Departamento de Derecho Constitucional, Ciencia Política y de la Administración

Programa de Doctorado en Derecho, Ciencia Política y Criminología

TESIS DOCTORAL

Régimen jurídico de la toma de decisiones automatizadas y el uso de sistemas de inteligencia artificial en el marco del derecho a la protección de datos personales

Presentada por:

Adrián Palma Ortigosa

Dirigida por:

Prof. Dr. Lorenzo Cotino Hueso

Octubre 2021

ÍNDICE

ABREVIATURAS

AEPD	Agencia Española de Protección de Datos.
AN	Audiencia Nacional.
CNIL	Commission nationale de l'informatique et des libertés.
CEPD	Comité Europeo de Protección de Datos.
CRISP-DM	Cross-industry standard process for data mining.
CTBG	Consejo de Transparencia y Buen Gobierno.
Directiva 95/46	Directiva 95/46/CE del Parlamento Europeo y del Consejo, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respeta al tratamiento de datos personales y a la libre circulación de estos datos.
EIPD	Evaluación de Impacto de Protección de Datos.
EEUU	Estados Unidos de América.
FMI	Fondo Monetario Internacional.
GAIP	Comisión de Garantía del derecho de acceso e información púbica de Cataluña.
GT29	Grupo del Artículo 29.
IA	Inteligencia Artificial.
ICO	Information Commissioner´s Office.
LOPD de 2018	Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.
OCDE	Organización para la cooperación y desarrollo económico.
PRAI	Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.
RGPD	Reglamento General de Protección de Datos.
STC	Sentencia del Tribunal Constitucional.
STJUE	Sentencia del Tribunal de Justicia de la Unión Europea.
STS	Sentencia del Tribunal Supremo.
TC	Tribunal Constitucional.

TJUE	Tribunal de Justicia de la Unión Europea.
TS	Tribunal Supremo.
UE	Unión Europea.

INTRODUCCIÓN

I . OBJETO DE LA TESIS

El objetivo general que pretendemos con el estudio de esta tesis es analizar y clarificar el régimen jurídico aplicable a las decisiones automatizadas con relación al derecho fundamental a la protección de datos personales. No sólo se abordan las cuestiones relacionadas con las decisiones estrictamente automatizadas sino también aquellas decisiones que sin ser plenamente automatizadas, se utilizan como apoyo a la decisión que finalmente adopta un humano.

La automatización del proceso decisorio a través de algoritmos plantea toda una serie de retos jurídicos que afectan directamente a los derechos y libertades de las personas. La normativa en materia de protección de datos se muestra como una de las principales herramientas con las que cuenta el ordenamiento jurídico para dar respuesta a dichos retos.

Nuestra tesis central es que si bien el derecho a la protección de datos personales no puede encarar todos los desafíos que presenta el uso de sistema de inteligencia artificial en la toma de decisiones automatizadas, la fuerte consolidación del régimen jurídico de este derecho tanto a nivel europeo como nacional favorece un marco adecuado de protección. Así, la estructura de este régimen jurídico se compone de varios pilares: i) el reconocimiento de toda una serie de derechos en favor de los titulares de los datos, ii) el diseño de todo un conjunto de principios aplicables a todos los tratamientos de datos personales, iii) la obligación impuesta a las organizaciones que tratan datos personales de desplegar toda una serie de medidas que demuestren y aseguren el cumplimiento de la norma y, iv) la existencia de una sólida red de autoridades de control encargadas de velar precisamente por el respeto de ese derecho.

Por tanto, el aspecto relevante de este trabajo es analizar dicho régimen jurídico focalizando toda nuestra atención en la incidencia que presenta la irrupción de la automatización del proceso decisorio en el derecho fundamental a la protección de datos mediante el uso de algoritmos. Para ello, tomamos como referencia las dos grandes fases que están presentes durante el ciclo de vida de los sistemas de toma de decisiones

automatizadas, estas son: i) la fase de desarrollo, cuyo objetivo principal es diseñar y elaborar el sistema de toma de decisiones y, ii) la fase de despliegue, la cual tiene como finalidad la puesta en marcha del sistema en el entorno donde este desplegará sus efectos. Los distintos elementos jurídicos que componen la normativa de protección de datos se analizan e interpretan tomando como referencia la realidad presente en este tipo de sistemas. Se facilita así la proposición de soluciones que respetan el difícil equilibrio entre la protección adecuada de los derechos de las personas y el desarrollo de estas nuevas tecnologías guiando con ello hacia nuevos enfoques futuros de regulación.

II. METODOLOGÍA

Son tres los métodos de investigación esenciales sobre los que se ha sustentado la elaboración de este trabajo.

En primer lugar se ha llevado a cabo un *método analítico general* del fenómeno de la toma de decisiones automatizadas. Esto nos ha obligado a profundizar en temas que se escapan de lo estrictamente jurídico y que abordan las principales cuestiones técnicas relacionadas con el desarrollo y despliegue de los sistemas algorítmicos, los cuales, son posteriormente utilizados para la toma de decisiones. Esta primera fase del trabajo ha resultado esencial ya que nos ha permitido posteriormente afrontar de forma más adecuada las implicaciones de esta realidad tecnológica en el ordenamiento jurídico objeto de este estudio.

En segundo lugar, a través del *método jurídico descriptivo* se ha realizado un análisis del conjunto de reglas específicas y generales que la normativa de protección de datos prevé con relación a las decisiones automatizadas. Entre estas reglas se ha estudiado el conjunto de derechos en favor de los particulares sometidos a dichas decisiones, así como el conjunto de deberes que deben asumir los responsables del tratamiento de datos que decidan implantar estos sistemas en sus organizaciones.

En tercer lugar también se ha empleado para la elaboración de este trabajo el *método jurídico propositivo.* Así, se interpretan los preceptos que versan sobre la materia indicada y además se pretende una clarificación de aquellos elementos que o bien no están del todo definidos en el cuerpo legal o bien su redacción enfrenta otros bienes e intereses en juego que requieren de una interpretación equilibrada. Precisamente este último aspecto resulta esencial ya que el derecho fundamental a la protección de datos no es un derecho absoluto, la interpretación que se realiza de esta

normativa y las propuestas que se materializan en este estudio tienen en cuenta el resto de intereses y derechos implicados que residen en favor de las organizaciones que hacen uso de este tipo de algoritmos.

Este estudio centra su atención en la normativa de protección de datos, sin embargo, a lo largo del trabajo se analizan otros ámbitos jurídicos ya que el uso de sistemas automatizados por parte de todo tipo de organizaciones está fuertemente implantado en un gran número de sectores como el laboral, el financiero, seguros, sanidad, judicial, penal, etc. Dicho esto, algunos apartados específicos de esta tesis se focalizan en el uso de los sistemas automatizados en el sector público y en las plataformas digitales tomando como referencia la normativa desarrollada hasta la fecha por parte del derecho de la Unión Europea. En ambos casos, las especiales características que presentan las organizaciones de estos sectores y las implicaciones legales que genera el uso de algoritmos en el proceso decisorio en dichos contextos justifican un tratamiento más pormenorizado de estos ordenamientos jurídicos.

III. ESTRUCTURA DE LA TESIS

Este estudio se divide en cinco capítulos y termina con la exposición de las conclusiones y la presentación de la bibliografía utilizada.

El primer capítulo se centra en el fenómeno de los sistemas de toma de decisiones automatizadas. Para ello se exponen las razones que han llevado a que ahora y no antes las organizaciones cada vez acudan en mayor medida a la implementación de los algoritmos en los procesos decisorios. A través de un enfoque técnico se sintetizan y explican las fases que comprenden el ciclo de vida de los sistemas automatizados.

En el segundo capítulo se realiza una aproximación a un conjunto de elementos y nociones básicas que aparecen indicados en la normativa de protección de datos y que resultan relevantes a la hora de afrontar las implicaciones jurídicas que presenta el uso de sistemas de toma de decisiones automatizadas. Para ello, primeramente se analizan los conceptos de decisión y elaboración de perfiles. Seguidamente se realiza un acercamiento a las distintas tipologías de datos que están presentes durante el ciclo de vida de estos sistemas y que pueden tener incidencia en la normativa de protección de datos. Por último, se estudia el encaje legal de los principales agentes y sujetos

presentes durante las fases que engloban dicho ciclo de vida de los sistemas de toma de decisiones automatizadas. Las nociones extraídas de este capítulo se utilizarán como base para el desarrollo de los siguientes capítulos y apuntalarán el contenido legal del objeto de esta tesis.

El tercer capítulo pone el foco de atención en aquellas medidas de responsabilidad activa que han de establecer los responsables del tratamiento cuando pretendan desarrollar o desplegar sistemas de toma de decisiones automatizadas. En función del riesgo que presenten para los derechos y libertades de los individuos los distintos tratamientos de datos implicados, los responsables deberán establecer unas u otras medidas y salvaguardas. Como se verá, la mayoría de los tratamientos que están presentes durante el ciclo de vida de los sistemas de decisiones automatizados requerirán la implementación de prácticamente la totalidad de medidas de responsabilidad activa que prevé la normativa de protección de datos y que analizamos en esta tesis.

El cuarto capítulo se dedica al examen del conjunto de principios del tratamiento que reconoce la normativa de protección de datos. Se analizan uno a uno cada una de las reglas mínimas que todo responsable ha de ejecutar cuando lleva a cabo un tratamiento de datos. Concretamente se estudia la incidencia de estos principios durante el desarrollo y despliegue de los sistemas automatizados.

El quinto capítulo se destina al estudio de los derechos que la normativa de protección de datos reconoce en favor de los particulares. Se analiza cada uno de estos derechos y su incidencia en la toma de decisiones automatizadas. Se presta especial atención al derecho a no ser sometido a decisiones plenamente automatizadas relevantes. Este capítulo se cierra mostrando algunas de las deficiencias que ofrece la normativa de protección de datos frente al uso de sistemas automatizados y se proponen algunas soluciones para compensar estos problemas.

