autónoma y las cuestiones que ello plantea para los marcos vigentes en materia de derechos de autor, patentes, modelos de utilidad o diseños industriales asociados, comportan nuevos desafíos, especialmente en relación con la aplicación y su protección a través de dichos marcos, así como sobre su titularidad, derechos, contenido, ejercicio y responsabilidad, máxime ante creaciones e invenciones que, por su propia naturaleza, han venido siendo asociadas, tradicional y exclusivamente, al intelecto y la capacidad inventiva o creativa humana, y cuyos marcos reguladores no contemplan nuevas realidades donde se podría producir una disociación absoluta entre el creador y la inteligencia humana interviniente, bien en su diseño, construcción, entrenamiento, funcionamiento, aplicación y/o materialización.

Las capacidades actuales y potenciales de estos sistemas inteligentes para la generación de contenidos "creativos" e "innovadores" exige reflexionar jurídicamente sobre si las mismas son o deben ser susceptibles de protección por los marcos reguladores vigentes de propiedad intelectual -basados en el denominado "derecho de autor" o en el "copyright" propio de la *common law*- y de propiedad industrial y, en caso de serlo, quién debe o debería, o puede o podría ser considerado creador, autor, inventor y/o titular de todos o algunos de los derechos asociados a dicha condición y su régimen de protección.

Los resultados "creativos" e "innovadores" en el contexto de la inteligencia artificial suelen ser consecuencia de un proyecto complejo y conjunto donde confluyen muchas personas de perfiles muy distintos en función de su naturaleza -diseñadores, desarrolladores, ingenieros, científicos de datos, historiadores, etc.-, una fuerte inversión y distintos elementos y componentes que integran un sistema inteligente, como *hardware*, *software*, algoritmos o datos para entrenamiento u operación.

La envergadura de este tipo de proyectos, recursos e inversión que requieren están motivando que sus promotores puedan hallarse interesados en la protección de sus resultados mediante derechos exclusivos, si bien, parece que este interés de los mismos puede estar más focalizado inicialmente en la protección del medio o sistema empleado que en la protección del resultado en sí mismo, sobre el cual estaría más interesado posiblemente el licenciatario o usuario autorizado.

A modo de ejemplo, Schlackman[944] analiza esta cuestión en relación con el cuadro de *The next Rembrandt* y concluye que, con independencia de que ING pueda detentar el derecho de autor, probablemente a ésta no le interesa perseguir copias imperfectas de su obra, sino que más bien estará interesada en la protección del algoritmo que la ha hecho posible.

Una de las cuestiones esenciales que estos sistemas plantean es reflexionar sobre si el marco jurídico vigente en materia de propiedad intelectual e industrial debe y puede resultar de aplicación a obras aparentemente fruto de la creatividad y del ingenio de un sistema inteligente y, en tal caso, si regula adecuadamente su protección o, si por el contrario, se hace necesaria su revisión para introducir cambios en el mismo, en caso de considerar estos marcos los más apropiados para su protección cuando realmente se consideran protegibles o, incluso, valorar la creación de una protección especial.

En mi opinión, anticipando ya alguna de las conclusiones de mi análisis, considero que la clave es diferenciar entre la "capacidad creadora" de un sistema inteligente, lo que considero incuestionable, y la "capacidad creativa", entendida como capacidad de creación de obras e invenciones originales y novedosas fruto del ingenio, concebido éste como facultad humana de la que carecerían los sistemas inteligentes, creados en el ejercicio de dicha facultad por los seres humanos.

Pretendo analizar estas cuestiones a lo largo de este capítulo, si bien, del mismo modo en el que he abordado distintos aspectos y conceptos esenciales de manera previa a su análisis en los distintos capítulos precedentes de esta investigación, considero necesario revisar, previa y sucintamente, los conceptos clave asociados al objeto de este capítulo para, posteriormente, abordar su análisis con profundidad.

Del mismo modo, me he permitido diferenciar los diversos regímenes de protección existentes en distintos ordenamientos jurídicos, esto es, de propiedad intelectual y de propiedad industrial, así como las obras creadas por o con la intervención de sistemas

[944] SCHLACKMAN, S. (2018). "Who owns the copyright of the Next Rembrandt?". *Art Law Review*. 2018. Disponible en: https://alj.orangenius.com/the-next-rembrandt-who-holds-the-copyright-in-computer-generated-art

inteligentes que podrían hallarse dentro del ámbito de protección de cada uno de ellos y las que quedarían inicialmente fuera.

Por último, he abordado el análisis de estas cuestiones diferenciando básicamente las obras e invenciones creadas por sistemas inteligentes con intervención humana de las creadas autónomamente, así como las implicaciones en esta materia sobre el uso y tratamiento de los contenidos utilizados por estos sistemas para su operación o entrenamiento y posibles responsabilidades en que se podría incurrir derivadas de la vulneración de estos derechos.

2. Propiedad intelectual: Marco jurídico básico

De inicio, no todo resultado, contenido u "obra" creada u obtenida por el funcionamiento de un sistema inteligente puede considerarse una obra creativa o del ingenio a los efectos de estos marcos reguladores, ya sea de origen o con intervención humana, o de origen artificial o algorítmico.

La propiedad intelectual es definida por la Real Academia Española como el derecho de explotación exclusiva sobre las obras literarias o artísticas, que la ley reconoce a su autor durante un cierto plazo. No hace referencia expresa a las obras científicas o técnicas.

La Ley de Propiedad Intelectual española, aprobada mediante Real Decreto Legislativo 1/1996, de 12 de abril[945] -en lo sucesivo LPI-, regula su objeto y contenido, pero no contiene una definición expresa del concepto "propiedad intelectual".

En particular, en su artículo 10.1 establece que, son objeto de propiedad intelectual, todas las creaciones originales literarias, artísticas o científicas expresadas por cualquier medio o soporte, tangible o intangible, actualmente conocido o que se invente en el futuro, comprendiéndose entre ellas distintas categorías como libros, folletos, impresos, escritos, discursos y alocuciones, conferencias, informes forenses, explicaciones de cátedra y

[945] Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia

cualesquiera otras obras de la misma naturaleza, composiciones musicales, obras dramáticas y dramático-musicales, las coreografías, las pantomimas y, en general, las obras teatrales, obras cinematográficas y cualesquiera otras obras audiovisuales, esculturas y obras de pintura, dibujo, grabado, litografía y las historietas gráficas, tebeos o cómics, así como sus ensayos o bocetos y las demás obras plásticas, sean o no aplicadas, proyectos, planos, maquetas y diseños de obras arquitectónicas y de ingeniería, los gráficos, mapas y diseños relativos a la topografía, la geografía y, en general, a la ciencia, obras fotográficas y las expresadas por procedimiento análogo a la fotografía o los programas de ordenador.

Y por lo que se refiere a su contenido, el artículo 2 de la norma precitada establece que la propiedad intelectual está integrada por derechos de carácter personal y patrimonial, que atribuyen al autor la plena disposición y el derecho exclusivo a la explotación de la obra, sin más limitaciones que las establecidas en la Ley.

El concepto jurídico de propiedad intelectual del que parte la norma precitada, si incluye expresamente las obras científicas y los programas de ordenador, entre otras obras susceptibles de calificación como tales, sin perjuicio de otras más complejas, no incluidas expresamente en la LPI española, pero igualmente susceptibles de categorización como obras protegidas por la misma, como web o videojuego.

Por su parte, el Parlamento Europeo define la propiedad intelectual, sobre la base de los artículos 114 y 118 del Tratado de Funcionamiento de la Unión Europea (TFUE), como el conjunto de derechos exclusivos sobre las creaciones intelectuales y se divide en dos ramas: La propiedad industrial, que incluye los inventos (patentes), las marcas, los dibujos y modelos industriales y las indicaciones geográficas, y los denominados derechos de autor, que abarcan las obras literarias y artísticas[946]. De nuevo, no se recoge expresamente en esta definición básica las obras científicas.

[946] Recuperado de: https://www.europarl.europa.eu/factsheets/es/sheet/36/la-propiedad-intelectual-industrial-y-comercial#:~:text=La%20propiedad%20intelectual%20es%20el,exclusivos%20sobre%20las%20creaciones%20intelectuales.&text=El%20Tratado%20de%20Funcionamiento%20de,propiedad%20intelectual%20(art%C3%ADculo%20118). Consultado el 03.01.2021

La *Organización Mundial de la Propiedad Intelectual* (OMPI) asocia la propiedad intelectual (PI) a las creaciones del intelecto, desde las obras de arte hasta las invenciones, los programas informáticos, las marcas y otros signos comerciales [947].

En España, la propiedad intelectual *stricto sensu* y su sistema de derecho de autor se hallan regulados básicamente en LPI precitada, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia.

Por su parte, la propiedad industrial se regula diferenciadamente en España en la legislación específica sobre Patentes, Diseño Industrial, Topografía de Elementos Semiconductores (Chips) y Marcas, que serán abordadas posteriormente.

La propiedad intelectual se regula esencialmente a nivel internacional en el Convenio de Berna para la protección de las obras literarias y artísticas, en el Tratado de la Organización Mundial de la Propiedad Intelectual (OMPI) sobre derecho de autor y en el Tratado de la OMPI sobre Interpretación o Ejecución y Fonogramas, sin perjuicio de destacar por su contenido específico, documentos como el reciente *Documento Temático revisado de la OMPI, de 29 de mayo de 2020, sobre las Políticas de Propiedad Intelectual y la inteligencia artificial*, que será objeto de tratamiento en posteriores apartados.

Por su parte, la UE ha creado un marco regulador disperso en materia de protección de la creatividad y la innovación, con el objetivo último de garantizar la promoción de la innovación y la creatividad, así como el acceso al conocimiento y a la información.

En materia de marcas, dibujos y modelos, destacar la Directiva 2015/2436/CE del Parlamento Europeo y del Consejo, de 16 de diciembre de 2015, relativa a la aproximación de las legislaciones de los Estados miembros en materia de marcas, el Reglamento (UE) 2017/1001 del Parlamento Europeo y del Consejo, de 14 de junio de 2017, sobre la marca de la Unión Europea, la Directiva 98/71/CE, de 13 de octubre de 1998 que aproxima las disposiciones nacionales sobre la protección jurídica de los dibujos y modelos, el Reglamento (CE) n.º 6/2002, de 12 de diciembre de 2001, que ha sido objeto de modificaciones y que instituye un sistema comunitario de protección de los

[947] Recuperado de: https://www.wipo.int/publications/es/details.jsp?id=4528. Consultado el 03.01.2021

dibujos y modelos, la Decisión 2006/954/CE del Consejo y el Reglamento (CE) n.º 1891/2006 del Consejo, ambos de 18 de diciembre de 2006, con la finalidad de vincular el sistema de registro de los dibujos o modelos de la Unión al sistema internacional de registro de dibujos y modelos industriales de la Organización Mundial de la Propiedad Intelectual (OMPI).

En materia de derechos de autor y derechos conexos, sin perjuicio de las propuestas más actuales que se abordarán a lo largo de análisis de este capítulo, significar la Directiva 2001/29/CE, de 22 de mayo de 2001, relativa a la armonización de determinados aspectos de los derechos de autor y derechos afines a los derechos de autor en la sociedad de la información, la Directiva (UE) 2019/790 del Parlamento Europeo y del Consejo, de 17 de abril de 2019, sobre los derechos de autor y derechos afines en el mercado único digital y por la que se modifican las Directivas 96/9/CE y 2001/29/CE[948] - pendiente de transposición en España y ya transpuesta en Alemania, Francia e Italia-, la Directiva (UE) 2019/789 (la Directiva relativa a la radiodifusión vía satélite y la distribución por cable), de 17 de abril de 2019, la Directiva 2017/1564, de 13 de septiembre de 2017, sobre ciertos usos permitidos de determinadas obras y otras prestaciones protegidas por derechos de autor y derechos afines en favor de personas ciegas, con discapacidad visual o con otras dificultades para acceder a textos impresos, facilita el acceso a los libros y otro material impreso en formatos adecuados y su circulación en el mercado interior, el Reglamento (UE) 2017/1128 del Parlamento Europeo y del Consejo, de 14 de junio de 2017, relativo a la portabilidad transfronteriza de los servicios de contenidos en línea en el mercado interior, que tiene por objeto garantizar que los consumidores compradores o abonados de películas, retransmisiones deportivas, música, libros electrónicos y juegos puedan disfrutar de estos contenidos en sus desplazamientos a otros Estados miembros de la Unión, la Directiva 2011/77/UE por la que se modifica la Directiva 2006/116/CE relativa al plazo de protección del derecho de autor y de determinados derechos afines, la Directiva 91/250/CEE que imponía a los Estados miembros el deber de proteger los programas informáticos a través de derechos de autor en la medida en que están considerados como obras literarias en el sentido del Convenio de Berna para la Protección de las Obras Literarias y Artísticas, y que se codificó mediante la Directiva 2009/24/CE del Parlamento Europeo y del Consejo, de 23

[948] DO L 130 de 17.5.2019, p. 92

abril 2009, sobre la protección jurídica de programas de ordenador[949], la Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos[950], la Directiva 2014/26/UE relativa a la gestión colectiva de los derechos de autor y derechos afines y a la concesión de licencias multiterritoriales de derechos sobre obras musicales para su utilización en línea en el mercado interior, la Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativa a los datos abiertos y la reutilización de la información del sector público[951]. Asimismo, destacar la Resolución, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica[952], entre otras normas.

Del mismo modo, significar otras normas en materia de patentes, secretos industriales y lucha contra la falsificación como la Directiva (UE) 2016/943 del Parlamento Europeo y del Consejo relativa a la protección de los conocimientos técnicos y la información empresarial no divulgados (secretos comerciales) contra su obtención, utilización y revelación ilícitas[953], la Directiva 2004/48/CE relativa al respeto de los derechos de propiedad intelectual y el Reglamento (UE) n.º 608/2013, relativo a la vigilancia por parte de las autoridades aduaneras del respeto de los derechos de propiedad intelectual, que establece las normas de procedimiento destinadas a las autoridades para vigilar el respeto de los derechos de propiedad intelectual en relación con las mercancías que estén sujetas a vigilancia aduanera o a controles aduaneros.

Y, en materia de protección de datos personales y libre circulación de datos de no personales, en los aspectos que puedan estar asociados el tratamiento de información personal mediante estos sistemas, el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos)[954], y el Reglamento (UE) 2018/1807 del Parlamento Europeo y del Consejo, de 14

[949] DO L 111 de 5.5.2009. P. 16
[950] DO L 77 de 27.3.1996. P. 20
[951] DO L 172 de 26.6.2019. P. 56
[952] DO C 252 de 18.7.2018. P. 239
[953] DO L 157 de 15.6.2016. P. 1
[954] DO L 119 de 4.5.2016. P. 1.

de noviembre de 2018, relativo a un marco para la libre circulación de datos no personales en la Unión Europea[955].

Por último, los países europeos no regulan de manera uniforme la propiedad intelectual e industrial, encontrándonos países con marcos construidos sobre el sistema tradicional de derechos de autor, pero con disparidades en la regulación de aspectos específicos, como España y Alemania, así como con países que han construido sus marcos en esta materia sobre los sistemas de *copyright* propios de la *common law*, como Reino Unido.

La mayoría de marcos reguladores precitados fueron elaborados y aprobados en un contexto tecnológico muy diferente al actual, en el que algunas de las tecnologías más vanguardistas incluso ya existían, si bien, adolecían de un desarrollo y aplicación extendida que permitiera vislumbrar los conflictos y retos que pueden plantear en múltiples aspectos y, en especial, respecto de la suficiencia y adecuación de los propios marcos jurídicos vigentes para regular las nuevas o inminentes realidades y conflictos que pueden suscitarse alrededor de las mismas.

La tecnología digital ha venido siendo utilizada desde hace décadas en la creación artística y en la invención como herramienta o medio al servicio del creador o inventor que, en definitiva, debería ser el objetivo esencial de la propia tecnología, es decir, constituir un medio para solucionar problemas, satisfacer necesidades y mejorar la vida del ser humano.

La cuestión es que actualmente ya existen sistemas dotados de inteligencia artificial que, no es que actúen de manera plenamente autónoma, independiente, libre y por su cuenta como argumentan algunos todavía hoy, sino que, son capaces, con cierto grado de autonomía o automatismo, y bajo instrucciones y/o entrenamiento humano, de llegar a crear obras, contenidos "creativos" e "invenciones" de forma *cuasi-autónoma*, utilizando procedimientos que emulan la inteligencia humana y con capacidad de tomar determinadas decisiones asociadas al proceso creativo o inventivo de manera relativamente independiente.

[955] DO L 303 de 28.11.2018. P. 59

De este modo, la inteligencia artificial está revolucionando el mundo de la creación artística y de la innovación, permitiendo la creación de nuevos modelos creativos, artísticos e inventivos.

Este escenario plantea distintos retos y cuestiones de responsabilidad en estas materias, conforme he comentado anteriormente, especialmente en relación con el uso de contenidos, creaciones o invenciones como datos para su tratamiento y entrenamiento por los algoritmos, como en relación con la autoría y protección de los resultados creados u obtenidos mediante sistemas inteligentes, y la titularidad y contenido de los derechos sobre los mismos.

La UE ha ido más allá y el legislativo unionista adoptó la Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial[956], con el objetivo de establecer la hoja de ruta para la protección uniforme de los derechos de propiedad intelectual e industrial en la UE, especialmente ante la dificultad de su trazabilidad y su aplicación a los resultados generados por la inteligencia artificial. Esta Resolución será objeto de análisis en posteriores apartados.

3. Inteligencia artificial y propiedad intelectual.

El potencial de la inteligencia artificial para el avance científico, tecnológico e industrial es incuestionable pero también para el avance cultural, artístico y social.

La inteligencia artificial, como he expuesto en la introducción, genera importantes retos jurídicos en relación con la propiedad intelectual, especialmente y entre otros, para determinar si las creaciones generadas por sistemas inteligentes pueden considerarse susceptibles de protección jurídica por la vía de los derechos de autor, así como sobre

[956] Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial (2020/2015(INI))

quién o qué ha de ser considerado autor y titular de dichos derechos sobre las creaciones, así como tipología de estos derechos y contenido de los mismos.

La inteligencia artificial, ya sea "débil" o "fuerte", en los términos en los que fueron definidos en el capítulo I de esta investigación, se relaciona con la propiedad intelectual en tres aspectos o dimensiones principales:

- La titularidad y protección de los sistemas de inteligencia artificial, incluyendo el *software, hardware*, algoritmos, bases datos o la propia información que los integran, que permiten generar creaciones intelectuales.

- La titularidad y protección de las creaciones llevadas a cabo por sistemas inteligentes.

- La titularidad y protección de los datos y contenidos de los que se nutren los sistemas inteligentes para operar o entrenarse.

La inteligencia artificial considerada "fuerte" es la que se desplegará y aplicará antes o después en todos estos ámbitos, y quizás esté más cerca de lo que podamos pensar para la creación de música o noticias, pero la "débil" ya está siendo aplicada en estas dimensiones y plantea múltiples cuestiones y retos jurídicos.

3.1. Titularidad y protección de los sistemas de inteligencia artificial, incluyendo programas de ordenador, algoritmos, *hardware*, bases datos o la propia información, que permiten generar creaciones intelectuales.

La titularidad y protección de los sistemas inteligentes vía propiedad intelectual, en general, se halla ya contemplada por el ordenamiento jurídico vigente en España y en la UE, no como conjunto en sí mismo -lo que sería susceptible de análisis más profundo incluso en el ámbito de la propiedad industrial- sino singularmente, respecto de los distintos elementos y componentes que los integran, por lo que, conforme al objeto y alcance limitados de esta investigación, trataré sucintamente esta cuestión, para focalizarme principalmente en el resto de ámbitos o dimensiones, esto es, la titularidad y

protección de los resultados creativos generados por los mismos y, de los datos y contenidos utilizados por estos sistemas para su funcionamiento y entrenamiento.

Los sistemas de inteligencia artificial pueden estar conformados por distintos elementos y componentes con distinto grado de novedad y especificidad en función del tipo de sistema, como *hardware*, *software*, algoritmos, bases de datos e información, y cada uno de los mismos se halla regulado y protegido por distintos marcos jurídicos en España y la UE.

Además, la relevancia de cada uno de estos elementos respecto del resultado obtenido puede ser muy variada en función del tipo de sistema concebido y resultado pretendido, y pueden tratarse de componentes generales o comerciales, o específicos creados *ad hoc* para ese sistema y para la creación de esos resultados como, por ejemplo, determinados dispositivos robóticos (*hardware*), impresoras especiales 3D (*hardware*), *software*, datos, etc.

De hecho, el *hardware* que ejecuta y materializa la creación "lógica" de un sistema inteligente puede tener un rol esencial en la misma, especialmente por su precisión y ejecución diferencial. El *hardware* en este sentido aportaría las habilidades mecánicas y físicas, que pueden ser únicas, para la materialización de lo que la inteligencia artificial define y crea previamente con esta intención.

Los programas de ordenador en España -al igual que en la UE-, se hayan protegidos por la LPI española como obra intelectual, regulada en los artículos 95 a 104 de la misma.

Del mismo modo, la protección del *software* por esta vía de protección se haya igualmente reconocida en el artículo 4 del Tratado OMPI sobre Derecho de Autor[957], en el artículo 10 del ADPIC[958] y el artículo 1 de la Directiva 2009/24/CE del Parlamento Europeo y del

[957] Tratado OMPI sobre Derecho de Autor, adoptado en Ginebra el 20.12.1996. BOE núm. 148, de 18.6.2010.
[958] Acuerdo TRIPS. Acuerdo sobre los aspectos de Derechos de Propiedad Intelectual relacionados con el Comercio, inclusive el comercio de mercancías falsificadas). Anexo 1C del Acuerdo de Marrakech por el que se establece la Organización Mundial del Comercio, firmado en Marrakech (Marruecos), el 15 de abril de 1994.

Consejo de 23 de abril de 2009 sobre la protección jurídica de programas de ordenador[959], transpuesta al ordenamiento jurídico español pero con notables diferencias en distintos aspectos entre lo dispuesto en ésta y el marco jurídico español resultante de dicha transposición. Esta protección no se extiende a sus aspectos funcionales.

El *software* no sería patentable en España en la actualidad, salvo aquél susceptible de aplicación industrial, sin perjuicio de la previsible evolución de esta protección y su inclusión en el marco de las patentes y modelos de utilidad, en la medida que se considera tradicionalmente por una parte de la doctrina como una protección imperfecta, incompleta y obsoleta[960], si bien, la industria no tiene un posicionamiento único al respecto.

No obstante, la LPI española prevé expresamente la compatibilidad de ambos regímenes en su artículo 96.3, disponiendo que cuando los programas de ordenador formen parte de una patente o un modelo de utilidad gozarán, sin perjuicio de lo dispuesto en la LPI, de la protección que pudiera corresponderles por aplicación del régimen jurídico de la propiedad industrial. En este contexto, el *software* dispondría de una doble protección en virtud de ambos regímenes.

Del mismo modo, el artículo 8 de la Directiva 24/CE del Parlamento Europeo y del Consejo, de 23 de abril de 2009, sobre la protección jurídica de programas de ordenador, establece que sus disposiciones se entenderán sin perjuicio de cualesquiera otras disposiciones jurídicas tales como las relativas a los derechos de patente.

En definitiva, cuando el *software* forme parte de una patente o modelo de utilidad y sea original, la protección mediante el derecho de autor se acumulará a la protección mediante propiedad industrial, lo que podrá suceder, por ejemplo, en el caso de determinadas máquinas o robots de aplicación industrial patentados y dotados de sistemas de inteligencia artificial integrados y asociados a sus procesos.

[959] Directiva 2009/24/CE del Parlamento Europeo y del Consejo de 23 de abril de 2009 sobre la protección jurídica de programas de ordenador. DO L de 5 mayo 2009. Pp. 16-22.
[960] GALLEGO SÁNCHEZ, E. (2019). "La patentabilidad de la inteligencia artificial. La compatibilidad con otros sistemas de protección". *La Ley Mercantil*. Nº 59, de 1 de junio 2019. Wolters Kluwer 2019. P. 15.

El sistema de protección del *software* es más permisivo en estos aspectos en EE.UU. respecto de la patentabilidad del mismo[961], y la UE ha flexibilizado también su postura para la posible protección del *software* como patente[962] cuando se trate del denominado *Computer Implemented Inventions*, esto es, cuando venga implementado y forme parte de una invención y produzca un efecto técnico, siempre y cuando el conjunto cumpla con todos los demás requisitos de patentabilidad.

Los algoritmos[963] no se hayan protegidos inicialmente como tales por la legislación de propiedad intelectual española y de la UE, tal y como refleja el propio Considerando 11 de la Directiva 2009/24/CE, sobre la protección jurídica de los programas de ordenador, el cual señala que "en la medida en que la lógica, los algoritmos y los lenguajes de programación abarquen ideas y principios, estos últimos no están protegidos con arreglo a la presente Directiva".

En la misma línea se expresa el artículo 1.2 de dicha Directiva, que señala que "Las ideas y principios en los que se base cualquiera de los elementos de un programa de ordenador, incluidos los que sirven de fundamento a sus interfaces, no estarán protegidos mediante derechos de autor con arreglo a la presente Directiva", del mismo modo que el artículo 96.4 LPI española.

---

[961] STROWEL, A. Y UTKU, S. (2016). *The trends and current practices in the area of patentability of computer implemented inventions within the EU and the U.S.* Disponible en: http://ec.europa.eu/newsroom/document.cfm?doc_id=41192. 2016. Pp. 10 y ss; KUR, A. Y DREIER, T. (2013). *European Intellectual Property Law*. EE, Cheltenham, UK, Northampton, MA, USA. 2013. P. 139; MARTÍNEZ, C. (2016). "Expanding Patents in the Digital World: The exemple of Patents in Software", en SEUBA, X.; GEIGER, C. Y PENIN, J. (Eds). (2018). "Intellectual Property and Digital Trade in the Age of Artificial Intelligence and Big Data*"*. *CEIPI-ICTSD*. Issue Number 5. 2018. P. 57.

[962] SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret*. Barcelona. Enero 2018. Pp. 8-9.

[963] El concepto de algoritmo puede definirse de manera general como un conjunto de operaciones ordenadas y finitas que permiten hallar la solución a un problema o, de manera más específica, como un conjunto de instrucciones que se le dan a una máquina, en un lenguaje concreto, para que realice una serie de operaciones determinadas con el fin de obtener un resultado. Es decir, el algoritmo podría ser una receta de cocina para combinar de manera ordenada unos ingredientes y obtener un resultado. Navas Navarro lo define como "el procedimiento para encontrar la solución a un problema mediante la reducción del mismo a un conjunto de reglas", es decir, la secuencia de instrucciones o estructura algorítmica que necesita y utiliza el sistema inteligente que especifica las acciones a ejecutar para resolver el problema. Ver NAVAS NAVARRO, S. (2017). "Derecho e inteligencia artificial desde el diseño. Aproximaciones", en NAVAS NAVARRO, S. (Dir.) *inteligencia artificial. Tecnología y Derecho.* Editorial Tirant lo Blanch. Valencia 2017. P. 24.

El derecho de autor no protege las ideas, los procedimientos, las metodologías o los conceptos matemáticos como tales.

Otra cuestión es su protección indirecta, en la medida que el algoritmo pueda plasmarse por escrito, ya que la normativa española precitada extiende la protección de los programas de ordenador -*software*- a toda la documentación técnica y los manuales de uso de un programa, donde podría recogerse el citado algoritmo o instrucciones asociadas al mismo.

Los algoritmos tampoco serían susceptibles de protección inicial vía propiedad industrial, esto es, como patentes, en la medida que el Convenio de Múnich sobre Concesión de Patentes Europeas, de 5 de octubre de 1973, elimina esta posibilidad en su artículo 52.2 a), que excluye de forma expresa la protección de los "descubrimientos y teorías matemáticas". Del mismo modo, el artículo 52.2 c) de dicha norma tampoco permite patentar "planes, principios y métodos para el ejercicio de actividades intelectuales, en materia de juegos o en el campo de las actividades económicas, así como los programas de ordenador". Y, por último, la Ley 24/2015, de 24 de julio, de Patentes, española recoge idénticas limitaciones.

No obstante, los algoritmos podrían protegerse conforme al artículo 52.3 del Convenio sobre la Patente Europea conforme comentaré en el próximo apartado y, en cualquier caso, serían susceptibles de protección a mi juicio, en caso de concurrir los requisitos necesarios para ello, como secreto empresarial al amparo de lo previsto en la Ley 1/2019, de 20 de febrero, de Secretos Empresariales, que transpone la Directiva (UE) 2016/943 del Parlamento Europeo y del Consejo, de 8 de junio de 2016, relativa a la protección de los conocimientos técnicos y la información empresarial no divulgados (secretos comerciales) contra su obtención, utilización y revelación ilícitas. Y todo ello sin perjuicio de otras vías de protección y reacción ante determinadas conductas, como los marcos jurídicos de competencia desleal o los marcos penales reguladores de los delitos contra el secreto de empresa, entre otros.

La jurisprudencia estadounidense ha sido algo más flexible al abordar algunos supuestos y ha confirmado en algún caso aislado la posibilidad de proteger un algoritmo asociado a

un método de negocio determinado[964] a través de las patentes, partiendo de la regla general de la imposibilidad de proteger los algoritmos mediante esta vía.

Por su parte, el *hardware* quedaría protegido principalmente por la legislación española en materia de patentes y de protección del diseño industrial, esto es, la Ley de Patentes[965] y la Ley de Protección Jurídica del Diseño Industrial[966] españolas.

Los "chips" o topografía de elementos semiconductores se hayan regulados y protegidos por la Ley 11/1988, de 3 de mayo, de protección jurídica de las topografías de los productos semiconductores[967] y, de otro, el Decreto 1465/1988, de 2 de diciembre, por el que se aprueba el reglamento para la ejecución de la Ley 11/1988, de 3 de mayo, para la protección jurídica de las topografías de los productos semiconductores[968].

Las bases de datos, a su vez, se hallan protegidas doblemente en la precitada LPI y la información se hallaría protegida por distintos marcos, especialmente conforme a su naturaleza, esto es, aquella relativa a creaciones intelectuales a través de la LPI, la relativa a datos personales por la normativa general y sectorial en Europa y España en materia de protección de datos personales, en particular, por el RGPD y por la LOPDGDD, y, en caso de información empresarial confidencial y secreta y relacionada con sus innovaciones, por la Ley de Secretos Empresariales precitada, entre muchas otras de carácter general y sectorial.

La protección de algunos sistemas de inteligencia artificial que hacen posible la creación de obras intelectuales e invenciones están siendo objeto de solicitud y obtención de protección global como patentes en distintos países, especialmente EE.UU., lo que ha motivado que algunos creativos y artistas humanos hayan hecho público su rechazo a estos sistemas ante su capacidad para crear esculturas, pintar oleos, escribir guiones, crear modelos 3D, diseñar personajes, animar imágenes, creación de escenarios[969] o incluso

[964] SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret.* Barcelona. Enero 2018. P. 11
[965] Ley 24/2015, de 24 de julio, de Patentes. BOE 25 Julio 2015
[966] Ley 20/2003, de 7 de julio, de Protección Jurídica del Diseño Industrial. BOE 8 Julio 2003
[967] BOE 05.05.1988
[968] BOE 08.12.1988
[969] Algunos ejemplos publicados: "Electronic Arts patenta la creación de escenarios mediante redes neuronales". Recuperado de: https://www.anaitgames.com/noticias/electronic-arts-patenta-la-creacion-de-escenarios-mediante-redes-neuronales. Consultado el 30.04.2021; "Warner Bros patents Shadow of

desarrollar videojuegos, si bien, de nuevo, reiterar que la tecnología no viene a sustituir al ser humano sino a complementarlo y mejorarlo, por lo que, en mi opinión, debe ser aceptada como un instrumento para la consecución de los fines pretendidos.

El posible registro y protección como patente de los sistemas inteligentes fue abordada por la Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial[970], para lo cual, según la misma, además de concurrir los requisitos de actividad inventiva y novedad mundial, se debe exigir una descripción exhaustiva de la tecnología subyacente, lo cual, en la práctica, puede suponer dificultades en determinados sistemas de inteligencia artificial, especialmente por la complejidad de sus razonamientos y funcionamiento.

De nuevo, como he anticipado, la inteligencia artificial debe ser acometida como un medio para el ser humano más que como una amenaza.

En mi opinión, la creatividad humana nunca podrá ser sustituida por una máquina, si bien, las capacidades que comporta la tecnología deberán ser aprovechadas adecuadamente por los creadores humanos como un medio o instrumento para su actividad profesional y empresarial, así como un complemento cada vez más necesario.

3.2. La titularidad y protección de creaciones de sistemas inteligentes

Respecto de la titularidad y protección de creaciones llevadas a cabo por sistemas inteligentes, la pregunta inicial a formularse en esta materia es obvia: ¿Puede un sistema inteligente ser creativo o tener capacidad creativa? La capacidad creadora, sin duda, la tiene.