En definitiva, primeramente se realiza una aproximación técnica al fenómeno de la toma de decisiones automatizadas a través de algoritmos (Capítulo I). Posteriormente se lleva a cabo un acercamiento a las nociones y elementos básicos presentes en la normativa de protección de datos y que resultan elementales para este trabajo (Capítulo

II). Finalmente se efectúa un análisis pormenorizado de los tres pilares elementales sobre los que pivota la normativa de protección de datos, estos son: las medias de responsabilidad activa basadas en el enfoque del riesgo (Capítulo III), los principios del tratamiento (Capítulo IV) y los derechos de los interesados (Capítulo V).

IV. JUSTIFICACIÓN DE LA ELECCIÓN DEL TEMA DE TESIS

Las razones que nos han llevado a encarar este trabajo son de diversa índole.

En primer lugar, desde el punto de vista jurídico, hasta la fecha existe una falta de regulación general que afronte las repercusiones jurídicas del uso de sistemas automatizados por parte de las organizaciones. La protección que brinda el actual ordenamiento jurídico aparece muy limitada y compartimentada en diferentes derechos fundamentales más o menos desarrollados por diversas normas tal y como ocurre con la prohibición de no discriminación, el derecho de acceso a la información pública, el debido proceso, etc. Uno de esos regímenes jurídicos es la protección de datos personales. Tal y como hemos señalado anteriormente, este derecho fundamental, pese a que no puede encarar todos y cada uno de los riesgos que presentan la incorporación de algoritmos al proceso de toma de decisiones automatizadas, sí que puede ofrecer importantes remedios e instrumentos para proteger a las personas y desplegar un marco que favorezca la seguridad jurídica. La motivación por cubrir algunas de las lagunas legales que presenta el ordenamiento jurídico en el contexto de la toma de decisiones automatizadas nos ha llevado a encarar esta tesis tomando como punto de partida la protección de datos personales.

En segundo lugar, desde el punto de vista personal, varios motivos han condicionado la elección del tema. Por una parte, la formación del doctorando en los estudios ligados al Derecho de la Unión Europea y Derecho Público le ha empujado a cuestionarse las garantías jurídicas que pueden devenir de la protección de datos en este ámbito jurídico. Por otra parte, la incorporación del doctorando al proyecto de investigación que dirige el director de esta tesis titulado “La regulación de la trasformación digital y la economía colaborativa” catapultaron la elección de la temática

de este trabajo[1]. En este sentido, la clara orientación de los miembros de este grupo por el derecho público, concretamente el derecho constitucional y el derecho administrativo, han impulsado al doctorando al desarrollo de apartados específicos de esta tesis relacionados con la incidencia de la toma de decisiones automatizadas en el ámbito de las Administraciones Públicas. Disciplina del derecho hacia el que el doctorando pretende dirigir sus futuras investigaciones. Finalmente, la propia constatación fáctica del proceso decisorio automatizado mediante algoritmos por parte del doctorando en su día a día a través de por ejemplo los recomendadores de contenido presentes las plataformas digitales o el acceso a los servicios públicos por medio de los *chat bots* causaron en el doctorando una motivación especial por entender este fenómeno y sus repercusiones jurídicas.

V. BIBLIOGRAFÍA Y FUENTES DE INFORMACIÓN UTILIZADAS

Las fuentes que se ha utilizado en esta tesis doctoral pueden dividirse esencialmente en cinco grupos, estos son: doctrina científica, documentos e informes de diversas organizaciones, resoluciones de autoridades de protección de datos, jurisprudencia y textos normativos.

En primer lugar, por lo que respecta a la doctrina, los estudios analizados engloban un elenco muy diverso de temáticas. Con relación a la protección de datos personales, los trabajos analizados van desde manuales de referencia y estudios clásicos de este derecho fundamental hasta alcanzar aquellos más novedosos que analizan las últimas reformas legales de esta normativa. Junto al derecho fundamental a la protección de datos se han estudiado multitud de trabajos que analizan las repercusiones jurídicas del proceso decisorio en otros contextos como la transparencia algorítmica, la inteligencia artificial explicable, el debido proceso en la toma de decisiones o la prohibición de discriminación algorítmica. Finalmente se han recopilado y analizado todo un conjunto de artículos y obras centradas específicamente en los elementos técnicos relacionados con el desarrollo de sistemas algorítmicos. Estos últimos trabajos se caracterizan por el uso de un léxico muy específico y técnico lo que ha supuesto para el doctorando un plus de complejidad a la hora de abordar este fenómeno.

[1] La presente tesis doctoral se ha realizado en el marco de una beca para personal investigador en formación (PIF) dentro del proyecto titulado "La regulación de la transformación digital y la economía colaborativa" PROMETEO/2017/064 de la Generalitat Valenciana cuyo investigador principal es Lorenzo Cotino Hueso.

En segundo lugar, en los últimos años han proliferado todo tipo de documentos no vinculantes por parte de diversos organismos públicos y privados en el panorama nacional e internacional que versan sobre el uso de sistemas automatizados en el proceso decisorio. En esta tesis se han analizado y estudiado los más relevantes. Pese a ello, la cascada de información generada recientemente resulta inabarcable.

En tercer lugar se han analizado los principales documentos emitidos por las autoridades de control de protección de datos. Entre otras, han resultado muy relevantes las resoluciones del Grupo del Artículo 29, el Comité Europeo de Protección de Datos, la Agencia Española de Protección de Datos, la autoridad de protección de datos británica (ICO) o la *Autoritat Catalana de Protecció de Dades*.

En cuarto lugar, por lo que se refiere a la jurisprudencia, pese a que esta tesis se ha centrado en las sentencias dictadas por parte del Tribunal de Justicia de la Unión Europea y los tribunales nacionales que han tratado temas de protección de datos, también se han analizado otras resoluciones judiciales dictadas en terceros países tanto europeos como no europeos que versan sobre la materia objeto de este trabajo.

Finalmente, y por lo que respecta a las fuentes normativas utilizadas, cabe destacar el Reglamento General de Protección de datos y la Ley Orgánica 3/2018, de 5 de diciembre de Protección de Datos Personales y garantía de los derechos digitales. Ambos textos han sido la base sobre la que se ha desarrollado esta tesis. Junto a estas dos normas, a lo largo del trabajo se han analizado otras legislaciones sectoriales que han resultado relevantes para el desarrollo del mismo. Cabe destacar en este sentido la labor legislativa proveniente esencialmente de la Unión Europea que en lo últimos años esta ha desarrollado a través de la regulación del mercado único digital, el papel de las grandes plataformas digitales y el uso de sistemas de inteligencia artificial en la adopción de decisiones. A modo de ejemplo podemos destacar la propuesta de Reglamento europeo sobre la regulación de los sistemas de inteligencia artificial o la Carta de Derechos Digitales Española (ambos textos del año 2021). Debido al calado y la novedad que presentan algunas de estas reformas o propuestas legales, el doctorando ha tenido que ir actualizando y moldeando a golpe de novedad legislativa la estructura y configuración de la tesis doctoral que se presenta.

INTRODUCTION

I . OBJECT OF THE THESIS

The general objective of this dissertation is to analyse and clarify the legal regime applicable to automated decisions in relation to the fundamental right to the protection of personal data. Not only issues related to strictly automated decisions are addressed, but also those decisions that, without being fully automated, are used to support the decision that is finally made by a human.

The automation of the decision-making process through algorithms raises a series of legal challenges that directly affect the rights and freedoms of individuals. Data protection law is one of the main tools available to the legal system to respond to these challenges.

Our central thesis is that although the right to the protection of personal data cannot address all the challenges presented by the use of artificial intelligence systems in decision-making, the strong consolidation of the legal regime of this right at both the European and national levels favours an adequate framework of protection. Thus, the structure of this legal regime is composed of several pillars: (i) the recognition of a series of rights in favour of data subjects, (ii) the design of a set of principles applicable to all processing of personal data, (iii) the obligation imposed on organisations processing personal data to deploy a series of measures to demonstrate and ensure compliance with the rule, and (iv) the existence of a strong network of supervisory authorities in charge of ensuring that this right is respected.

Therefore, the relevant aspect of this work is to analyse this legal regime by focusing all our attention on the impact of the irruption of the automation of the decision-making process on the fundamental right to data protection through the use of algorithms. To do so, we take as a reference the two main phases that are present during the life cycle of automated decision-making systems: i) the development phase, whose main objective is to design and elaborate the decision-making system, and ii) the deployment phase, whose purpose is the implementation of the system in the

environment where it will generate its effects. The different legal elements that make up data protection regulations are analysed and interpreted taking as a reference the reality present in this type of systems. This facilitates the proposal of solutions that respect the difficult balance between the adequate protection of the rights of individuals and the development of these new technologies, thereby guiding regulatory approaches towards a new future.

II. METHODOLOGY

There are three essential research methods on which this work has been based.

Firstly, a *general analytical approach* to the phenomenon of automated decision-making has been carried out. This has obliged us to delve into matters that are not strictly legal in nature and which deal with the main technical issues related to the development and deployment of algorithmic systems, which are subsequently used for decision-making. This first phase of the work has been essential, as it has enabled us to deal more adequately with the implications that this technological reality has for the data protection legal framework.

Secondly, through the *descriptive legal method* an analysis has been made of the set of specific and general rules that data protection regulations provide for in relation to automated decisions. Among these rules, we have studied the set of rights in favour of individuals subject to such decisions, as well as the set of duties that data controllers who decide to implement these systems in their organisations must assume.

Thirdly, we have also used the *propositional legal method in the* preparation of this work. Thus, the provisions that deal with the matter indicated are interpreted and, in addition, an attempt is made to clarify those elements that are either not entirely defined in the legal body or whose wording confronts other assets and interests at stake that require a balanced interpretation. This last aspect is particularly important since the fundamental right to data protection is not an absolute right, the interpretation of this regulation and the proposals that materialise in this study take into account the rest of the interests and rights involved, which favour the organisations that use this type of algorithms.

This study focuses its attention on data protection regulations, however, other legal fields are also analysed throughout the work given the fact that the use of automated systems by all types of organisations is strongly implemented in a large

number of sectors such as labour, finance, insurance, health, law enforcement, etc. That said, some specific sections of this thesis focus on the use of automated systems in the public sector and on digital platforms, taking as a reference the regulations developed to date by European Union law. In both cases, the special characteristics of organisations in these sectors and the legal implications generated by the use of algorithms in the decision-making process in these contexts justify a more detailed treatment of these legal systems.

III. STRUCTURE OF THE THESIS

This study is divided into five chapters and ends with the main conclusions and a presentation of the bibliography.

The first chapter focuses on the phenomenon of automated decision-making systems. It explains the reasons why organisations are increasingly implementing algorithms in decision-making processes. Using a technical approach, the phases that comprise the life cycle of automated systems are summarised and explained.

The second chapter addresses a set of basic elements and notions established by data protection regulations that are relevant when it comes to facing the legal implications of the use of automated decision-making systems. To this end, the concepts of decision and profiling are first analysed. This is followed by an approach to the different types of data that are present during the life cycle of these systems and which may have an impact on data protection regulations. Finally, the legal framework of the main agents and subjects present during the phases that encompass the life cycle of automated decision-making systems is studied. The notions extracted from this chapter will be used as a basis for the development of the following chapters and will underpin the legal content of the object of this thesis.

The third chapter focuses on the active accountability measures that controllers must put in place when they intend to develop or deploy automated decision-making systems. Depending on the risk posed to the rights and freedoms of individuals by the different data processing operations involved, controllers will have to put in place different measures and safeguards. As will be seen, most of the processing operations

that are present during the lifecycle of automated decision-making systems will require the implementation of virtually all of the active accountability measures foreseen by data protection law and analysed in this thesis.

The fourth chapter examines the set of processing principles recognised by data protection law. Each of the minimum rules that a controller has to implement when carrying out data processing are analysed one by one. Specifically, the impact of these principles during the development and deployment of automated systems is studied.

The fifth chapter studies the rights that data protection law recognises in favour of individuals. Each of these rights and their impact on automated decision-making is analysed. Special attention is paid to the right not to be subject to a decision based solely on automated processing. This chapter closes by showing some of the shortcomings of data protection law in relation to the use of automated systems and proposes some solutions to compensate for these problems.

In short, first of all, a technical approach is made to the phenomenon of automated decision-making by means of algorithms (Chapter I). Subsequently, the basic notions and elements present in data protection regulations, and which are key to the development of this work, are addressed (Chapter II). Finally, a detailed analysis is made of the three basic pillars on which data protection law is structured, namely: the measures of active accountability based on the risk approach (Chapter III), the principles of processing (Chapter IV) and the rights of data subjects (Chapter V).

IV. JUSTIFICATION OF RESEARCH TOPIC

The reasons that have led us to undertake this work are diverse.

Firstly, from a legal point of view, there is to date a lack of general regulation that addresses the legal implications of the use of automated systems by organisations. The protection provided by the current legal system appears very limited and compartmentalised in different fundamental rights more or less developed by different rules such as the prohibition of non-discrimination, the right of access to public

information, due process, etc. One of these legal regimes is the protection of personal data. As we have pointed out above, while this fundamental right cannot address each and every risk posed by the incorporation of algorithms into the automated decision-making process, it can offer important remedies and tools to protect individuals and to deploy a framework that favours legal certainty. The motivation to fill some of the legal gaps in the legal system in the context of automated decision-making has led us to approach this thesis with the protection of personal data as a starting point.

Secondly, from a personal point of view, several reasons have conditioned the choice of the subject. On the one hand, the PhD student's training in studies linked to European Union Law and Public Law has led him to question the legal guarantees that may arise from data protection in this legal sphere. On the other hand, the incorporation of the PhD student into the research project directed by the supervisor of this thesis entitled "The regulation of digital transformation and the collaborative economy" was decisive in making the choice of the subject matter of this work[2]. In this sense, the clear orientation of the members of this group towards public law, specifically constitutional law and administrative law, has led the PhD student to develop specific sections of this thesis related to the impact of automated decision-making in the field of public administrations. This is the discipline of law towards which the PhD student intends to direct his future research. Finally, the factual observation of the automated decision-making process by algorithms by the PhD student in his daily life through, for example, the content recommenders present in digital platforms or the access to public services through chat bots generated in the PhD student a special motivation to understand this phenomenon and its legal repercussions.

V. BIBLIOGRAPHY AND SOURCES OF INFORMATION USED

The sources used in this doctoral thesis can be divided into five groups: scholarly works, documents and reports from various organisations, resolutions of data protection authorities, case law and regulatory texts.

Firstly, the scholarly works analysed cover a very diverse range of subjects. With regard to the protection of personal data, the works analysed range from reference

[2] This thesis has been carried out with the support of a grant within the project entitled "La regulación de la transformación digital y la economía colaborativa" PROMETEO/2017/064 of the Generalitat Valenciana whose principal investigator is Lorenzo Cotino Hueso.

handbooks and classic studies of this fundamental right to the most recent ones that analyse the latest legal reforms of this regulation. Alongside the papers and books focusing on the fundamental right to data protection, many works that analyse the legal repercussions of the decision-making process in other contexts such as algorithmic transparency, explainable artificial intelligence, due process in decision-making or the prohibition of algorithmic discrimination have also been studied. Finally, a whole set of articles and works that specifically focus on the technical elements related to the development of algorithmic systems have been compiled and analysed. These latter works are characterised by the use of a very specific and technical lexicon, which has added an element of further complexity for the PhD student when dealing with this phenomenon.

Secondly, in recent years there has been a proliferation of all kinds of non-binding documents from various public and private bodies at both the national and international levels on the use of automated systems in decision-making processes. This thesis has analysed and studied the most relevant ones. Despite this, the cascade of information generated recently is not exhaustive.

Thirdly, the main documents issued by data protection supervisory authorities were analysed. Among others, the resolutions of the Article 29 Working Party, the European Data Protection Board, the Spanish Data Protection Authority, Information Commissioner´s Office or the Catalan Data Protection Authority have been very relevant for this work.

Fourthly, as regards case law, although this thesis has focused on the judgments handed down by the Court of Justice of the European Union and national courts that have dealt with data protection issues, other judicial decisions handed down in third countries, both European and non-European, that deal with the subject matter of this work have also been analysed.

Finally, with regard to the regulatory sources used, the General Data Protection Regulation and Organic Law 3/2018 of 5 December on the Protection of Personal Data and the guarantee of digital rights should be highlighted. Both texts have been the basis on which this thesis has been developed. Together with these two regulations, other sectorial regulatory instruments that have been relevant for the development of the work have been analysed throughout the dissertation. In this sense, it is worth highlighting the legislative developments that have taken place at the European Union level in recent years with regard to the regulation of the single digital market, the role of large digital

platforms and the use of artificial intelligence systems in decision-making. By way of example, we can highlight the Artificial Intelligence Act Proposal or the Spanish Digital Rights Charter (both texts from 2021)[3]. Due to the depth and novelty of some of these reforms or legal proposals, the PhD student has had to constantly and continuously update and shape the structure and configuration of the doctoral thesis presented here.

[3] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. Resolution adopted on April 21, 2021.

CAPÍTULO I. LA AUTOMATIZACIÓN DE LAS DECISIONES

En la consecución de todo tipo de tareas los seres humanos siempre se han valido de una gran variedad de herramientas. Así, nuestros ancestros, empujados por los instintos más básicos de supervivencia rápidamente desarrollaron todo tipo de materiales para cubrir sus necesidades primarias; ropajes de animales para abrigarse, el uso de metales para cazar o los primeros arados rudimentarios son muestra de ello. Con el paso de los años nuevas técnicas se desarrollaron y permitieron a las personas poder construir, navegar o trasladarse de unos lugares a otros. Hoy en día, un simple ordenador con conexión a internet permite realizar multitud de faenas impensables hasta hace pocos años. Estos instrumentos en muchas ocasiones simplemente han facilitado la realización de dichas labores. En cambio, en otros casos, estos han llegado a sustituir la mano del hombre permitiendo que las tareas más complejas, costosas o incluso imposibles de realizar para los humanos hayan pasado a ejecutarse por máquinas y objetos inmateriales.

Pese a que este proceso de mecanización de las tareas ha estado presente a lo largo de las historia, cabe destacar dos fases donde se ha visto fuertemente impulsado. La primera tuvo lugar en la revolución industrial allá por el siglo XVIII. La misma supuso un antes y un después en el uso e invención de un gran número de artilugios que favorecieron una rápida transición que vino a sustituir los trabajos manuales por el uso de máquinas en la fabricación de todo tipo de bienes, así como en el transporte de mercancías y pasajeros. El segundo periodo de automatización de tareas y procesos comenzó aproximadamente en la mitad del siglo XX con el desarrollo de las primeras computadoras. En este sentido, y ya en la década de los noventa, la estandarización de los ordenadores permitió su accesibilidad a importantes sectores de la sociedad favoreciendo su implantación tanto en el sector privado como en el público. Ello ha derivado en la sustitución gradual del hombre por la máquina en multitud de labores[4].

Pues bien, un vestigio más de ese fenómeno donde la mecanización de las tareas realizadas por el humano ha pasado a ejecutarse por las máquinas lo encontramos con cada vez más frecuencia en los procesos de toma de decisiones de las organizaciones tanto públicas como privadas. De esta manera, si hasta la fecha era un humano el

[4] Como ejemplo, en España, la Ley 11/2007, de 22 de junio, de acceso electrónico de los ciudadanos a los Servicios Públicos fijó las bases esenciales de implantación de todos los elementos necesarios para ofrecer un servicio público a los ciudadanos a través de medios electrónicos.

encargado de decidir si una persona recibía un crédito o se le otorgaba una subvención. Ahora, la incorporación del factor máquina de forma total o parcial en estos procesos decisorios pasa a convertirse en el funcionamiento habitual en muchos ámbitos. Es decir, al incorporar estas herramientas en la toma de decisiones se accede a un ámbito que normalmente estaba vetado a las máquinas por entenderse que la complejidad de las mismas era sólo alcanzable a la consciencia humana, ya fuera por imposibilidad técnica de la máquina o incluso social, esto es, resultaba intolerable que una máquina tomara dicha decisión.

Dicho lo anterior, y aunque a día de hoy cada vez existen más sectores donde este proceso de mecanización es más patente, encontramos una serie de espacios donde las características de los mismos favorecen la incorporación de estos instrumentos.