Mordor's 'Nemesis System'", Recuperado de: https://www.kitguru.net/gaming/matthew-wilson/warner-bros-patents-shadow-of-mordors-nemesis-system/. Consultado el 30.04.2021.

[970] https://www.europarl.europa.eu/doceo/document/TA-9-2020-0277_ES.html

Esta cuestión constituye uno de los aspectos que mayor debate ha suscitado entre la comunidad científica, artística y cultural, y plantea diversas cuestiones jurídicas a resolver, especialmente, en relación con los conceptos tradicionales de creatividad, autoría y originalidad sobre los resultados obtenidos de su aplicación, esto es, sobre las creaciones "artificiales", "algorítmicas" o "sintéticas"[971].

A continuación, abordaré esta cuestión diferenciando entre obras creadas mediante sistemas de inteligencia artificial con intervención humana y las obras de origen estrictamente artificial, sintético o algorítmico, esto es, sin intervención humana.

Las preguntas posteriores obligadas serían las siguientes: ¿Debemos proteger las obras artificiales, algorítmicas o sintéticas?, ¿cómo?, ¿quién sería su titular?, ¿el marco regulador vigente permite esta protección mediante el derecho de autor?, ¿es el marco adecuado para su protección?, en caso contrario, ¿cómo proteger esos resultados?

Esta cuestión se está planteando ya en la práctica, especialmente en las artes creativas y, en especial, en varias de sus disciplinas y categorías, como el arte plástico o visual, escultura, fotografía, literatura, música, videojuegos o cine, pero también se ha suscitado en las disciplinas científicas.

De inicio, me permito citar al tecnólogo Tung[972], como punto de vista a considerar para su posterior análisis jurídico: "Las máquinas no son dueñas de lo que hacen". Y me permito adicionar a dicha afirmación lo siguiente: "Las máquinas no son dueñas de lo que hacen, ni deberían ni podrían serlo", entre otras razones, en la medida que carecen de personalidad y capacidad jurídica, de conciencia y de criterio, de autonomía y libertad plena, se hallan sujetas en su funcionamiento a reglas, instrucciones y restricciones predefinidas en su código y deberían estar sujetas durante todo su ciclo de vida al control

[971] Interesantes las reflexiones sobre lo se considera "arte" y el denominado arte sintético en ADSUARA, A. (2014). "*De otro(s) mundo (s)*". Editado por Sendemá Editorial y Escuela Superior de Arte y Tecnología (ESAT). Valencia 2014.

[972] TUNG, J.R. (2016). *Who Owns the Creation of an Artificial Intelligence?* 22 de agosto de 2016. Disponible en: https://blogs.findlaw.com/technologist/2016/08/who-owns-the-creation-of-an-artificial-intelligence.html.

y supervisión humana, por lo que quizás, éste podría ser uno de los puntos de partida más adecuados para abordar estas cuestiones desde un punto de vista jurídico.

3.2.1. Creaciones intelectuales como objeto de protección

Siguiendo un enfoque reflexivo y analítico de todas estas cuestiones, conforme he anticipado en la introducción de este capítulo, considero necesario abordar, de manera previa, en que consiste la "creatividad" -como capacidad o cualidad para crear o inventar, hasta fechas recientes, propia y exclusiva del ser humano, supuestamente-, para abordar posteriormente los requisitos para su protección a través del derecho de autor.

Según el Diccionario de la Real Academia Española, "creatividad"[973] puede concebirse como la facultad o capacidad de crear, esto es, de producir algo de la nada.

No obstante, se trata de un concepto completado a nivel jurídico por el contexto en el que se produce la creación y matizado por la legislación que regula su posible protección.

De inicio, una cosa es creativa no sólo por ser original sino por que aporta un sentido a su creador o a la sociedad en general.

Desde el punto de vista jurídico, una "creación"[974] sería el producto del esfuerzo intelectual de una persona que representa un incremento del conocimiento o el progreso de carácter científico, técnico o estético.

Conforme profundizamos en los distintos marcos reguladores en España que protegen las creaciones, nos encontramos, de un lado, con el concepto creación intelectual que ofrece la LPI[975] española.

[973] Recuperado de https://dle.rae.es/creatividad. Consultado el 12.12.2020.
[974] Recuperado de https://dpej.rae.es/lema/creaci%C3%B3n-industrial
[975] Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia. BOE 22 abril 1.996.

La norma regula en sus artículos 10 y siguientes que creaciones son susceptibles de protección como propiedad intelectual, en particular y como abordé anteriormente, todas las creaciones originales literarias, artísticas o científicas expresadas por cualquier medio o soporte, tangible o intangible, actualmente conocido o que se invente en el futuro y, entre otras, los libros, folletos, impresos, epistolarios, escritos, discursos y alocuciones, conferencias, informes forenses, explicaciones de cátedra y cualesquiera otras obras de la misma naturaleza, las composiciones musicales, con o sin letra, las obras dramáticas y dramático-musicales, las coreografías, las pantomimas y, en general, las obras teatrales, las obras cinematográficas y cualesquiera otras obras audiovisuales, las esculturas y las obras de pintura, dibujo, grabado, litografía y las historietas gráficas, tebeos o cómics, así como sus ensayos o bocetos y las demás obras plásticas, sean o no aplicadas, los proyectos, planos, maquetas y diseños de obras arquitectónicas y de ingeniería, los gráficos, mapas y diseños relativos a la topografía, la geografía y, en general, a la ciencia, las obras fotográficas y las expresadas por procedimiento análogo a la fotografía, los programas de ordenador, las obras derivadas o las bases datos.

Todavía hoy no se recogen expresamente en estos marcos creaciones que integran elementos como la interactividad que, sin duda, se hayan igualmente protegidas como propiedad intelectual, como web o videojuego, a mi juicio, obras complejas compuestas por múltiples obras.

Asimismo, el artículo 5 de la LPI establece que, únicamente puede considerarse “autor” de una creación a la persona natural -no artificial-, sin perjuicio de que las personas jurídicas puedan beneficiarse del régimen de protección previsto en la misma en los casos expresamente previstos en esta norma.

Por su parte, el Código Civil español establece en su artículo 429 que es la Ley sobre Propiedad Intelectual la que determina las personas a quienes pertenece ese derecho, la forma de su ejercicio y el tiempo de su duración. En casos no previstos ni resueltos por dicha ley especial se aplicarán las reglas generales establecidas en el Código Civil sobre la propiedad.

Los requisitos fundamentales que debe reunir una creación intelectual para ser considerada susceptible de protección por derechos de autor, conforme a la Ley de Propiedad Intelectual española, son esencialmente dos: La obra debe ser el resultado de un acto creativo de una persona natural y ese resultado debe ser original.

El marco regulador español en materia de propiedad intelectual parece claro en este sentido, sin perjuicio de determinadas particularidades que luego analizaré con mayor profundidad.

Por su parte, en relación con las invenciones y la propiedad industrial, sin perjuicio de su tratamiento con mayor detalle en el próximo apartado, la Ley de Patentes[976] española establece en su artículo 4 que únicamente se considerarán patentables las invenciones que sean nuevas -novedad, es decir, no comprendidas en el estado de la técnica-, que impliquen actividad inventiva y que sean susceptibles de aplicación industrial.

En su artículo 10.1 la Ley de Patentes española establece que el derecho a la patente pertenece al inventor o a sus causahabientes, no especificando la condición de persona natural del mismo, como sí expresamente contempla la LPI española, si bien, conforme a una interpretación histórica, sistemática y contextual del precepto, parece razonable considerar que se está refiriendo a personas, no a sistemas inteligentes.

No obstante, el precepto no requiere una condición específica al inventor, al que se le reconoce incluso un derecho esencial a ser mencionado en la solicitud y que puede ejercer frente al solicitante y titular de la patente.

Sin embargo, la ausencia de personalidad jurídica y de capacidad para ser titular de derechos y obligaciones por parte de los sistemas inteligentes, imposibilita de antemano dicha titularidad.

En cualquier caso, abordaré esta cuestión con mayor profundidad en el siguiente apartado, existiendo distintos autores y alguna resolución aislada que consideran que determinadas

[976] Ley 24/2015, de 24 de julio, de Patentes. BOE 25 Julio 2015

invenciones generadas por concretos sistemas inteligentes serían susceptibles de ser protegidas a través de un patente.

En relación con lo anterior, la precitada Ley de Protección Jurídica del Diseño Industrial[977] española, regula en sus artículos 5 a 12 los requisitos de las creaciones protegibles como tal, exigiéndose que las mismas sean nuevas y posean carácter singular. No obstante, su registro y protección queda reservado a personas físicas o jurídicas.

Por su parte, el marco jurídico que regula los denominados "chips", de un lado, la precitada Ley 11/1988, de 3 de mayo, de protección jurídica de las topografías de los productos semiconductores[978] y, de otro, el Decreto 1465/1988, de 2 de diciembre, por el que se aprueba el reglamento para la ejecución de la Ley 11/1988, de 3 de mayo, para la protección jurídica de las topografías de los productos semiconductores[979], reservan su registro y protección a personas físicas y jurídicas, siempre que sea el resultado del esfuerzo intelectual de su creador y no sea un producto corriente en la industria de semiconductores. Los sistemas de inteligencia artificial carecen de dicha condición, por ahora.

La cuestión es si, a pesar de la claridad del marco regulador vigente en España en materia de propiedad intelectual, se podría sostener de algún modo la susceptibilidad de protección bajo el mismo de las creaciones generadas por un sistema inteligente.

De inicio, sin perjuicio de lo que expondré más adelante y anticipando mis conclusiones, la respuesta sería negativa conforme al tenor literal vigente del mismo, salvo que se asocie a las personas físicas que intervengan en el proceso para su generación, en la medida que se exige inexorablemente un origen y autoría humana.

[977] Ley 20/2003, de 7 de julio, de Protección Jurídica del Diseño Industrial. BOE 8 Julio 2003
[978] BOE 05.05.1988
[979] BOE 08.12.1988

3.2.2. Los sistemas de inteligencia artificial como creadores de obras.

Los sistemas de inteligencia artificial son ya capaces en la actualidad de crear, sin prácticamente intervención humana, obras literarias, pinturas, videojuegos, cortometrajes, recetas de cocina o sinfonías[980].

De inicio, estos resultados podrían formalmente encajar en la definición de "obra" del artículo 2 del Convenio de Berna[981] o del artículo 10 de la LPI española.

Por lo que se refiere a obras literarias, en 2016 un grupo de investigadores de la *Universidad de Hakodate* en Japón crearon una serie de algoritmos pensados para escribir una obra literaria de ficción, entrenando el mismo e introduciendo las características básicas de los personajes. El resultado fue *El día que un ordenador escribe una novela*, completada por aquéllos y presentada a un concurso literario en el que quedó finalista[982].

Por lo que se refiere a obras fotográficas, en la actualidad es difícil distinguir entre una fotografía realizada por un sistema de inteligencia artificial y la realizada por un fotógrafo profesional debido a sistemas como "Creatism", un proyecto de los laboratorios de inteligencia artificial de Google, que consiste en un sistema inteligente basado en *Machine Learning*, que navega autónomamente a través de *Street View* y es capaz de reconocer si un paisaje tiene interés fotográfico y hacer los enfoques y ajustes necesarios para obtener una fotografía que podría ser indistinguible de la obtenida por un fotógrafo profesional.

Recientemente, ingenieros y programadores de la Universidad Tecnológica de Dalian en China y de la Ciudad Universitaria de Hong Kong han desarrollado una inteligencia

---

[980] A modo de ejemplo, durante el desarrollo de esta investigación, se está desarrollando un proyecto mediante inteligencia artificial para la terminación de la décima sinfonía que Beethoven dejó proyectada pero inacabada, basándose en estudios de historiadores, musicólogos, compositores e informáticos que partieron de los bocetos que dejó de la misma, su obra y procesos creativos del compositor alemán.

[981] Convenio de Berna para la Protección de Obras Literarias y Artísticas, revisado en París el 24.07.1971, aprobado por Instrumento de Ratificación de 02.07.1973. BOE nº 260, de 30.10.1974.

[982] ZAVIA, M.S. (2016). "Inscriben a una inteligencia artificial en un concurso literario. Queda finalista". Publicado en *Gizmodo*. 24.03.2016. https://es.gizmodo.com/inscriben-a-una-inteligencia-artificial-en-un-concurso-1766836828. Consultado el 19.02.2021

artificial que podría convertir una película o serie con subtítulos (no puede procesar la voz) en un comic para leer en papel. Es decir, crear obras *crossmedia*.

Por lo que se refiere a obras pictóricas, los sistemas inteligentes son capaces de crear las mismas con una calidad, técnica y características propias de algunos de los grandes pintores de la historia de la humanidad, incluso con características propias.

"e-David" es una máquina que crea pinturas utilizando un algoritmo de optimización visual que toma fotografías con su cámara y dibuja pinturas originales a partir de las mismas.

The Next Rembrandt, anteriormente comentado, fue un proyecto de inteligencia artificial limitado y absolutamente dirigido para producir un resultado, esto es, una impresión tridimensional de un nuevo lienzo de Rembrandt.

Los autores del sistema crearon un programa capaz de analizar minuciosamente la obra del pintor neerlandés del Barroco, para lo que nutrieron al mismo con más de 300 obras del autor -ya en el dominio público y del propio autor-, para su escaneo y análisis detallado de su técnica, incluyendo aspectos como trazo, composición, perspectiva, ángulos, geometría y aspectos estilísticos de la época, para generar una obra nueva con soporte en la base de datos obtenida.

En 2018 el retrato realizado por un sistema inteligente titulado *Edmond de Belamy, de La Famille de Belamy* se vendió en New York por 432.500 dólares por una de las principales casas de subastas a nivel mundial[983], de modo que, con independencia de su consideración como obra intelectual, el resultado obtenido tiene un valor y es, desde luego y de manera incuestionable, objeto de propiedad y de tráfico económico.

La creación se llevó a cabo mediante redes generativas antagónicas o *Generative Adversarial Networks* -GAN por sus siglas en inglés-, que constituye una de las técnicas

[983] COHN, G. (2018), "El arte de IA en Christie's se vende por $ 432,500". Publicado en *The New York Times.* 25.10.2018. Disponible en: https://www.nytimes.com/2018/10/25/arts/design/ai-art-sold-christies.html. Consultado el 19.02.2021.

más avanzadas en el desarrollo de la "creatividad" de los sistemas inteligentes, tanto cualitativa como cuantitativamente.

Esta técnica parte de un sistema de arquitecturas que reproduce las redes neuronales humanas y que fue nutrido con una selección de las mejores obras pictóricas de los siglos XIV a XX. El sistema generado tenía la capacidad para aprender de la experiencia, establecer patrones y alcanzar conclusiones, por lo que con todas esas herramientas y la base de datos generada se creó la obra bajo una inspiración múltiple (no única).

Como amante del arte y, también, por supuesto, del arte digital, podríamos incluso plantearnos de antemano una pregunta previa: ¿*Puede considerarse arte u obra artística*?

Según el artista estadounidense Andy Warhol "Si te sales con la tuya, es arte".

La siguiente cuestión que podríamos plantearnos sería si podemos considerarlo una expresión creativa o creación intelectual a efectos jurídicos y su protección como tal, con independencia de su calidad artística. La respuesta puede ser muy distinta.

Por lo que se refiere a obras más complejas como videojuegos, el programa "Angelina" creado por Michael Cook, estudiante de doctorado en el Grupo de Creatividad Computacional del *Imperial College of London* en Reino Unido, desarrolla videojuegos a partir de frases y parámetros básicos asignados por su creador.

La pregunta adicional asociada a todos estos supuestos es a quién y qué se pretende proteger, ¿el resultado, el algoritmo que ha permitido su creación o el sistema globalmente concebido que integra hardware, software, algoritmos o datos?

A modo de ejemplo, en futuras creaciones pictóricas, la precisión de trazo, intensidad, profundidad o estilo que pueda conllevar un "maridaje" perfecto de *hardware* complejo y *software* específico, pueden ser lo determinante para obtener la ejecución y materialización de un resultado pictórico concreto, posterior, obviamente, a estudios previos, la inclusión de un volumen de datos adecuado, un entrenamiento consecuente y una concepción artificial.

Reiterar que, en ocasiones, la doctrina se focaliza en el sistema inteligente como la suma de dos elementos esenciales como lo es el *software* y los algoritmos, pero no podemos obviar la relevancia de otros, en ocasiones, determinante del resultado, como el *hardware* o los datos, o el propio entrenamiento o autoaprendizaje, en función del tipo de sistema y resultado a obtener.

La cuestión es que las técnicas, metodologías y tecnologías sobre las que se sustentan los sistemas inteligentes no cesan en desarrollarse y aumentar sus capacidades.

Los sistemas GAN precitados están siendo superados por nuevos procedimientos de creación artificial más sofisticados como, por ejemplo, los CAN -*Creative Adversarial Network*- desarrollada por el laboratorio de IA de la *Universidad Rutgers de Nueva Jersey* en el marco de un proyecto dirigido por el profesor Ahmed Elgammal.

Este nuevo sistema no pretende generar obras de arte basadas en obras preexistentes sino desarrollar un proceso creativo que dé lugar a obras de carácter único y distinto a todo lo anterior.

El sistema se nutriría de obras ya creadas, pero se le adicionarían conceptos artísticos que permitirían al sistema desarrollar un sentido estético, y se le formaría para construir un procedimiento de creación similar al del ser humano, pero con desviación artificial de estilos preestablecidos, con el objetivo de permitirle alcanzar un determinado nivel de originalidad, si bien, delimitando dicha desviación con respecto a modelos de arte comercial, de modo que no provoque el rechazo del público.

Las preguntas que, de nuevo, me surgen inmediatamente son: ¿Y esto es arte? ¿Existe creatividad redirigida y parametrizada que pueda ser considerada original? ¿Podemos considerar que se trata de una obra derivada?

A mi juicio, estos sistemas comportan una mayor autonomía que los sistemas GAN, pero sigue siendo una autonomía relativa, determinada y determinante del resultado, el cual se haya alejado, consecuentemente, de una manifestación artística libre, independiente y creativa del propio sistema inteligente.

Los sistemas inteligentes precisarán de un estudio previo y generación del conocimiento necesario para su diseño y definición previa a su desarrollo, en base a los objetivos pretendidos y capacidades de los que serán dotados. Estos sistemas operarán técnicamente en todos estos ámbitos siguiendo un proceso de producción que, de manera muy básica y general a efectos ilustrativos, podemos diferenciar en tres etapas:

- Diseño y desarrollo: Diseño del sistema y programación, mediante la elaboración del código informático por parte de los programadores humanos que constituye en sí mismo una obra.

- Aprendizaje: Aprendizaje automático y, en su caso, aprendizaje profundo, que utiliza una estructura en capas de algoritmos que permite que la máquina aprenda y tome decisiones por sí misma. Este aprendizaje está automatizado por lo que frecuentemente prescindiría de la aportación humana directa. En esta fase se suscitaría la primera cuestión respecto de quién sería el titular de los resultados "relativamente impredecibles" creados por la máquina, especialmente mediante aprendizaje profundo.

- Creación: Producción o creación por el sistema de un resultado, de un producto (artístico, literario, etc.).

La cuestión esencial a reflexionar es si una inteligencia artificial puede ser creativa.

Dejando a un lado el concepto jurídico de "creatividad" por el momento, una primera aproximación desde una óptica científica al mismo y partiendo el posicionamiento al respecto del experto en inteligencia artificial López de Mántaras[984], según éste, la creatividad computacional consistiría en desarrollar *software*, basado en técnicas de inteligencia artificial, capaz de exhibir un comportamiento que podríamos considerarlo "creativo". Actualmente se está utilizando para componer música o producir artes plásticas.

[984] LÓPEZ DE MÁNTARAS, R. (2021). "¿Pueden las máquinas ser creativas?". *La Vanguardia.* 31.05.2021. Recuperado de: https://www.lavanguardia.com/ciencia/20210531/7484405/maquinas-creativas.html. Consultado el 31.05.2021.

Según el mismo, una idea creativa, en general, podría definirse como una combinación novedosa y valiosa de ideas conocidas, considerando que es una forma avanzada de resolución de problemas que involucra memoria, analogía, aprendizaje y razonamiento, y que podría ser emulada.

No obstante, considera que alcanzar un alto nivel de creatividad esta fuera del alcance de la inteligencia artificial en la actualidad, dado que requiere no solamente generar combinaciones novedosas y de ideas conocidas, sino que requiere inventar conceptos e ideas radicalmente diferentes, lo que supondría ser disruptivo y esto es algo que un sistema inteligente no puede hacer actualmente.

Como he anticipado anteriormente, a mi juicio, la denominada creatividad computacional o algorítmica no pretende sustituir la creatividad humana sino potenciar la misma mediante la creación asistida, la cual se puede ver aumentada gracias a la tecnología.

Un ejemplo de plataformas incipientes en estas áreas es la plataforma colaborativa desarrollada por el Instituto de Investigación en IA en el marco del proyecto europeo "PRAISE" para el aprendizaje de habilidades creativas que incluye hombre y máquina, en la necesaria simbiosis a la que he hecho referencia en distintos apartados de esta investigación.

Desde un punto de vista científico, los sistemas inteligentes no pueden ser creativos, carecen de intención y de consciencia, aunque para el precitado López de Mántaras, no considera que deban ser las razones fundamentales para negar su potencial creativo.

Sin embargo, algunos expertos, como Margaret Ann Boden[985], especialista en ciencia cognitiva e inteligencia artificial, considera que los sistemas informáticos tienen comportamientos creativos y que la creatividad no es exclusiva de la inteligencia humana. De hecho, diferencia tres clases de creatividad que pueden estar presentes en los algoritmos de inteligencia artificial, en concreto, la creatividad combinatoria, la exploratoria y la transformadora.

[985] BODEN, M. A. (2009). *La mente creativa: Mitos y mecanismos*. Gedisa. Barcelona 2009

Lo que es indiscutible es que la inteligencia artificial es un medio también para la creación de obras artísticas, literarias, musicales, científicas o de cualquier otra naturaleza, y constituye uno de los argumentos que defienden distintas plataformas como el colectivo de artistas e investigadores que integran *Obvious Art*, reflejado en su "Manifesto[986]".

Entre otros aspectos, significan que los algoritmos de aprendizaje automático pueden potenciar su creatividad natural y que la noción de creatividad es extremadamente difícil de encapsular, ya que parece ser un proceso que implica una serie de factores que aún no están definidos adecuadamente.

Entre otros objetivos de este colectivo, pretende promover un nuevo nivel de colaboración entre un artista y su herramienta, "donde las manos del artista y las de la máquina se unan en la búsqueda de un nuevo tipo de estética y un marco conceptual más profundo". Como destacan en sus principios, la tecnología siempre ha estado al servicio de las ambiciones humanas como la mejor herramienta para superar nuestros límites.

Como expuse anteriormente, mediante algoritmos de Redes Generativas Antagónicas -RGAs o GANs por sus siglas en inglés-, se ha conseguido crear y pintar retratos que no existen en realidad, retratos sintéticos, que se han convertido, incluso, en codiciados objetos de propiedad y tráfico mercantil.

El marco jurídico vigente en España en materia de propiedad intelectual, como abordaré con posterioridad, impide igualmente reconocer ese atributo al sistema inteligente, por asociar la creatividad exclusivamente al ser humano para su protección a través del mismo.

3.2.3. Requisitos esenciales para su protección a través de la propiedad intelectual

Los requisitos esenciales para que una obra o creación intelectual pueda ser considerada susceptible de protección por derechos de autor conforme a la normativa española, son,

[986] Disponible en: http://obvious-art.com/wp-content/uploads/2020/04/MANIFESTO-V2.pdf. Consultado el 09.03.2021

de un lado, que la obra sea el resultado de un acto creativo de una persona natural y, de otro, que ese resultado sea original, conforme a lo dispuesto en el artículo 10 de la LPI española y en el artículo 2 del precitado Convenio de Berna, al igual que en otros ordenamientos jurídicos.

Los conceptos obra y autor en la normativa de propiedad intelectual española se presentan como indisociables, no hay obra sin autor, ni autor sin obra[987], y se vincula la originalidad de la obra con la persona del autor.

De inicio, de manera previa a abordar el primero de los requisitos de protección, esto es, el relativo a que la obra sea el resultado de un acto creativo de una persona natural, así como a analizar la posibilidad de atribuir la condición de autor a un sistema inteligente y de considerar una creación intelectual generada por el mismo como obra protegida por derechos de autor, me permito abordar el segundo de los requisitos, que es el de su originalidad.

El concepto de originalidad no está expresamente definido en la legislación española, ni en la europea, ni en los convenios internacionales en esta materia. No existe una definición legal.

Se trata de un concepto evolutivo no armonizado, ni a nivel internacional ni de la UE como destaca Saiz García[988], citando a Derclaye y Margoni.

La originalidad como criterio exigido en una obra protegible bajo el derecho de autor se ha consolidado a nivel de la UE y ha sido abordado en múltiples pronunciamientos del Tribunal de Justicia de la Unión Europea[989] como requisito para determinar la existencia de obra y protección, así como su vinculación con el autor.

---

[987] XALABARDER, R. (2020). "Inteligencia artificial y Propiedad Intelectual", en CERRILLO I MARTÍNEZ, A. Y PEQUERA POCH, M. (Coord). *Retos jurídicos de la inteligencia artificial*. Aranzadi S.A.U. (Thomson Reuters). Navarra 2020. P. 209.

[988] SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret.* Barcelona. Enero 2018. P. 18.

[989] Sentencias TJUE en los asuntos Infopaq C-5/08, de 16 de junio 2009, FAPL C-403/08, de 4 de octubre de 2011 y Painer C-145/100, de 1 de diciembre de 2011.

La doctrina y jurisprudencia[990] españolas han ido estableciendo los criterios para determinar la concurrencia de este requisito de originalidad, básicamente de carácter objetivo asociado a su carácter novedoso,[991] de modo que la creación sea susceptible de diferenciación exterior de las que ya se conocen dentro de su género, esto es, concibiendo "obra original" como "obra novedosa", exigiendo únicamente que esa novedad objetiva tenga una relevancia mínima sin necesidad de que sea sustancial, a diferencia de las patentes, que requiere una novedad absoluta a nivel mundial. Y ello sin perjuicio de una cierta altura creativa.

En definitiva, en España, en general, la originalidad no se determina en base a criterios estrictamente subjetivos relativos a que se trate de una creación propia de su autor, que sí se consideran respecto de programas de ordenador como en otras jurisdicciones, tal y como luego abordaré.

Este posicionamiento difiere del construido por el Tribunal de Justicia de la Unión Europea (TJUE) en relación con los criterios subjetivos a considerar para valorar la concurrencia de estos requisitos, lo que genera cierta inseguridad jurídica y a juicio de algunos autores, como la precitada Saiz García, una posible afectación del principio de libre circulación de mercancías y de la libertad de prestación de servicios en el seno de la UE.

Según el TJUE, se consideran obras originales aquellas expresadas por cualquier medio o soporte tangible o intangible, actualmente conocido o que se invente en el futuro, que sean fruto de la creación intelectual propia de su autor. En definitiva, obras que sean "manifestación de la actividad creativa" de su autor y no una mera copia de obras anteriores.

[990] Sentencia del Tribunal Supremo español de 26.04.2017 (Sala Primera). F.J. 9.3.
[991] BERCOVITZ, R. (2017). *Comentarios a la Ley de Propiedad Intelectual*. 4.ª Ed. Editorial Tecnos, 2017. Pp. 162 y siguientes.

A mi juicio, la "expresión creativa" del autor debería ser igualmente considerada a la hora de determinar la originalidad de una obra, cuestión sobre la que ha tenido ocasión de pronunciarse el TJUE[992] en diversas ocasiones.

En cualquier caso, el requisito de originalidad en el derecho español se haya estrechamente ligado al primero de los requisitos expuestos, esto es, el de creación humana, por lo que éste excluye la protección jurídica bajo derechos de autor de las obras creadas por un sistema inteligente.

Ese vínculo indisoluble entre autor y obra protegible al que aludía anteriormente también se halla reconocido en otros países como EE.UU., donde recientemente se ha cuestionado de nuevo ante la negativa de su Copyright Office[993] a registrar el *selfie* de un mono que manipuló equipamiento del fotógrafo británico de naturaleza David Slater.

La doctrina en la que se fundamenta esta decisión podría ser aplicada, con ciertas reservas conforme aludiré posteriormente, a las obras creadas por sistemas inteligentes, los cuales, en base a la misma, no podrían considerarse autores.

Según el *Copyright Office Compendium*[994] estadounidense, para la protección de una obra, la autoría sólo puede ser humana, por lo que niega la protección a obras sin autoría

---

[992] Sentencias de 16 de julio de 2009 (Caso Infopac International A/S c. Danske Dagblades Forening, C-5/08), de 22 de diciembre de 2010 (Caso Bezpe-nostní softwarová asociace - Svaz softwarové ochrany (BSA) c. Ministerstvo kultury, C-393/09), de 1 de diciembre de 2011 (Caso Eva-Maria Painer c. Axel Springer AG y otros, C-145/10) y de 12 de septiembre de 2019 (caso Cofemel - Sociedade de Vestuário,S.A. c. G-Star Raw CV; C-683/17).

[993] PALLANTE, M. (2017), "From monkey selfies to open source: The essential interplay of creative culture, technology, copyright office practice, and the law". *Washington Journal of Law, Technology & Arts*. Vol. 12, Issue 2 Winter 2017. Disponible en: http://digital.law.washington.edu/dspace-law/handle/1773.1/1703. Consultado el 09.03.2021.

[994] U.S. Copyright Office, Compendium of U.S. Copyright Office Practices. Chapter 313.2. 3ª Edición 2021: *As discussed in Section 306, the Copyright Act protects "original works of authorship" 17 U.S.C. § 102(a) (emphasis added). To qualify as a work of "authorship" a work must be created by a human being. See Burrow-Giles Lithographic Co., 111 U.S. at 58. Works that do not satisfy this requirement are not copyrightable. (…) Similarly, the Office will not register works produced by a machine or mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author. The crucial question is "whether the 'work' is basically one of human authorship, with the computer [or other device] merely being an assisting instrument, or whether the traditional elements of authorship in the work (literary, artistic, or musical expression or elements of selection, arrangement, etc.) were actually conceived and executed not by man but by a machine".* Traducción libre del autor de esta investigación: "Como se explica en la Sección 306, la Ley de Derechos de Autor protege las "obras originales de autor" 17 U.S.C. § 102(a) (énfasis añadido). Para ser considerada una obra de "autoría", una obra debe ser creada por un ser humano. Véase Burrow-Giles Lithographic Co., 111 U.S. en 58. Las obras que no cumplen este requisito no son susceptibles de ser protegidas por derechos de autor. (...) Del mismo

humana, así como a las producidas por máquinas o procesos meramente mecánicos que operan aleatoria o automáticamente sin ninguna aportación creativa o intervención de un autor humano.

Este compendio de prácticas de la Oficina de Derechos de Autor de los Estados Unidos constituye el manual administrativo del Registro de Derechos de Autor estadounidense, de conformidad con lo previsto en el Título 17 del *United States Code* y en el Capítulo 37 del *Code of Federal Regulations*.

El mismo establece expresamente que no se registrarán las obras producidas por una máquina o un mero proceso mecánico que funcione de forma aleatoria o automática sin ninguna aportación creativa o intervención de un autor humano.

En este sentido, este compendio significa que la cuestión crucial es si la "obra" es básicamente de autoría humana, siendo el ordenador u otro dispositivo un mero instrumento auxiliar, o si los elementos tradicionales de autoría de la obra (expresión literaria, artística o musical o elementos de selección, arreglo, etc.) fueron realmente concebidos y ejecutados no por el hombre sino por una máquina.

En definitiva y como conclusión inicial, para que una obra pueda ser susceptible de protección como propiedad intelectual en general y por el sistema de derechos de autor en particular, será necesario que el acto creativo originario que genere la obra sea llevado a cabo por una persona física, que sea novedosa y que, siguiendo los criterios jurisprudenciales precitados del TJUE, que incorpore su expresión artística y su impronta personal.

Hechas estas conclusiones iniciales sobre la originalidad como el segundo de los requisitos objeto de análisis, a continuación, abordaré con detalle el primero de los requisitos precitados, esto es, la autoría humana.

---

modo, la Oficina no registrará las obras producidas por una máquina o un mero proceso mecánico que funcione de forma aleatoria o automática sin ninguna aportación creativa o intervención de un autor humano. La cuestión crucial es "si la 'obra' es básicamente de autoría humana, siendo el ordenador [u otro dispositivo] un mero instrumento auxiliar, o si los elementos tradicionales de autoría de la obra (expresión literaria, artística o musical o elementos de selección, arreglo, etc.) fueron realmente concebidos y ejecutados no por el hombre sino por una máquina".

La protección de las obras creadas mediante o por sistemas de inteligencia a través del derecho de autor exige autoría humana.

El principio de autoría sobre el que se sustenta el sistema continental europeo de derecho de autor impide el reconocimiento de un derecho y protección por esta vía, si el resultado u obra no es fruto de la inteligencia humana.