En primer lugar, es ampliamente extendida la idea de que los sistemas de decisiones automatizadas tendrán un fuerte impacto en aquellos procesos que exijan decisiones repetitivas y sencillas[5]. En este sentido, este proceso será más incisivo en aquellos espacios donde sea relativamente fácil estandarizar las decisiones que ha de adoptar la máquina ante los inputs que esta última recibe. Ello va desde el filtrado de currículums vitae para un proceso de selección de trabajo, pasando por la automatización de las centralitas de atención al cliente y terminando con la asignación automatizada de citas en los accesos a urgencias de los hospitales[6]. El elemento de valor que en estos casos otorgaría a las organizaciones es esencialmente el de eficiencia. Dicha eficiencia favorece en muchos casos la correcta asignación de los escasos recursos de los que disponen las organizaciones tanto públicas como privadas.

En segundo lugar, la total o parcial automatización también se debe al propio desarrollo de las sociedades cada vez más dependientes del uso de nuevas tecnologías. Se podría decir que la automatización de los procesos ha llevado a la mecanización de la

[5] La toma de decisiones automatizadas encaja perfectamente en la resolución de casos cotidianos y rutinarios. En: CHIAO, V: "Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice", International Journal of Law in Context, 2019, 15, págs..133. Disponible en: https://www.cambridge.org/core/journals/international-journal-of-law-in-context/article/fairness-accountability-and-transparency-notes-on-algorithmic-decisionmaking-in-criminal-justice/635E1CB265F4F94335D2CAEBDC4D68EE

[6] Autoritat Catalana de Protecció de dades. *Intel·ligència Artificial. Decisions Automatitzades a Catalunya*, 2020, pág.40

resolución de los problemas que los primeros generan. Es decir, se combaten los problemas generado por una máquina a través del control de otra máquina[7].

En tercer lugar, las decisiones automatizadas son utilizadas por las organizaciones no para sustituir tareas que hasta la fecha venían realizando los humanos sino para resolver problemas que el humano le es imposible o muy difícil resolver debido a la complejidad de las mismas. Podemos citar por ejemplo el papel de los motores de búsqueda a la hora de autocompletar los resultados o la función de *spam* que realizan los correos electrónicos. Desde el punto de vista legal, en el plano europeo algunas propuestas normativas apuestan por el uso de técnicas automatizadas como medida para hacer frente al control de contenidos ilícitos que se vierten en el interior de las grandes plataformas digitales. Entre otras, podemos destacar el Reglamento Europeo sobre la lucha contra la difusión de contenidos terroristas en línea[8] o la Propuesta de Reglamento sobre un Mercado Único de Servicios Digitales (Ley de Servicios Digitales)[9]. A su vez y no de forma tan explícita, dicha mecanización también es patente en la Directiva (UE) 2019/790 del parlamento europeo y del consejo de 17 de abril de 2019 sobre los derechos de autor[10]. En todas estas disposiciones no existe una obligación de utilizar sistemas automatizados, sin embargo, teniendo en cuenta que la cantidad de información que transita internamente por estas plataformas es incontrolable y existe una obligación de estas plataformas para que eviten o traten de evitar la circulación de contenido ilícito, es muy probable que dichas plataformas acaben utilizando sistemas automatizados o semi automatizados de moderación y control del contenido. Se está por tanto autorizando a que sean máquinas las que lleven a cabo

[7] Un importante número de los delitos cibernéticos que se cometen en nuestros días son realizados con ayuda de sistemas de inteligencia artificial que adoptan continuamente decisiones. Para combatir estos ilícitos la media más idónea y prácticamente única en muchos casos es la de desarrollar otros sistemas de inteligencia artificial que hagan frente a los problemas generados. En: VALLS PRIETO,J: "Combating terrorist financing with artificial intelligence systems", *Repositorio universidad de Granada*, 2020, pág.9. Disponible en: https://digibug.ugr.es/handle/10481/58937

[8] Artículo 5 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.

[9] Artículos 2.o), 12.1, 15.2.c) , 17, 23.1.c), 24.c).26,1.c),29 de la Propuesta DEL PARLAMENTO EUROPEO Y DEL CONSEJO sobre un Mercado Único de Servicios Digitales (Ley de Servicios Digitales) y por la que se modifica la Directiva 2000/31/CE. Accesible en: https://eur-lex.europa.eu/legal-content/es/TXT/?qid=1608117147218&uri=COM%3A2020%3A825%3AFIN

[10] Artículo 17. 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.
Accesible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32019L0790

principalmente estas tareas ante la imposibilidad física o psíquica que tienen los humanos[11].

En cuarto lugar, y aunque las tareas puedan ser realizadas por un humano, la dificultad o el grado de incertidumbre que puede derivarse de las mismas hace que se desconfíe del criterio del humano y se potencie la supuesta objetividad de la máquina. Ello es fácilmente reconocible en aquellos sistemas que son utilizados para valorar la probabilidad de que un determinado hecho suceda o no. Dicha probabilidad emitida por la máquina se traduce en asignaciones de riesgo de impagos de un crédito[12], probabilidad de que una persona vuelva a reincidir o posibilidad de ganar en un juicio[13]. Se tratan así de sistemas de pronóstico o predictivos[14]. Es decir, estos sistemas ayudan a reducir el coste de la incertidumbre y la predicción que está presente en la adopción de decisiones que adoptan las organizaciones[15].

En quinto lugar, las decisiones total o parcialmente automatizadas también tienen encaje en todos aquellos procesos donde se requiera de respuestas rápidas para

[11] Esta incapacidad para que un humano realice ciertas tareas se deriva por ejemplo de la imposibilidad de detectar todos los posibles contenidos que contengan o inciten al odio en redes sociales. Así, de promedio, en cada minuto se publican 350,000 tuits en Twitter, se suben, 300 horas de video a YouTube y en Facebook, se publican 510,000 comentarios, se actualizan 293,000 estados y se cargan 136,000 fotos. En. MACDONALD,S; GIRO CORREJA,S Y WATKIN A,L: *International Journal of Law in Context* , 2019, 15, págs.183–197,pág.184. Disponible en: https://www.cambridge.org/core/journals/international-journal-of-law-in-context/article/regulating-terrorist-content-on-social-media-automation-and-the-rule-of-law/B54E339425753A66FECD1F592B9783A1

[12] HURLEY, Y ADEBAYO, J: "Credit scoring in the era of big data*"*, *The yale journal of law & technology*, 2017, Vol 18, págs.148-216. Disponible en: https://digitalcommons.law.yale.edu/yjolt/vol18/iss1/5 .

[13] KEHL, D, GUO,P y KESSLER,S: "Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing". *Internet & Society, Harvard Law School*, 2017. Disponible en: https://dash.harvard.edu/handle/1/33746041

[14] LATZER, M; HOLLNBUCHNER, K; JUST, N. Y SAURWEIN, F: "The economics of algorithmic selection on the Internet". WorkingPaper – Media Change & Innovation Division. University of Zurich, Zurich, 2014,pág.9. Disponible en: http://www.mediachange.ch/media/pdf/publications/economics_of_algorithmic_selection.pdf .
Otro ejemplo puede devenir de la asignación adecuada de los recursos para colectivos específicos. Así, por ejemplo, la policía de Durham, ciudad situada al noreste de Inglaterra, utiliza un modelo algorítmico denominado Hart cuyo objetivo es que personas que han cometido un delito sean susceptibles de formar parte de un programa conocido como "Check Point". Este módulo formativo y de asesoría se propone a estas personas como alternativa al procesamiento o enjuiciamiento del delito cometido. El programa ofrece intervenciones para abordar las razones subyacentes por las que cometieron el delito con el objetivo de que no vuelvan a reincidir. Depende del riesgo que establezca el algoritmo, las personas se beneficiarán o no de este programa. En: OSWALD,M; GRACE,J URWIN,S; BARNES,G: "Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality". *Information & Communications Technology Law*, Volume 27, 2018. Disponible en**:** https://www.tandfonline.com/doi/citedby/10.1080/13600834.2018.1458455?scroll=top&needAccess=true

[15]OCDE, *Artificial Intelligence in Society*, Paris, 2019, págs. 36 y 37.
Disponible en: *https://doi.org/10.1787/eedfee77-en*

resolver problemas que se presentan frecuentemente en el funcionamiento de las organizaciones. La celeridad exigida a la hora de tomar decisiones hace que los humanos difícilmente puedan trabajar con tanta velocidad convirtiéndose las máquinas a través de sus decisiones en la herramienta ideal. Así, por ejemplo, el personal presente en los controles fronterizos de los aeropuertos se ve sometido a una presión muy alta por el poco tiempo que dispone para seleccionar a las personas que posiblemente sean sospechosas, ello ha llevado a que en muchos aeropuertos acaben implantándose sistemas de reconocimiento facial.[16].

En último lugar, en otras ocasiones, los sistemas de decisiones automatizadas se convierten en una de las herramientas principales en las que se sostiene la organización. La obtención de ingresos es un fin legítimo que habilita a las empresas a utilizar estas herramientas con fines meramente económicos. Así, tales sistemas son utilizados en distintos ámbitos y con distintos roles. Por ejemplo, resulta habitual utilizar estos sistemas como método de clasificación de los distintos productos o bienes que se ofrecen en una determinada plataforma[17]. En muchos casos son incorporados en departamentos o fases concretas respecto del global de la organización como ocurre en el sector bancario o de seguros, mientras que en otros casos, el funcionamiento de estos sistemas y los datos que los alimentan se convierten en uno de los elementos centrales que soportan el sostenimiento de la organización. Ejemplo paradigmático lo podemos encontrar en plataformas sociales como Facebook o Instagram y el marketing personalizado. En estos supuestos, el objetico principal de estos sistemas y las decisiones que toman pretenden que los particulares permanezcan en estas redes el mayor tiempo posible con el correspondiente aumento de ingresos por la publicidad mostrada[18].

---

[16]European Union Agency for Fundamental Rights, *Preventing unlawful profiling today and in the future: a guide*, 2018, pág.56. Disponible en: https://fra.europa.eu/en/publication/2018/preventing-unlawful-profiling-today-and-future-guide

[17] Los servicios de intermediación en línea como Amazon o motores de búsqueda como Google basan parte de su negocio en ofrecer bienes y servicios actuando como intermediarios entre los proveedores de dichos bienes y servicios y los consumidores que accede a tales plataformas. La clasificación de los bienes y servicios que elaboran estas plataformas para mostrar tales productos son realizadas por sistemas automatizados. Véase. 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.
Accesible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32019R1150

[18] Para Andrés Boix Palop es difícil pensar que estas plataformas cambien de modelo a corto plazo. En: BOIX PALOP, A: "El control de las fake news y la calidad del debate público en las sociedades democráticas". La página definitiva. 21/03/2019. Disponible en:

Los ejemplos indicados no dejan de ser un reflejo de una realidad mucho mayor[19]. Así, la automatización de los procesos decisorios en las organizaciones está alcanzando ámbitos hasta la fecha inimaginables gracias al desarrollo de nuevas tecnologías como la inteligencia artificial que permiten la adopción de decisiones con un grado de precisión más que aceptable. Es precisamente ese grado de acierto generado por estas nuevas técnicas el que está empujando a las organizaciones a sustituir total o parcialmente el papel del humano en favor de los sistemas automatizados. En este sentido, dejar que una máquina tome decisiones que hasta la fecha venía realizando una persona es un proceso al que los humanos nos costará adaptarnos, sin embargo, dicho proceso es imparable. Resulta por ello necesario realizar un estudio de las implicaciones legales que acarrea la incorporación del factor máquina en el proceso decisorio, concretamente, nosotros nos centraremos en las consecuencias jurídicas que tienen especial incidencia en el derecho fundamental a la protección de datos personales.

Dicho lo anterior, en este capítulo inicial comenzaremos analizando primeramente los elementos básicos sobre los que descansan estas nuevas tecnologías que permiten la toma de decisiones y las razones por las que ahora han alcanzado su eclosión. Seguidamente estudiaremos las fases esenciales envueltas en el proceso decisorio, esto es, desde que se diseñan tales sistemas hasta que finalmente despliegan sus efectos. Realizado ese acercamiento o estudio técnico de este fenómeno, en capítulos posteriores, ya sí que nos adentraremos en las consecuencias jurídicas del mismo.

I. ELEMENTOS QUE INTEGRAN EL DESARROLLO Y DESPLIEGUE DE LOS SISTEMAS AUTOMATIZADOS: TECNOLOGÍA, DATOS Y PERSONAS

Para que la toma de decisiones automatizada sea eficiente se requiere de una serie de elementos que han de estar presentes tanto en el desarrollo inicial como en la implantación posterior de estos sistemas. Dichos factores son básicamente tres; en primer lugar los *datos*, esto es, el alimento que se utiliza para entrenar a los sistemas. En segundo lugar la *tecnología* utilizada, es decir, el conjunto de herramientas y técnicas

http://www.lapaginadefinitiva.com/aboix/?p=1672

[19] OCDE, *Artificial Intelligence in Society*, Paris, 2019, págs 49 y ss.

que procesan los datos. Por último y como elemento final destacan las *personas encargadas* de desarrollar y programar estos sistemas.

Resulta por tanto necesario analizar cada uno de estos elementos y su papel ya que posteriormente serán claves para entender mejor no sólo el funcionamiento de los sistemas automatizados sino lo más relevante en este estudio, es decir, las implicaciones jurídicas de estos sistemas.

1. Tecnologías presentes en los sistemas de toma de decisiones automatizadas

En las siguientes líneas analizaremos las principales técnicas en las que se apoyan los mencionados sistemas. Por un lado explicaremos el concepto de algoritmo que resulta ser la herramienta básica en la que se apoyan estos sistemas. Después estudiaremos todo un elenco de tecnologías y técnicas que se valen de los algoritmos y mediante las cuales, los sistemas automatizados son elaborados para adoptar decisiones.

A) Soporte de los sistemas. Los algoritmos

Se entiende por algoritmo la secuencia finita de reglas formales -operaciones lógicas e instrucciones- que permiten obtener un resultado de la entrada inicial de información[20]. De esta manera, ante la entrada de un determinado input, el algoritmo seguirá toda una serie de pautas previamente establecidas y emitirá un resultado o salida. Trasladado al mundo de las decisiones automatizadas, ante la entrada de un determinado dato, el algoritmo procesará ese dato y, en base al dato introducido y a las reglas prefijadas, se adoptará una decisión.

Existen multitud de algoritmos que pueden ser utilizados por las organizaciones o científicos de datos. La elección de unos u otros dependerá esencialmente del problema que se pretenda resolver y de los datos que estén disponibles. A su vez, dependiendo de la técnica utilizada, algunos algoritmos requerirán de una mayor o menor programación.

[20] Consejo de Europa. *European ethical charter on the use of artificial intelligence in judicial systems and their environment*, 2018, pág.69. Disponible en: https://www.unodc.org/ji/en/resdb/data/european_ethical_charter_on_the_use_of_artificial_intelligence_in_judicial_systems_and_their_environment_adopted_at.html
.

En este momento inicial nos interesa diferenciar solamente los algoritmos deterministas de los algoritmos no deterministas. En los primeros, todas las instrucciones y reglas que debe seguir el algoritmo para adoptar un resultado están definidas de antemano por los programadores. En los segundos en cambio, no aparecen marcadas todas las pautas durante el procesamiento de la información sino que más bien, el algoritmo goza de cierta autonomía a la hora de determinar el resultado, esto es, en el momento de adoptar la decisión. Los algoritmos deterministas son especialmente idóneos en aquellos escenarios poco variables y fácilmente definibles ya que todas las situaciones a las se enfrentará el sistema pueden predefinirse de antemano. Por el contrario, cuando el entorno en el que se pretende que el algoritmo adopte decisiones sea excesivamente dinámico e inesperado, los algoritmos deterministas dejan de ser efectivos y son suplidos por los no deterministas ya que la autonomía presente en estos últimos cubre esa incertidumbre presente en dichos escenarios permitiendo su adaptación a los cambios inesperados y constantes que sufre el entorno, los cuales, no pueden ser predefinidos de antemano.

Por ejemplo, un algoritmo integrado en un programa de calculadora de un ordenador se correspondería con un algoritmo determinista. Puede ser muy complejo, pero todas las posibles eventualidades a las que se enfrentará el algoritmo están previamente definidas e instrumentalizadas debido a que el entorno es totalmente predecible. Así, ante la entrada de cualquier operación numeral introducida por un usuario, el sistema arrojará un resultado previamente programado y predefinido.

Por otro lado, los algoritmos que operan en las plataformas digitales como puede ser YouTube necesariamente deben utilizar algoritmos no deterministas ya que resulta imposible o sumamente complejo secuenciar y establecer todas las instrucciones y reglas de todos los posibles escenarios o situaciones a las que podría hacer frente el algoritmo cuando opera en el entorno con el que interactúa. Son precisamente este tipo de algoritmos, apoyados en técnicas como el aprendizaje automático, los que en los últimos tiempos cada vez están adquiriendo mayor relevancia por la precisión de sus resultados en estos entornos no sólo complejos sino también cambiantes.

B) Tecnologías en las que se basan los sistemas de toma de decisiones automatizadas

Junto al concepto de algoritmo, es necesario hacer mención y desgranar toda una serie de técnicas y tecnologías que se valen precisamente de los algoritmos para poder ser operativas, las cuales, están presentes en la mayoría de los sistemas actuales de toma de decisiones automatizadas. Es momento de analizarlas.

b.1. Inteligencia artificial

La inteligencia artificial -en adelante IA- es una de las ciencias más recientes. Su origen data de los años 1950 cuando Alan Turing realizó una primera aproximación a este concepto a través de lo que se conoció como el Test de Turing[21]. En dicha prueba se valoraba la "inteligencia de la máquina". De esta manera, el sistema debía mantener una conversación de cinco minutos con un humano. Si la persona durante ese tiempo no lograba discernir si el diálogo era generado por un humano o por una máquina, entonces, esta última habría superado la prueba. Aunque esta teoría ha sufrido importantes desarrollos y el concepto de inteligencia artificial aglutina muchos más ámbitos, dicho estudio sentó el primer antecedente conocido de lo que hoy conocemos por inteligencia artificial[22].Dicho lo anterior, no fue hasta 1956 cuando por primera vez se hizo alusión expresa al concepto de inteligencia artificial. Concretamente, en el verano de ese mismo año, un grupo de científicos argumentaron a través del desarrollo y apoyo de varias teorías la posibilidad de que en un futuro una máquina pudiera llegar a realizar tareas que hasta esa fecha solo estaban reservadas para los seres humanos debido a la complejidad que ello podía comportar[23].

Con altibajos, esta ciencia ha ido avanzando y nutriéndose de todo un conjunto de teorías y estudios aplicados a sistemas computacionales basados en disciplinas tan variopintas como la filosofía, las matemáticas, la economía, la neurociencia, la

[21] TURING,A,M: "Computing Machinery and Intelligence" , *Mind*, vol. 59, (1 October 1950), págs. 433–460. Disponible en: https://www.cs.ox.ac.uk/activities/ieg/e-library/sources/t_article.pdf

[22] Sir irnos más lejos, hace poco Google lanzó un asistente virtual de voz en el que se demostraba que los humanos, al interactuar con el mismo, no se percataban que mantenían un conversación con una máquina. En este supuesto la máquina solicita una reserva en una peluquería para un corte de pelo. El trabajador de este establecimiento en ningún momento se percató que estaba conversando con un sistema automatizado. Accesible en: https://www.youtube.com/watch?v=NW7bTGGmDNw&ab_channel=ElMundo

[23]. En: MCCARTHY,J; MINSKY,M; ROCHESTER ,N ; SHANNON,C: "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence", 1955. Disponible actualmente en: AI Magazine Volumen 27 Number 4, 2006. https://ojs.aaai.org//index.php/aimagazine/article/view/1904

psicología, la ingeniaría computacional o la cibernética computacional[24]. A partir de 1980 muchas de las teorías en las que se basada esta ciencia empezaron a incorporarse a dispositivos industriales. Hoy día, la IA se ha convertido en uno de los motores de la actual revolución digital en la que nos encontramos sumidos. Así, conviene indicar que la IA no está compuesta por una pieza única de tecnología sino que más bien está formada por un conglomerado de varias como pueden ser el reconocimiento de patrones, la representación de conocimientos, la inferencia automática, *machine learning*, etc.