Conforme significa Saiz García[995], ni aun produciéndose una perfecta emulación del cerebro humano por un sistema de inteligencia artificial, el resultado producido exclusivamente por la máquina podría calificarse como obra del ingenio ni daría lugar al nacimiento del derecho de autor.

Y esto no es exclusivo de este sistema de protección, sino que distintas legislaciones nacionales sustentadas en el sistema del *common law* o *copyright*, como la estadounidense precitada[996], exigen el origen humano de las obras susceptibles de protección por el mismo.

Otros ordenamientos no recogen expresamente esta exigencia, si bien, la misma se deriva de los propios fundamentos sobre los que se sustenta el derecho de autor y su reconocimiento en el artículo 27.II de la *Declaración Universal de los Derechos Humanos*, sin perjuicio de las referencias indirectas contenidas en los mismos, como por ejemplo en el Derecho alemán, que se refiere a las obras protegibles como "creaciones personales del espíritu" - *persönliche geistige Schöpfung*-, del que carecen los sistemas inteligentes.

No obstante, debido al momento en que fueron concebidos algunos de estos regímenes y sistemas, es obvio que no podían ni tan siquiera considerar entonces una creación inteligente no humana, ni desde luego la realidad actual y el potencial del desarrollo y aplicación de la inteligencia artificial, sus capacidades y posibilidades, que cuestionan

[995] SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret.* Barcelona. Enero 2018. P. 15.
[996] El § 306 Compendium of the U.S. Copyright Office Practices, al que he aludido anteriormente, exige expresamente para inscribir registralmente una obra que haya sido creada por un ser humano.

posturas doctrinales taxativas e inflexibles respecto de la imposibilidad de que una máquina pueda ser creadora intelectual de obra alguna[997].

Hasta fechas recientes la capacidad natural de crear parecía exclusiva del ser humano. Ahora nos enfrentamos a una capacidad artificial de crear, creada y conferida por el ser humano.

De hecho y reflexionando más allá de los marcos vigentes, si nos cuestionamos qué es el "intelecto" y aun considerándolo un atributo exclusivamente humano, no podemos negar la capacidad creadora de los sistemas inteligentes, si bien, no fruto de un intelecto propio, sino del creado y dependiente del intelecto humano, es decir, fruto de un intelecto artificial creado por el intelecto humano.

Ello podría ser la base de reflexión para analizar si los sistemas tradicionales de propiedad intelectual son los marcos adecuados para la protección jurídica de las creaciones llevadas a cabo por sistemas inteligentes y, en caso afirmativo, para analizar la protección más adecuada de las obras de origen artificial, algorítmico o sintético, la titularidad y tipología de los derechos sobre las mismas, la diferenciación entre creador y autor, así como la atribución de ésta última condición exclusivamente a la persona física o jurídica detrás de la obra, ante la imposibilidad de hacerlo a un ente sin personalidad jurídica y sin posibilidad de ser titular y ejercer derechos y obligaciones -sin perjuicio de que pudiere reconocérsele la condición de creador en base a una capacidad atribuida al mismo por el intelecto humano-.

El derecho de autor actual se construyó sobre la base de un derecho de la personalidad del que nunca podría ser titular un sistema inteligente, carente de personalidad jurídica y toda capacidad para ser titular y ejercer derechos y obligaciones, como he analizado en anteriores capítulos.

Dejando a un lado parcialmente estas cuestiones por ahora, y prosiguiendo con el análisis de ese vínculo indisoluble entre autor y obra protegible al que he aludido anteriormente,

[997] ULMER, E. (1980). *Urheber- und Verlagsrecht*. 3ª Ed. Springer. Berlin-Heidelberg-New York. 1980. P. 128.

éste debe ser matizado en supuestos particulares, expresamente previstos en distintos ordenamientos jurídicos.

La propia legislación española de propiedad intelectual precitada -LPI-, prevé algunos supuestos donde se produce una cierta disociación entre los conceptos precitados y se protegen obras y la titularidad de derechos sobre las mismas a personas distintas al autor o no intervinientes directamente en la creatividad o aportación intelectual -siempre asociada al ser humano- como consecuencia, principalmente, de la existencia de una inversión de recursos o aportación organizativa o empresarial.

A modo de ejemplo, los artículos 8 y 97 de la LPI española respecto de la obra colectiva, a la que luego aludiré con mayor detenimiento.

Y estas excepciones no son exclusivas de la legislación española y de los regímenes sustentados en el denominado "derecho de autor".

Por ejemplo, EE.UU. regula las obras creadas por encargo -*works made for hire*- reguladas en el artículo 201 (b) de su *United States Code - Copyright* -USCC por sus siglas en inglés-, de las que se considera "autor" al comitente de dichas obras, sea una persona física o jurídica, sin perjuicio de la exigencia de que el creador -que no ostentará la condición de autor a efectos jurídicos- debe ser una persona física, generando una nueva ficción jurídica para atribuir dicha condición al comitente.

La cuestión es que la inteligencia artificial ya es capaz de producir en la actualidad resultados menos dirigidos y más "creativos", con escasa intervención humana y será capaz incluso de hacerlo sin intervención humana salvo en lo que se refiere a la creación del propio sistema inteligente, y esto por el momento, en la medida que los sistemas de inteligencia artificial más avanzados podrían tener la capacidad de crear otros sistemas inteligentes y autoprogramarse.

3.2.4. Análisis global, reflexiones y posibles soluciones.

3.2.4.1. Análisis global.

Durante mi análisis, estudio y reflexiones sobre el marco español así como de otros marcos como el estadounidense, el británico y el alemán, me planteé distintas alternativas para su análisis más profundo y reflexión, al objeto de valorar, de un lado, la posibilidad de incardinar las creaciones de origen artificial o sintético como objeto de protección por los derechos de autor u otras categorías, de otro, la posibilidad de considerar autor no sólo al "alguien" sino a "algo", esto es, un sistema inteligente y, de otro, como ya he expuesto anteriormente, la posibilidad de considerar si la creatividad en sí misma no es un atributo exclusivo de la mente o intelecto humano directo, sino que podría ser atribuido a un sistema inteligente.

De inicio, abrir el derecho de autor a las obras creadas por sistemas inteligentes chocaría frontalmente contra los fundamentos sobre los que se construyeron y sustentan estos sistemas tradicionales.

La claridad de algunos ordenamientos jurídicos comentados anteriormente, como el español, el británico o el estadounidense no deja lugar a dudas.

No obstante, los tratados internacionales de derechos de autor no recogen una definición específica del concepto "autor", ni una restricción expresa y formal a la condición humana del mismo, si bien, es mayoritario el posicionamiento de la doctrina respecto de la necesaria concurrencia de dicha condición.

A modo de ejemplo, el Convenio de Berna para la Protección de las Obras Literarias y Artísticas de 1886 alude al concepto "nacionalidad" como criterio para determinar la protección del autor que, en las construcciones doctrinales, a su vez, se asocia al de ciudadano, y éste a la condición de persona jurídica.

El problema es que esta construcción empieza a cuestionarse, especialmente al encontrarnos ya robots de inteligencia artificial que ostentan la condición de "ciudadano", por ejemplo, el robot *Sophia*, al que he aludido anteriormente, al que se le ha reconocido la "ciudadanía" Saudí, como expuse.

A pesar de las modificaciones posteriores de este instrumento internacional, es obvio que el contexto histórico, tecnológico, económico y social en el que se redactó difiere radicalmente del actual y no podía vislumbrar las cuestiones que se nos plantean en la actualidad o que nos plantearemos en un futuro cada vez más próximo.

La *Declaración Universal de los Derechos Humanos de las Nacionales Unidas de 1948* reconoce, en su artículo 27.2, el derecho que toda persona tiene a la protección de los intereses morales y materiales que le correspondan por razón de las producciones científicas, literarias o artísticas de las que sea autora, referido a personas físicas o naturales, de modo que tampoco tienen cabida en el mismo los sistemas inteligentes que carecen de dicha condición, ni de personalidad jurídica ni de la condición de persona artificial.

A nivel europeo, el concepto de autor vinculado formal y expresamente a las personas físicas también se contempla expresamente en las directivas que regulan los programas de ordenador y las bases de datos, sin perjuicio de que pueda considerarse como tal una persona jurídica, por ejemplo, en los supuestos de obra colectiva, sin perjuicio de que al final, detrás de la obra cuyos derechos pertenezcan a una persona jurídica, siempre deba existir una persona física que aportó su creatividad para generar la obra.

En concreto, el artículo 2 de la Directiva 2009/24/CE del Parlamento Europeo y del Consejo, de 23 de abril de 2009, sobre la protección jurídica de programas de ordenador, establece que se considerará autor del programa de ordenador a la persona física o grupo de personas físicas que lo hayan creado o, cuando la legislación de los Estados miembros lo permita, a la persona jurídica que sea considerada titular del derecho por dicha legislación, sin perjuicio de que las legislaciones de los distintos Estados miembro reconozcan las obras colectivas, en cuyo caso, las personas físicas o jurídicas que, según dichas legislaciones hayan creado el programa, serán consideradas autor, conforme ocurre en España.

Por su parte, el artículo 4.1 de la Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos, recoge un concepto similar.

La legislación de derechos de autor en Reino Unido, esto es, la *Copyright, Design and Patents Acts 1988* (CDPA) regula, en su artículo 9.1, el principio general de que debe considerarse autor de una obra a la persona que la haya creado, si bien, en el apartado 3 del precitado artículo, contempla las obras creadas por ordenador, considerando autor de las mismas a la persona que hubiere realizado "los arreglos" necesarios para la creación de dichas obras. Según su artículo 178, se considera que han sido "generadas" por un programa de ordenador, es decir, que el resultado creativo ha sido generado por el mismo, en los casos en que no exista un autor humano, considerando que estos autores carecerían de derechos morales según su artículo 81.

Esta norma reconoce la creación no humana, artificial o sintética y su protección por los derechos de autor, aunque de manera disociada a la máquina, de manera que la autoría no se otorgaría al sistema inteligente sino a la persona o personas que realizaron los arreglos para obtener ese resultado creativo, es decir, a diseñadores, y programadores que hubieran participado en su programación.

Algunos expertos[998] consideran que, según esta norma, en lo referente a obras creadas por ordenador, el creador será el programa, pero a los efectos de la titularidad de derechos serán considerados autores aquellas personas que hayan realizado los arreglos necesarios para que el programa genere la obra.

Comparto parcialmente este posicionamiento, en la medida que considero que este marco reconoce la protección del resultado de la máquina como creativo y susceptible de protección por derechos de autor, confiriendo a aquella la condición de "creador" aunque la condición de "autor" se le confiera a la persona detrás de la misma y de sus tareas de procesamiento para obtener el resultado creativo, que será quien ostente todos los derechos de autor sobre el mismo, es decir lo que denomino el "autor" mediato, frente al "creador inmediato" que sería el sistema inteligente, si bien, considero que es más adecuado identificarlo como tal, esto es, como "creador" o "sistema o medio creativo", y no atribuirle condición formal de "autor", vacía de todo contenido o derecho.

[998] SANJUÁN, N. (2019). "Inteligencia artificial y propiedad intelectual*"*. *Actualidad Jurídica Uría Menéndez*. 52-2019. Pp. 82-90.

EE.UU., a pesar de su tradición y construcción de su régimen de propiedad intelectual en el denominado sistema del *copyright*, regula estos aspectos de una manera muy próxima al sistema de derechos de autor que rige en la mayor parte de los Estados miembros de la UE, incluyendo España, en la línea que he referido anteriormente, esto es, no reconociendo la creación de obras protegibles por derechos de autor sin la intervención humana.

Al igual que el Convenio de Berna, la legislación estadounidense no lo contempla expresamente en su *United States Code - Copyright o* USCC -por sus siglas en inglés- pero si es requerido por su Oficina de Derecho de Autor -*US Copyright Office* o USCO por sus siglas en inglés- como requisito para registrar una obra para ser protegida por el *copyright*, la cual considera, conforme destacan expertos como Nerea Sanjuán[999], que el derecho de autor solo protege los frutos de la labor intelectual que están basados en los poderes creativos de la mente, citando en este sentido los casos Trade-Mark Cases, 100 U. S. 82 (1879) y Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 58 (1884) del Tribunal Supremo de EE. UU.

Asimismo, como he comentado previamente, según la *Copyright Office Compendium*[1000] estadounidense, para la protección de una obra, la autoría sólo puede ser humana, por lo que niega la protección a obras sin autoría humana, así como a las producidas por máquinas o procesos meramente mecánicos que operan aleatoria o automáticamente sin ningún aportación creativa o intervención de un autor humano.

Por lo que se refiere a España, como igualmente he referido anteriormente, el marco regulador vigente es claro en esta materia, de modo que únicamente será considerado "autor" de una obra protegida mediante la legislación de propiedad intelectual la persona física, sin perjuicio de que, en determinados supuestos, algunas personas jurídicas puedan beneficiarse del régimen de protección asociado a estos derechos.

[999] SANJUÁN, N. (2019). "Inteligencia artificial y propiedad intelectual". *Actualidad Jurídica Uría Menéndez*. Op.cit. Pp. 82-94.
[1000] Cit. U.S. Copyright Office, Compendium of U.S. Copyright Office Practices. Chapter 313.2. 3ª Edición 2021.

Y, es más, en España, así como en otros países de nuestro entorno, el derecho a la producción y creación literaria, artística, científica y técnica se ha elevado a la categoría de derecho fundamental recogido en el artículo 20.1.b) de la Constitución española, esto es, reconocido a las personas y protegido constitucionalmente.

De manera consecuente a todo ello, se reconocen al autor una serie de derechos morales y/o patrimoniales, de los que un sistema inteligente nunca podría ser titular en la medida que, en la actualidad, carece de personalidad jurídica y de la capacidad para ser titular de derechos y obligaciones, conforme al Derecho vigente en España y la UE.

Ello no obsta a que, en determinados supuestos particulares expresamente previstos, como he referido anteriormente, la legislación de propiedad intelectual española reconozca la condición de "autor" o los derechos patrimoniales, económicos o de explotación asociados a dicha condición a una persona distinta al autor, que puede ser incluso una persona jurídica, como en el caso de las obras colectivas o de los creadores de programas de ordenador contratados laboralmente por una entidad, a la que pertenecerían los derechos patrimoniales sobre el mismo, conforme analizaré con más detalle a continuación.

El artículo 8 de la Ley de Propiedad Intelectual española, establece que, salvo pacto en contrario, los derechos sobre la obra colectiva[1001] corresponderán a la persona física o jurídica que la edite y divulgue bajo su nombre, que no la condición de "autor".

Sin embargo, cuando la obra sea un programa de ordenador, la norma española regula un tratamiento distinto.

Respecto de programas de ordenador, el artículo 97.1 de la Ley de Propiedad Intelectual española considera "autor" del programa de ordenador la persona o grupo de personas naturales que lo hayan creado, o la persona jurídica que sea contemplada como titular de los derechos de autor en los casos expresamente previstos por esta norma.

[1001] Obra creada por la iniciativa y bajo la coordinación de una persona natural o jurídica que la edita y divulga bajo su nombre y está constituida por la reunión de aportaciones de diferentes autores cuya contribución personal se funde en una creación única y autónoma, para la cual haya sido concebida sin que sea posible atribuir separadamente a cualquiera de ellos un derecho sobre el conjunto de la obra realizada.

El precepto citado prevé en su apartado 2º que la condición de autor se atribuirá a la persona física o jurídica que edite o divulgue un programa como obra colectiva bajo su nombre, es decir, no sólo será meramente titular originario de los derechos de autor, sino que el precepto le atribuye la propia condición de "autor" -a diferencia del régimen general precitado de la obra colectiva-, en una nueva ficción jurídica para acomodar y dar solución a una cuestión económica y social.

Asimismo, el apartado 4º de este precepto establece que, cuando un trabajador asalariado cree un programa de ordenador -en el ejercicio de las funciones que le han sido confiadas o siguiendo las instrucciones de su empresario-, la titularidad de los derechos de explotación correspondientes al programa de ordenador así creado, tanto el programa fuente como el programa objeto, corresponderán, exclusivamente, al empresario -ya sea persona física o jurídica, salvo pacto en contrario y sin perjuicio de los derechos morales que correspondan sobre el programa a los trabajadores.

Este último precepto citado, establece en apartado 5º que la protección se concederá a todas las personas naturales y jurídicas que cumplan los requisitos establecidos en esta ley para la protección de los derechos de autor.

En definitiva, se trata de una nueva ficción jurídica creada por el legislador para atribuir los beneficios de la protección jurídica a través del derecho de autor de este tipo de obras a un tercero distinto de sus creadores o "autores" materiales, pero partiendo siempre de que detrás exista una creación o intelecto humano.

Quizás esa ficción jurídica para resolver problemas reales, podría ser una de las soluciones a barajar en el futuro para abordar la cuestión relativa a la protección de las obras creadas por sistemas inteligentes sin intervención humana, en la medida que en el diseño y programación del mismo que posibilite el resultado siempre haya intervenido el ser humano. Otra cosa distinta sería cuando la intervención humana se diluya todavía más en el futuro y nos podamos encontrar con sistemas inteligentes creados por otros y creaciones sin intervención humana.

3.2.4.2. Reflexiones iniciales.

De inicio, conforme a lo expuesto y analizado, en España y en la UE únicamente se reconoce protección mediante el derecho de autor a las obras que hayan sido creadas por una persona natural, derivada de la creatividad e inteligencia humana, exigiéndose la personalidad de su autor.

No obstante, considero necesario diferenciar entre las obras creadas por sistemas inteligentes como medio o instrumento al servicio del ser humano y las obras creadas por sistemas inteligentes sin intervención humana o con una intervención irrelevante, conforme analizaré en las conclusiones finales, en la medida que el grado de intervención humana y la autonomía efectiva son aspectos esenciales para determinar la adecuación de los marcos jurídicos actuales para la protección de las creaciones llevadas a cabo por sistemas inteligentes.

3.2.4.3. Posibles soluciones: Posicionamiento doctrinal y otras alternativas de protección

Las posturas doctrinales más conservadoras sustentadas en el concepto tradicional de autoría, sobre el que se ha construido la mayoría de los sistemas jurídicos actuales, incluyendo el español, parten de que solo una persona física puede crear una obra intelectual, por lo que las obras creadas por sistemas inteligentes no son protegibles por el derecho de autor tradicional ni por la propiedad intelectual en general.

La doctrina más autorizada es pacífica en ese sentido. Bercovitz[1002] afirma que sería absurdo tan siquiera especular con la posibilidad de una obra de ingenio cuya autoría no correspondiera a un ser humano.

[1002] BERCOVITZ, R. (2017). *Comentarios a la Ley de Propiedad Intelectual*. 4.ª Ed. Editorial Tecnos, 2017. P. 113

La mayoría de los países europeos con un régimen jurídico construido sobre el sistema tradicional del derecho de autor como Francia, Alemania, Italia o Portugal, regulan los derechos de autor en términos similares, como he referido.

Xalabarder[1003] considera que en estos supuestos "no hay 'autor' porque no hay obra", y pone como ejemplo de este paradigma el sistema "Iamus Computer" para la composición de música clásica o "Aaron", un sistema inteligente que pinta sobre lienzos creando "resultados" pictóricos sin aportación o intervención humana.

Para esta autora, ni las composiciones ni las pinturas creadas por estos sistemas son "obras" porque no hay obra sin autor persona física, ni autor sin obra, además de considerar la exigencia del marco regulador español, el cual requiere para su protección que las obras sean el resultado de "elecciones creativas y libres" del autor que, en estos supuestos, no pueden asociarse al sistema inteligente al no existir persona.

No obstante, según la misma, los resultados podrían beneficiarse de la posible protección como derechos conexos mediante una interpretación flexible de la norma, de modo que las creaciones pictóricas podrían protegerse como fotografía y los derechos de propiedad intelectual, más limitados en alcance y en el tiempo, corresponderían a su realizador que "no necesariamente debe ser directamente una persona natural."

Como he referido anteriormente, Reino Unido introdujo en 1988 una definición del atributo "generado por computadora" en su legislación de derechos de autor -CDPA por sus siglas en inglés-, concibiendo como tal la generada por ordenador en un contexto en el que no haya ningún autor humano de la obra generada. Según la autora precitada, se atribuiría la titularidad de los derechos de autor sobre las creaciones generadas por un sistema informático y se consideraría autor a la persona que hubiera llevado a cabo las tareas necesarias para la creación de la obra.

Este marco y reflexión consecuente, plantea la cuestión relativa a si deberían entenderse incluidos en este concepto, las personas que entrenaron al sistema o incluso a los programadores que crearon el sistema (*hardware* y *software*), o exclusivamente el

[1003] XALABARDER, R. (2020). "Inteligencia artificial y Propiedad Intelectual" en Cerrillo i Martínez, A. y Pequera Poch, M. (Coord). *Retos jurídicos de la inteligencia artificial*. Op. Cit. P. 215.

programa correspondiente sobre el que se sustenta. La ley de Irlanda es muy similar a aquélla.

Distintas legislaciones internacionales hablan de la necesidad de que se trate de "obras de la mente" o incluso en otros países se habla de "obras espíritu", como he referido, como la alemana y la libanesa.

En este sentido, una posible solución a la protección de las obras creadas por sistemas inteligentes autónomos podría ser otorgar un derecho afín, vecino o conexo o *sui generis* nuevo como proponen autores Sanjuán Rodríguez[1004] que incentive la inversión y desarrollo de estos sistemas con estas finalidades.

En este contexto se podría plantear la atribución de la titularidad originaria de los derechos de autor a un tercero distinto al verdadero autor -artificial, algorítmico, material o inmediato-, que podría ser la persona física o jurídica que hubiera coordinado la creación de la obra y su divulgación como propia, incluyendo a los programadores informáticos que hubieran desarrollado el sistema generador de la obra.

No obstante, esta opción no la considero viable, tal cual, conforme al marco jurídico vigente, en la medida que requeriría necesariamente la revisión del concepto de autor previsto en la LPI española, para posteriormente valorar su inclusión en el concepto de obra colectiva anteriormente precitada, lo que ha sido rechazado a nivel doctrinal por autores como Bercovitz[1005].

Sanjuán Rodríguez toma como referencia, en relación con los derechos afines, el derecho del editor que publica una obra inédita que esté en el dominio público y que pueda ser individualizada por su composición tipográfica, presentación y demás características

[1004] SANJUÁN, N. (2019). "Inteligencia artificial y propiedad intelectual". *Actualidad Jurídica Uría Menéndez*. Op. Cit. Pp. 90-91.
[1005] BERCOVITZ, R. (2017). *Comentarios a la Ley de Propiedad Intelectual*. 4.ª Ed. Editorial Tecnos, 2017. P. 113.

editoriales -ex artículo 129.2 de la LPI española-, citando en apoyo a su hipótesis a Ramalho[1006].

Asimismo, para su posible protección como derecho *sui generis* toma como referencia los artículos 133 y siguientes de la LPI española reguladores de su protección, la cual se sustenta en reconocer la inversión empresarial y humana -ya sea de medios financieros, tiempo, esfuerzo y otros-, en determinado tipo de productos novedosos con cierto grado de creación intelectual, aunque no alcance el grado de originalidad para su protección por derechos de autor, conforme al artículo 12 de la precitada normativa.

Sin embargo, en este contexto, la protección jurídica se orienta a la persona natural o jurídica que toma la iniciativa y asume el riesgo de efectuar las inversiones sustanciales orientadas a la obtención, verificación o presentación del resultado. Y los sistemas inteligentes carecen de dichas condiciones, por lo que no puede atribuírseles personalidad jurídica. De manera consecuente a todo ello, se evidencia de nuevo la necesidad de modificar el marco jurídico vigente para poder barajar esta opción.

Del mismo modo, Saiz García[1007] también manifiesta la conveniencia de la protección de estas obras bajo otras categorías dentro de la propiedad intelectual, en particular, como derecho afín, vecino o conexo al derecho de autor o incluso como el antedicho derecho *sui generis*.

En mi opinión, considero que todas estas opciones no resultan viables conforme al marco jurídico vigente, especialmente en la medida que, en estos supuestos, se requiere algún tipo de intervención humana a nivel intelectual para su protección, por lo que ambos planteamientos exigirían inicialmente una redefinición del concepto de autor y del artículo 5 de la LPI española. Además, como ésta última experta citada analiza, debería considerarse titular a la persona natural o jurídica que haya adoptado todas las medidas

[1006] RAMALHO, A. (2018). "Ex Machina, Ex Auctore? Machines that create and how EU copyright law views them". Disponible en: http://copyrightblog.kluweriplaw.com/2018/11/12/ex-machina-ex-auctore-machines-that-create-and-how-eu-copyrightlaw-views-them/.
[1007] SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret.* Op.cit. P. 15.

necesarias para llevar a cabo la obra, si bien, la cuestión es sobre qué obra ¿el programa desarrollado por los programadores o la obra creada sin su intervención?

Una interpretación flexible del artículo 8 de la Ley de Propiedad Intelectual española podría permitir atribuir esa titularidad a quienes hayan coordinado la creación y difusión del programa de ordenador que haya generado la obra, por ser ésta una extensión de aquella, lo que llevaría a una protección suplementaria que no parece que tenga justificación económica en la mayoría de los supuestos que se pueden plantear en la práctica, en particular, a personas titulares del hipotético derecho *sui generis* sobre creaciones generadas por inteligencia artificial, que lo serían sobre todas aquellas obras que el sistema inteligente hubiera podido crear para un tercero, licenciatario del derecho de uso o explotación.

A nivel internacional, juristas como Jane Ginsburg[1008], concluyen que un autor es el "ser humano que ejerce un juicio subjetivo al componer la obra y que controla su ejecución".

A la vista de este análisis expuesto, considero que, en la medida que el derecho de autor y los derechos conexos están concebidos en la actualidad para la protección de contenidos generados directa o indirectamente por personas físicas y que los sistemas inteligentes más avanzados carecen de personalidad jurídica alguna -no pudiendo ser titulares de derechos y obligaciones-, se requeriría una revisión de los aspectos esenciales del derecho de autor vigente para la inclusión expresa de la protección de las creaciones llevadas a cabo por sistemas inteligentes sin intervención o aportación humana, con un nuevo marco que definiera la titularidad de los derechos asociados a la misma y un específico marco de protección, restringido y, como algunos de los autores citados proponen, más acotado en el tiempo.

No obstante, cualquier iniciativa normativa en este sentido debe ser adecuado y prudente para no incurrir en una regulación excesiva que pueda provocar un efecto contrario o alejado de los objetivos de fomento y desarrollo pretendidos, como ocurrió con la

[1008] GINSBURG, J. C. (2003). "The Concept of Authorship in Comparative Copyright Law". 52 *Depaul L. Rev.* 1063, 1066. 2003.

regulación de los derechos *sui generis* sobre las bases de datos en la UE en comparación con otros sistemas más flexibles[1009].

Por lo que se refiere a algunos de los principales argumentos normativos y doctrinales que se esgrimen para su protección, analizados por Gervais[1010], se encuentran: El valor indudable del resultado creado -como reflejé al exponer algunos sistemas específicos-, la garantía de la competencia en el mercado -de modo que las producciones de los sistemas inteligentes deben protegerse porque de otro modo serían bienes gratuitos que competirían con creaciones de pago distorsionando el mercado-, el valor de una creación protegida por el derecho de autor ajeno a su calidad o mérito, o los seres humanos como autores sustitutos al tener los derechos sobre el código y, en consecuencia, también el resultado.

Otra de las razones esgrimidas para justificar la protección de los derechos de autor para las producciones de sistemas inteligentes es que las máquinas, al igual que los humanos, producen obras derivadas de obras preexistentes con derechos de autor, como lo ejemplifica el algoritmo de composición musical al que he hecho referencia con anterioridad. Sin embargo, la exigencia de que la obra derivada, para ser protegida por derechos de autor, también deba ser una obra original de autor, convierte este planteamiento en un argumento en contra.

Los argumentos normativos y doctrinales contra su protección se centran la necesidad de autoría humana como base del derecho de autor, requisito de la originalidad, concepto de obra derivada y en la responsabilidad asociada a las creaciones, no atribuible a un sistema inteligente.

No podemos negar que estos sistemas utilizan sus propios conocimientos para crear, sin perjuicio de que los derechos siempre correspondan al ser humano. La dificultad es su protección y la atribución de la titularidad de los derechos sobre dichas creaciones, conforme a los marcos vigentes, en función también del contexto.

[1009] *Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases*. European Union. 2018.
[1010] GERVAIS, D.J. (2020). "The machine as author". *Iowa Law Review*, 105 Rev.2053. 2020. Disponible: https://ilr.law.uiowa.edu/assets/Uploads/ILR-105-5-Gervais.pdf. Consultado el 10.03.2021.

Además, no debemos obviar que esta realidad podría entrar en conflicto con el futuro de la creatividad humana y si podría "quitar" espacio a creadores profesionales, que es uno de los debates que se están planteando ya en sectores como el del videojuego ante el aluvión de solicitudes de registro de patentes de sistemas inteligentes relacionados con el sector, especialmente en EE.UU., dada su relativa flexibilidad y cierta permisibilidad de admisión y otorgamiento de protección respecto de Europa.

Ni tampoco olvidar que la expresión artística, literaria y técnica es un medio para el ejercicio de derechos humanos fundamentales, como lo es la libertad de expresión, el derecho de información, la crítica o la libertad de cátedra.

Algunos expertos y autores como el precitado Gervais[1011], consideran que no deben protegerse las obras artificiales mediante el derecho de autor y que deben pasar directamente al dominio público, en la medida que los sistemas inteligentes actúen de forma autónoma y ya no sea una herramienta o medio en manos del ser humano. En ese sentido, favorecería el acceso a dichas creaciones y se fomentaría una mayor competencia creativa, pero ¿sería desleal frente a creadores humanos?

Esta solución plantea desventajas, como expone la precitada Xalabarder, como incertidumbre legal y dificultad de diferenciación entre obras personales o artificiales, inexistencia de plazos de protección, no sujeción a límites o excepciones que se aseguren el ejercicio de derechos fundamentales a la libertad de expresión, de información, al acceso a la cultura, al derecho a la educación o investigación, entre otros.

Otro enfoque muy distinto, sería partir de lo previsto en el precitado artículo 429 del Código Civil español, el cual establece que, debe ser la ley especial (LPI) la que establezca a que "personas" pertenecen los derechos de autor, la forma de ejercicio y el tiempo de duración, si bien, en los supuestos no previstos ni resueltos por la misma se deben aplicar las reglas generales establecidas en el Código Civil sobre la propiedad. En estos supuestos, la titularidad de derechos debería situarse en la esfera del propietario o,

[1011] GERVAIS, D.J. (2020). "The machine as author". *Iowa Law Review*, 105 Rev.2053. 2020. Disponible: https://ilr.law.uiowa.edu/assets/Uploads/ILR-105-5-Gervais.pdf. Consultado el 10.03.2021.

en su caso, del licenciatario/usuario cuando exista, ante la ausencia de personalidad jurídica y capacidad de obrar del propio sistema.

Por último, respecto de estas creaciones "artificiales", "sintéticas" o "algorítmicas" sin intervención o aportación humana, donde no existiría autoría, ni obra, ni protección, existen algunos marcos reguladores a nivel internacional que no podemos obviar, como he referido a lo largo de mis reflexiones precedentes, que ya contemplan su protección, disociando obra y autor, considerando la existencia de lo primero pero no de lo segundo y atribuyendo la creación a persona distinta del autor, por ejemplo, la precitada UK Copyright Act, en su artículo 9 (3) que establece que en los casos obras literarias, dramáticas, musicales o artísticas generadas por ordenador, se considerará autor a la persona que haya tomado las medidas necesarias para la creación de la obra. Nueva Zelanda, Hong Kong, Sudáfrica o India tienen disposiciones similares.

En consecuencia, en aquellos ordenamientos jurídicos donde no sean susceptibles de protección por la vía de derechos de autor, las creaciones artificiales se sujetarían a la regulación y protección contractual, a la protección técnica que proceda conforme a su naturaleza y al derecho de propiedad, sin perjuicio de su protección por otros marcos generales.

Si la conclusión en este sentido resulta irrefutable conforme a lo previsto por el marco jurídico vigente y posicionamiento mayoritario de la doctrina sobre esta cuestión, la pretensión de protección de las creaciones llevadas a cabo por sistemas inteligentes en España mediante el régimen de propiedad intelectual, exigiría reformular el concepto de autor y modificar el marco jurídico vigente en esta materia, para incorporar la posibilidad de contemplar autores o "creadores" no humanos como algunos ordenamientos empiezan a incorporar progresivamente -como el británico anteriormente citado-, y todo ello al margen de la persona física o jurídica en la que recaiga la titularidad y ejercicio de todos o algunos de los derechos asociados sobre la creación protegida.