Pues bien, fruto precisamente de esa constelación de tecnologías a la que previamente aludíamos, el concepto de IA ha contemplado distintos enfoques en función del objetivo que se ha pretendido dar a los sistemas tildados de inteligentes[25].

Por un lado tendríamos el *enfoque centrado en el comportamiento humano.* De tal manera que un sistema se considera inteligente cuando realiza actuaciones que para un humano son consideradas o requieren de un elemento de inteligencia. Es precisamente este enfoque el que normalmente está más presente en el imaginario común cuando pensamos en el concepto de inteligencia artificial ya que asociamos la imagen de una máquina o robot humanoide realizando acciones propias de una persona.

Por otro lado, encontramos *el enfoque basado en la racionalidad*, esto es, se considera inteligente aquel sistema que actúa de forma racional[26]. Entendiendo racionalidad como la capacidad del sistema de elegir la mejor acción posible para alcanzar un objetivo propuesto cuando dicho sistema interactúa con el entorno en el que se pone en funcionamiento[27]. Así, a través de sus sensores, estos sistemas son capaces de percibir los datos presentes en el entorno, procesarlos y adoptar una u otra decisión.

Ambos enfoques están presentes en muchas de las definiciones que distintas autoridades y grupos de expertos han elaborado a la hora de fijar directrices o buenas

[24] RUSSELL, S Y NORVING, P: *Inteligencia Artificial. Un enfoque moderno*, Ed. Pearson Educación, 2ªed., Madrid, 2004, págs. 6 a 19.
[25] RUSSELL, S Y NORVING, P: *Inteligencia Artificial. Un enfoque moderno,* op.cit., pág.2
[26] Para muchos, el concepto de inteligencia artificial habría tenido más coherencia si se hubiera denominado racionalidad computacional, teniendo en cuenta los objetivos que busca esta tecnología. RUSSELL, S Y NORVING, P: *Inteligencia Artificial. Un enfoque moderno,* op.cit., pág.20.
[27] Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Una definición de la inteligencia artificial: Principales capacidades y disciplinas científicas*, 2019. pág.3.

prácticas en sus respectivos ámbitos para aquellas organizaciones que pretendan diseñar y hacer uso de estos sistemas inteligentes[28].

A modo de ejemplo, el concepto de inteligencia artificial basado en el enfoque del comportamiento humano podemos encontrarlo en las definiciones realizadas por diversas autoridades internacionales. Así, la autoridad de protección de datos británica establece que:

> *La Inteligencia Artificial (IA) es un término general para una gama de tecnologías basadas en algoritmos que a menudo intentan imitar el pensamiento humano para resolver tareas complejas*[29].

Por su parte, el Consejo de Europa define inteligencia artificial como:

> *Conjunto de métodos, teorías y técnicas científicas cuyo objetivo es reproducir, mediante una máquina, las habilidades cognitivas de los seres humanos. Los desarrollos actuales buscan que las máquinas realicen tareas complejas previamente realizadas por humanos*[30].

En el otro lado, el enfoque basado en la racionalidad está más presente en las definiciones establecidas por las instituciones de la Unión Europea[31]: De esta manera, el Grupo de Expertos de la Comisión Europea indica que:

> *Los sistemas de inteligencia artificial (IA) son sistemas de software (y en algunos casos también de hardware) diseñados por seres humanos que, dado un objetivo complejo, actúan en la dimensión física o digital mediante la percepción de su entorno a través de la obtención de datos, la interpretación de*

[28] En 2019 se identificaron al menos 84 documentos de distintas organizaciones públicas y privadas que contemplan principios y directrices para la inteligencia artificial que contiene pautas éticas para la IA. El 88% de estos documentos se han promulgado a partir de 2016. En.: JOBIN,A ; IENCA,M Y VAYENA,E., "The global landscape of AI ethics guidelines", *Nature Machine Intelligence*, Vol 1, Septiembre 2019, págs.392–394.

[29] Information Commissioner's Office. *Explaining decision made with AI | Part 1: The basics of explaining AI*, 2019, pág.4

[30] Consejo de Europa. *European ethical charter on the use of artificial intelligence in judicial systems and their environment*, op., cit. 2018, págs .69-70.

[31] También la OCDE se ha inclinado por este enfoque al definir el concepto de inteligencia artificial como *un sistema basado en máquina que puede, para un conjunto dado de objetivos definidos por el hombre, hacer predicciones, recomendaciones o decisiones que influyen en entornos reales o virtuales. Lo hace mediante el uso de entradas de máquina y / o basadas en humanos para: i) percibir entornos reales y / o virtuales; ii) abstraer tales percepciones en modelos a través del análisis de manera automatizada (por ejemplo, con ML, o manualmente); y iii) usar la inferencia modelo para formular opciones de información o acción. Los sistemas de IA están diseñados para operar con diferentes niveles de autonomía.* En OCDE, *Artificial Intelligence in Society*, Paris, 2019, págs. 23 y 24.

los datos estructurados o no estructurados que recopilan, el razonamiento sobre el conocimiento o el procesamiento de la información derivados de esos datos, y decidiendo la acción o acciones óptimas que deben llevar a cabo para lograr el objetivo establecido[32].

Por su parte el Parlamente Europeo y el Consejo lo definen 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[33].

En tercer lugar encontramos otra serie de organizaciones que han adoptado una definición funcional. Es decir, más que precisar que ha de entenderse por inteligencia artificial hacen referencia a toda una serie de características que, estando presentes en sistemas automatizados, se le asigna el concepto de IA y las consecuencias jurídicas que ello supone. Es la definición que hasta la fecha ha adoptado el Gobierno de los Estados Unidos, así como la Comisión Europea.

Definición Funcional de Inteligencia Artificial	
Gobierno de los EEUU[34]	Comisión Europea[35]
-Sistema artificial que realiza tareas en circunstancias variables e impredecibles sin supervisión humana significativa, o que puede aprender de la experiencia y mejorar el rendimiento cuando se expone a conjuntos de datos.	-Enfoques de aprendizaje automático, incluyendo el aprendizaje supervisado, no supervisado y de refuerzo, incluyendo el aprendizaje profundo; - Enfoques basados en la lógica y el conocimiento, incluida la representación

[32] Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Directrices éticas para una IA fiable,* 2019, pág.50.

[33] 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. Resolución del Parlamento Europeo, de 20 de octubre de 2020.
Disponible en: https://www.europarl.europa.eu/doceo/document/A-9-2020-0186_ES.html#title1

[34] En el Memorándum de la Casa Blanca dirigido a las agencias federales sobre cómo regular las aplicaciones de inteligencia aparece una definición de inteligencia artificial. 17 de noviembre de 2020.
Disponible en: https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf
Dicha definición tiene su origen en la Section 238(g) of the John S. McCain National Defense Authorization Act for Fiscal Year 2019, Pub. L. No. 115- 232, 132 Stat. 1636, 1695 (Aug. 13, 2018) (codified at 10 U.S.C. § 2358, note). Pág.331.
Disponible en: https://www.congress.gov/115/plaws/publ232/PLAW-115publ232.pdf

[35] Artículo 3.1 y Anexo I de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

-Sistema artificial desarrollado en software de computadora, hardware físico u otro contexto que resuelve tareas que requieren percepción, cognición, planificación, aprendizaje, comunicación o acción física similar a la humana. -Sistema artificial diseñado para pensar o actuar como un ser humano, incluidas las arquitecturas cognitivas y las redes neuronales. -Un conjunto de técnicas, incluido el aprendizaje automático, diseñado para aproximar una tarea cognitiva. -Sistema artificial diseñado para actuar racionalmente, que incluye 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 conocimiento, la programación (lógica) inductiva, las bases de conocimiento, los motores de inferencia y deducción, el razonamiento (simbólico) y los sistemas expertos; -Enfoques estadísticos, estimación bayesiana, métodos de búsqueda y optimización.

En nuestra opinión, desde un punto de vista jurídico, nos inclinamos por aquellos conceptos que se basan en el enfoque racional o que adoptan una definición funcional de inteligencia artificial. Si únicamente nos centramos en el enfoque humano de la inteligencia artificial muchos sistemas que adoptan decisiones automatizadas quedarían fuera del ámbito de aplicación de los mismos debido a que no siempre los objetivos que tienen estas tecnologías corresponden con habilidades o imitaciones humanas. En este sentido, el concepto de inteligencia artificial que se pretenda implantar en los sistemas jurídicos ha de ser lo suficientemente flexible y uniforme para que pueda adaptarse adecuadamente a los progresos técnicos en los que se encuentra sumida la IA permitiendo que desarrolladores y aplicadores de estas técnicas conozcan el marco jurídico aplicable. Tal y como ha señalado el Parlamente Europeo, *para permitir que exista un enfoque regulador unificado y, por tanto, seguridad jurídica, tanto para los ciudadanos como para las empresas, es necesaria una interpretación*

común en la Unión de conceptos tales como inteligencia artificial, robótica, tecnologías conexas y reconocimiento biométrico[36].

Para finalizar con este apartado, basta indicar que todas las definiciones a las que hemos aludido tienen bajo su prisma los llamados *sistemas de inteligencia artificial débil*. Esto es, sistemas inteligentes diseñados para realizar tareas específicas y concretas. Dichos sistemas son los que actualmente están presentes en nuestro día a día y se muestran como eficientes. Quedan a priori aún lejos los *sistemas de inteligencia artificial fuertes* que prometen realizar tareas cognitivas sofisticadas similares a las realizadas por los seres humanos o excesivamente complejas[37].

b.2. Data mining

Por *data mining* o minería de datos entendemos el conjunto de procesos de forma más o menos pautada que tiene como objetivo principal extraer información desconocida y potencialmente útil presente en un conjunto de datos[38]. Una vez se obtiene esa información esta es posteriormente utilizada por las organizaciones para generar conocimiento que les permitirá resolver problemas de la vida real.