No obstante, en mi opinión, considero necesaria una profunda reflexión previa, que excede del objeto y alcance limitados de esta investigación, sobre si el régimen jurídico a aplicar a estas creaciones artificiales debe ser, de antemano, el propio de los derechos

de autor u otras categorías de derechos de propiedad intelectual o de otra naturaleza, reconocidas en el ordenamiento jurídico europeo y español.

Desde mi punto de vista, si nos abstraemos de los fundamentos sobre los que se construyó el derecho de autor en el pasado, conforme a su finalidad, desde una interpretación lógica, sistemática e integradora, *a priori*, una obra creada por un sistema inteligente debería ser susceptible de ser considerada protegible como creación intelectual en la medida que la intervención humana hoy está presente de manera directa o indirecta en el mismo, si bien, la autoría y titularidad de los derechos tradicionalmente asociados a la creación -todos o algunos de ellos- nunca podrían ser reconocidas al sistema inteligente conforme a la legislación vigente, especialmente, ante la propia ausencia de personalidad jurídica y de capacidad para ser titular de derechos y obligaciones de estos sistemas, por lo que debería ser atribuida a la persona o personas físicas que estén detrás de dicha creación.

El resultado creativo de un sistema inteligente, en mi opinión, necesariamente debería ser protegible, vía derecho de autor y/u otras categorías, sin perjuicio de que la condición de autor recaiga en un tercero, y sin perjuicio de que pueda considerarse al sistema inteligente "creador" inmediato o instrumental, y al creador mediato atribuirle la condición de "autor" y/o otorgarle los derechos morales y patrimoniales sobre el resultado. Además, deberá abordarse el sujeto titular de estos derechos en función de la tipología de sistema, uso y participación humana, dado que pueden confluir los diseñadores, desarrolladores y entrenadores del sistema con el propio propietario, así como con el usuario/licenciatario del mismo que lo utilice como medio para la creación de obras, cuya disposición y contratación del sistema en cuestión puede tener como finalidad principal la producción de creaciones susceptibles de explotación y comercialización por parte del mismo, el cual puede tener un papel relevante o no en el proceso creativo.

La revisión de los marcos jurídicos quizás debería ir más allá y abordar una realidad innegable que son las "creaciones intelectuales artificiales, sintéticas o algorítmicas" -tangibles o intangibles- y los "creadores artificiales o sintéticos" -mejor que utilizar el término "autores".

Entre otras vías alternativas de protección al derecho de autor, se encontraría la propiedad ordinaria o común, sin perjuicio de su protección por otros cauces como, por ejemplo, a través de la propiedad industrial o la competencia desleal.

Ortego Ruiz[1012] plantea incluso extrapolar la posibilidad de dotar de personalidad jurídica a los sistemas inteligentes, a los derechos de autor, para poder así reconocerles la autoría de sus creaciones y la consecuente protección jurídica de las mismas, si bien, como he analizado en los capítulos anteriores, actualmente estos sistemas carecen de personalidad jurídica y parece que su reconocimiento ha salido, por el momento, de la agenda del legislador europeo, por lo que no considero jurídicamente viable este planteamiento en la actualidad, en base al estado de la propia tecnología, el grado de desarrollo de la inteligencia artificial aplicable y los marcos jurídicos vigentes, dejando a un lado los argumentos éticos y sociales.

Pero considero que no podemos conformarnos con los criterios tradicionales, dado que el marco jurídico vigente en España, como he apuntado, podría ofrecer algunas soluciones satisfactorias a cuestiones concretas si partimos desde un enfoque distinto, considerando que la calificación jurídica de las obras protegibles y prestaciones resultantes de la creación y producción artificial depende del grado de intervención o participación humana.

3.2.5. Reflexiones finales

Como he expuesto en mis reflexiones iniciales y en congruencia con el análisis realizado, para analizar la posible adecuación y aplicación de los marcos jurídicos vigentes para proteger las creaciones llevadas a cabo por sistemas inteligentes, considero necesario diferenciar el uso y procesos llevados a cabo por los mismos en relación con el grado de intervención humana, para determinar su protección por estos regímenes.

[1012] ORTEGO RUIZ, M. (2018). "El concepto de autor en la era de los robots". *Anuario de Propiedad Intelectual* 2017. Editorial Reus. 2018. Pp. 431 y 432.

3.2.5.1. La inteligencia artificial como medio o instrumento para la creación y producción de obras protegidas

La inteligencia artificial puede ser utilizada como medio o instrumento para la expresión creativa y artística de autores y artistas. Un ejemplo de ello podría ser "Flow Machines"[1013], un sistema de composición musical basado en inteligencia artificial que opera como asistente para su composición, con el objetivo de mejorar la creatividad artística.

El sistema de inteligencia artificial actuaría como medio o instrumento tecnológico para ayudar a la creación e interpretación sin afectar a la protección de éstas, por lo que el proceso creativo se asocia a las personas físicas que han creado el sistema y sus resultados.

La mayoría de los resultados generados por sistemas inteligentes han sido diseñados, preparados, alimentados, seguidos y validados por equipos humanos con la finalidad de obtener los mismos.

En consecuencia, en estos supuestos, en función del grado de intervención humana y la posible participación del titular del sistema inteligente en la explotación del resultado, el marco vigente podría contemplar ya soluciones para la adecuada protección de los derechos de autor y conexos de todas las partes involucradas y de las creaciones e interpretaciones resultantes, de un lado, los de los creadores y, en su caso, intérpretes o ejecutantes, y de otro, los del titular del sistema, mediante las reglas generales previstas en el caso de concurrencia de varios autores, esto es, obra colectiva, obra en colaboración, obra derivada o compuesta.

La no intervención humana o su reducción a niveles no relevantes requeriría otro tipo de protección distinta, ante la imposibilidad de que los marcos vigentes en materia de propiedad intelectual puedan dispensar la misma, conforme he analizado en los apartados anteriores.

[1013] https://www.flow-machines.com/

En este sentido, me permito significar el informe de la UE dando respuesta al *Draft Issues Paper on Intellectual Property and Artificial Intelligence*[1014] de la WIPO de 13 de diciembre de 2019, en el que expresamente propone que los debates comiencen con cuestiones fundamentales sobre la posible concesión de derechos de autor o derechos afines no sólo a los productos generados por la inteligencia artificial, sino también a los producidos con la ayuda de la misma.

No obstante, la protección a través del derecho de autor requerirá la concurrencia de los requisitos analizados con anterioridad, esto es, la originalidad y que constituya un acto creativo de una persona natural.

El principio de autoría propio del ordenamiento jurídico español, basado en la teoría personalista, exige que la titularidad originaria del derecho de autor corresponda al autor efectivo -persona física- de la obra.

Conforme a este principio, el titular originario del derecho de autor es la persona que efectivamente ha creado la obra, lo que constituye la regla general en el sistema español, así como en la mayoría de ordenamientos jurídicos, tanto sustentados en el derecho de autor como en el *copyright*.

No obstante, si la obra fuera el resultado de la colaboración de varios autores, la titularidad originaria del derecho pertenecería a todos ellos.

Como he analizado en los apartados anteriores, la propia LPI española contempla excepciones a la titularidad originaria humana y al principio general de autoría, en particular, la obra colectiva prevista en el artículo 8 de la misma, en la que las personas jurídicas pueden ser titulares originarios del derecho de autor en base a una ficción jurídica. Del mismo modo, otros países contemplan otras excepciones como las obras por encargo -*works made for hire*- analizadas previamente.

[1014] Disponible en: https://www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/call_for_comments/pdf/org_european_union.pdf. Consultado el 10.03.2021.

Estas excepciones son aplicables para la protección de obras creadas por equipos mediante sistemas de inteligencia artificial. A continuación, expondré con más profundidad algunas de estas excepciones:

a) Obra colectiva

El sistema de propiedad intelectual español, al igual que el francés o el italiano integran en su marco vigente la obra colectiva[1015]. Se considera titular del derecho a la persona natural o jurídica que "edita y divulga" la obra bajo su nombre. La norma italiana refiere esta titularidad y autoría a la persona que "organiza y dirige" la creación de una obra colectiva. No obstante, otros ordenamientos como Alemania, se rigen por el principio de autoría *stricto sensu*.

La obra colectiva supone una planificación previa por una persona para organizar, coordinar e integrar un conjunto de aportaciones concebidas aisladamente y con dicho propósito, para la obtención de un resultado final concreto.

Algunas obras creadas mediante sistemas inteligentes encajarían a la perfección dentro de este modelo colaborativo, en los que el derecho sobre la obra resultante se atribuiría a la persona que la edita y divulga bajo su nombre.

La protección de estas obras requerirá originalidad en los términos analizados, que sea creada por iniciativa y bajo la coordinación de una persona natural o jurídica que la edita y divulga bajo su nombre y estar constituida por la reunión de distintas aportaciones de varios autores cuya contribución personal se fusiona en una creación única, sin que sea posible atribuir un derecho *pro indiviso* por separado a cualquiera de ellos sobre el conjunto de la obra realizada.

[1015] Artículo 8 LPI española, artículo 113.2 y 5 del *Code de la propriété intellectuelle* francés -CPI- y 7 de la LDA italiana.

Sin embargo, esta vía podría chocar con la falta de armonización del concepto de originalidad y de obra colectiva dentro de la propia UE, como he expuesto anteriormente y destacan autores como Saiz García[1016].

b) Las obras por encargo.

Las denominadas *works made for hire,* a las que he hecho referencia anteriormente y que se regulan en ordenamientos jurídicos como el estadounidense, suponen la atribución de la titularidad originaria y exclusiva del derecho al empresario respecto de las obras creadas por sus empleados en el ámbito de la actividad para la que han sido contratados, y al comitente respecto de concretas obras realizadas por terceros encargadas por aquél en virtud de pacto escrito que así lo autorice. Es decir, como significan autores como Denicola[1017] o Ginsburg[1018], la persona natural o jurídica que promovió la creación de la obra es la que sería considerada *ope legis* titular originario de la obra a pesar de no haber participado en su creación.

c) Titularidad de obras generadas con programas de ordenador.

Distintos ordenamientos jurídicos contemplan también un régimen especial para estas obras.

El ordenamiento jurídico español únicamente regula específicamente la titularidad de los programas de ordenador, sin diferenciar entre programas y obras generadas por los programas, por lo que no contempla normas específicas en relación con la titularidad de los resultados de los programas de ordenador.

[1016] SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret*. Op.cit. P. 26.
[1017] DENICOLA, R. (2016). "Ex Machina: Copyright Protection for Computer-Generated Works". *Rutgers University Law Review* 251. 2016. Pp. 275 y ss.
[1018] GINSBURG, J. Y TREPPOZ, E. (2015). *International Copyright Law, U.S. and E.U. Perspectives*. Elgar. Cheltenham-Northampton 2015. Pp. 538 y ss.

Las obras generadas con programas de ordenador, al hilo de lo expuesto anteriormente y siguiendo también a la precitada Saiz García[1019], podemos diferenciarlas en tres categorías:

- Obras generadas con la asistencia de un programa de ordenador o *computer-aided works* -CAW por sus siglas en inglés-, en las que el programa constituye una herramienta o instrumento para el autor.

- Obras generadas por el ordenador independientemente o *computer-generated works* -CGW por sus siglas en inglés-.

- Obras con intervención humana relevante, pero con participación activa por parte del ordenador.

En general, en los casos en los que la intervención humana es relevante y que utiliza los programas de ordenador como instrumento o medio para la creación de una obra, se podría aplicar las reglas generales del derecho de autor, conforme expuse anteriormente.

En los casos donde no exista esa intervención humana o fuera irrelevante, quedaría fuera del derecho de autor conforme al marco jurídico vigente en España, salvo que el resultado creado por el sistema coincidiera con el objeto de protección de algún derecho conexo, por ejemplo, una base de datos. Algunos autores como Andrew[1020] y Denicola[1021] consideran que aquellas obras creadas por el ordenador, pero con cierta intervención humana relevante, las asimilan a las generadas de manera independiente por éste.

En Reino Unido, como he referido anteriormente, la Section 9 (3) de su CPDA incorporada en 1988, cuando todavía no se planteaba la problemática de las obras

[1019] SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret.* Op. Cit. P. 27.
[1020] ANDREW, J. W. (1997). "From Video Games to Artificial Intelligence: Assigning Copyright to Works generated by increasingly sophisticated Computer Programs". 25 *AIPLA Quarterly Journal*, vol. 25, nº1, 1997. Pp. 131 y ss.
[1021] DENICOLA, R. (2016). "Ex Machina: Copyright Protection for Computer-Generated Works". 69 *Rutgers University Law Review* 251. 2016. P. 283.

de origen artificial, atribuye la autoría de las obras generadas por ordenador a la persona que ha hecho los arreglos necesarios para que exista la obra[1022]. Y en el mismo sentido se regula en la legislación irlandesa, en particular, en la Section 2 (1) de la *Irish Copyright and Related Rights Act 2000*, como he referido anteriormente.

No obstante, a pesar de lo avanzado del marco regulador en Reino Unido, existe un debate actual interno sobre la regulación vigente de estas obras, el concepto de autor y la posible necesidad de actualizar la legislación británica de derechos de autor para adaptarse a las realidades de la inteligencia artificial[1023], especialmente ante la dificultad de establecer una distinción clara entre las obras de autoría humana -en las que la IA sería un medio o instrumento- y las obras generadas por ordenador sin autor humano. Asimismo, se está discutiendo sobre la posible limitación de la protección y el reconocimiento exclusivamente de derechos económicos.

Otros países fuera de Europa como Nueva Zelanda, Hong-Kong, Sudáfrica y la India, también contemplan esta posibilidad. Todos estos marcos atribuyen la titularidad del resultado a la persona que lleva a cabo los arreglos necesarios para la creación de la obra.

Prosiguiendo con mis reflexiones finales, la autoría, como he analizado, exigiría una aportación creativa principalmente centrada en el proceso más que en el resultado, en la medida que una obra se protege por ser una creación original propia de su autor con independencia de su valor artístico o económico en el mercado.

Si asociamos el principio de autoría con el criterio principalmente subjetivo de originalidad que pueda ser exigido en función de cada ordenamiento jurídico, se deben

[1022] Las opciones interpretativas en la doctrina, entre otros, Saiz García, oscilan entre otorgar el derecho a quien dirige el entrenamiento del sistema hasta que adquiere autonomía, a quien lo adquiere y lo pone en funcionamiento, o a quien financia su fabricación, procesamiento, etc. hasta que el mismo sale al mercado o a quien asume la dirección de todo el proceso creativo (que puede ser una persona distinta de quien presta la financiación.

[1023] BOND, T. Y BLAIR, S. (2019). "Artificial Intelligence & copyright: Section 9(3) or authorship without an author". *Journal of Intellectual Property Law & Practice*. Vol. 14. Nº 6. Oxford University Press. 2019. Editorial 1.

considerar tres premisas esenciales para considerar protegible la obra a través del derecho de autor:

a) De un lado, la intervención humana en la generación de una obra creada mediante un sistema de inteligencia artificial no debe limitarse a realizar actividades meramente técnicas, mecánicas, de acompañamiento, en definitiva, irrelevantes desde un punto de vista creativo e intelectual para la obtención del resultado, dado que en tal caso no podría considerarse original, como destaca Saiz García[1024]. Según esta autora, estas tareas podrían incluir la alimentación del sistema inteligente con datos (no su selección), el escaneado de elementos físicos para su traducción digital o la recarga de la impresora 3D.

b) De otro, la concepción y la ejecución de la obra para su protección, que pueden reunirse en la misma persona o disociarse.

   En el supuesto de creaciones por sistemas inteligentes estaríamos hablando de su disociación, en la que la concepción sería humana y la ejecución sería a cargo del sistema, con o sin intervención humana, sin perjuicio de que las crecientes capacidades de sistemas más avanzados puedan llegar a permitir la concentración la concepción y ejecución en el propio sistema. En este sentido, el grado de intervención humana y la libertad creativa puede ser diferente.

   Una iniciativa creativa o encargo con instrucciones poco específicas a su ejecutor deja en manos de éste el margen de autonomía y margen de libertad creativa para concretar la forma expresiva de la idea inicial, por lo que la autoría recaería en la persona que la desarrolle y concrete. Cuantas más instrucciones concretas se proporcionen al ejecutor sobre la forma expresiva de la idea inicial, menos margen tiene éste y más se aproxima a una obra colaborativa. Si estas instrucciones no dejan margen de creatividad al ejecutor, la autoría podría recaer en el comitente, en función del género y tipo de obra.

[1024] SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret.* Op.cit. P. 20.

Si trasladamos a los sistemas inteligentes estas reflexiones, nos llevaría a concluir que a mayor participación, intervención e instrucciones del equipo humano -principalmente ingenieros y programadores-, mayor es la probabilidad de imprimir originalidad a la obra resultante creada por el sistema situada en la esfera de aquél.

Si las instrucciones no son concretas y es el sistema quien las determina y ejecuta de autónoma, no se consideraría actividad relevante, por lo que no concurrirían los requisitos precitados para su protección, sin perjuicio de que pueda gozar de una presunción de originalidad.

c) Y de otro, la intervención humana y la originalidad asociada requerida para la protección de la obra también puede concurrir en la disposición y selección de los datos y contenidos que integren la obra, que puede constituir el factor determinante para dotarla de originalidad y, en consecuencia, ser objeto de protección por derechos de autor.

Si es el sistema inteligente el que decide bajo su criterio la selección y disposición de los contenidos que integran la obra final, con inexistencia o intervención humana meramente técnica o irrelevante, no concurriría el elemento de originalidad y, por tanto, no sería susceptible de protección mediante el derecho de autor.

La posible aportación creativa o intelectual, según el contexto, sería doble, de un lado, en la concepción de la obra y, de otro, en su ejecución -expresión material o formal-, rechazándose la autoría o coautoría a la persona que hace una aportación meramente técnica en su ejecución por muy valiosa que sea. Esto nos llevaría igualmente a negar la protección creativa de un resultado por parte de un sistema inteligente si la aportación técnica del mismo no se considera verdaderamente creativa, sino simplemente *instruccionada* por el creador del sistema, titular de los derechos de autor sobre el mismo.

Otra cosa distinta sería cuando los resultados de la obra generada por el sistema inteligente y la creatividad de los mismos tienen su origen en el creador del sistema y no tanto en su usuario, por ejemplo, en sistemas para la creación de animaciones o videojuegos. En estos supuestos, una de las soluciones más adecuadas en función del contexto podría ser su consideración como obra compuesta, considerando que la

titularidad de los derechos sobre la obra resultante, que incorpore obras preexistentes sin la colaboración del autor de éstas, correspondería al usuario del sistema conforme a lo dispuesto en el artículo 9 de la LPI, sin perjuicio de los derechos que le correspondan al autor de aquéllas y su preceptiva autorización.

La creatividad y producción artificial de obras y prestaciones protegidas es una realidad innegable como he expuesto.

La creación y creatividad artificial existen, pero, en mi opinión, ésta última no se conformaría como una capacidad o facultad de creación autónoma, independiente, libre y ajena al ser humano, sino habitualmente como medio o instrumento del mismo para crear todo tipo de obras, desde las más similares a las que podría crear un creador humano sin la utilización de estas herramientas, o a las absolutamente distintas y sólo susceptibles de creación con capacidades propias de sistemas inteligentes.

La mayoría de supuestos actuales, en base al estado de desarrollo y aplicación de la inteligencia artificial, considero que nos sitúa ante creaciones humanas sustentadas en sistemas inteligentes, en esa necesaria simbiosis hombre-máquina, cuyo fruto sería una obra artificial, algorítmica o sintética, por lo que considero incluso más apropiado hablar, más que de creatividad artificial, de creatividad sintética, la cual, por el momento, podría tener cabida en el marco regulador vigente en España en materia de propiedad intelectual, como medio o instrumento de los creadores humanos.

La cuestión más compleja se plantea en relación con aquellos sistemas inteligentes que no son un mero medio o instrumento para creadores humanos, sino sistemas que pueden tener capacidad para concebir, crear, interpretar y producir resultados de forma autónoma y sin intervención ni aportación humana, más allá de la creación del sistema.

Como he analizado anteriormente, las creaciones por parte de estos sistemas no tienen un acomodo actual en el marco jurídico vigente a nivel español y de la UE en materia de derecho de autor.

No obstante, en mi opinión, podríamos considerar que existe siempre una cierta intervención humana desde el momento del diseño y concepción de estos sistemas con

finalidades determinadas, aunque no concretas, aunque sea considerada de relevancia mínima y no excluya la categorización del resultado como artificial o algorítmico. No existiría esta intervención en sistemas inteligentes "creativos" creados por otros sistemas.

3.2.5.2. La protección de obras de origen artificial, sintético o algorítmico

La protección de obras creadas por sistemas de inteligencia artificial de manera autónoma, donde no existe intervención humana o la misma es irrelevante o meramente mecánica o técnica, no se haya prevista en la LPI española ni en la normativa de la UE. En consecuencia, de inicio no podrían protegerse los resultados obtenidos mediante este régimen.

Si no se acomete una previa revisión de los fundamentos sobre los que se sustenta el actual derecho de autor para incardinar este tipo de obras en su espectro de protección, nos encontramos ante resultados que pasarían al dominio público, siendo protegibles mediante el derecho general de propiedad, sin perjuicio de su protección por otros marcos como los reguladores de la competencia desleal.

El principal riesgo de este escenario podría ser el efecto desincentivador de la inversión y la innovación empresarial en este tipo de sistemas y una consecuente desventaja competitiva con otros países que han optado por establecer cierta protección a este tipo de obras en sus respectivas leyes de propiedad intelectual como, por ejemplo, Hong Kong, India y Nueva Zelanda.

Ante este escenario, algunas de las posibles opciones de protección en España a reflexionar, propuestas por distintos autores, serían las siguientes:

- Valorar la protección de estas obras por los derechos conexos o vecinos al derecho de autor ya contemplados en el ordenamiento jurídico español.

Conforme he referido en apartados anteriores, esta opción ha sido analizada por distintos autores, con distintos posicionamientos. La precitada Saiz García[1025] considera que el derecho conexo que mayor paralelismo guardaría con las obras creadas por sistemas inteligentes sería el derecho *sui generis* del fabricante de bases de datos.

Según la misma y, coincidiendo con otros autores citados anteriormente, las bases de datos realizadas autónomamente por un sistema de inteligencia artificial podrían ser protegibles por el derecho *sui generis* en cuestión, así como las fotografías con origen en estos sistemas como derecho conexo o las obras audiovisuales por el derecho exclusivo del productor. En cualquier caso, considero necesaria la previa reforma legislativa de los marcos vigentes para adecuar la titularidad a entes no humanos dado que, de otro modo, no podrían protegerse por esta vía.

- La creación de un nuevo derecho *sui generis* para proteger estas obras.

Esta opción ha sido igualmente propuesta por distintos expertos como Sanjuán Rodríguez, tal y como he expuesto anteriormente, si bien, exigiría igualmente una profunda revisión previa sobre su pertinencia y su impacto que podría tener un nuevo derecho de esta naturaleza en distintos ámbitos, especialmente ante experiencias negativas como el derecho *sui generis* sobre bases de datos, conforme han puesto de manifiesto autores como Ramalho, Perry y Margoni o la propia Saiz García[1026], que hace un interesante análisis de esta opción en el artículo citado al pie, en el que la autora se inclina más hacia una posible protección de las obras creadas por sistemas inteligentes a través de la propiedad industrial, bajo un sistema registral de inscripción constitutiva.

El fundamento de un derecho exclusivo de esta naturaleza parece que debería estar centrado en la protección de la inversión y esfuerzo empresarial en estos sistemas,

[1025] SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret.* Barcelona. Op. cit. P. 31.
[1026] SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret*. Op. cit. P. 33.

si bien, en ese caso, su titular debería ser quien realizó la inversión, participe o no en el proceso de generación del resultado.

Si conforme a este planteamiento, aquella persona o entidad es la que debe conservar los derechos exclusivos sobre todas y cada una de las obras creadas por sistemas inteligentes cuando puedan ser comercializados a terceros, el sistema propuesto carecería de sentido alguno en el tráfico mercantil. Además, si se atribuye al propietario - sea o no inversor- se desnaturalizaría de nuevo este fundamento. Realmente si lo que se pretende proteger son las obras que crean los sistemas inteligentes autónomamente, el propietario del sistema, el arrendatario o el usuario legítimo del sistema debería poder autorizar los usos que terceros lleven a cabo de las obras creadas por estos sistemas, no quién invirtió sus esfuerzos y recursos para la creación del sistema, que probablemente lo creó con el objetivo de explotarlo comercialmente a sus adquirentes o licenciatarios -propietarios o usuarios-, que a su vez lo adquieren con el objeto de utilizarlo para sí y/o explotar los resultados que del mismo se generen.

A mi juicio, en función del contexto, el valor de la inversión puede hallarse más en la funcionalidad del sistema inteligente más que en los concretos resultados producidos por el mismo, por lo que la creación de una protección y derecho sobre estos resultados no necesariamente debe constituir un aspecto verdaderamente determinante para la inversión en este tipo de sistemas, en función del contexto, obviamente. En esta misma línea, destacar a Saiz García que cita el estudio realizado por Keisner, Raffo y Wunsch-Vincent[1027].

En cualquier caso, la creación de una protección específica exigiría reflexionar y definir con precisión la misma.

- No modificar el marco actual del derecho de autor ni crear un derecho exclusivo. Dejar su regulación y protección por el propio mercado con utilización del régimen general de protección de la propiedad y de la explotación de cualquier producto en

[1027] KEISNER, A.; RAFFO, J. Y WUNSCH-VINCENT, S. (2015). *Breakthrough technologies- Robotics, innovation and intellectual property*. Economic Research. Working Paper No. 30. WIPO. 2015.

un sistema de libre competencia, llevándose a cabo las mismas a través de los marcos contractuales de regulación y protección para su uso y explotación. Autores como Ramalho[1028], proponen una solución de dominio público que integre un derecho exclusivo de divulgación.

Además, el propietario del sistema dispondría igualmente, entre otros, de los marcos reguladores del secreto empresarial y de la competencia desleal para proteger sus intereses. Y, por último, en caso de concurrir los requisitos legales específicos, el propietario del sistema inteligencia que pretendiese explotar los resultados producidos por el mismo dispondría también, como hasta la fecha, de otras vías para la protección de su inversión como el derecho de marcas, el diseño industrial, las patentes o los modelos de utilidad, y tanto del sistema como de las obras creadas por el mismo.

- Revisar el sistema de protección del derecho de autor *strictu sensu* y sus fundamentos para dar cabida a la protección de obras creadas por los sistemas inteligentes, con o sin creación de un nuevo derecho, pero revisando y armonizando el concepto de originalidad y autoría, especialmente respecto de la obra colectiva en este tipo de creaciones. Del mismo modo, esta revisión debería considerar la posibilidad de disociar la titularidad de los derechos morales de los económicos o de explotación.

3.3. Titularidad y protección de los datos y contenidos de los que se nutren los sistemas inteligentes para operar y entrenarse

La tercera cuestión más relevante que plantea la inteligencia artificial es la relativa a la titularidad y protección de los datos y contenidos de los que se nutren los sistemas inteligentes para operar y, en su caso, entrenarse, que pueden incluir contenidos y

[1028] RAMALHO, A. (2017). "Will Robots Rule the (Artistic) World? A Proposed Model for the Legal Status of Creations by Artificial Intelligence Systems". *Journal of Internet Law*. Julio 2017. SSRN: https://ssrn.com/abstract=2987757. Pp. 16-20.

creaciones protegidas por derechos de autor o por otros derechos, como patentes, diseño industrial, marcas, privacidad, imagen o secretos empresariales.

¿Qué ocurriría si el sistema inteligente se nutriera y entrenara con obras, imágenes, diseños protegidos? En todos estos supuestos, las facultades exclusivas de los titulares de los derechos indicados para autorizar o prohibir su uso y explotación podrían ser un obstáculo para el desarrollo y aplicación de la inteligencia artificial en este campo, y su uso podría constituir una infracción de derechos y determinaría el nacimiento de responsabilidades que podrían ser de distinta naturaleza, tanto civiles como incluso penales.

La LPI española regula en sus artículos 138, siguientes y concordantes, en relación con lo previsto en la Ley de Enjuiciamiento Civil[1029] española, distintos medios de protección de los titulares de derechos, entre otros, los de cesación y reclamación de los correspondientes daños y perjuicios.

Del mismo modo, la utilización de otros contenidos protegidos sin autorización como patentes, diseños industriales, marcas, datos personales, imagen o secretos empresariales podría dar lugar a similares responsabilidades.

Los datos constituyen el *input* esencial para el funcionamiento de los sistemas de inteligencia artificial y, en especial, para al aprendizaje automático y profundo, pero su utilización puede vulnerar distintos derechos ya protegidos por diferentes marcos normativos como los precitados, por ejemplo, los de privacidad, propiedad intelectual o secretos empresariales.

La cuestión plantea posibles conflictos de derechos que, debiendo resolverse a favor de sus titulares -máxime cuando se trate de derechos fundamentales-, puede impactar seriamente en el despliegue y aplicación de la inteligencia artificial en ámbitos específicos.

[1029] Ley 1/2000, de 7 de enero, de Enjuiciamiento Civil. BOE 8.01.2000

Sin duda, el uso y la libre circulación de datos se haya asociada a la libre circulación de conocimientos, si bien, deben considerarse los distintos intereses en juego y definir marcos equilibrados que no desincentiven la creatividad, la innovación y la inversión en las mismas. Sobre estos aspectos, me permito destacar el estudio y reflexiones de Ottolia[1030] sobre la libre circulación de datos para aplicaciones relacionadas con la inteligencia artificial, en especial, de un lado, sobre la posibilidad de un nuevo derecho de exclusiva sobre los datos que el precitado autor critica, el cual fue introducido por la propia Comisión Europea[1031] y objeto de debate en la comunidad científica[1032] -bajo el argumento de que un nuevo derecho de esta naturaleza permitiría la circulación de datos, a diferencia de los que sucede cuando se los protege a través del secreto, fomentando la misma- y, de otro, sobre la propuesta de una nueva excepción relativa a su aprovechamiento por parte de la inteligencia artificial con fines comerciales, como limitación de los derechos sobre los bienes inmateriales (obras intelectuales, bases de datos y datos personales). Del mismo modo, destacar el análisis que lleva a cabo en la obra referenciada, sobre los posibles impactos de la libre circulación de datos en la protección de la competencia.

El objeto y alcance de esta investigación me impiden profundizar en estos aspectos de indudable trascendencia, sin perjuicio de las consideraciones efectuadas sobre algunas de las recientes propuestas regulatorias en el ámbito de la UE con impacto en esta materia.

Los ordenamientos jurídicos de otros países, a diferencia de la UE, tradicionalmente han permitido la aplicación de excepciones y limitaciones para el aprovechamiento por parte de la IA de datos y creaciones, como el *fair use* en EE.UU. y los usos legalmente contemplados en los marcos vigentes en la UE y España.

La UE dispone de un marco jurídico que contempla importantes limitaciones generales a dicha circulación, especialmente en materia de propiedad intelectual y, en particular, el derecho *sui generis* sobre bases de datos, en materia de secretos empresariales y en materia de protección de datos personales, sin perjuicio de las barreras específicas que

[1030] OTTOLIA, A. (2018). *Derecho, Big Data e inteligencia artificial*. Editorial Tirant lo Blanch y G. Giappichelli Editore. Valencia-Torino. 2018. Pp.114-121.
[1031] Comunicación de la Comisión de 6 de mayo de 2015 bajo el título Estrategia para el mercado único digital en Europa. COM (2015).

tengan un origen contractual o corporativo, en virtud de códigos de conducta, éticos y de autorregulación.

En ese sentido, la UE dio un paso más relacionado con todo ello y, a través de la Directiva (UE) 970/2019[1033] sobre los derechos de autor y derechos afines en el mercado único digital, pendiente de transposición en España, impuso a los Estados miembros la obligación de introducir un límite legal para fines de minería de textos y datos[1034] en determinados contextos.

Asimismo, en la fecha de cierre de esta investigación, se está tramitado la Propuesta de Reglamento para la gobernanza de los datos en la UE[1035], que complementará la Directiva sobre datos abiertos, de junio de 2019[1036].

El objeto y alcance de esta investigación me impiden abordar el análisis de esta Directiva, si bien, la exigencia de autorización de sus fines o por parte de entidades distintas a las contempladas en la misma, podría suponer un obstáculo para el desarrollo y aplicación de la inteligencia artificial, para la innovación y para el desarrollo económico y cultural de la UE, en la medida que puede posicionarla en desventaja competitiva frente a otros países con regímenes más permisivos o laxos, como EE.UU., Japón o China.

4. Inteligencia artificial y patentes

Conforme he anticipado en el anterior apartado, la inteligencia artificial también plantea distintas cuestiones en materia de propiedad industrial, en especial, para el derecho de

[1033] Directiva (UE) 2019/790 del Parlamento Europeo y del Consejo, de 17 de abril de 2019, sobre los derechos de autor y derechos afines en el mercado único digital y por la que se modifican las Directivas 96/9/CE y 2001/29/CE. OJ L 130, 17.5.2019. Pp. 92-125

[1034] El artículo 2.2) de la Directiva (UE) 2019/790 define la minería de textos y datos como toda técnica analítica automatizada destinada a analizar textos y datos en formato digital a fin de generar información que incluye, sin carácter exhaustivo, pautas, tendencias o correlaciones.

[1035] Documento de Trabajo de los Servicios de la Comisión Resumen del Informe de la Evaluación de Impacto que Acompaña al Documento Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la Gobernanza Europea de Datos (Ley de Gobernanza de Datos). 25.11.2020. SWD/2020/296 final.

[1036] Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativa a los datos abiertos y la reutilización de la información del sector público.
PE/28/2019/REV/1. OJ L 172, 26.6.2019. Pp. 56-83.

patentes y del diseño industrial, y similares a las analizadas en relación con la propiedad intelectual, bajo tres ámbitos o dimensiones distintas:

- El registro y protección como patente o modelo de utilidad de los sistemas de inteligencia artificial y los elementos que lo integran, cuestión que ya he analizado en sus aspectos más generales en el anterior apartado.

- El registro y protección de las invenciones creadas por sistemas inteligentes como patentes o modelos de utilidad.

- La protección de los datos y contenidos de los que se nutren los sistemas inteligentes para operar y entrenarse, en especial los protegidos por la propiedad industrial, incluyendo marcos reguladores de patentes y modelos de utilidad, *chips*, diseño industrial o marcas, sobre lo que igualmente he realizado mis consideraciones generales al abordar esta cuestión en el apartado anterior.

Respecto de la primera de las cuestiones, simplemente destacar la necesidad para la protección de una creación del intelecto humano como patente que, junto al carácter de invención y novedad, debe concurrir su carácter técnico como requisito implícito, partiendo de inicio de que los métodos matemáticos y los programas de ordenador quedarían fuera de este ámbito de protección.

De las resoluciones de la Oficina Europea de Patentes emanadas hasta la fecha podría predicarse, conforme destacan algunos autores como Gallego Sánchez y Lievens[1037], que el carácter técnico de la inteligencia artificial podría consistir tanto en el hecho de su aplicación a un campo de la tecnología, como en su adaptación a una implementación técnica específica, lo que podría abrir su posible protección por esta vía.

Mi análisis se centrará en la segunda de estas tres dimensiones previamente expuestas, y que constituye uno de los principales retos que plantea la inteligencia artificial en relación

[1037] GALLEGO SÁNCHEZ, E. (2019). "La patentabilidad de la inteligencia artificial. La compatibilidad con otros sistemas de protección". *La Ley Mercantil*, nº 59, de 1 de junio 2019. Wolters Kluwer 2019. P. 8.

con las invenciones, en particular, la creación de invenciones por sistemas inteligentes con intervención o participación humana, en colaboración o sin intervención humana.

Respecto del primero de los escenarios, considero que la utilización de los sistemas inteligentes por el ser humano como medio o instrumento para crear invenciones susceptibles de protección por esta vía no comporta, *a priori*, mayor consideración que la aplicación del marco vigente en materia de propiedad industrial para su registro y protección a favor del inventor, sin perjuicio de la mayor o menor flexibilidad en este sentido por parte de algunos ordenamientos jurídicos, conforme he expuesto al analizar el apartado anterior.

No obstante, el informe de la UE dando respuesta al *Draft Issues Paper on Intellectual Property and Artificial Intelligence*[1038] de la WIPO de 13 de diciembre de 2019, propone que el documento incluya la patentabilidad de las invenciones asistidas por la IA como invenciones asistidas por ordenador.

Considero pues, que es el segundo de los escenarios propuestos dónde se plantean los retos más relevantes y está generando mayor debate en relación con la inteligencia artificial, por lo que pretendo responder a la siguiente pregunta ¿Se puede reconocer la condición de inventor a un sistema inteligente?

La cuestión ha sido ya planteada a la *Oficina Europea de Patentes* en el denominado "Caso Dabus", aunque los aspectos de fondo que motivaron la misma no fueron abordadas para su resolución, en particular, la relativa a la posible condición de inventor de los sistemas inteligentes.

"Dabus" es el nombre de la inteligencia artificial creada por Stephen Thaler que combina redes neuronales completas con supuesta capacidad de invención.

El supuesto planteado se inició a partir de dos solicitudes de patentes presentadas por la entidad *Artificial Inventor Project* compuesta por un equipo de expertos en patentes que pretendían comprobar la posibilidad de que distintas oficinas de propiedad industrial se

[1038] Disponible en: https://www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/call_for_comments/pdf/org_european_union.pdf. Consultado el 10.03.2021.

pronunciaran expresamente sobre la posible condición de inventor de un sistema inteligente, en este caso, el software denominado "Dabus".

La *Oficina Europea de Patentes* (EPO -por sus siglas en inglés), mediante sendas decisiones de 27 de enero de 2020, y la *United Kingdom Intellectual Property Office* (UKIPO -por sus siglas en inglés) rechazaron la solicitud inicial. Posteriormente también lo ha hecho la *Oficina de Patentes de EE.UU.* -UPSTO por sus siglas en inglés-.

Las Decisiones de la EPO denegaron las solicitudes de patente europea nº EP 18 275 163 y EP 19 275 174, en la que el solicitante alegaba ser en empleador y posteriormente el sucesor de la máquina "Dabus", creadora de la invención sin intervención humana.

En las mismas, la EPO concluyó que los sistemas de inteligencia artificial no pueden ser titulares de derechos en la medida que carecen de personalidad jurídica y de capacidad jurídica y de obrar similar a la de las personas físicas o jurídicas, por lo que los sistemas de inteligencia artificial no pueden ser titulares de derechos derivados de la condición de inventor. Además, consideró que los mismos, al carecer de personalidad jurídica, tampoco pueden ser considerados empleados ni pueden transmitir los derechos de patente. No obstante, no abordó la posible condición de inventor de los sistemas inteligentes, es decir su capacidad para inventar.

La denegación también por la UKIPO fue recurrida y resuelta recientemente, en particular el pasado 21 de septiembre de 2020 por el Tribunal Superior de Inglaterra y Gales (Tribunal de Patentes), confirmando la decisión de la UKIPO y su rechazo.

Sin embargo, en fechas coincidentes con la finalización de esta esta investigación, se ha publicado[1039] la concesión en Sudáfrica de la primera patente del mundo para una invención generada por inteligencia artificial (Dabus), sin un inventor humano tradicional, conformándose como una patente titularidad del propietario del sistema

[1039] Publicado en: https://artificialinventor.com/first-patent-granted-to-the-artificial-inventor-project/.

inteligente con inclusión de la inteligencia artificial que la ideó (Dabus) como el "inventor"[1040].

Y, del mismo modo, en fechas igualmente coetáneas, se ha dictado una Sentencia en primera instancia (no firme y susceptible de apelación) por parte del Tribunal Federal de Australia, de 30.07.2021, en el asunto Thaler v Commissioner of Patents [2021] FCA 879[1041], en la que el juez -no entrando a resolver en esta sede cuestiones de fondo y limitándose a resolver si la presentación de la solicitud fue válida-, resuelve la nulidad de la resolución de archivo dictada por la *Commissioner of Patents* respecto de la solicitud de patente presentada, considerando, en opinión del mismo, que la condición de "inventor" reconocido por la Ley "puede ser un sistema o dispositivo de inteligencia artificial", si bien, "ese inventor no humano no puede ser ni un solicitante de una patente ni un concesionario de una patente[1042]. En el caso de Australia, la legislación local de aplicación no regula y define la condición de "inventor".

En cualquier caso, dejando a un lado estas resoluciones aisladas a nivel internacional, adoptadas conforme a marcos jurídicos específicos y, en cualquier caso, susceptibles de impugnación o recurso, como expuse al analizar distintas cuestiones de propiedad intelectual, ya tenemos experiencias previas de protección de creaciones intelectuales llevadas a cabo por sistemas inteligentes en países como China, como he referido anteriormente. En las resoluciones precitadas, se reconoce la condición de "inventor" al sistema, pero no la de solicitante, concesionario o titular.

El creador[1043] de "Dabus" defiende la condición de inventor de la máquina.

La cuestión a analizar es si un sistema inteligente podría considerarse o tener la condición jurídica de "inventor", al igual que lo hice en el apartado anterior respecto de la condición

---

[1040] Patent Journal. Vol. 54. Nº 07. Julio 2021. P. 255. Recuperado de: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/viewer.html?pdfurl=https%3A%2F%2Fiponline.cipc.co.za%2FPublications%2FPublishedJournals%2FE_Journal_July%25202021%2520Part%25202.pdf&clen=19466154&chunk=true. 5

[1041] Texto íntegro de la resolución publicado en: https://artificialinventor.com/a-federal-court-in-australia-has-held-ai-generated-inventions-are-patentable/.

[1042] Traducción libre del autor de esta investigación. La versión original en inglés se halla disponible en el enlace indicado en la cita anterior.

[1043] ABBOTT, R. (2019) "The Artificial Inventor Project". *WIPO Magazine*. Diciembre 2019. Disponible en: https://www.wipo.int/wipo_magazine/en/2019/06/article_0002.html. Consultado el 10.03.2021.

de "autor", para los marcos jurídicos vigentes, y en el que se evidenció su imposibilidad entre otros aspectos, ante la exigencia de personalidad jurídica y, salvo excepciones, de condición de persona natural. Otra cosa distinta es reconocer el carácter de "creador" al sistema, lo que considero no discutible, que no de "autor".

Para analizar esta cuestión, debemos partir de dos consideraciones.

La primera, de carácter técnico, relativa a las capacidades de las que se halle dotado el sistema inteligente, esto es, si se trata de un sistema de inteligencia artificial "débil" o "fuerte", siendo esto último más un objetivo que una realidad en la actualidad, dado podría suponer un funcionamiento creativo e inventivo del sistema inteligente al margen de toda participación o intervención humana.

La segunda, de carácter jurídico, relativa a determinar si concurren los requisitos exigidos por el marco legal vigente para tener dicha condición.

En primer lugar, la legislación de patentes no establece inicialmente como requisito para reconocer una invención, que ésta proceda del proceso creativo de una persona física, jurídica o de un objeto o ente sin personalidad jurídica, como así destacan distintos autores como Ríos López[1044].

Según esta autora, "las invenciones generadas por un sistema de inteligencia artificial son susceptibles de ser protegidas a través de una patente, pues, según las Directrices de la EPO, la inteligencia artificial se asimila a las invenciones implementadas en ordenador y, aunque los algoritmos como tales, no se pueden proteger, si se integran en un sistema que tenga una aplicación práctica, un uso de medios técnicos o dispositivos, sí se puede proteger".

La Ley de Patentes[1045] española, conforme regula en su artículo 4, únicamente establece como requisitos para que las invenciones sean patentables su novedad mundial, que impliquen actividad inventiva y que sean susceptibles de aplicación industrial. Y su

[1044] RÍOS LÓPEZ, Y. (2020). "Inteligencia artificial y Patentes: ¿Hacia un 'Inventor Artificial'?", en CERRILLO I MARTÍNEZ, A. Y PEQUERA POCH, M. (Coord). "*Retos jurídicos de la inteligencia artificial*". Aranzadi S.A.U. (Thomson Reuters). Navarra 2020. P. 231.
[1045] Ley 24/2015, de 24 de julio, de Patentes. BOE 25.07.2015

artículo 10 establece que el derecho de patente pertenece al inventor o a sus causahabientes, siendo transmisible conforme a Derecho.

Del mismo modo, el precitado artículo contempla que, si la invención hubiere sido realizada por varias personas de forma independiente, el derecho a la patente pertenecerá en común a todas ellas, así como una presunción registral de que el solicitante estará legitimado para ejercer el derecho a la patente.

Como se desprende del tenor literal del precepto, dejando a un lado interpretaciones de la norma en su conjunto y su contexto, no se requiere una condición específica al inventor, al que se le reconoce, en el artículo 14 de la ley precitada, un derecho esencial a ser mencionado en la solicitud y que puede ejercer frente al solicitante y titular de la patente.

Sin embargo, se trataría de un derecho personalísimo, inalienable, irrenunciable que nace con la invención, y es aquí donde se plantea un primer obstáculo al carecer los sistemas inteligentes de personalidad jurídica y de la capacidad de ser sujetos de derechos y obligaciones conforme al ordenamiento jurídico vigente en España y la UE.

En consecuencia, para poder valorar la condición de inventor de un sistema inteligente se precisaría, a mi juicio, de una necesaria reformulación del derecho de patentes vigente para dar cabida a una nueva categoría de "inventor" -sin posibilidad de ejercicio de los derechos asociados a dicha condición-, que recaería en la persona física propietaria, licenciataria o usuaria del mismo, o a la persona que dirige el proceso creativo de la invención. Estas son las cuestiones que se están abordando a nivel internacional en relación con el precitado "Caso Dabus", con resoluciones distintas en base a marcos jurídicos e interpretaciones diferentes.

En cualquier caso, la cuestión adicional a plantear es si debería tratarse de un sistema lo suficientemente autónomo como para ser considerado creador inventivo y si debería concurrir la ausencia de intervención humana, de modo que el sistema no sea un mero instrumento o medio.

Siguiendo la precitada Ríos López, podrían plantearse para su análisis cuatro posibles situaciones respecto de invenciones técnicas generadas de forma autónoma e independiente, sin intervención humana:

a) La patente no designa inventor alguno, lo que es obligatorio conforme al Convenio de Patente Europea.

b) La patente identifica como inventor a una persona natural, que inicialmente debería ser el propietario o licenciatario del sistema inteligente.

c) La patente identifica al inventor artificial (con el riesgo de obtener una resolución similar a las dictadas en el "Caso Dabus"*).*

d) La patente no pudiera ser reconocida por no existir un inventor "válido" según el ordenamiento jurídico vigente.

A mi juicio, la opción adecuada debería ser la c), esto es, identificar al sistema como inventor -al igual que al abordar estas cuestiones en relación con su protección vía propiedad intelectual, planteaba su posible consideración como "creador"- , si bien, para ello sería necesario reformular previamente la condición de "inventor" en los marcos vigentes, la titularidad de los derechos sobre la invención y el monopolio de explotación sobre la misma que, en principio, deberían recaer en el propietario o licenciatario del sistema, que son las personas que han invertido para el desarrollo o adquisición del resultado -no del sistema- con dicho propósito y que podrían considerarse "sucesor" o "causahabiente" en el sentido previsto en el artículo 60 del Convenio de Patente Europea y en el artículo 10 de la Ley de Patentes española.

Obviamente, estas reflexiones podrían chocar frontalmente contra los pilares básicos que sustentan los marcos vigentes de propiedad industrial, evidenciado en los pronunciamientos realizados hasta la fecha en el "Caso Dabus".

Otra opción sería atribuir la condición de inventor al creador del sistema, pero atribuyendo los derechos de explotación a la persona física o jurídica licenciataria del sistema, en caso de generación de la invención bajo la licencia conferida a éste.

Esta opción comentada, sería la más acorde a los requerimientos regulatorios vigentes que exigen que el inventor, creador real de la idea inventiva, conste en el documento de patente, partiendo de una interpretación histórica e integradora de la norma.

En este sentido, me permito significar de nuevo el informe de la UE dando respuesta al *Draft Issues Paper on Intellectual Property and Artificial Intelligence*[1046] de la WIPO de 13 de diciembre de 2019, en el que expresamente se propone que la cuestión de la invención/titularidad incluya cuestiones fundamentales relativas a la identificación de las invenciones generadas o asistidas por inteligencia artificial por parte de las oficinas de PI, la posibilidad de nombrar a una persona jurídica como inventor y las posibles consecuencias para la sociedad de conceder derechos de invención a la inteligencia artificial.

La atribución de la condición de inventor a un ente sin personalidad jurídica al que hacía referencia en anteriores apartados, esto es, un sistema de inteligencia artificial, supuestamente autónomo e independiente, y sin posibilidad de ser titular de derechos y obligaciones, requeriría dicha reformulación, así como explorar y valorar otras vías, por ejemplo, la posible atribución de personalidad jurídica a determinados sistemas inteligentes en función de sus capacidades, conforme han propuesto algunas voces autorizadas y ha sido objeto de análisis anteriormente por mi parte -que no comparto por los motivos allí expuestos- o, de otro, no atribuirle personalidad jurídica, esto es, considerarlo un ente sin personalidad jurídica, pero atribuirle una capacidad jurídica y de obrar limitadas, incluso la condición de "creador" que, de nuevo y en cualquier caso exigiría, en mi opinión, la reformulación anteriormente indicada.

De nuevo y con remisión a mi análisis sobre la personalidad jurídica electrónica de estos sistemas, resultaría discutible y supuestamente inviable la creación de sistemas realmente

[1046] Disponible en: https://www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/call_for_comments/pdf/org_european_union.pdf. Consultado el 10.03.2021.

autónomos y absolutamente independientes por razones éticas, jurídicas y de seguridad -máxime ante los futuros marcos europeos en estos aspectos en tramitación-, y especialmente ante el necesario control y supervisión humana, como así incluso se refleja en el informe de la AIPPI de 14 de febrero de 2020 en respuesta al precitado *Draft Issues Paper on Intellectual Property and Artificial Intelligence* de la WIPO.

Sin embargo, la opción o conclusión final de no reconocer la invención y quedar excluida la misma por la condición de su creador inmediato no parece que sea la más favorecedora para la innovación y el avance tecnológico y, en especial para el desarrollo y aplicación de la inteligencia artificial.

En consecuencia y hasta una revisión de los marcos jurídicos vigentes, a mi juicio, la solución más adecuada, cuando resulte aplicable, será la asociación a la persona física o jurídica detrás del proceso inventivo, considerando al sistema el medio o instrumento creativo utilizado para generar la invención, lo que seguirá sin dar una solución adecuada a las invenciones por parte de sistemas inteligentes sin intervención humana.

5. La propuesta europea.

5.1. Objetivo

La UE dio un paso más para la armonización y protección de los derechos de propiedad intelectual e industrial en su seno, mediante la aprobación, conforme a los establecido en el artículo 118 del Tratado de Funcionamiento de la Unión Europea (TFUE), de la Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial[1047].

El Parlamento Europeo significó en dicha Resolución el avance significativo que está suponiendo la inteligencia artificial y las oportunidades y desafíos inherentes que

[1047] https://www.europarl.europa.eu/doceo/document/TA-9-2020-0277_ES.html

comporta y, en particular, en materia de derechos de propiedad intelectual e industrial, lo que evidencia y justifica, de nuevo, el por qué decidí incorporar estas cuestiones en esta investigación, inicialmente focalizada en los aspectos éticos, jurídicos y de seguridad, su relación y la responsabilidad jurídica derivada de daños causados por los sistemas inteligentes.

De nuevo, la Resolución se produce en el contexto de la incesante preocupación del legislador europeo por competir y liderar la inteligencia artificial a nivel mundial, "recuperando" una soberanía digital e industrial que considera mermada y salvaguardarla, y por ello considera necesario, entre otros esfuerzos, garantizar la competitividad, promover y proteger la innovación y reformar estructuralmente la política industrial de la UE.

En este sentido, el Parlamento Europeo consideró en esta Resolución que el establecimiento de un marco regulador totalmente armonizado en la UE en materia de inteligencia artificial podrá convertirse en una "referencia normativa a escala internacional", apostando por el "reglamento" como instrumento normativo para conseguir este objetivo en lo sucesivo.

El Parlamento Europeo significó la importancia de la técnica legislativa a utilizar para ello, de modo que los nuevos marcos reguladores estén preparados para soportar el rápido ritmo de desarrollo de la inteligencia artificial y ser objeto de seguimiento continuo para su adaptación constante, refrendando así la necesaria evolución del legislador para adaptarse a una realidad en constante cambio sustentada en la tecnología, a la que he hecho distintas referencias a lo largo de esta investigación, que exige crear marcos adaptables, flexibles y evolutivos, o lo que he venido denominando "responsive".

5.2. Seguridad jurídica y confianza

Conforme se recoge en las Resoluciones coetáneas sobre ética y responsabilidad jurídica en materia de inteligencia artificial, objeto de análisis en los anteriores apartados, esta Resolución coetánea destaca también la necesidad de crear un nuevo marco regulador en

estos otros aspectos, que proporcione la seguridad jurídica pretendida que contribuya al desarrollo tecnológico y a generar la necesaria confianza en la fiabilidad y seguridad de la inteligencia artificial, con el objetivo de garantizar el necesario equilibrio entre las salvaguardas públicas, el respeto de otros derechos y libertades fundamentales, el incentivo y fomento de la creatividad, el intercambio y la innovación, así como la protección de la inversión en recursos y esfuerzos en la misma.

En definitiva, marcos que deben proteger de manera eficaz y eficiente a la sociedad en general, así como a los creadores y a las empresas. En este sentido, el Parlamento Europeo instó a la Comisión y a los Estados miembros que ofrecieran su apoyo a las *startups* y a las PYMES a través del Programa sobre el Mercado Único y los centros de innovación digital con la finalidad de proteger sus productos.

5.3. Patentabilidad

La denominada "patente europea" está regulada en el Convenio de Munich sobre Concesión de Patentes Europeas, de 5 de octubre de 1973, en su versión consolidada tras la entrada en vigor del Acta de revisión de 29 de noviembre de 2000.

El Convenio sobre Patentes Europeas tiene su origen en un proyecto de la Comunidad Económica Europea que tenía como finalidad unificar el procedimiento de concesión de patentes en la CEE. El citado Convenio, que entró en vigor con el Tratado de Cooperación en Materia de Patentes, fue revisado en el año 2000 (CPE-2000).

El Convenio es aplicable a todos los Estados de la UE, incluyendo países como Albania, Islandia, Liechtenstein, Antiguo República Yugoslava de Macedonia, Mónaco, Noruega, San Marino, Serbia, Suiza y Turquía.

Además de este Convenio, se debe tomar en consideración el llamado Tratado de Cooperación en Materia de Patentes (en lo sucesivo, PCT por sus siglas en inglés), firmado en Washington en 1970 y que entró en vigor en 1978, de forma casi coetánea con el Convenio sobre Patente Europea. En realidad, el PCT no concede patentes, sino que regula un procedimiento único para su solicitud y tramitación en los diversos Estados.

Por ello, una vez finalizado el procedimiento previsto en el PCT, el solicitante debe traducir y presentar la solicitud en cada uno de los países en que desee la protección para que se conceda o deniegue la patente según las concretas leyes nacionales.

Los modelos y algoritmos computacionales en los que se basa la inteligencia artificial, calificados como métodos matemáticos conforme al precitado Convenio sobre la Patente Europea, no pueden patentarse como tales, conforme abordé en los apartados precedentes, salvo que se utilicen con fines técnicos en el contexto de inventos técnicos. Y los programas de ordenador tampoco son inicialmente patentables, por lo que su protección jurídica se contempla a través del derecho de autor.

La Resolución objeto de análisis considera que la inteligencia artificial y las tecnologías conexas se basan en *software*, esto es, en la creación y en la ejecución de programas informáticos, los cuales, conforme a su naturaleza, están sujetos a un régimen de protección de derechos de autor y, por consiguiente, no son patentables.

No obstante, el Parlamento europeo consideró expresamente en esta Resolución que los métodos matemáticos y los programas informáticos podrán ser protegidos mediante patentes en virtud del artículo 52.3 del *Convenio sobre la Patente Europea* cuando se utilicen como parte de un sistema de inteligencia que contribuya a producir un efecto técnico suplementario, lo que abre la puerta a su potencial de protección como patente, que el legislador europeo considera necesario evaluar con detalle.

Conforme igualmente recogió la Resolución objeto de análisis, la protección de la inteligencia artificial como patente exigirá el cumplimiento los requisitos establecidos por los marcos reguladores vigentes, especialmente actividad inventiva y novedad mundial, pero también debe exigir una descripción exhaustiva de la tecnología subyacente, lo que en la práctica puede suponer dificultades en determinados sistemas de inteligencia artificial, especialmente por la complejidad de sus razonamientos.

En cualquier caso, el legislador europeo propuso clarificar el marco jurídico para la protección de la innovación relacionada con el desarrollo de la inteligencia artificial y las tecnologías conexas, así como sobre la aplicación de los derechos de propiedad intelectual e industrial a los materiales, contenidos y datos generados por estas tecnologías, lo que

es una necesidad a la vista de las conclusiones que he expuesto, tras los análisis y reflexiones realizadas.

En este sentido, el Parlamento Europeo significó la necesidad de diferenciar entre las creaciones humanas con la ayuda de la inteligencia artificial y las creaciones generadas por la inteligencia artificial de forma autónoma, conforme ha sido objeto de tratamiento y análisis en los apartados precedentes.

5.4. Creación de obras con ayuda o por sistemas de inteligencia artificial

El Parlamento Europeo dejó clara su postura, de nuevo, respecto de la ausencia de personalidad jurídica de los sistemas de inteligencia artificial.

Como he referido anteriormente, el legislativo europeo puso su foco en la necesaria diferenciación entre creaciones humanas asistidas por la inteligencia artificial y las creaciones generadas por la inteligencia artificial, siendo estas últimas las que plantean nuevos retos jurídicos en materia de protección de los derechos de propiedad intelectual e industrial, en especial, no sólo sobre la titularidad, la condición de inventor o la remuneración, sino también otras relacionadas con la posible concentración del mercado.

Asimismo, el Parlamento Europeo significó la necesidad de distinguir entre los derechos de propiedad intelectual e industrial para el desarrollo de tecnologías o sistemas de inteligencia artificial, esto es, para la protección de sistemas, licencias de desarrollo o sobre *frameworks*, y los derechos concedidos a creaciones generadas mediante inteligencia artificial.

En el caso de que los sistemas de inteligencia artificial se usen exclusivamente como herramienta para ayudar a un "autor-inventor" en el proceso de creación, conforme he analizado en los anteriores apartados, el marco de derechos de propiedad intelectual e industrial aplicable sería el actual, correspondiendo a éste, tanto los morales como los de explotación.

En este sentido, el Parlamento Europeo considera que las creaciones técnicas generadas mediante sistemas de inteligencia artificial, esto es, como medio o instrumento, deben ser igualmente protegidas por el marco de jurídico de los derechos de propiedad intelectual e industrial con el objetivo de fomentar la inversión en esta forma de creación y la seguridad jurídica ciudadanos, empresas e inventores, siendo éstos últimos los principales usuarios de estos sistemas.

El reto surge cuando dichos sistemas son capaces de crear una obra de manera "autónoma".

El Parlamento Europeo abordó y aclaró esta cuestión en la línea de las conclusiones expuestas anteriormente, considerando que estas obras no deben poder acogerse a la protección dispensada por los derechos de autor, en la medida que debe respetarse el principio de originalidad vinculado a la persona física y el hecho de que el concepto de creación intelectual comporta la personalidad de su autor.

No obstante, ya previó que, en caso de estimarse que estas obras pudiesen ser susceptibles de acogerse a la protección mediante de derechos de autor en el futuro, insta a la Comisión que lo aborde jurídicamente desde un enfoque horizontal, basado en pruebas y neutro a nivel tecnológico.

Asimismo, el Parlamento Europeo abordó otra de las cuestiones clave objeto de análisis, reflexión y conclusión en los apartados precedentes, recomendando que la titularidad de los derechos, en su caso, se atribuya exclusivamente a las personas físicas o jurídicas que crearon la obra.

5.5. Evaluaciones de impacto

Las instituciones europeas son muy conscientes de los retos que plantea la inteligencia artificial para los marcos actuales de propiedad intelectual e industrial, y no sólo como objeto de los mismos, como medio o instrumento para su generación o como sistema de tratamiento de contenidos protegidos por los mismos para su operación o entrenamiento,

sino también como instrumento para la gestión, respeto y protección de los derechos de propiedad intelectual e industrial.

En este sentido, el Parlamento Europeo instó a que se lleve a cabo una evaluación de impacto con un objeto y alcance específicos respecto a la protección de los derechos de propiedad intelectual e industrial en el contexto del desarrollo de la inteligencia artificial.

Recomendó una evaluación sectorial y por tipo de implicaciones en estos derechos, pero bajo un enfoque que tuviera en consideración aspectos como el grado de intervención humana, la autonomía de la inteligencia artificial, la importancia del papel y el origen de los datos y el material protegido por derechos de autor utilizados y la posible participación de otros elementos relevantes.

El grado de intervención humana y la autonomía efectiva, como he expuesto al analizar estos aspectos en los anteriores apartados, son esenciales para determinar la adecuación de los marcos jurídicos actuales para la protección de las creaciones llevadas a cabo por sistemas inteligentes.

No obstante, en mi opinión, sin perjuicio de los contextos específicos, debe partirse de un concepto global de sistema de inteligencia artificial para abordar todas estas cuestiones, en la medida que, como sistema, puede integrar *software*, algoritmos, bases de datos y datos, pero también *hardware* de alto nivel y precisión creado *ad hoc*, el cual puede ser determinante para la creación artística o inventiva a nivel físico, por ejemplo hologramas, esculturas 3D o pinturas, y, en consecuencia, especialmente determinante por la relevancia que éste y las personas físicas o jurídicas detrás de él tengan para la obtención del resultado creativo o inventivo, su protección y la titularidad de los derechos sobre el mismo.

Estas evaluaciones, a juicio del Parlamento Europeo, deberán focalizarse en el impacto e implicaciones de la inteligencia artificial y las tecnologías conexas en el actual régimen en materia de derecho de patentes, protección de marcas, dibujos y modelos, derechos de autor y derechos afines, y en especial, en la protección jurídica de las bases de datos y los programas informáticos, así como en la protección de los conocimientos técnicos y la

información empresarial no divulgada (secretos comerciales) contra su obtención, utilización y revelación ilícitas.

Por último, el Parlamento Europeo significó el potencial de la inteligencia artificial también como medio para mejorar el respeto y protección de los derechos de propiedad intelectual e industrial, a lo que me permito adicionar no sólo su valor preventivo y detectivo, sino correctivo y de investigación, tanto para recabar y preservar la prueba, como para identificar la autoría, ubicación y perseguibilidad con instrumentos eficaces.

5.6. Ingeniería inversa

La Resolución destacó igualmente la preocupación por los retos jurídicos que plantea de la ingeniería inversa como excepción a la protección de los derechos de autor de los programas informáticos y la protección de los secretos comerciales, que son esenciales para la innovación y la investigación, por lo que deberán ser considerados en el desarrollo de los sistemas de inteligencia artificial, lo que significa incorporar mecanismos de protección, seguridad y cumplimiento de los mismos en su diseño (*Security by design*).

5.7. Libre acceso

El Parlamento Europeo incorporó en su Resolución uno de los aspectos más importantes de la misma en relación con el acceso libre a la innovación en materia de inteligencia artificial, considerando que las tecnologías de inteligencia artificial deben ser de libre acceso para fines educativos y de investigación, como métodos de aprendizaje más efectivos.

Y debo significar que "libre" no significa inicialmente "gratuito", ni tampoco que el Parlamento apueste exclusivamente por la construcción de sistemas sobre *software* de código libre y *open source* en contraposición con el denominado *software* propietario. Lo que se pretende es la accesibilidad al conocimiento y la tecnología para determinados

usos, que puede ser compatible con las distintas modalidades de licencia y explotación del *software* y sistemas de información.

Considero ésta una cuestión que requerirá más reflexiones en lo sucesivo en la medida que, entre otros aspectos, puede ser un obstáculo a la inversión privada en función del modelo y, por consiguiente, a la innovación y mejora, salvo que se incentive por otros cauces, sin perjuicio de que comparta totalmente esta filosofía y modelo en el contexto educativo y de la investigación.

No obstante, en el apartado 15 "in fine" de la Resolución, el Parlamento destaca la importancia de facilitar el acceso a los datos y de compartir los datos, de las normas abiertas y de las tecnologías basadas en fuentes abiertas, a la vez que se fomenta la inversión y se impulsa la innovación, de manera alineada con las estrategias europeas de gobierno de los datos.

5.8. Inteligencia artificial para proteger derechos y luchar contra los *deep fakes*

Conforme expuse al analizar los principales riesgos y retos que plantea la inteligencia artificial a nivel ético, el Parlamento Europeo significó especialmente los retos que también plantea la misma para los derechos de propiedad intelectual e industrial en relación con la creación de *deep fakes* y su preocupación ante la posibilidad de manipulación de la ciudadanía para desestabilizar democracias.

Por ello, el Parlamento Europeo exige un aumento de la sensibilización y formación mediática, así como la aceleración del uso de la propia inteligencia artificial para luchar contra los mismos, mediante la verificación de hechos o informaciones.

En este sentido, todos estos aspectos han sido objeto de regulación específica en la posterior Propuesta de Reglamento del Parlamento Europeo y del Consejo, de 21 de abril de 2021[1048] sobre inteligencia artificial.

[1048] COM (2021) 206 final 2021/0106 (COD)

Del mismo modo, destaca la utilidad de los sistemas de inteligencia artificial para el rastreo del uso de obra protegidas por derechos de autor y mejorar su protección, que constituye uno de los ámbitos y dimensiones en los que mayor impacto puede tener la inteligencia artificial, como he destacado en los anteriores apartados.