En la búsqueda de ese conocimiento escondido en las grandes bases de datos el *data mining* surgió como una técnica capaz de superar los métodos tradicionales de análisis de datos como la estadística. Ello supuso un salto cualitativo en el descubrimiento de información oculta[39]. Así, gracias a esta técnica, un analista, a la hora de analizar datos y tratar de obtener conclusiones no tendrá que partir siempre de una hipótesis inicial sino que más bien dejará que el propio análisis la genere. Esto último

[36] Considerando 7 de 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. Resolución del Parlamento Europeo, de 20 de octubre de 2020.

[37] Australian Human Rights Commission. *Human rights and Technology. Discussion Pape*r, 2019, pág. 60.

[38] Desde un punto de vista académico el data mining es una etapa dentro de un proceso más amplio conocido como la *extracción de conocimiento en bases de datos* (*Knowledge Discovery in Databases* o KDD). En MOLINA FÉLIX, L C: *"Data mining:* torturando a los datos hasta que confiesen", *UOC*, 2002, pág.3.

[39] En muchos establecimientos la minería de datos ha sustituido a las encuestas de evaluación de los clientes como método para detectar las posibles preferencias de estos. En: HUEI,KO,Y ; YU HSU,P; , MING SHIEN CHENG,M ; RUEI JHENG; CHAO LUO,Z: "Costumer Retention Prediction with CCN". En: TAN, Y SHI,Y: *Data Mining and Big Data*. 4th International Conference, DMBD, 2019, pág.105.
Disponible en: https://link.springer.com/chapter/10.1007%2F978-981-32-9563-6_11

presenta sus riesgos ya que las correlaciones que puede generar el sistema pueden ser erróneas y por tanto, aplicadas en la vida diaria, generar decisiones injustas[40].

Aunque no existe una hoja de ruta común para los analistas de datos a la hora de llevar a cabo un proceso de *data mining*. Es generalmente aceptada la existencia de una serie de fases más o menos definidas. Estas son[41]: i) *recopilación de datos*, ya sean datos o no estructurados, ii) *pre procesamiento de datos*, donde los datos recopilados se limpian y transforman en un formato estandarizado para que se puedan procesar. Esta fase es elemental debido a que muchos de los datos que se recopilan no presentan un formato adecuado para su posterior procesamiento y, iii) *procesamiento de los datos*, en la cual se utilizan distintas técnicas y algoritmos para obtener el conocimiento, creando así el modelo que posteriormente se implantará en la organización. Como ha señalado la doctrina, a través de este proceso los datos son torturados a través de distintas herramientas con el objetivo de que confiesen el conocimiento que esconden[42]. Con la intención de estandarizar en parte el proceso de *data mining*, se han desarrollado metodologías que establecen una serie de pautas a seguir por parte de los analistas de datos. La metodología que históricamente ha sido más utilizada es el CRISP-DM[43].

Por otro lado, y en función del problema que se pretende resolver y el tipo de análisis que se realice con los datos podemos distinguir dos tipos de métodos o tareas:

Tareas o métodos predictivos: son aquellos cuyo objetivo es estimar valores futuros o desconocidos de algunas variables de interés a partir de otra variable independiente. Dichas tareas son esencialmente la clasificación y la regresión.

La tarea de clasificación se utiliza para categorizar los datos de entrada en el sistema, como por ejemplo, clasificar un correo electrónico como *spam*. Es decir, los datos son previamente etiquetados y posteriormente son procesados. Una vez el sistema es capaz de extraer las principales características de los grupos de datos etiquetados, el sistema puede clasificar correctamente nuevos datos que él nunca antes había visto.

---

[40] MEESE, J. & JAGASIA, P. Y ARVANITAKIS, J: "Citizen or consumer? Contrasting Australia and Europe's data protection policies". *Internet Policy Review*, 2019, 8(2). pág.34.

[41] AGGARWAL,CH: *Data Mining: The Textbook*. Ed. Springer. 2015, pág.5.

[42] LIANE,C: "*A Taxonomy and Classification of Data Mining"*, *SMU Science and Technology Law Review*, 2013,pág.4. Disponible en: https://scholar.smu.edu/scitech/vol16/iss2/4/

[43] Esta metodología describe una serie de pasos específicos que han de seguir los analistas de datos que pretenden obtener conocimiento de tales datos. Concretamente los pasos a seguir son: comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación e implantación. Este modelo ha sufrido distintas actualizaciones. En: CHAPMAN, P; CLINTON,J; KERBER,R; KHABAZA,T; REINARTZ,T; SHEARER,C ; WIRTH,R: *Step-by-step data mining guide*, 2000.

En segundo lugar, *la tarea de regresión* puede utilizarse para estimar la salida deseada, por ejemplo, estimar las ventas de un producto en un lapso de tiempo. Esta tarea es útil para resolver problemas que exigen como resultado el cálculo de una probabilidad.

Tareas o métodos descriptivos. Su principal objetivo es revelar y explicar los patrones que están presentes en el conjunto de datos que se analiza. Se distinguen dos tareas, la de agrupamiento y la de asociación.

La *tarea de agrupamiento* supone dividir un conjunto de datos en varios grupos basados en ciertas similitudes predeterminadas, mostrando a su vez una clara diferenciación entre cada uno de los grupos[44]. Ejemplo: Analizar comportamientos comunes del conjunto de clientes de una organización. Para ello, la tarea de agrupamiento trata de detectar cada grupo de clientes que presenta similares características en los datos. Esto puede llevar a una empresa a que, a la hora de analizar el conjunto de datos que tiene de sus clientes, ofrezca productos distintitos a cada grupo de clientes en función de las características similares que pueden presentar los grupos detectados en los datos.

Por otro lado, *la tarea de asociación* se utiliza para descubrir relaciones existentes dentro del conjunto de datos que se analiza. Un ejemplo habitual de este tipo de tareas lo encontramos en el análisis de los tickets de compra de los clientes de un supermercado ya que a través de esta información se puede encontrar relaciones que ayuden a la tienda a comercializar sus productos de manera más efectiva en el futuro. De manera que se asocie un producto con otro.

El objetivo final de la minería de datos no es otro que el de, una vez obtenido el conocimiento, utilizarlo para todo tipo de finalidades que puedan ser útiles para las propias organizaciones que llevan a cabo el proceso o terceras que puedan considerarlo interesante.

Para ejemplificar lo indicado aludimos a un supuesto ficticio. Imaginemos que un supermercado aplica la tarea de asociación al conjunto de datos de compra históricos que ostenta y descubre que los clientes de una determinada área de la ciudad suelen adquirir de forma conjunta cacahuetes y yogures cuando visitan dicho establecimiento.

[44] GORUNESCU,F: *Data Mining: Concepts, Models and Techniques*. Ed. Springer, 2011, pág.22

Una vez obtenido ese conocimiento a través del *data mining*, esta empresa puede aprovecharlo para aumentar las ventas de dichos productos. Ahora bien, la forma de implantar ese conocimiento en la organización puede ser muy variada y las consecuencias derivadas de tal implantación distintas. Opciones: i). Se decide situar ambos productos en estanterías cercanas en el establecimiento comercial. ii). Se promocionan anuncios publicitarios donde ambos productos aparecen conjuntamente. iii). A través de la compra on-line se implanta un sistema de recomendación por el cual, ante la compra de cacahuetes, automáticamente son recomendados los yogures. iv). A través de la compra on-line, al acceder una persona cuya IP procede del área de la ciudad sobre la que se detectó la correlación, ante la adquisición de cacahuetes, se procede automáticamente a la recomendación de los yogures y, además, a un precio superior al ofrecido a otros clientes que no pertenecen a esa área de la ciudad a la que pertenece esa IP.

Este supuesto ficticio refleja las infinitas opciones que tienen las organizaciones de aplicar el conocimiento obtenido por el *data mining* y las también infinitas consecuencias que en su caso puede generar para terceros, así como las posibles normas jurídicas aplicables a cada supuesto[45].

b.3 Aprendizaje automático

Por aprendizaje automático o *machine learning* nos referimos al estudio y técnicas que tienen como objetivo aprender de un conjunto de datos existentes. En este proceso de aprendizaje son desarrollados algoritmos capaces de procesar grandes cantidades de datos y ofrecer resultados sobre dichos datos otorgando a los sistemas donde se implanta esta técnica la capacidad de aprender sin ser programados explícitamente. En función del método a través del cual los sistemas pueden aprender del conjunto de datos podemos distinguir distintas técnicas de aprendizaje automático,

[45] En la esfera de las Administraciones Públicas, las conclusiones a través del data mining pueden utilizarse para ofrecer servicios públicos personalizados. En: VELASCO RICO, C: "Personalización, proactividad e inteligencia artificial. ¿Un nuevo paradigma para la prestación electrónica de servicios públicos?". *IDP. Revista d'Internet, Dret i Política*, Núm. 30,2020, págs. 8, 9 y 10. Texto disponible en: https://doi.org/10.7238/10.7238/idp.v0i30.3226

estas son: aprendizaje supervisado, aprendizaje no supervisado, aprendizaje por refuerzo, aprendizaje semi-supervisado y aprendizaje federado[46].

Por lo que se refiere a los sistemas basados en *aprendizaje supervisado* su principal objetivo es deducir una función a partir de los datos de entrenamiento. De esta manera, el algoritmo es entrenado con un conjunto de datos de entradas y salidas previamente etiquetados. Ello le permite al algoritmo generalizar tales datos y, por tanto, ante la incorporación al sistema de nuevos datos desconocidos, este último puede ser capaz de reconocer e interpretar correctamente tales datos catalogándolos en una de las etiquetas previamente indicadas. Es decir, el conjunto de datos que utilizamos tiene un valor de salida que se desea predecir, valor de salida que previamente se ha etiquetado. A modo de ejemplo: queremos que el algoritmo aprenda a reconocer imágenes de gatos. Para ello, lo entrenamos con imágenes de gatos y otros animales, etiquetando aquellas imágenes de gatos como gatos y el resto de imágenes de animales como no gatos. Es decir, le decimos al algoritmo que imagen es un gato y que imagen no es un gato. A medida que el algoritmo es entrenado con dichas imágenes este irá generalizando poco a poco las características básicas de las mismas que hemos etiquetado como gatos. Finalmente, cuando el sistema visualice nuevas imágenes de gatos no presentes en los datos de entrenamiento, el sistema debería ser capaz de discernir entre una imagen de un gato respecto de otras imágenes de animales.