5.9. Requisitos de la inteligencia artificial

La Resolución acaba recordando y significando los principales requisitos que deben reunir los sistemas de inteligencia artificial como transparencia, supervisión humana, respeto de requisitos relativos a tecnología de código abierto en contratación pública y de la interconectividad de los servicios digitales y no discriminación.

Con esta última finalidad, el Parlamento Europeo consideró que, cuanto mayor sea el volumen posible de datos no personales disponibles para el entrenamiento y aprendizaje automático, más se podrá reducir el riesgo de sesgo, por lo que insta a la Comisión a que reflexione sobre el uso de datos de dominio público con estos fines y destaca la importancia de la aplicación efectiva en su integridad de la *Estrategia para el Mercado Único Digital* para la mejora de la accesibilidad e interoperabilidad de los datos no personales en la UE.

No obstante, conforme el Parlamento Europeo significó igualmente, la *Estrategia Europea de Datos* debe garantizar el equilibrio entre el fomento del intercambio de datos, la protección de los derechos de propiedad intelectual e industrial y de los secretos comerciales, así como la protección de la intimidad y los datos personales.

El Parlamento Europeo avaló en su Resolución la creación de un espacio único europeo de datos pero que debe considerar el marco de protección en la UE en materia de propiedad intelectual, así como la información sobre el uso de datos protegidos por derechos de propiedad intelectual e industrial, especialmente en el marco de las relaciones entre plataformas y empresas.

Asimismo, puso en valor el análisis que está llevando a cabo la Comisión para adoptar medidas legislativas en materia de datos no personales, la posible revisión de la Directiva

sobre las bases de datos y una posible aclaración de la aplicación de la Directiva relativa a la protección de los secretos comerciales.

6. Responsabilidades relacionadas con la creatividad, las invenciones y la innovación

6.1. Responsabilidad en el ámbito de la propiedad intelectual

La vulneración de derechos de propiedad intelectual puede generar no sólo una responsabilidad civil sino también responsabilidad penal en los casos más graves conforme a los artículos 270 a 272 del Código Penal español, tal y como expuse en el capítulo VI, sin perjuicio en determinados contextos, de la posibilidad de sanción administrativa.

La responsabilidad civil en el ámbito de la propiedad intelectual se regula en el Libro Tercero de la LPI bajo el título "De la protección de los derechos reconocidos en esta Ley" y, en especial en sus artículos 138 a 143 (Título Primero), reguladores de las acciones y procedimientos.

Los preceptos indicados confieren distintos derechos al titular de los derechos sobre cualquier creación protegida conforme a la LPI que constituya o integre un sistema inteligente o que sea utilizado por el mismo.

El artículo 138 de la LPI española establece que se podrá instar el cese de la actividad infractora o ilícita, así como exigir la indemnización de los daños materiales y morales causados en los términos previstos en los artículos 139 y 140.

La cesación de actividad podrá incluir la suspensión de la explotación o actividad infractora, la remoción o el precinto de instrumentos para facilitar la supresión o neutralización no autorizada de cualquier dispositivo técnico de protección de obras o prestaciones, incluyendo el *software*, aunque no fuera su único uso, e incluso la destrucción de estos instrumentos en el caso de protección de programas de ordenador.

Adicionalmente, se podrá instar la publicación o difusión, total o parcial, de la resolución judicial o arbitral en medios de comunicación a costa del infractor, así como la adopción de medidas cautelares para asegurar la efectividad de sus derechos conforme a lo establecido en el artículo 141, que pueden incluir, entre otras, intervención y depósito de ingresos, suspensión de actividad de reproducción, distribución y comunicación pública, secuestro de material empleado así como de los instrumentos, dispositivos y productos utilizados en determinados contextos.

El precepto citado también considera responsable de la infracción a quien induzca a sabiendas la conducta infractora, quien coopere con la misma -siendo conocedor de la conducta infractora o contando con indicios razonables para conocerla-, y quien teniendo un interés económico directo en los resultados de la conducta infractora, cuente con una capacidad de control sobre la conducta del infractor.

En todos estos supuestos, el sistema inteligente podrá vulnerar los derechos precitados, de lo que podría derivarse una responsabilidad, no respecto del mismo, ante la ausencia de personalidad jurídica y de capacidad para ser sujeto de derechos y obligaciones, sino a alguno de los agentes relacionados con su diseño, desarrollo, funcionamiento, entrenamiento o uso, conforme al contexto específico.

De este modo, en el caso de sistema de inteligencia artificial que integre capacidad de autoaprendizaje y que se interaccione con su entorno, que pueda recopilar contenidos protegidos mediante derechos de propiedad intelectual sin tener predefinida limitación alguna, por ejemplo a través de Internet, y que los utilice para su aprendizaje y entrenamiento para generar los resultados pretendidos conforme a la finalidad para la que fue concebido, la responsabilidad debería recaer inicialmente en la esfera del fabricante/productor, por ser quien tiene la capacidad de control, así como en el proveedor. Pero también podría recaer en el propietario, el operador y/o el usuario, en la medida que participa y se beneficia de todo ello, máxime si además utiliza conscientemente herramientas y utilidades incorporadas en el interfaz del sistema con esta finalidad.

Sin embargo, los contextos concretos pueden ser muy variados pero inabordables en el marco del objeto y alcance limitados de esta investigación y enfoque horizontal.

Asimismo, los contenidos pueden ser facilitados por el propio operador o usuario o predefinir los mismos para su búsqueda e incorporación por el sistema inteligente, en cuyo caso, la responsabilidad recaería en la esfera de éste, en la medida que las herramientas del sistema inteligente no hayan sido creadas con esta exclusiva finalidad.

Otros de los supuestos será el de los entrenadores del sistema que son los que pueden incorporar los contenidos protegidos al mismo, en cuyo caso, deberán responder contractual o extracontractualmente en caso de contenidos con infracción de derechos, para lo que deberán incorporarse las correspondientes cláusulas de salvaguarda en los contratos de servicios de entrenamiento y formación o de suministro de contenidos por parte de los fabricantes/productores del sistema -o propietarios/licenciatarios- y los mismos.

De conformidad con lo previsto en el artículo 140 de la LPI, el titular del derecho infringido tendrá derecho a una indemnización por los daños patrimoniales y morales sufridos, que comprenderá no sólo el valor de la pérdida que haya sufrido, sino también el de la ganancia que haya dejado de obtener a causa de la violación de su derecho. La indemnización podrá incluir, en su caso, los gastos de investigación en los que se haya incurrido para obtener pruebas razonables de la comisión de la infracción objeto del procedimiento judicial, lo que en el caso de sistemas inteligentes pueden resultar de una cuantía elevada.

La indemnización por daños y perjuicios se fijará, a elección de la persona perjudicada, conforme a alguno de los siguientes criterios:

a) Las consecuencias económicas negativas, entre ellas la pérdida de beneficios que haya sufrido la persona perjudicada y los beneficios que el infractor haya obtenido por la utilización ilícita. En caso de daño moral, procederá su indemnización, aunque no se haya probado la existencia de perjuicio económico, para cuya valoración se atenderá a las circunstancias de la infracción, gravedad de la lesión y grado de difusión ilícita de la obra protegida.

b) La cantidad que hubiera percibido el perjudicado como remuneración, en su caso, si el infractor hubiera pedido autorización para utilizar el derecho de propiedad intelectual en cuestión.

Por lo que se refiere a la prescripción de la acción para reclamar los daños y perjuicios, de conformidad con lo previsto en el precitado artículo 140.3 prescribirá a los cinco años desde que la persona legitimada pudo ejercitarla.

Por último, significar que en determinados contextos legalmente previstos pueden derivarse responsabilidades administrativas a instancias de la Comisión de Propiedad Intelectual, al amparo del artículo 195 de la LPI, el Real Decreto 1889/2011, de 30 de diciembre, por el que se regula el funcionamiento de la Comisión de Propiedad Intelectual[1049] y otras normas relacionadas. El incumplimiento de requerimientos de retirada de contenidos declarados infractores puede dar lugar a sanciones de entres 150.001 hasta 600.000 euros.

6.2. Responsabilidad en el ámbito de la propiedad industrial

Del mismo modo que pueden derivarse determinadas responsabilidades en materia de propiedad intelectual, en el ámbito de la propiedad industrial también pueden exigirse no sólo responsabilidades penales derivadas de los delitos en estas materias junto con las responsabilidades civiles derivadas de los mismos, conforme hice referencia al abordar el capítulo VI, sino también responsabilidades civiles, la cuales se hayan reguladas en los artículos 40, siguientes y concordantes de la Ley 17/2001, de 7 de diciembre, de Marcas, en los artículos 71, siguientes y concordantes de la Ley 24/2015, de 24 de julio, de Patentes, en los artículos 52, siguientes y concordantes de la Ley 20/2003, de 7 de julio, de Protección Jurídica del Diseño, y ellos artículos 721, siguientes y concordantes de la Ley 1/2000, de 7 de enero, de Enjuiciamiento Civil española.

[1049] BOE 31.12.2011

6.3. Responsabilidad en el ámbito del secreto de empresa

Las innovaciones, información o conocimiento, incluido el tecnológico, científico, industrial, comercial, organizativo o financiero calificable como secreto de conformidad con lo previsto en la precitada Ley 1/2019, de 20 de febrero, de Secretos Empresariales[1050] y no protegido por los marcos anteriores, se haya protegido por esta norma. En este marco, podrían también protegerse los algoritmos.

No obstante, la violación de los secretos de empresa también se haya contemplada en el Código Penal[1051] y podría ser constitutiva de delito en determinados contextos.

En este sentido, la utilización de contenidos protegidos por este marco con infracción de la obligación de secreto regulada en el mismo por o mediante sistemas inteligentes, siempre y cuando intervenga dolo o culpa, podrá causar daños efectivos, que deberán ser indemnizados conforme a la misma.

El artículo 4 de la *Ley de Secretos Empresariales* considera el secreto empresarial como objeto del derecho de propiedad y transmisible.

La misma regula específicamente la defensa y acciones en caso de violación de los secretos empresariales en sus artículo 8 y 9, que puedan incluir, entre otras pretensiones, la declaración de la violación del secreto empresarial, la cesación o prohibición de los actos de violación, la prohibición de fabricar, ofrecer, comercializar o utilizar mercancías infractoras o de su importación, exportación o almacenamiento con dichos fines, remoción, aprehensión de mercancías, atribución de la propiedad sobre las mismas, publicación o difusión de la sentencia y, especialmente, la indemnización de los daños y perjuicios, si ha intervenido dolo o culpa del infractor conforme regula sus artículos 9.1.g) y 10.

[1050] BOE 21.02.2019
[1051] Artículo 278 y siguientes del Código Penal

Los preceptos citados establecen dicha responsabilidad siempre que concurra culpa o dolo, estableciendo que la indemnización será adecuada en relación con la lesión realmente sufrida como consecuencia de la violación.

El artículo 10 de la precitada norma regula el cálculo de dichos daños y perjuicios. En particular establece que, al fijarse la indemnización de los mismos, se tendrán en cuenta todos los factores pertinentes, como son los perjuicios económicos, incluido el lucro cesante, que haya sufrido el titular del secreto empresarial, el enriquecimiento injusto obtenido por el infractor y, cuando proceda, otros elementos que no sean de orden económico, como el perjuicio moral causado al titular del secreto empresarial por su obtención, utilización o revelación ilícitas. También podrán incluirse, en su caso, los gastos de investigación en los que se haya incurrido para obtener pruebas razonables de la comisión de la infracción objeto del procedimiento judicial.

El precepto contempla de manera alternativa, la fijación, según los casos, de una cantidad a tanto alzado en concepto de indemnización de daños y perjuicios, atendiendo, al menos y entre otros aspectos, al importe que la parte demandada habría tenido que pagar al titular del secreto empresarial por la concesión de una licencia que le hubiera permitido utilizarlo durante el período en el que su utilización podría haberse prohibido.

En relación con el cálculo y liquidación de los daños y perjuicios, la Ley de Secretos Empresariales se remite a lo dispuesto en el artículo 73 de la Ley de Patentes.

Los litigios civiles para la depuración de las responsabilidades conforme a esta normativa, se resolverán por los jueces y tribunales del orden jurisdiccional civil y según conforme corresponda conforme a la precitada Ley de Enjuiciamiento Civil española, correspondiendo la competencia al Juzgado de lo Mercantil correspondiente al domicilio del demandado o, a elección del demandante, el Juzgado de lo Mercantil de la provincia donde se hubiera realizado la infracción o se hubieran producido sus efectos, conforme regula sus artículos 12 y 14.

La legitimación activa corresponderá al titular del secreto empresarial y sus licenciatarios legítimos.

La norma también prevé distintas previsiones especiales en materia de diligencia de comprobación de hechos, acceso a fuentes de prueba, su aseguramiento y medidas cautelares, de indudable interés y relevancia, especialmente en caso de hallarse involucrado un sistema inteligente.

Las acciones de defensa de los secretos empresariales prescribirán por el transcurso de tres años desde el momento en que la persona legitimada tuvo conocimiento de la persona que realizó la violación del secreto empresarial.

7. Consideraciones finales

El automatismo y/o autonomía de la que pueda estar dotado un sistema de inteligencia artificial puede permitir la creación de obras e invenciones susceptibles de protección como patentes, modelos de utilidad, diseño industrial o derechos de autor con o sin intervención humana.

En relación con esto último, la posibilidad de que un sistema de inteligencia artificial puede asistir al autor para la creación de nuevas obras o, incluso, crear por sí mismo y de forma "autónoma" obras originales, exige una reflexión profunda sobre el marco vigente en España y Europa sobre derechos de autor y afines.

En mi opinión, la utilización de sistemas de inteligencia artificial como asistente, medio o herramienta para la creación de nuevas obras e invenciones por parte de su usuario, no supone un nuevo escenario no contemplado por los marcos vigentes en materia de propiedad intelectual e industrial, con ciertas particularidades.

Sin embargo, la creación "autónoma" por parte de un sistema de inteligencia artificial de una obra o invención nueva -artificial, algorítmica o sintética- puede plantear múltiples cuestiones en relación con su protección, autoría y titularidad de derechos, que no pueden ser resueltas adecuadamente por los marcos jurídicos vigentes.

El derecho de autor fue concebido y construido para la protección de creaciones generadas directa o indirectamente por seres humanos y los sistemas inteligentes más avanzados

carecen de dicha condición y personalidad jurídica alguna -no pudiendo ser titulares de derechos y obligaciones-. Lo mismo ocurre con el derecho de patentes.

Por ello, la protección de creaciones así como de invenciones por sistemas inteligentes requeriría, de considerarse ésta la vía la adecuada para su protección, una revisión de los aspectos esenciales sobre los que se sustenta el derecho de autor así como el de patentes, para la inclusión expresa de la protección de las creaciones e invenciones llevadas a cabo por sistemas inteligentes sin intervención o aportación humana (o no significativa), con la definición de un nuevo marco que revise la condición de autor o inventor, derechos asociados, su titularidad y contenido, y su específico marco de protección, restringido y posiblemente más acotado en el tiempo.

A nivel internacional, la mayoría de ordenamientos jurídicos, incluyendo el español, considera que las creaciones protegibles por derecho de autor deben proceder del intelecto de una persona física, sin perjuicio de las ficciones legales abordadas en las que se atribuye dicha autoría y titularidad de derechos asociados a terceros, ya sean personas físicas o jurídicas.

En materia de propiedad industrial y, en particular, en relación con patentes, modelos de utilidad y diseños industriales, he extraído similares conclusiones sobre la necesidad de revisar los marcos jurídicos actuales para poder brindar una adecuada protección a las nuevas realidades por este cauce, incluyendo la posible revisión de la condición de "inventor".

El derecho de patentes debe adaptarse también a los nuevos paradigmas que ya supone la inteligencia artificial y los que comportará su desarrollo y aplicación al ámbito de las invenciones.

En breve si no ya, nos podemos encontrar con creaciones humanas diseñadas por sistemas inteligentes y creaciones artificiales diseñadas por seres humanos. La interacción es bidireccional, la fusión incuestionable.

Las creaciones artificiales, algorítmicas o sintéticas, esto es, creaciones llevadas a cabo por sistemas inteligentes sin intervención humana, podrían ser protegidas por distintas

vías como he analizado, pero inicialmente no tendrían su acomodo en los marcos reguladores vigentes en España y en la UE en materia de derechos de autor.

La revisión de los marcos vigentes requerirá una reflexión y evaluación previa del impacto a nivel económico, social, moral y también jurídico.

Hasta que dispongamos de un nuevo marco, la creación de sistemas inteligentes, de los elementos que lo integren como el *software*, algoritmos, *hardware*, chips, datos y bases de datos, así como las creaciones resultantes de la utilización de estos sistemas como medio o instrumento por los seres humanos para llevar a cabo dicha creación se seguirán protegiendo por los marcos normativos vigentes de propiedad intelectual e industrial, protección de datos, imagen, competencia desleal o secretos empresariales.

Asimismo, los datos de los que se nutren los sistemas para su aprendizaje seguirán estando igualmente protegidos salvo para las finalidades y utilización en los sectores y por las entidades recogidas en los marcos regulativos reguladores de la minería de textos y de datos, y cuya infracción en su uso comportará las responsabilidades correspondientes.

Y, por último, respecto de las creaciones artificiales o sintéticas por sistemas inteligentes avanzados sin intervención ni participación humana, podrán igualmente ser protegidos por la vía técnica y contractual, sin perjuicio de su sujeción y protección por la vía del derecho de propiedad y, en función de su naturaleza y contenido, por los demás marcos reguladores precitados.

La reflexión adicional es valorar la posibilidad de adoptar una postura conservadora para mantener los marcos vigentes con pequeñas adaptaciones, lo que considero inviable para dar cabida a las nuevas realidades, en la medida que habría que abordar las bases y pilares esenciales sobre los que se fundamenta el derecho de autor y de patentes a nivel nacional, europeo e internacional, o bien crear un nuevo marco específico que regule una nueva categoría de “autores”, “creadores”, artistas, intérpretes, ejecutantes, productores o inventores “artificiales” o “sintéticos”, reiterando que siempre que estos marcos se consideren la vía más adecuada para su protección, aspecto que constituye un nuevo hilo que debería ser objeto de una investigación específica que excede del objeto y alcance de esta investigación. En definitiva, sería necesaria una profunda reforma legislativa del

marco regulador vigente, de sus bases y fundamentos esenciales, especialmente sobre el propio concepto de autor o, como he expuesto y a mi juicio, del concepto de creador.

Las propuestas regulatorias en este sentido, deberían también revisar los conceptos de obra, régimen de protección de estas creaciones la titularidad, derechos y contenido de los mismos, así como incluso los conceptos de creador, autor, inventor, artista, intérprete, ejecutante, editor o productor.

Capítulo IX

Conclusiones

Mis reflexiones, consideraciones, posicionamiento, argumentos y conclusiones sobre cada uno de los aspectos abordados a lo largo de esta investigación han sido expuestas a lo largo de su tratamiento específico, incorporando adicionalmente unas consideraciones finales en cada uno de los capítulos que integran la misma, a las que me remito.

Sin perjuicio de ello, como conclusiones generales y finales, me gustaría significar las siguientes:

PRIMERA.- La inteligencia artificial está impulsando cambios disruptivos en el mundo en que vivimos que no pudieron ser previstos por los ordenamientos jurídicos vigentes en el momento en que fueron concebidos. Además, se trata de un conjunto de tecnologías en constante cambio y evolución, tanto en su desarrollo y capacidades como en su aplicación.

El desarrollo y aplicación de la inteligencia artificial en todos los ámbitos y sectores de nuestra vida avanza a un ritmo vertiginoso y exponencial, creando nuevas realidades para las que el Derecho vigente no tiene todas las respuestas y soluciones.

SEGUNDA.- La inteligencia artificial y el conjunto de tecnologías con las que puede interactuar y operar constituyen el principal medio para el desarrollo de nuestra sociedad y la economía digital, para el crecimiento industrial, así como para garantizar nuestra salud, la mejora de nuestra calidad de vida, la sostenibilidad y la consecución del frecuentemente "utópico" bien común, y posiblemente sea el único medio para garantizar nuestra propia supervivencia y existencia, en un futuro que contrarreste el egoísmo, individualidad y estupidez innata que también puede llegar a caracterizar al ser humano.

TERCERA.- El Derecho vigente no regula la inteligencia artificial ni permite proporcionar respuestas adecuadas a los distintos retos que plantea y a las cuestiones que pueden suscitarse en relación con la misma.

Conforme al análisis efectuado, nos encontramos ante una realidad compleja, en constante cambio y distinta de la realidad sobre la que se construyeron los marcos jurídicos vigentes, y ante una previsible realidad futura todavía más compleja y alejada de éstos, por lo que aplicar el marco actual a la variedad de situaciones que se plantean y plantearán en la práctica relacionados con sistemas de inteligencia artificial constituye una tarea arduamente complicada y, en muchos casos, simplemente imposible.

Los marcos reguladores de la responsabilidad extracontractual y de responsabilidad por productos defectuosos en España y la UE no pueden proporcionar respuestas adecuadas a la totalidad de contextos y situaciones que se pueden plantear en relación con daños ocasionados por el funcionamiento o uso de sistemas inteligentes.

Los marcos reguladores de la seguridad de productos y sistemas en España y la UE tampoco contemplan ni regulan específicamente estos sistemas.

Y tampoco los marcos reguladores de la propiedad intelectual e industrial permiten atribuir la condición de autor o inventor a un sistema inteligente ni permiten resolver la totalidad de situaciones que se pueden plantear en relación con creaciones algorítmicas o mediante sistemas inteligentes.

En definitiva, deben revisarse las instituciones actuales para su adaptación a un nuevo contexto.

CUARTA.- El Derecho debe ordenar y regular una realidad tan compleja como la inteligencia artificial en aquello que sea necesario, así como los riesgos potenciales que puede suponer para los principales bienes y derechos cuya protección constituye el principal objetivo de los ordenamientos jurídicos, y ello de manera proactiva y no meramente reactiva.

QUINTA.- El Derecho precisa nuevos enfoques y técnicas jurídicas para regular esta realidad y ofrecer soluciones a los nuevos problemas, legislando desde una visión estratégica y un enfoque global, dinámico, adaptativo, flexible, evolutivo e integrador, de modo que prevea su adecuación, revisión y actualización constante a la realidad que pretende regular.

El Derecho que rige nuestras vidas, relaciones e interacciones debe incorporar una regulación avanzada adecuada al contexto actual e inminente, y no quedar anclado en instituciones, fundamentos y realidades pasadas, cuya interpretación forzada pretende dar solución a nuevas necesidades y problemas que no existían en el momento de su concepción.

El Derecho debe contemplar realidades, regular relaciones y ofrecer soluciones a los conflictos que se puedan suscitar, garantizando un marco equilibrado, que considere todos los intereses en juego y que proporcione seguridad para todos los operadores que integran estas nuevas realidades, y tanto a diseñadores, desarrolladores, fabricantes, desplegadores, comercializadores, proveedores y operadores, como a los usuarios de sistemas de inteligencia artificial, promoviendo, facilitando y garantizando, de un lado, la innovación y el desarrollo tecnológico sostenible y basado en valores, necesidades y finalidades humanas y, de otro, garantizando la seguridad, los derechos y el bienestar presente y futuro del ser humano.

De nuevo, si una de las manifestaciones de la inteligencia es la capacidad de adaptarse a los cambios, el Derecho y su creador debe ser más inteligente que nunca para asegurar los principios, valores, derechos y libertades por los que nos regimos, mediante nuevos marcos acordes a los nuevos contextos, nuevas necesidades, nuevas formas de organización, nuevas formas de relación, interacción y resolución de conflictos, y desde una visión global.

La inteligencia artificial puede suponer un gran poder en manos de quien lo gestione y más grande todavía si tiene acceso a los datos de los que se nutre. Ello comporta una gran responsabilidad y la necesaria reflexión sobre si el Derecho debe anticiparse (o al menos intentarlo) a la tecnología y prever su aplicación y evolución futura en base a su potencial y posibilidades, especialmente en su relación e interacción con datos y otras tecnologías.

La definición previa de conceptos como “inteligencia artificial” o “autonomía” es esencial para acometer su regulación, especialmente ante la necesidad de acotar y condicionar esta última.

La actual inteligencia artificial considerada “débil” podría evolucionar hacia una inteligencia artificial más avanzada y “fuerte”, con mayor autonomía, capacidades y posible impredecibilidad, que debe ser adecuadamente regulada y limitada hoy, especialmente por los niveles de riesgo inherente asociado a sus propias capacidades y características.

Una mera reflexión general sobre lo que tecnologías como la computación cuántica podrán suponer de llegar a ser una realidad, considero que no podemos ni cuestionarnos ya la necesidad de que gobiernos y legisladores se adelanten a la realidad.

¿Qué ocurrirá cuando la supremacía cuántica sea una realidad y pueda estar en manos privadas, y probablemente en unas pocas? ¿Cómo limitarlo? ¿Cómo compartirlo? ¿Y la inteligencia artificial ante una capacidad de computación cuántica que pueda ser generada o gestionada por la misma? ¿Quién responde si quién controla esta tecnología tiene el poder de acceder, suspender, alterar o destruir con una sola acción un gobierno, una economía, un sistema crítico, una red de comunicaciones o toda una región mundial?

SEXTA.- El Derecho y la tecnología deben hallarse alineados con los objetivos humanos para conseguir el futuro que pretendamos.

Durante los últimos años hemos presenciado una deshumanización de la sociedad correlativa a la instauración de la tecnología como medio de relación e interacción social. Sin embargo, especialmente en los últimos años parece que algunos han empezado a ser conscientes de dicha deshumanización para volver a centrarse en la persona y en la satisfacción de las necesidades y deseos de la misma, y ello a pesar de la creciente digitalización de la sociedad, forzada en los últimos meses por la pandemia mundial que estalló en 2020 y contra la que hoy luchamos con más ciencia que tecnología, y con más reactividad que proactividad.

SÉPTIMA.- El Derecho debe garantizar también el libre desarrollo de la personalidad, la identidad humana y los derechos fundamentales e impedir que se reduzca la autonomía personal.

El mundo va a seguir cambiando y cada vez a un ritmo más rápido y exponencial gracias al desarrollo y, sobre todo, la aplicación incesante de las distintas tecnologías en todos los ámbitos, cuyas capacidades en constante aumento y potenciadas por su interacción, están provocando lo que algunos denominan "tsunami digital", que derribará muros hasta ahora infranqueables y que afectará y desestabilizará los pilares éticos y jurídicos que sostienen y garantizan el Estado de Derecho y los derechos y libertades de las personas.

Ello nos obliga a cambiar y además a una velocidad nunca vista. Debemos ser más adaptativos que nunca si queremos mantener nuestros principios, valores, derechos y libertades. Y ello exige estrategia, inteligencia y planificación para que no se deriven efectos negativos más allá de los inherentes y necesarios y así explotar todas sus ventajas.

OCTAVA.- El papel de los gobiernos y la cooperación internacional serán determinantes para abordar y gestionar adecuadamente las bondades, retos y riesgos que plantea la inteligencia artificial, mediante enfoques globales y actuaciones oportunas en todos los ámbitos.

El valor que supone para el ser humano la tecnología es incuestionable, si bien, comporta retos y riesgos que deben ser adecuadamente identificados y gestionados mediante el diseño de estrategias y planes de actuación, que no sólo deben conllevar actuaciones en el ámbito político sino jurídico, económico y social.

La demora en estas actuaciones, lejos de contribuir a la evitación o mitigación de estos riesgos puede potenciar el impacto de los mismos tanto cualitativa como cuantitativamente, en la medida que el avance tecnológico es imparable y estas actuaciones deben acompañarlo al mismo ritmo en crecimiento constante con el que se produce.

NOVENA.- La ética, la seguridad y el Derecho se encuentran inevitablemente unidos, especialmente al abordar la responsabilidad jurídica derivada del funcionamiento y uso de realidades tan complejas y de tan alto impacto para el ser humano como la inteligencia artificial.

La ética, la seguridad y el Derecho conforman la base para el desarrollo y aplicación de la inteligencia artificial y el desarrollo tecnológico en general, asegurando su confiabilidad, así como la innovación y la competitividad para la creación de productos y servicios seguros que garanticen el cumplimiento ético y jurídico y el respeto de los principios y derechos fundamentales.

Así se recoge expresamente en las citadas *Directrices éticas para una inteligencia artificial* fiable, elaboradas por el *Grupo de expertos de alto nivel sobre inteligencia artificial*, que se sustentan en tres pilares fundamentales, esto es, licitud, ética y seguridad que, como he referido, constituyen las tres dimensiones desde las que he abordado esta investigación para llegar a los aspectos de la responsabilidad por daños causados por o mediante la inteligencia artificial.

No puede concebirse una tecnología transformadora, una sociedad, una economía o el propio Derecho sin una ética. La ética conforma la base de todo ello sobre la que se deben construir los marcos regulatorios, las sociedades y las economías.

La seguridad, el cumplimiento del marco regulador y la responsabilidad son exigencias éticas de la inteligencia artificial.

Cualquier propuesta reguladora futura nacional o europea debería partir de las directrices éticas precitadas, objeto de un pretendido consenso internacional, cualquiera que sea el nivel preliminar de riesgo asociado al sistema inteligente, sin perjuicio de diferenciar, de un lado, los requerimientos éticos y jurídicos esenciales exigibles a cualquier sistema y, de otro, los verticales y específicos para determinados sistemas en función de sus capacidades y características, nivel de riesgo, sector y usos potenciales, especialmente los razonablemente previsibles, debidos o indebidos.

La ética debería ser exigible y vinculante en sus principios y valores esenciales, no sólo corporativa o contractualmente, mediante la exigencia de adhesión y cumplimiento de determinados marcos o códigos, sino jurídicamente.

La ética exige a su vez la seguridad, constituyendo elementos que integran la diligencia debida contractual y extracontractual, cuya ausencia puede conllevar el nacimiento de

responsabilidades jurídicas. La ética y la seguridad contribuyen a conformar el marco obligacional y la diligencia debida, junto con los marcos legales y contractuales.

DÉCIMA.- Necesidad de un marco jurídico global e integrado, proactivo y reactivo, que contemple los principios y normas éticas esenciales, la seguridad y la protección de los derechos fundamentales.

Conforme he expuesto, considero necesario y urgente la construcción de un marco jurídico global, regulador de los principios, normas y obligaciones éticas y jurídicas que deben regir el diseño, desarrollo, capacidades, despliegue, aplicación, uso y mantenimiento de la inteligencia artificial durante todo su ciclo de vida, así como el régimen de responsabilidad y mecanismos de respuesta ante daños, que garanticen la confianza y seguridad para todas las partes implicadas y asegure la constante innovación y competencia en inteligencia artificial a nivel mundial.

Las resoluciones y propuestas del Parlamento Europeo de octubre de 2020 reflejan un necesario proceso de reflexión y maduración previa sobre la inteligencia artificial y un profundo análisis y discusión sobre sus aspectos clave en el ámbito ético, jurídico, de propiedad intelectual y de seguridad.

El Parlamento Europeo evidenció en las mismas su firme convicción sobre la necesidad de un nuevo marco regulador que contemple las obligaciones éticas y de seguridad, así como de responsabilidad civil, apostando definitivamente en este último aspecto por una responsabilidad objetiva para los sistemas de inteligencia artificial de alto riesgo, manteniendo una responsabilidad subjetiva para el resto.

La posterior Propuesta de Reglamento de 21 de abril de 2021 pretende erigirse como el marco regulador de la inteligencia artificial en el seno de la UE y norma de referencia a nivel internacional, si bien, su contenido se aleja de una regulación completa de la misma. La misma aborda la inteligencia artificial de manera parcial, en congruencia con la opción legislativa previamente valorada, en la medida que se focaliza en la denominada inteligencia artificial “débil” y, en especial, en la prohibición de determinados sistemas inteligentes de riesgo inadmisible y en la regulación muy detallada y técnica de los sistemas inteligentes calificados conforme a la misma como de alto riesgo, no regulando

el resto, salvo determinadas obligaciones de transparencia para sistemas específicos, para los que opta por la creación de códigos de conducta de adscripción voluntaria.

UNDÉCIMA.- El ordenamiento jurídico vigente en España y la UE imposibilita reconocer personalidad jurídica a los sistemas de inteligencia artificial más avanzados en este momento, cuestión que se planteó especialmente como posible opción para resolver las cuestiones de responsabilidad, pero que considero superada y alejada de la agenda política europea ante el posicionamiento al respecto de las instituciones europeas.

De manera consecuente y junto con el resto de motivos y fundamentos expuestos en mi investigación, no puede imputarse responsabilidad civil, penal o administrativa a un sistema inteligente.

DUODÉCIMA.- La inteligencia artificial deberá ser regulada desde un enfoque global y bajo técnicas legislativas distintas y adaptadas a las nuevas realidades, mediante la combinación adecuada del precitado *hard* y *soft law* e instrumentos jurídicos dinámicos, adaptativos, evolutivos, *responsive*, basados en la ética, la seguridad y el respeto a los principios constitucionales y derechos fundamentales.

En base a los retos jurídicos y riesgos potenciales que comporta, del análisis llevado a cabo de los marcos jurídicos actuales, de las propuestas en tramitación y de cómo deberían ser los futuros marcos reguladores de una realidad tan compleja y de tan alto impacto potencial para el ser humano en todos los ámbitos, considero que debemos ya plantearnos la posibilidad de crear un Derecho de la inteligencia artificial que aborde la misma hoy y su evolución en el futuro, de ahí el título de esta investigación.

De hecho, las distintas propuestas analizadas a lo largo de esta investigación evidencian ya, en mi opinión, la existencia de un derecho embrionario en fase de diseño.

Hasta la disposición de marcos jurídicos específicos, este derecho lo estamos construyendo en la práctica a nivel corporativo -mediante la creación de políticas y marcos, estándares, códigos éticos y de conducta corporativos de obligatoria adhesión, ya sea interna y/o externa-, a nivel contractual -mediante la adecuada regulación de los derechos, obligaciones, responsabilidades, garantías, seguridad, marcos de referencia,

acuerdos de nivel de servicio y otros aspectos que configuran y regulan las relaciones jurídicas vinculantes para las partes participantes en un contrato y/o exigencia de adhesión a determinados estándares de terceros-, así como a nivel sectorial -mediante la creación de códigos éticos, códigos de buenas prácticas o de autorregulación-.

Bibliografía

ABBOTT, R. (2019). "The Artificial Inventor Project". *WIPO Magazine*. Diciembre 2019. Disponible en: https://www.wipo.int/wipo_magazine/en/2019/06/article_0002.html. Consultado el 10.03.2021.

ADSUARA, A. (2014). *De otro(s) mundo (s)*. Editado por Sendemá Editorial y Escuela Superior de Arte y Tecnología (ESAT). Valencia.

ALARCÓN, G. (2014). "El soft law y el sistema de fuentes*"*. *Tratado sobre la Ley General Tributaria: Homenaje a Álvaro Rodríguez*. Vol. 1, Tomo I. Editorial Aranzadi Thomson Reuters.

ÁLVAREZ OLALLA, P. (2019). "Responsabilidad Civil en la circulación de vehículos autónomos", en Monterrosso Casado, E. (Dir.), *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch 2019.

AMAT LLOMBART, P. (2020). "La protección de las personas físicas en relación al tratamiento de sus datos personales. Condiciones para el ejercicio de sus derechos y en el marco de la normativa comunitaria y española", en Bello Janeiro, D. (Coord.) *Nuevas tecnologías y responsabilidad civil*. Editorial Reus 2020.

ANDRAE, A. (2017). *Total Consumer Power Consumption Forecast*. Nordic Digital Business Summit.

ANDREW, J. W. (1997). "From Video Games to Artificial Intelligence: Assigning Copyright to Works generated by increasingly sophisticated Computer Programs". *AIPLA Quarterly Journal*. Vol. 25, nº1.

ARNTZ, M.; GREGORY, T. Y ZIERAHN, U. (2016). "The risk of automation for jobs in OECD countries: A comparative analysis". *Employment and Migration Working Papers*, nº 189. OECD Social. Disponible en: https://doi.org/10.1787/5jlz9h56dvq7-en. Consultado el 25.02.2021.

ARUTE, F.; ARYA, K.; BABBUSH, R. ET AL. (2019). "Quantum supremacy using a programmable superconducting processor". *Nature*, n.º 574. Disponible en: https://doi.org/10.1038/s41586-019-1666-5

ASARO, P. (2007). "Robots and responsibility from a legal perspective". *Proceedings of the IEEE conference of robotics and automation; Workshop on roboethics*. IEEE. Roma. Recuperado de https://peterasaro.org/writing/ASARO%20Legal%20Perspective.pdf. Consultado el 11.01.2021

ATIENZA NAVARRO, Mª. L. (2020). "Una aproximación a la Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a la responsabilidad civil por el funcionamiento de los sistemas de inteligencia artificial de 27 de abril de 2020". *Revista Aranzadi de derecho y nuevas tecnologías*. ISSN 1696-0351, Nº. 54.

AYERBE, A. (2020). *La ciberseguridad y su relación con la inteligencia artificial*. ARI 128/2020. Real Instituto Elcano.

BADILLO ARIAS, J.A. (2019). "Responsabilidad civil y aseguramiento obligatorio de los robots". En MONTERROSSO CASADO, E. (Dir.). *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch.

BALLARÍN, P. (2021). *Desafíos de la ciberseguridad en la inteligencia artificial (IA). Análisis del informe de ENISA*. ODISEIA.

BARLETT, R.; MORSE, A.; STANTON, R. Y WALLACE, N. (2019). *Consumer-Lending Discrimination in the FinTech Era*. National Bureau of Economic Research.

BARREDO, A.; DÍAZ-RODRÍGUEZ. N. ET AL. (2017), "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsable AI". *Information Fusion*. Science Direct. Vol. 58. Elsevier.

BARRIO ANDRÉS, M. (2018). *Derecho de los Robots*. Wolters Kluwer España, S.A. Madrid. (Edición Kindle).

BARRIO ANDRÉS, M. (2020). *Manual de Derecho Digital*. Tirant lo Blanch. Valencia.

BAUZÁ REILLY, M. (1994). "Responsabilidad civil en materia informática". *Actualidad Informática Aranzadi*, n.º 12.

BELLMAN, R. (1978). *An Introduction to Artificial Intelligence: Can Computers Think?*. Boyd & Fraser Publishing Company.

BENJAMINS, R.; BARBADO, A. Y SIERRA, D. (2019). "Responsible AI by Design in Practice*". Proceedings of the Human-Centered AI: Trustworthiness of AI Models & Data (HAI) track at AAAI Fall Symposium*. DC.

BENJAMINS, R. (2020). "Los líderes mundiales de la inteligencia artificial ética". *Thing Big / Empresas*. 08.01.2020. Disponible en: https://empresas.blogthinkbig.com/los-lideres-mundiales-de-la-inteligencia-artificial-etica/. Consultado el 24.02.2021

BENJAMINS, R. Y SALAZAR, I. (2020) *El mito del algoritmo*. Editorial Anaya Multimedia. Madrid.

BERCOVITZ, R. (2017). *Comentarios a la Ley de Propiedad Intelectual*. 4.ª Ed. Editorial Tecnos.

BERCOVITZ RODRÍGUEZ-CANO, R. (1992). "Artículo 28". En Bercovitz Rodríguez-Cano, R; Salas Hernández, J. y Bercovitz Rodríguez-Cano. A. *Comentarios a la Ley General para la defensa de los consumidores y usuarios*. Civitas. Madrid. 1992.

BERRIMAN, R. Y HAWKSWORTH, J. (2017). *Will robots steal our jobs? The Potential Impact of Automation on the UK and other major economies*. PWC. Disponible en: https://www.pwc.co.uk/economic-services/ukeo/pwcukeo-section-4-automation-march-2017-v2.pdf. Consultado el 25.02.2021.

BERTOLINI, A. (2013). "Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules*". Law, Innovation and Technology* 5 (2). DOI: http://dx.doi.org/10.5235/17579961.5.2.214

BIJKER, W.; HUGHES, T. Y PINCH, T. (1987). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. Cambridge: MIT Press.

BILLEAUD, J. Y KRISHER, T. (2020). "Se acusa al conductor de respaldo en el accidente autónomo mortal de Uber en Arizona" *Associated Press*. Publicado el 16 de septiembre de 2020. Recuperado de: https://apnews.com/article/homicide-arizona-transportation-archive-phoenix-fdd1574ac6a3c418d4f2b569b797dc16. Consultado el 28.12.2021.

BLACKISTON, D. ET AL. (2021). "A cellular platform for the development of synthetic living machines". *Science Robotics,* vol. 6, edición 52, eabf1571. DOI: 10.1126 / scirobotics.abf1571. Disponible en: https://robotics.sciencemag.org/content/6/52/eabf1571.Consultado el 31.03.2021.

BOBBIO, N. Y RUIZ DE MIGUEL, A. (1990). *Contribución a la teoría del derecho*. Editorial Debate.

BODEN, M. A. (2009). *La mente creativa: Mitos y mecanismos*. Gedisa. Barcelona.

BOND, T. Y BLAIR, S. (2019). "Artificial Intelligence & copyright: Section 9(3) or authorship without an author". *Journal of Intellectual Property Law & Practice*. Vol. 14. Nº 6. Oxford University Press. 2019.

BONFANTI, M. E. (2020). *Artificial Intelligence and Cybersecurity: A Promising but Uncertain Future*. Real Instituto Elcano. ARI 139/2020.

BONFANTI, M.E. Y KOHLER, K. (2020). "Artificial Intelligence for Cybersecurity". *CSS Analyses in Security Policy,* nº 265. ETH Zurich.

BONNEFON, J.F.; SHARIFF A. Y RAHWAN, I. (2016). "El dilema social de los vehículos autónomos". *Science*. Vol. 352, n.º 6293.

BOSTROM, N. Y YUDKOWSKY, E. (2011). "The Ethics of Artificial Intelligence", en RAMSEY, W. Y FRANKISH, K. (2011). *Cambridge Handbook of Artificial Intelligence*.

Cambridge University Press. Recuperado de: http://www.nickbostrom.com/ethics/artificial-intelligence.pdf. Consultado el 20.02.2021

BOSTROM, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press. Oxford.

BOXBYTE. (2012). "El origen de: El Cómputo en la Nube". *FayerWayer*, 06.01.2012. Recuperado de https://www.fayerwayer.com/2012/01/el-origen-de-el-computo-en-la-nube/. Consultado el 24.11.2020

BROZEK, B. Y JAKUBIEC, M. (2017). "On the legal responsibility of autonomous machines". *Artificial Intelligence and Law*, vol. 25, n.º 3. DOI 10.1007/s10506-017-9207-8.

BRUNDAGE, M. ET AL. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. Oxford University. Recuperado de: https://maliciousaireport.com/. Consultado el 18.02.2021

BRYNJOLFSSON, E.; MITCHELL, T. Y ROCK, D. (2018). "What can machines learn, and what does it mean for occupations and the economy?". *AEA Papers and Proceedings*, vol. 108.

BRYSON, J.J.; DIAMANTIS, M.E. Y GRANT, T.D. (2017). "Of, for, and by the people: the legal lacuna of synthetic persons". *Artif Intell Law*.

BUCHANAN, B. (2020). *A National Security Research Agenda for Cybersecurity and Artificial Intelligence.* Center for Security and Emerging Technology -CSET-.

BUCHHOLZ, R. A. Y ROSENTHAL, S. B. (2002). "Technology and Business: Rethinking the Moral Dilemma*". Journal of Business Ethics*. N.º 41.

BULLEMORE, V.R. Y MACKINNON, J. (2004). "Fin y Función del Derecho Penal y de la pena: las teorías de la pena". *Anales Facultad de Derecho Universidad de Chile*. N.º 1.

BUOLAMWINI, J. Y GEBRU, T. (2018). *Gender shades: Intersectional accuracy disparities in commercial gender classification*. Proceedings of the 1st Conference on Fairness, Accountability and Transparency. *PMLR*. Vol. 81.

BUOLAMWINI, J. Y GEBRU, T. (2018). "Proceedings of the 1st Conference on Fairness, Accountability and Transparency". *PMLR*. Vol. 81.

CALO, R. (2016). "La robótica y las lecciones del derecho cibernético". *Revista de privacidad y derecho digital*. N.º 2.

CAMACHO CLAVIJO, S. (2017). "La subjetividad "cyborg". En NAVAS NAVARRO, S. (Dir.). *Inteligencia artificial, tecnología, derecho*. Editorial Tirant lo Blanch. Valencia

CARBAJOSA, A. Y PÉREZ, J. (2020). "Ciberataque a un hospital alemán en tiempos de pandemia". *El País*, 04.10.2020. Recuperado de: https://elpais.com/internacional/2020-10-03/ciberataque-a-un-hospital-aleman-en-tiempos-de-pandemia.html. Consultado el 16.02.2021.

CARRASCOSA LÓPEZ, V. (2000). *La contratación informática: El nuevo horizonte contractual*. Editorial Comares.

CASTÁN TOBEÑAS, J. (1988). *Derecho civil español, común y foral*. 14a. Ed., Editorial Reus, Madrid, T.1V.

CASTELLÓN, J. Y LÓPEZ, M.M. (2017). "Crisis y ciberespacio: Hacia un modelo integral de respuesta en el Sistemas de Seguridad Nacional". Cuadernos de Estrategia 185. *Ciberseguridad: La cooperación público-privada*. Instituto Español de Estudios Estratégicos. Departamento de Seguridad Nacional (DSN). Ministerio de Defensa.

CASTRILLÓN, O.D.; RODRÍGUEZ, M. Y LEYTON, J.D. (2008). "Ética e inteligencia artificial ¿Necesidad o urgencia?". *International Institute of Informatics and Systemics*. Recuperado de: http://www.iiis.org/CDs2008/CD2008CSC/CISCI2008/PapersPdf/C054TM.pdf. Consultado el 2.03.2021.

CERRILLO I MARTÍNEZ, A. Y PEQUERA POCH, M. (Coord). (2020). *Retos jurídicos de la inteligencia artificial*. Aranzadi S.A.U. (Thomson Reuters). Navarra.

CHARNIAK, E. Y MCDERMOTT, D. (1985). *Introduction to Artificial Intelligence*. Addison Wesley.

CHEN, B.X. Y METZ, C. (2019). "Google's Duplex Uses A.I. to Mimic Humans (Sometimes)". *The New York Times*, 22.05.2019. Disponible en: https://www.nytimes.com/2019/05/22/technology/personaltech/ai-google-duplex.html. Consultado el 23.02.2021.

COHN, G. (2018). "El arte de IA en Christie's se vende por $ 432,500". *The New York Times*, 25.10.2018. Disponible en: https://www.nytimes.com/2018/10/25/arts/design/ai-art-sold-christies.html. Consultado el 19.02.2021.

COOMBS, T. (2018), "Artificial Intelligence & Cybersecurity for Dummies". *IBM*. Disponible en: https://hosteddocs.ittoolbox.com/ai_cybersecurity_dummies.pdf

CORTINA, A. (2015). *Ética mínima. Introducción a la filosofía práctica*. Editorial Tecnos, 17ª Edición. Madrid.

COTINO, L. (2020). "SyRI, ¿a quién sanciono? 'Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público y la sentencia holandesa de febrero de 2020'". *La Ley Privacidad*, n.º 4. Wolters Kluwer.

CUESTA AGUADO, P.M. (2020). "Inteligencia artificial y responsabilidad penal". *Revista penal México*. ISSN 2007-4700. Nº. 16-17, 2019-2020.

DARÍO VERGEL, S. (1994). "Responsabilidad civil derivada de la informática". *Revista Informática y Derecho*, vol. 4. Aranzadi-UNED. Mérida.

DELCKER, J. (2018). "Europe divided over robot 'personhood'". *Politico.eu*, 11.04.2018. Disponible en: https://www.politico.eu/article/europe-divided-over-robot-ai-artificial-intelligence-personhood/. Consultado el 21.01.2021.

DENICOLA, R. (2016). "Ex Machina: Copyright Protection for Computer-Generated Works". *Rutgers University Law Review*. Vol. 69:251.

DIAZ ALABART, S. (2018). Conferencia "Robótica y Responsabilidad Civil". Real Academia de Jurisprudencia y Legislación. 31.05.2018.

DÍAZ-LIMÓN, J.A. (2016). "*Daddy's car:* La inteligencia artificial como herramienta facilitadora de derechos de autor". *Revista La Propiedad Inmaterial*, n.º 22. Universidad Externado de Colombia. DOI: http://dx.doi.org/10.18601/16571959.n22.06.

DÍEZ-PICAZO, L. (2011). *Fundamentos del Derecho Civil Patrimonial. V. La Responsabilidad Civil Extracontractual*. Editorial Civitas-Thomson Reuters. Navarra.

DOMÍNGUEZ PECO, E.M. (2018). "Los Robots en el Derecho Penal", en BARRIO ANDRÉS, M. (Director). *Derecho de los Robots*. Wolkers Kluwer. Madrid.

DREYFUS, H. (1972). *What computers can't do: The limits of artificial intelligence*. Editorial HarperCollins. Londres.

DUNCAN, J. (2016). "Robots and computers will commit more crime than humans by 2040, expert warns". Publicado en *Mailonline*. 08 de septiembre de 2016. Recuperado de: https://www.dailymail.co.uk/news/article-3780314/Robots-computers-commit-crime-humans-2040-expert-warns.html. Consultado el 15.12.2020.

EBERS, M. (2016). "La utilización de agentes electrónicos inteligentes en el tráfico jurídico jurídico; ¿Necesitamos reglas especiales en el Derecho de la responsabilidad civil?". *Indret: Revista para el Análisis del Derecho*. N.º 3.

ENCABO VERA, M.A. (2012). *Derechos de la personalidad*. Marcial Pons. Madrid.

ERCILLA GARCÍA, J. (2018). "Aproximación a una Personalidad Jurídica Específica para los robots". *Revista Aranzadi de derecho y nuevas tecnologías*. N.º 47.

ERCILLA GARCÍA, J. (2018). *Normas de derecho civil y robótica. Robots inteligentes, personalidad jurídica, responsabilidad civil y regulación*. (Primera edición, 2018). Thomson Reuters Aranzadi.

EYKHOLT, K.; EVTIMOV, I. ET AL. (2018). "Robust Physical-World Attacks on Deep Learning Visual Classification*". CVPR*, 10.04.2018. Disponible en: https://arxiv.org/pdf/1707.08945.pdf.

FEATHERSTONE, M. (2000). "Velocidad y violencia: sacrificio en Virilio, Derrida y Girard". Publicado en *Anthropoethics: The Journal of Generative Anthropology*. Antropoética VI. Nº 2. Disponible en: http://anthropoetics.ucla.edu/ap0602/virilio/. Consultado el 20.12.2020.

FERNÁNDEZ, J. (2020). "Los ciberataques amenazan con colapsar los hospitales: 'Sería terrorífico'". *Expansión*, el 24.10.2020. Recuperado de: https://www.expansion.com/economia-digital/companias/2020/10/24/5f915917e5fdea64298b45e1.html. Consultado el 17.02.2021.

FERRER, I. (2021). "El Gobierno holandés dimite en bloque por el escándalo en las ayudas al cuidado de los niños". *El País*, 16.01.2021.

FINLAYSON, S. G., BOWERS, J. ET AL. (2019). "Adversarial attacks on medical machine learning". *Science Journal*, vol. 363, Issue 6433. DOI: 10.1126/science.aaw4399

FJELD, J.; HILLIGOSS, H; ACHTEN, N.; DANIEL, M.L. ET AL. (2019). *Principled artificial intelligence: a Map of Ethical and Rights-Based Approaches*. Berkman Klein Center. Recuperado de: https://ai-hr.cyber.harvard.edu/images/primp-viz.pdf. Consultado el 14.02.2021

FJELD, J. ET ALT. (2020). *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI*. Berkman Klein Center for Internet & Society.

FOOT, P. (1967). "The Problem of Abortion and the Doctrine of the Double Effect". *Oxford Review*, n.º 5. Recuperado de: http://pitt.edu/~mthompso/readings/foot.pdf.

FORD, M. (2015). *Rise of the Robots*. Basic Books. Nueva York.

FRAZONI, M. (1988). *Culpa presenta e responsabilitá del debitore*. Cedán, Padova.

FREY, C. B. Y OSBORNE, M.A. (2013). "The future of employment: How susceptible are jobs to automation". *Oxford Martin Programme on Technology and Employment*, 17.09.2013. Disponible en https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf?link=mktw. Consultado el 25.02.2021.

FUSSELL, S. (2019). "Consumer Surveillance Enters Its Bargaining Phase". *The Atlantic*. Disponible en: www.theatlantic.com/technology/archive/2019/06/alexa-googleincognito-mode-not-real-privacy/590734. Consultado el 02.03.2021.

GALLEGO SÁNCHEZ, E. (2019). "La patentabilidad de la inteligencia artificial. La compatibilidad con otros sistemas de protección". *La Ley Mercantil*, n.º 59. Wolters Kluwer.

GALYARDT A. ET AL (2019). *Artificial Intelligence and Cyber Intelligence: An Implementation Guide*. Carnegie Mellon University. Recuperado de: https://resources.sei.cmu.edu/asset_files/EducationalMaterial/2019_011_001_548767.pdf. Consultado el 04.02.2021.

GARCÍA MEXÍA, P.L. (2016). "Lex robótica y derecho digital". *Revista de privacidad y derecho digital*. N.º 2.

GARCÍA SÁNCHEZ, M. D. (2020). "Inteligencia artificial y oportunidad de creación de una personalidad electrónica". *Ius et Scientiam: Revista Electrónica de Derecho y Ciencia*, vol. 6, n.º 2. ISSN 2444-8478. Editorial Universidad de Sevilla.

GARCÍA TERUEL, R.M. (2021). "El derecho de daños ante la inteligencia artificial y el machine learning: una aproximación desde las recomendaciones del Parlamento Europeo

y del Grupo de Expertos de la Comisión Europea", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón*. Tomo II. Ed. Aranzadi-Thomson Reuters. 2021.

GARCÍA VARELA, R. (2008). "La responsabilidad por hecho ajeno". *Diario La Ley*, n.º 6874, Sección Columna, Ref. D-27. Editorial La Ley.

GARCÍA-PRIETO, J. (2018). "¿Qué es un robot?". En BARRIO ANDRÉS, M. (Dir.). *Derecho de los robots*. Wolters Kluwer, Madrid.

GARDNER, H. (2004). *Frames of Mind: The Theory of Multiple Intelligences*. Nueva York: Basic Books.

GERVAIS, D.J. (2020). "The machine as author". *Iowa Law Review*, 105 Rev.2053. Disponible: https://ilr.law.uiowa.edu/assets/Uploads/ILR-105-5-Gervais.pdf. Consultado el 10.03.2021

GEIGER, C. Y PENIN, J. (Eds). (2018). "Intellectual Property and Digital Trade in the Age of Artificial Intelligence and Big data*"*. *CEIPI-ICTSD*. Issue Number 5.

GIL SALDAÑA, M. (2008). *El producto sanitario defectuoso en Derecho español*. Editorial Atelier. Barcelona. 2008.

GINSBURG, J. C. (2003). "The Concept of Authorship in Comparative Copyright Law". *Depaul Law* Review. Vol. 52. N.º 3, Issue 4, Summer 2003.

GINSBURG, J. Y TREPPOZ, E. (2015). *International Copyright Law, U.S. and E.U. Perspectives*. Elgar. Cheltenham-Northampton.

GLAVESKI, S. (2018). "El caso de la jornada laboral de 6 horas". *Harvard Business Review*, 11.12.2018. Disponible en: https://hbr.org/2018/12/the-case-for-the-6-hour-workday?language=es. Consultado el 21.02.2021.

GÓMEZ ÁGREDA, A.; MOLINO, J y otros (2019). *Usos militares de la inteligencia artificial, la automatización y la robótica*. Instituto Español de Estudios Estratégicos (ieee.es).

GÓMEZ CALLE, F. (2008) "Los sujetos de la responsabilidad civil: La responsabilidad por hecho ajeno". En REGLERO CAMPOS, L.F. (Coord.). *Tratado de Responsabilidad Civil.* Vol. I. 2008 (Parte General). Tomo I.

GÓMEZ SALADO, M.A (2018). "Robótica, empleo y seguridad social. La cotización de los robots para salvar el actual estado de bienestar". *Revista Internacional y Comparada de Relaciones Laborales y Derecho del Empleo*. Vol. 6, n.º 3, ISSN-e 2282-2313.

GÓMEZ-RIESCO, J. (2018). "Los robots y la responsabilidad civil". En Barrio Andrés, M. (Coord.), *Derecho de los Robots*. Wolkers Kluwer.

GONZÁLEZ CUSSAC, J.L. (2021) "Sobre el fundamento de la responsabilidad penal de las personas jurídicas", en GALÁN MUÑOZ, A. Y MENDOZA CALDERÓN, S. (Coord.). *Derecho penal y política criminal en tiempos convulsos: libro homenaje a la Profa. Dra. Maria Isabel Martínez González*. Tirant lo Blanch.

GONZÁLEZ GRANADO, J. (2016). *Derecho y Robots en la Unión Europea: Hacia una persona electrónica*. Taller de derechos. Disponible en: https://tallerdederechos.com/derecho-y-robots-en-la-union-europea-hacia-una-persona-electronica/ Consultado el 30 de diciembre de 2020

GONZÁLEZ, S.E. (2020). "La inteligencia de las sumas y las restas". *Inspiring Blog Tecnalia*, 23.07.2020. Disponible en: http://blogs.tecnalia.com/inspiring-blog/2020/07/23/la-inteligencia-las-sumas-restas/ Consultado el 02.01.2021.

GOTTFREDSON, L.S. (1994). "Mainstream Science on Intelligence". *Wall Street Journal*, 13.12.1994, ISSN: 0160-2896. Disponible en: http://www1.udel.edu/educ/gottfredson/reprints/1997mainstream.pdf. Consultado el 05.01.2021.

GRATTON, L. Y SCOTT, A. J. (2016). *The 100-Year Life: Living and Working in an Age of Longevity*. Bloomsbury Publishing. London.

GUIMÓN, P. (2017). "Un ciberataque paraliza 16 hospitales de Reino Unido y les exige dinero". *El País*, 12.05.2017. Disponible en:

https://elpais.com/tecnologia/2017/05/12/actualidad/1494602389_458942.html. Consultado el 16.02.2021.

GUTIÉRREZ SANTIAGO, P. (2012). "El 'daño' en la responsabilidad civil por productos defectuosos (Régimen Jurídico de sus clases, cobertura y limitaciones en la legislación de consumo española, a la luz del cuarto Informe de la Comisión Europea, de 8 de septiembre de 2011, sobre la Directiva 85/374/CEE)". *Diario La Ley*, n.º 7859, Sección Doctrina, Ref. D-201, Editorial La Ley.

HAGE, J. (2017). "Theoretical foundations for the responsability of autonomous agents". *Artif Intentell Law*, 31 de agosto de 2017.

KAPLAN, A. Y HAENLEIN M. (2018). *Siri, Siri, in my Hand, who's the Fairest in the Land? On the Interpretations, Illustrations and Implications of Artificial Intelligence.* Business Horizons.

HARWELL, D. (2018). "Google to drop Pentagon AI contract after employee objections to the 'business of war'". *The Washington Post*, 01.06.2018. Recuperado de: https://www.washingtonpost.com/news/the-switch/wp/2018/06/01/google-to-drop-pentagon-ai-contract-after-employees-called-it-the-business-of-war/. Consultado el 18.02.2021.

HAUGELAND, J. (1985) *Artificial Intelligence: The Very Idea*. MIT Press. Cambridge.

HERN, A. (2019). "New AI fake text generator may be too dangerous to release, say creators". *The Guardian*, 14.02.2019. Disponible en: https://www.theguardian.com/technology/2019/feb/14/elon-musk-backed-ai-writes-convincing-news-fiction. Consultado el 23.02.2021.

JANOFSKY, A. (2018). "AI Could Make Cyberattacks More Dangerous, Harder to Detect". *The Wall Street Journal*, 13.11.2018. Recuperado de https://www.wsj.com/articles/ai-could-make-cyberattacks-more-dangerous-harder-to-detect-1542128667?mod=article_inline. Consultado el 11.03.2021

JARVIS THOMSON, J. (1985). "The Trolley Problem". *The Yale Law Journal*. Vol. 94, n.º 6.

JEWELL, C. (2019). "Artificial intelligence: The new electricity". *WIPO Magazine*, junio 2019.

JI, X.; PAULSEN, B.D.; G.K.K. ET AL. (2021). "Mimicking associative learning using an ion-trapping non-volatile synaptic organic electrochemical transistor". *Nature Communications*. 12, 2480. https://doi.org/10.1038/s41467-021-22680-5

JIMÉNEZ, R. (2018). "Primer atropello mortal de un coche sin conductor". Publicado en *El País*, el 20.03.2018. Recuperado de: https://elpais.com/tecnologia/2018/03/19/actualidad/1521479089_032894.html. Consultado el 12.02.2021.

JIMENO MUÑOZ, J. (2017). *La responsabilidad civil en el ámbito de los ciberriesgos*. Fundación Mapfre, Madrid.

JOBIN, A.; IENCA, M. Y VAYENA, E. (2019). "The global landscape of AI ethics guidelines". *Nature Machine Intelligence*. https://arxiv.org/ftp/arxiv/papers/1906/1906.11668.pdf. Consultado el 14.02.2021.

JOHNSON, C Y TYSON, A. (2020). "People globally offer mixed views of the impact of artificial intelligence, job automation on society". *Pew Research Center*, 15.12.2020. Recuperado de: https://www.pewresearch.org/fact-tank/2020/12/15/people-globally-offer-mixed-views-of-the-impact-of-artificial-intelligence-job-automation-on-society/. Consultado el 22.02.2021.

JOYANES, L. (2017). "Ciberseguridad: La colaboración público-privada en la era de la cuarta revolución industrial (Industria 4.0 versus ciberseguridad 4.0)". Cuadernos de Estrategia 185. *Ciberseguridad: La cooperación público-privada*, marzo 2017. Instituto Español de Estudios Estratégicos. Departamento de Seguridad Nacional (DSN). Ministerio de Defensa.

KHARPAL, A. "Stephen Hawking says A.I. could be 'worst event in the history of our civilization'". Publicado en *CBNC* el 06.11.2017. Recuperado de: https://www.cnbc.com/2017/11/06/stephen-hawking-ai-could-be-worst-event-in-civilization.html. Consultado el 14.01.2021.

KEISNER, A.; RAFFO, J. Y WUNSCH-VINCENT, S. (2015). "Breakthrough technologies-Robotics, innovation and intellectual property". Economic Research. Working Paper. *WIPO*. N.º 30.

KINGSTON, J. (2016). "Artificial Intelligence and Legal Liability" (2016). En BRAMER M. Y PETRIDIS M. (Eds) *Investigación y desarrollo en sistemas inteligentes XXXIII*. SGAI. Springer, Cham. https://doi.org/10.1007/978-3-319-47175-4_20.

KOOPS, B.J. Y PIRNI, A. (2014). "Aspectos éticos y legales de la mejora de las capacidades humanas a través de la robótica". *Derecho, innovación y tecnología*. Número especial. IX, 2/2013

KUR, A. Y DREIER, T. (2013). *European Intellectual Property Law*. EE, Cheltenham, UK, Northampton, MA, USA.

KURZWEIL, R. (1990). *The Age of Intelligent Machines*. MIT Press.

LACOZ MORATINOS, G. (2020). "Análisis de la propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas". *Ius et Scientia*. Vol. 6, n.º 2.

LACRUZ MANTECÓN, M. L. (2018). "Potencialidades de los robots y capacidades de las personas". En ROGEL VIDE, C. (Coor.). *Los robots y el Derecho*. Editorial Reus, Madrid.

LANT, K. (2019). "Universal Basic Income: UBI Pilot Programs Around the World". *Futurism*. Disponible en: https://futurism.com/images/universal-basic-income-ubi-pilot-programs-around-the-world. Consultado el 20.02.2021

LARSON, A.J.; MATTU, S. Y KIRCHNER, L. (2016). "Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks".

ProPublica, 23.05.2016. Recuperado de: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

LARSON, J.; MATTU, S.; KIRCHNER, L. Y ANGWIN, J. (2016). "How we analyzed the COMPAS recidivism algorithm". *ProPublica*, 23.05.2016. Disponible en: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm

LASARTE ÁLVAREZ, C. (2013). *Manual sobre protección de consumidores y usuarios*. 5ª Editorial Dykinson. Madrid.

LASARTE ÁLVAREZ, C. (2016). *Compendio de derecho de la persona y del patrimonio*. Editorial Dykinson, Madrid.

LEE, KAI-FU (2018). *Superpotencias de la inteligencia artificial*. Versión Kindle. Deusto.

LEENDERS, G. (2019). "The Regulation of Artificial Intelligence - A Case Study of the Partnership on AI". *Medium*. Publicado el 13.04.2019. Recuperado de: https://becominghuman.ai/the-regulation-of-artificial-intelligence-a-case-study-of-the-partnership-on-ai-c1c22526c19f. Consultado el 14.01.2021.

LEIBNIZ, G.W. (1923). *Sämtliche Schriften und Briefe*. Deutsche Akademie der Wissenschaften zu Berlin. O. Reichl.

LEÓN, G. (2019). "Situación y perspectivas de las tecnologías y aplicaciones de inteligencia artificial". En Documento de Seguridad y Defensa 79. *La inteligencia artificial aplicada a la defensa*. Instituto Español de Estudios Estratégicos.

LESLIE, D.; BURR, C.; AITKEN, M.; COWLS, J.; KATELL, M. AND BRIGGS, M. (2021). *Artificial intelligence, human rights, democracy, and the rule of law: a primer*. The Council of Europe.

LLORET, J.A. (2019). "Estándares y seguridad en el uso humano de la IA". *Revista La Biblia de AI*, 21.10.2019. Recuperado de: https://editorialia.com/2019/10/21/estandares-y-seguridad-en-el-uso-humano-de-la-ia/. Consultado el 20.12.2020.

LÓPEZ DE MÁNTARAS, R. Y MESEGUER, P. (2017). *Inteligencia artificial*. Madrid, CSIC.

LÓPEZ DE MÁNTARAS, R. (2021). "¿Pueden las máquinas ser creativas?". *La Vanguardia*. 31.05.2021. Recuperado de: https://www.lavanguardia.com/ciencia/20210531/7484405/maquinas-creativas.html. Consultado el 31.05.2021.

LÓPEZ-TARRUELLA, A. (2006). *Contratos internacionales de software*. Tirant lo Blach. Valencia.

MANYIKA, J.; CHUI, M. Y OTROS (2017). *A future that works: Automation, employment, and productivity*. McKinsey Global Institute. Chicago.

MARÍN, S. (2019). "Ética e inteligencia artificial*"*. *Cuadernos de la Cátedra Caixabank de Responsabilidad Social Corporativa-IESE Business School*, n.º 42. IESE.

MARSDEN, C. ET AL. (2019) *Regulating disinformation with artificial Intelligence*. EPRS European Parliamentary Research Service.

MARTÍNEZ, C. (2018). "Expanding Patents in the Digital World: The exemple of Patents in Software", en SEUBA, X.; GEIGER. C. Y PENIN, J. (Eds). *Intellectual Property and Digital Trade in the Age of Artificial Intelligence and Big data*. CEIPI-ICTSD. Issue n.º 5.

MARTÍNEZ MERCADAL, J.J. (2018). "Vehículos autónomos y derecho de daños. La estructura clásica de la responsabilidad civil frente al avance de la inteligencia artificial". *Revista de la Facultad de Ciencias Económicas UNNE*, n.º 20. Universidad de la República de Montevideo Uruguay. Disponible en https://revistas.unne.edu.ar/index.php/rfce/article/view/3267. Consultado el 31.01.2021

MARTÍNEZ REY, M.A. Y PAZOS SIERRA, J. (2019). "La inteligencia artificial y el derecho: Pasado, presente y futuro", en MONTERROSSO CASADO, E. (Dir.), *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch.

MARTÍNEZ QUIRANTE, R. (2018). *Inteligencia artificial y armas letales autónomas. Un nuevo reto para Naciones Unidas*. Tres Ensayos. Asturias.

MCCARTHY, J. (2007). "What Is Artificial Intelligence". Consultado el 23 de noviembre de 2020. Recuperado de: http://www-formal.stanford.edu/jmc/whatisai/node1.html.

MCCARTHY, J.; MINSKY, M.; ROCHESTER, N.; SHANNON, C. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. 31.08.1955. Recuperado de: http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html. Consultado el 24.11.2020.

MCDERMOTT, J. P. (1982). "R1: A rule-based configurer of computer systems". *Artificial Intelligence*. Vol. 19. Issue 1. Elsevier.

MERCADER, J. R. (2017). *El futuro del trabajo en la era de la digitalización y la robótica*. Editorial Tirant lo Blanch. Valencia.

DE MIGUEL, R. (2021). "Un ciberataque obliga a Irlanda a cerrar el sistema informático de la sanidad pública". *El País*, 14.05.2021. Recuperado de: https://elpais.com/internacional/2021-05-14/un-ataque-cibernetico-en-irlanda-obliga-a-cerrar-el-sistema-informatico-de-la-sanidad-publica.html. Consultado el 30.05.2021

MITTELSTADT, B.; D., ALLO, P.; TADDEO, M.; WATCHER, S. Y FLORIDI, L. (2016). "The ethics of algorithms: Mapping the debate". *Big data & Society*. Vol. 3. N.º 2.

MOLINA MIRANDA, A. Y JUBERÍAS SÁNCHEZ, A. (2017). "Producto sanitario defectuoso". En Juberías Sánchez, A. (Coor.). *Medicamentos, productos sanitarios y protección del consumidor*. Editorial Reus. Madrid.

MONJE, C. (2018). "Salud El poder terapéutico de un robot-peluche". *El País*, 11.11.2018. Disponible en: https://retina.elpais.com/retina/2018/11/09/tendencias/1541790426_183947.html. Consultado el 21.02.2021.

MONTÉS PENADÉS, V.L. (2007). "La responsabilidad por dolo", en MORENO, J.A. (Coord.), *La responsabilidad civil y su problemática actual*. Dykinson.

MONTÉS PENADÉS, V.L. (2011). "El significado institucional y técnico de la idea de persona", en BLASCO GASCO, F. (Coord.). *Derecho Civil. Parte General. Derecho de la Persona*. Tirant Lo Blanch, Valencia.

MONTOYA, M (2005). "Derecho y política en el pensamiento de Bobbio: una aproximación". *Estudios Políticos.* Nº 26. Medellín. Enero-junio 2005.

MORIELLO, S. (2006*). "Los robots inteligentes tendrán tres niveles de conciencia". Levante EMV.* Recuperado de https://tendencias21.levante-emv.com/los-robots-inteligentes-tendran-tres-niveles-de-conciencia_a832.html#:~:text=Los%20niveles%20reactivo%2C%20deliberativo%20y,fecha%20ninguna%20definici%C3%B3n%20universalmente%20aceptada. Consultado el 28.01.2021.

MOZO SEOANE, A. (2018). "Robots e inteligencia artificial: Control de sus riesgos". *Revista General de Legislación y Jurisprudencia*, n.º 2. Editorial Reus. Madrid.

MUÑOZ VILLARREAL, A. Y GALLEGO CORCHERO, V. (2019). "Inteligencia artificial e irrupción de una nueva personalidad en nuestro ordenamiento jurídico ante la imputación de responsabilidad a los robots". En Monterrosso Casado, E. (Dir.). *Inteligencia artificial y riesgos cibernéticos. Responsabilidades y aseguramiento*. Tirant lo Blanch.

MUÑOZ VELA, J.M. (2019). "Los retos del Derecho en una Sociedad Digital", en GIMÉNEZ, I. (Coord.). *Retos de la sociedad digital y medios de pago*. Colección Tratados y Manuales de Economía. Editorial Civitas - Thomson Reuters Aranzadi. Navarra.

MUÑOZ VELA, J. M. (2019). Conclusiones finales en la Conferencia impartida el 25.03.2019 en la Universidad Politécnica de Valencia, organizada por el Instituto Tecnológico de Informática (ITI), bajo el título "Responsabilidad de consultoras, productoras y comercializadoras de software contable, de facturación o de gestión empresarial: Situación actual y tendencias regulatorias"

MUÑOZ VELA, J. M. (2020). Conclusiones finales en la Conferencia impartida el 18.12.2020 en la Facultad de Derecho de Valencia, organizada por su Departamento de Derecho Penal, bajo el título "Ciberdelincuencia y Ciberseguridad. Delitos intrusivos, hacking, cracking y otros".

MUÑOZ VELA, J.M. (2021). "Las estrategias delictivas en el ciberespacio se perfeccionan". *Valencia Plaza*, 17.02.2021. Recuperado de: https://valenciaplaza.com/estrategias-delictivas-ciberespacio-ciberataques

MURDOCH, W. J., SINGH, CH., KUMBLER, K., ABBASI-ASI, R Y YU, B. (2019). "Interpretable machine learning: definitions, methods, and applications". *PNAS*.

MURPHY, R. (2000). *Introduction to AI Robotics 2e (Intelligent Robotics and Autonomous Agents series)*. Prensa del MIT. Cambridge EE.UU.

NAVARRO-DOLMESTCH, R. Y VIDAL-TAMAYO I. (2020) "Sobre la justificación de aplicar el derecho penal a las entidades de inteligencia artificial". En AZUAJE PIRELA, M. Y CONTRERAS P. *Inteligencia artificial y Derecho: Desafíos y perspectivas*. Tirant lo Blanch.

NAVAS NAVARRO, S. (2017). "Derecho e inteligencia artificial desde el diseño. Aproximaciones", en NAVAS NAVARRO, S. (Dir.). *Inteligencia artificial, tecnología, derecho*. Editorial Tirant lo Blanch. Valencia.

NAVAS NAVARRO, S. (2016). "Smart Robots y otras máquinas inteligentes en nuestra vida cotidiana". *Revista CESCO Digital*. Nº 20.

NEVEJANS, N. (2016). *Study for Jury Committee European Civil Law Rules in Robotics Policy*. Department C: Citizens' Rights and Constitutional Affairs. European Parliament. Bruselas. Disponible en: http://www.europarl.europa.eu/committees/fr/supporting-analyses-search.html

NIETO MENGOTTI, J.P. (2016). *El Derecho Penal frente a los robots*. FIDE Papers, Madrid.

NIETO, M. (2021). "Conciencia e inteligencia artificial, un debate abierto". Disponible en: https://blogthinkbig.com/conciencia-inteligencia-artificial. Consultado el 09.02.2021.

NIEVA FENOLL, J. (2020). "Conferencia: inteligencia artificial y Proceso Penal". Universitas Fundación, 22 de julio de 2020. Disponible en: https://www.youtube.com/watch?v=5BrCNVTPp0o

NILSSON, N. (1998). *Artificial Intelligence: A New Synthesis*. Elsevier.

NILSSON, P. (2018). "How AI helps recruiters track jobseekers' emotions". *Financial Times*. Disponible en: https://www.ft.com/content/e2e85644-05be-11e8-9650-9c0ad2d7c5b5

NÚÑEZ ZORRILLA, M.C. (2018). "Los nuevos retos de la Unión Europea en la regulación de la responsabilidad civil por los daños causados por la inteligencia artificial". *Revista Española de Derecho Europeo,* n.º 66 (Abril-junio 2018). Editorial Civitas (Thomson Reuters). Navarra.

NÚÑEZ ZORRILLA, M.C. (2019). *Inteligencia artificial y responsabilidad civil*. Reus Editorial, Madrid.

OLIVER, N. (2019). "Governance in the Era of Data-driven Decisionmaking Algorithms". *Women Shaping Global Economic Governance*. CEPR.

OROZCO GONZÁLEZ, M. (2021) "Reflexiones acerca de la relación entre inteligencia artificial y robótica", en ATAZ LÓPEZ, J. Y COBACHO GÓMEZ, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón. Tomo III*. Ed. Aranzadi-Thomson Reuters.

ORTEGO RUIZ, M. (2018). "El concepto de autor en la era de los robots". *Anuario de Propiedad Intelectual 2017*. Editorial Reus.

ORTIZ FERNÁNDEZ, M. (2020). "Reflexiones acerca de la responsabilidad civil derivada del uso de la inteligencia artificial: los 'principios' de la Unión Europea". *Revista de Direito da ULP*. Vol. 14. Nº. 1.

ORTS BERENGUER, E. Y GONZÁLEZ CUSSAC, J.L. (2014). *Compendio de Derecho Penal. Parte General*. 4ª edición. Tirant lo Blanch.

OTTOLIA, A. (2018). *Derecho, Big data e inteligencia artificial*. Editorial Tirant lo Blanch y G. Giappichelli Editore. Valencia-Torino.

PAGALLO, U. (2017). "AI and bad robots. The criminology of automation", en HOLT, T. (Coord.). *The Routledge Handbook of Technology, Crime and Justice*. Ed. Routledge. Londres.

PALAZUELOS, F. (2017). "Elon Musk: `La inteligencia artificial amenaza la existencia de nuestra civilización". *El País*, 18.07.2017. Disponible en: https://elpais.com/tecnologia/2017/07/17/actualidad/1500289809_008679.html. Consultado el 25.02.2021

PALLANTE, M. (2017), "From monkey selfies to open source: The essential interplay of creative culture, technology, copyright office practice, and the law". *Washington Journal of Law, Technology & Arts.* Vol. 12, Issue 2. Disponible en: http://digital.law.washington.edu/dspace-law/handle/1773.1/1703. Consultado el 09.03.2021.

PALMA, J. Y MARÍN, R. (2008). *inteligencia artificial. Técnicas, métodos y aplicaciones.* McGraw-Hill -Interamericana de España.

PALMERINI, E. (2017). "Robótica y derecho: Sugerencias, confluencias, evoluciones en el marco de una investigación europea". *Revista de Derecho Privado,* n.º 32. Universidad externado de Colombia. Disponible en https://dialnet.unirioja.es/servlet/articulo?codigo=6073053

PANEZI, A. (2021). "IA: un enfoque ecosistémico para gestionar el riesgo y la incertidumbre", en García Mexia, P. y Pérez Bes, F. (Eds); Panezi. A. (Coord). (2021). *Artificial Intelligence and the Law*. Ed. La Ley (Wolters Kluwer). 2021. Edición electrónica (SMARTECA).

PARRA LUCÁN, M.A. (2011). *La protección del consumidor frente a los daños. Responsabilidad civil del fabricante y del prestador de servicios.* Colección Derecho del Consumo. Editorial Reus. Madrid.

PASCUAL, M.G. (2021). "Cuando el algoritmo se equivoca". *El País*, 27.06.2021.

PATEL, A. ET AL. (2019). "Security Issues, Dangers and Implications of Smart Information Systems*"*. *Sherpa Project*, D1.3. Recuperado de: https://dmu.figshare.com/articles/D1_3_Cyberthreats_and_countermeasures/7951292. Consultado el 02.02.2021.

PAUL, K. (2019). "Microsoft Japan tested a four-day work week and productivity jumped by 40%". *The Guardian*, 04.11.2019. Disponible en: https://www.theguardian.com/technology/2019/nov/04/microsoft-japan-four-day-work-week-productivity. Consultado el 21.02.2021.

PEGUERA, M. (2020). "En búsqueda de un marco normativo para la inteligencia artificial" en CERRILLO I MARTÍNEZ, A. Y PEQUERA POCH, M. (Coord). *Retos jurídicos de la inteligencia artificial*. Aranzadi S.A.U. (Thomson Reuters). Navarra.

PFEIFER, R. Y SCHEIER, C. (1999). *Understanding Intelligence*. Editorial MIT Press. Cambridge.

PLATERO ALCÓN, A. (2021). "Breves notas sobre el régimen de responsabilidad civil derivado de los sistemas de inteligencia artificial: especial referencia al algoritmo de recomendaciones de Netflix". En *Ius et Scientia*. Vol. 7. Nº 1. Universidad de Sevilla.

PLAZA PENADÉS, J. (2018). "Primeras reflexiones desde el Derecho sobre la inteligencia artificial". *Revista Aranzadi de Derecho y Nuevas Tecnologías*. N. 47, Editorial. Thomson Reuters (Legal) Limited.

POOLE, D., GOEBEL, R.G. Y MACKWORTH, A.K. (1998). *Computational Intelligence: A Logical Approach*. Oxford University Press. New York.

POSADA MAYA, R. (2019). "La responsabilidad penal de los agentes de inteligencia artificial: entre la ficción y una realidad que se aproxima". En PORTILLA CONTRERAS, G. Y VELÁSQUEZ F. (Dir.); POMARES CINTAS, E. Y FUENTES OSORIO, J.L. (Coord.). *Un juez para la democracia. Libro homenaje a Perfecto Andrés Ibáñez*. Dykinson.

PRIETO, M. (2017). "Equifax reconoce un ciberataque masivo que afecta a 143 millones de clientes". *Expansión*, 08.09.2017. Recuperado de: https://www.expansion.com/economia-digital/companias/2017/09/08/59b23dd822601dc97c8b4655.html. Consultado el 01.02.2021.

PRIGG, M. (2014). "Who goes there? Samsung unveils robot sentry that can kill from two miles away". *Daily Mail*, 16.09.2014. Recuperado de: https://www.dailymail.co.uk/sciencetech/article-2756847/Who-goes-Samsung-reveals-robot-sentry-set-eye-North-Korea.html. Consultado el 18.02.2021

QUINTERO OLIVARES, G. (2017). "La robótica ante el Derecho Penal: El vacío de respuesta jurídica a las desviaciones incontroladas". *Revista Electrónica de Estudios Penales y de la Seguridad*, n.º 1, ISSN:2531-1565.

RAINER, J.J. Y RODRÍGUEZ, L. (2019). "Perspectiva histórica y evolución de la inteligencia artificial". *Documento de Seguridad y Defensa 79: La inteligencia artificial aplicada a la defensa*. Instituto Español de Estudios Estratégicos.

RAJAN, A. (2021). "La inteligencia artificial supondrá un cambio "más profundo que el fuego, la electricidad o internet": Sundar Pichai, líder de Google". *BBC News*, 13.07.2021.

RAMALHO, A. (2017). "Will Robots Rule the (Artistic) World? A Proposed Model for the Legal Status of Creations by Artificial Intelligence Systems". *Journal of Internet Law*.

RAMALHO, A. (2018). "Ex Machina, Ex Auctore? Machines that create and how EU copyright law views them". *Wolters Kluwer*. Disponible en: http://copyrightblog.kluweriplaw.com/2018/11/12/ex-machina-ex-auctore-machines-that-create-and-how-eu-copyright-law-views-them/

RAO, A.S. Y VERMEIJ, G. (2017). *Sizing the Prize*. PwC. 27 de junio de 2017. Disponible en: https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf. Consultado el 02.03.2021.

RAPOSO, M.A. ET AL. (2019). *The future of road transport*. EUR 29748 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-14318-5. DOI:10.2760/668964.

RAMÓN FERNÁNDEZ, F. (2019). "Robótica, inteligencia artificial y seguridad: ¿Cómo encajar la responsabilidad civil?". *La Ley Digital* 9365. Wolters Kluwer.

RAMOS HERRANZ, I. (2006). "El estándar mercantil de diligencia: El ordenado empresario". *Anuario de derecho civil*. ISSN 0210-301X. Vol. 59. Nº 1.

REEDY, C. (2017). "Kurzweil Claims That the Singularity Will Happen by 2045". *Futurism*, el 10.05.2017.

RIBAS ALEJANDRO, J. (1994). "Informática y responsabilidad civil". *Actualidad Informática Aranzadi*, n.º 12.

RICH, E. Y KNIGHT, K. (1991). *Artificial Intelligence*. McGraw-Hill.

RIES, A. Y TROUT, J. (1993). *Las 22 leyes inmutables del marketing*. McGraw-Hill.

RÍOS LÓPEZ, Y. (2020). "Inteligencia artificial y Patentes: ¿Hacia un 'Inventor Artificial'?", en CERRILLO I MARTÍNEZ, A. Y PEQUERA POCH, M. (Coord). *Retos jurídicos de la inteligencia artificial*. Aranzadi S.A.U. (Thomson Reuters). Navarra.

ROBERT GUILLÉN, S. (2017). "Impresoras 3D y 4D", en Navas Navarro, S. (Dir.). *Inteligencia artificial, tecnología, derecho*. Valencia, Tirant lo Blanch.

ROBLES CARRILLO, M. (2020). "La gobernanza de la inteligencia artificial: contexto y parámetros generales". *Revista electrónica de estudios internacionales (REEI)*. Nº. 39 2020.

RODRÍGUEZ-ROSADO, B. (2020). "Los sistemas de responsabilidad contractual: Entre la responsabilidad por culpa y la strict liability". *Revista de Derecho Civil*. Vol. I, n.º 4. Disponible en: https://www.nreg.es/ojs/index.php/RDC/article/view/103. Consultado el 20.12.2020.

ROGEL VIDE, C. (2018). "Robots y personas". *Revista General de Legislación y Jurisprudencia*, n.º 1. Editorial Reus. Madrid.

ROSALES DE SALAMANCA, F. (2016) "¿Puede un robot ser sujeto de derecho?" Recuperado de https://notariofranciscorosales.com/puede-robot-sujeto-derecho/. Consultado el 05.02.2021.

RUBÍ, A. (2020). "Retos de la inteligencia artificial y adaptabilidad del derecho de daños", en CERRILLO I MARTÍNEZ, A. Y PEQUERA POCH, M. (Coord). *Retos jurídicos de la inteligencia artificial*. Aranzadi S.A.U. (Thomson Reuters). Navarra.

RUBIO, A. (2014) "Los efectos jurídicos del *soft law* en materia de igualdad efectiva. La experiencia española". *Anuario de filosofía del derecho AFD.* Nº 30, 2014. ISSN 0518-0872. Universidad de Granada.

RUBIO, I. (2020). "Inteligencia artificial valenciana para aconsejar si hospitalizar a un contagiado". *El País*, 18.03.2020. Disponible en: https://retina.elpais.com/retina/2020/03/17/innovacion/1584450877_681658.html. Consultado el 19.02.2021.

RUSSELL, STUART J. Y NORVIG, P. (2009). *Artificial intelligence: a modern approach*. Upper Saddle River, N.J. Prentice Hall. 3.ª Ed.

SAIZ, C. (2018). "Las obras creadas por Sistemas de inteligencia artificial y su protección por el derecho de autor". *InDret.* Barcelona.

SALVADOR CODERCH, P. S. Y RAMÓN GONZÁLEZ, S. (2008). "Defectos de producto". En SALVADOR CODERCH, P. S. Y GÓMEZ POMAR, F. (Coords). *Tratado de responsabilidad civil del fabricante*. Capítulo IV. Editorial Thomson Civitas. Navarra.

SÁNCHEZ DEL CAMPO, A. (2016). *Reflexiones de un replicante legal. Los retos jurídicos de la robótica y las tecnologías disruptivas*. Editorial Aranzadi, Navarra.

SÁNCHEZ DEL CAMPO, A. (2017). "Europa quiere regular a los robots*"*. *Diario La Ley*, n.º 4. Ed. Wolters Kluwer.

SÁNCHEZ, F. Y LÓPEZ, J. (2017). "Cooperación público-privada en la protección de infraestructuras críticas". Cuadernos de Estrategia 185. *Ciberseguridad: La cooperación público-privada*. Instituto Español de Estudios Estratégicos. Departamento de Seguridad Nacional (DSN). Ministerio de Defensa.

SÁNCHEZ-URÁN, Y. Y GRAU RUIZ, M.A. (2018). "El impacto de la robótica, en especial la robótica inclusiva, en el trabajo: Aspectos jurídicos-laborales y fiscales". Congreso Internacional sobre Innovación Tecnológica y Futuro del Trabajo. *Technological Innovation and the Future of Work: Emerging aspects worldwide*. Santiago de Compostela.

SANJUÁN, N. (2019). "Inteligencia artificial y propiedad intelectual*"*. *Actualidad Jurídica Uría Menéndez*, Nº 52-2019.

SANTOS GONZÁLEZ, M.J. (2017). "Regulación legal de la robótica y la inteligencia artificial: Retos del futuro". *Revista Jurídica de la Universidad de León*. N.º 4.

SANTOS BRIZ J. (1993). *La responsabilidad civil, Derecho sustantivo y derecho procesal*. 7ª Edición. Editorial Montecorvo, S.A. Madrid.

SARTOR, G. (2009). "Cognitive automata and the law: electronic contracting and the intentionality of software agents". *Artificial intelligence and law*. Vol. 17.4.

SCHAUB M. Y ZHAO A. (2020). "China Releases Big Plan for Autonomous Vehicles". *China Law Insight*.

SCHLACKMAN, S. (2018). "Who owns the copyright of the Next Rembrandt?". *Art Law Review*. Disponible en: https://alj.orangenius.com/the-next-rembrandt-who-holds-the-copyright-in-computer-generated-art

SEARLE, J. (1980). Minds, brains, and programs, *Behavioral and Brain Science 3 (3)*. Cambridge (UK)

SEARLE, J. (2000). *El misterio de la conciencia*. Editorial Paidos Ibérica, Madrid.

SELBST, A. D. (2020). "Negligence and AI's Human Users". *100 Boston University Law Review* 1315. Facultad de Derecho de UCLA. Documento de investigación de derecho público núm. 20-01. Disponible en SSRN: https://ssrn.com/abstract=3350508

SEOANE SPIEGELBERG, J.L. (2019) "La responsabilidad civil tras 130 años de vigencia del Código Civil". *Diario La Ley*, n.º 9537, Sección Tribuna. Wolters Kluwer.

SHULEVITZ, J. (2018). "Alexa, Should We Trust You?". *The Atlantic*. Recuperado de: www.theatlantic.com/magazine/archive/2018/11/alexa-how-will-you-change-us/570844.

SILVER, D., HUANG, A., MADDISON, C. ET AL. (2016). "Mastering the game of Go with deep neural networks and tree search". *Nature* 529. https://doi.org/10.1038/nature16961.

SOLÉ FELIU, J. (1997). *El concepto de defecto del producto en la responsabilidad civil del fabricante: ley 22/1994, de 6 de julio, de responsabilidad civil por los daños causados por productos defectuosos*. Tirant lo Blanch. Valencia.

SOLER MATUTES, P. (2004). *El contrato para la elaboración de programas de ordenador*. Pamplona.

SOLER MATUTES, P. (2004). *El contrato para la elaboración de programas de ordenador*. Pamplona, 2004.

SOLER MATUTES, P. (2006). "El contrato de desarrollo de software. La responsabilidad de las partes", en SOLER MATUTES, P. (Dir.). *Manual de Gestión y Contratación Informática*. Editorial Aranzadi.

SOURDIN, T. (2018). "Judge v Robot? Artificial Intelligence and judicial decision-making". *UNSW Law Journal*. Vol. 41 (4).

SPRING, J.M. ET AL. (2019). "Machine Learning In Cybersecurity: A Guide". *SEI-CMU Technical Report*, n.º 5. Software Engineering Institute-Carnegie Mellon University. Recuperado de: https://resources.sei.cmu.edu/asset_files/TechnicalReport/2019_005_001_633597.pdf. Consultado el 02.02.2021.

SPRING, J.M.; FALLON. J. ET AL. (2019). *Machine Learning in Cybersecurity*. Carnegie Mellon University.

STERBA, J. P. (2009). *Ethics: The Big Questions*. Reino Unido: John Wiley & Sons.

STIEG, C. (2020). "How this Canadian start-up spotted coronavirus before everyone else knew about it". *CNBC*, 03.03.2020. Disponible en: https://www.cnbc.com/2020/03/03/bluedot-used-artificial-intelligence-to-predict-coronavirus-spread.html. Consultado el 19.02.2021.

STONE, P. Y OTROS (2016). *Artificial Intelligence and life in 2030: The one hundred year study on artificial intelligence*. Universidad de Stanford.

STONE, Z. (2017). "Everything You Need To Know About Sophia, The World's First Robot Citizen". *Forbes*, 7.11.2017. Disponible en: https://www.forbes.com/sites/zarastone/2017/11/07/everything-you-need-to-know-about-sophia-the-worlds-first-robot-citizen/?sh=7dd2352f46fa. Consultado el 21.02.2021.

STROWEL, A. Y UTKU, S. (2016). "The trends and current practices in the area of patentability of computer implemented inventions within the EU and the U.S.". T*he European Commission*. Disponible en: http://ec.europa.eu/newsroom/document.cfm?doc_id=41192.

STUPP, C. (2019). "Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case". *The Wall Street Journal*, 30.08.2019. Recuperado de https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402. Consultado el 11.03.2021.

TEGMARK, M. (2018). *Vida 3.0: Qué significa ser humano en la era de la inteligencia artificial*. Penguin Random House Grupo Editorial.

TERRONES, A.L. (2018). "Inteligencia artificial y ética de la responsabilidad". *Cuestiones de filosofía*. Vol. 4. N.º 22.

TOLAN S., MIRON M., GOMEZ E. Y CASTILLO C. (2019)."Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia". Best Paper Award, *International Conference on AI and Law*.

TOLIDO, R. ET AL. (2019). *Reinventing Cybersecurity with Artificial Intelligence: The new frontier in digital security*. Capgemini. Recuperado de: https://www.capgemini.com/es-es/wp-content/uploads/sites/16/2019/07/AI-in-Cybersecurity_Report_20190710_V05.pdf. Consultado el 12.02.2021.

DE LA TORRE, I. (2018). "La disrupción tecnológica ya está aquí. Cómo afecta a las personas, los gobiernos y las empresas". En Cuadernos de Estrategia 199. *Gobernanza futura: Hiperglobalización, mundo multipolar y Estados menguantes*, diciembre 2018. Instituto Español de Estudios Estratégicos. Ministerio de Defensa.

TUNG, J.R. (2016). "Who Owns the Creation of an Artificial Intelligence?". *FindLaw*. 22 de agosto de 2016. Disponible en: https://blogs.findlaw.com/technologist/2016/08/who-owns-the-creation-of-an-artificial-intelligence.html.

ULMER, E. (1980). *Urheber- und Verlagsrecht*. 3ª Ed. Springer. Berlin-Heidelberg-New York.

VÁZQUEZ DE CASTRO, E. (2020). "Aproximación a la responsabilidad derivada de los riesgos de la inteligencia artificial en Europa". En Solar Cayón, J.I. *Dimensiones éticas y jurídicas de la inteligencia artificial en el marco del Estado de Derecho*. Cuadernos de la Cátedra de Democracia y Derechos Humanos. Universidad de Alcalá: Defensor del Pueblo. 2020.

VIDAL, M. (2019). *La era de la humanidad*. Versión Kindle. Ediciones Deusto. Barcelona.

VILLAMOR, N. (2021). Amy Webb: "Nos acercamos a un escenario catastrófico con la inteligencia artificial". *The Objective*, 09.07.2021.

VIRILIO, P. (1999). *Politics of the Very Worst*. New York. Semiotext (e).

WAGNER, B. (2018). Ethics as an Escape from Regulation: From ethics-washing to ethics-shopping?. En HILDEBRANDT, M. (Ed.), *Being Profiling: Cogitas ergo sum*. Amsterdam University Press.

WAGNER, G. (2019). "Robot Liability", en LOHSSE, S., SCHULZE R. Y STAUDENMEYER, D. (Eds.), *Liability for Artificial Intelligence and the Internet of Things*. Nomos. Baden-Baden.

WAKABAYASHI, D. (2018) "Self-driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam*". The New York Times*. Recuperado de: https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html.

WINDER D. (2019). "New Orleans Declares State Of Emergency Following Cyber Attack". *Forbes*, 14.12.2019. Recuperado de: https://www.forbes.com/sites/daveywinder/2019/12/14/new-orleans-declares-state-of-emergency-following-cyber-attack/?sh=65b824246a05. Consultado el 17.02.2021.

WINSTON, P. H. (1992). *Artificial intelligence*. Addison-Wesley, Reading, MA, 3ª Edición.

XALABARDER, R. (2020). "Inteligencia artificial y Propiedad Intelectual", en CERRILLO I MARTÍNEZ, A. Y PEQUERA POCH, M. (Coord). *Retos jurídicos de la inteligencia artificial*. Aranzadi S.A.U. (Thomson Reuters). Navarra.

XIN Y. ET AL. (2018). *Machine Learning and Deep Learning Methods for Cybersecurity*. IEEE. Disponible en: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8359287. Consultado el 03.03.2021.

YAFFE-BELLANY, D. (2019). "Computación cuántica explicada en unos minutos". *The New York Times*, 24.10.2019. Recuperado de:

https://www.nytimes.com/es/2019/10/24/espanol/ciencia-y-tecnologia/computacion-cuantica-google.html. Consultado el 24.02.2021.

YDEWALLE, G. Y DELHAYE, P. (1988). "La inteligencia artificial, la obtención del conocimiento y el estudio de la inteligencia humana". *Revista internacional de ciencias sociales*. Vol. 40, n.º 1.

YUSTE, R. ET AL. (2017). "Four ethical priorities for neurotechnologies and AI". *Nature*. Vol. 551. N.º 7679.

ZAMORANO E. (2020). "Los 'deepfakes' del porno: así han reaccionado las actrices". *El Confidencial*, 23.01.2020. Disponible en: https://www.elconfidencial.com/alma-corazon-vida/2020-01-23/deepfake-porno-sexualidad-internet-internet_2420819/. Consultado el 17.02.2021.

ZAVIA, M.S. (2016). "Inscriben a una inteligencia artificial en un concurso literario. Queda finalista". *Gizmodo*, 24.03.2016. Disponible en: https://es.gizmodo.com/inscriben-a-una-inteligencia-artificial-en-un-concurso-1766836828. Consultado el 19.02.2021

ZELEZNIKOW, J. (2017). "Can Artificial Intelligence and Online Dispute Resolution enhance efficiency and effectiveness in Courts". *International Journal for Court Administration*. Vol. 8. N.º 2. Disponible en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2999339.

ZHONG, HAN-SEN; WANG, HUI ET AL. (2020). "Quantum computational advantage using photons". *Science*, vol. 370, Issue 6523. DOI: 10.1126/science.abe8770.

ZURITA MARTÍN, I. (2021). "Gestión de riesgos y responsabilidad civil de los robots", en Ataz López, J. y Cobacho Gómez, J.A. (Coords). *Cuestiones clásicas y actuales del Derecho de daños. Estudio en Homenaje al Profesor Dr. Roca Guillamón.* Tomo III. Ed. Aranzadi-Thomson Reuters. 2021.