Otro ejemplo sería un modelo que utiliza un banco para establecer que usuarios pagarán o no pagarán un préstamo bancario. Para ello, la organización que quiere crear el modelo utiliza todo un conjunto de datos históricos de antiguos clientes que han o no han pagado en plazo el préstamo correspondiente. De manera que etiquetamos a cada cliente como buen o mal pagador. Cada uno de esos clientes presenta a su vez toda una serie de características que son relevantes a la hora de valorar su solvencia financiera. Una vez ingresamos las características que presenta cada cliente y lo etiquetamos como buen o mal pagador se entrena el algoritmo y este aprende de los datos extrayendo las principales características de los datos etiquetados como buenos y malos pagadores.

[46] Una definición de cada una de estas técnicas puede encontrarse en: Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, págs. 33 y 34.

Conjunto de datos históricos de clientes: Entrenamiento[47]

Ejemplos	**Características**			**Etiqueta**
	Sexo	Nivel de retribución	Modelo de coche	
Cliente 1	Hombre	1000 €	Renault Clio	Mal pagador
Cliente 2	Mujer	1200 €	Renault Megane	Mala pagadora
Cliente 3	Mujer	1850 €	Seat Ibiza	Buena pagadora
Cliente 4	Hombre	1300 €	Volkswagen Golf	Buen pagador

El objetivo final es que posteriormente, cuando se ingresen nuevos clientes no incorporados inicialmente, el sistema sea capaz de etiquetarlos como buen o mal pagador en función de las características que presentan tales clientes basado en las correlaciones previas que el sistema ha identificado en el conjunto de datos de entrenamiento. Este sistema será capaz de tomar decisiones posteriores sobre clientes nuevos a los que les asignará o no el préstamo.

Los sistemas de aprendizaje supervisado son tildados de predictivos debido a que su principal misión es la de predecir con precisión los datos nuevos que nunca ha visualizado el sistema tras aprender de los datos etiquetados.

Datos de nuevos clientes

Ejemplos	**Características**			**Etiqueta ¿?**
	Sexo	Nivel de retribución	Modelo de coche	
Cliente 1	Mujer	1200 €	Citroën C.5	¿Mala pagadora?
Cliente 2	Mujer	1500 €	Renault Cactus	¿Buena pagadora?
Cliente 3	Mujer	950 €	Seat Panda	¿?
Cliente 4	Hombre	1700 €	Audi A4	¿?

Pongamos otro ejemplo, queremos desarrollar un sistema que sea capaz de distinguir entre flores y árboles. Para ello, en primer lugar deberíamos de etiquetar todas las imágenes con las que contamos en nuestro conjunto de datos diferenciando entre flores e imágenes. Una vez que hemos etiquetado todas las imágenes, dichas imágenes se procesan por el algoritmo. Este último tiene como objetivo tratar de generalizar las principales características presentes en las imágenes de las flores por un lado y de los árboles por otro lado. Posteriormente, cuando presentemos una nueva imagen de un una flor o un árbol que no formó parte del entrenamiento, el algoritmo

[47] En el ejemplo aportado solo se utilizan 4 clientes con tres características. Sin embargo, en la vida real una entidad bancaria tiene en su haber muchos más ejemplos de clientes y además tienen en cuenta muchas más características a considerar como relevantes a la hora de valorar si una persona es mala o buena pagadora. Lo que complica la tarea de decidir las características más relevantes y el peso de cada una de ellas.

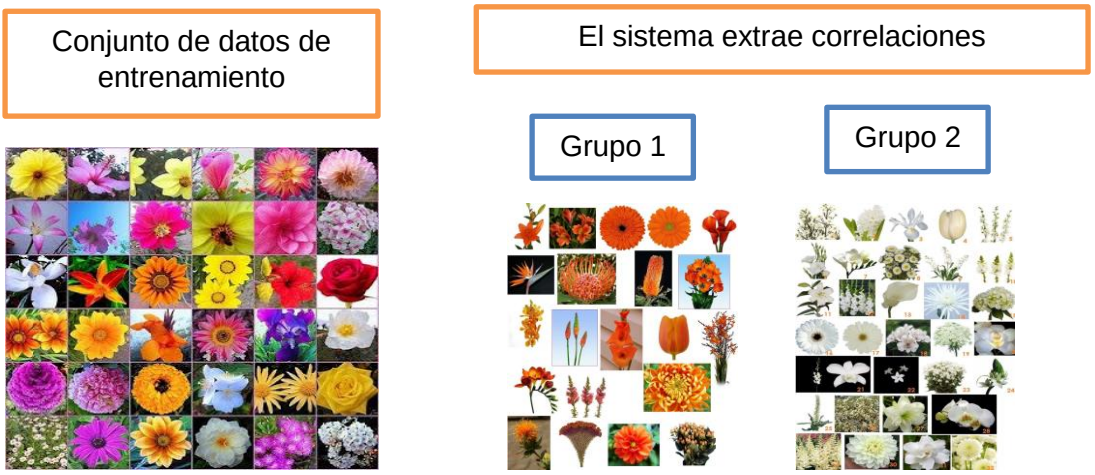
debería ser capaz de predecir correctamente si dicha la imagen corresponde a un árbol o a una flor. En nuestro caso ha predicho correctamente la imagen flor.

Elaboración propia. Aprendizaje no supervisado

Por otro lado encontramos el *aprendizaje no supervisado*, a diferencia del aprendizaje supervisado, en este método los datos que se utilizarán para que el sistema aprenda no se etiquetan previamente sino que es el propio algoritmo el que tratará de descubrir patrones y relaciones de interés presentes en el conjunto de datos. Es decir, el algoritmo se encargará de agrupar los datos en función de las características similares que presenten tales datos[48]. Este tipo de aprendizaje es sumamente útil cuando los diseñadores de los sistemas no conocen a ciencia cierta lo que pretenden buscar en los datos. Siguiendo con el ejemplo de los gatos, en este caso al algoritmo no se le indica previamente si los datos que se aportan son imágenes de gatos o no sino que el algoritmo se entrena con todas las imágenes disponibles y este es capaz de detectar aquellas imágenes que presentan características similares agrupándolas en distintos grupos. Los sistemas de aprendizaje no supervisado se definen como descriptivos ya que su principal objetivo es el de averiguar las características más relevantes presentes en los datos que se utilizan para entrenar el algoritmo.

Si continuamos con el ejemplo anterior de las flores y los árboles, en este caso al algoritmo de aprendizaje no supervisado no le decimos la etiqueta. Aquí dejamos que él mismo extraiga las principales características que pueda detectar y en su caso establezca grupos que puedan presentar asociaciones o correlaciones entre sí.

[48] The Norwegian Data Protection Authority. *Artificial intelligence and privacy*. Report, 2018, pág.9.



Elaboración propia. Aprendizaje no supervisado

A su vez, también encontramos el llamado *aprendizaje por* refuerzo. En este caso tampoco se proporcionan al algoritmo datos etiquetados sino que se deja libertad al sistema para que tome decisiones a lo largo del tiempo. De esta manera, cada vez que adopta una decisión, se le proporciona una señal de recompensa o castigo en función del acierto o error de la decisión. Visto así, el algoritmo aprende y maximiza sus resultados conforme interactúa con el entorno[49]. Un ejemplo de este enfoque aplicado a máquina puede ser el funcionamiento de una aspiradora robótica que aprende a evitar colisiones al recibir retroalimentación negativa cuando choca contra el mobiliario de una vivienda[50].

Seguidamente podemos aludir al conocido como *aprendizaje semi-supervisado*. Esta técnica se encuentra a medio camino entre el aprendizaje supervisado y el no supervisado[51]. Debido al alto coste que supone etiquetar datos. Este enfoque permite entrenar un algoritmo con una cantidad limitada de datos etiquetados y una considerable abundancia de datos sin etiquetar. Todo ello sin que el algoritmo pierda capacidad de aprendizaje. Resulta una técnica ideal en aquellas organizaciones que disponen de importantes masas de datos no etiquetadas. Es un tipo de aprendizaje automático que

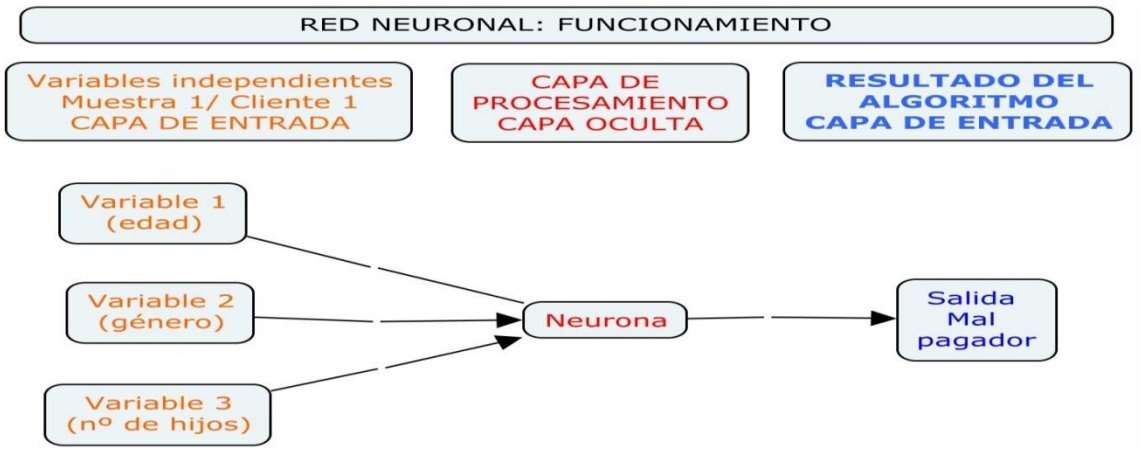
[49] La técnica del aprendizaje por refuerzo es idónea para entornos complejos. En: RUSSELL, S Y NORVING, P: *Inteligencia Artificial. Un enfoque moderno,* op.cit., pág. 868.

[50] PANESAR, A: *Machine Learning and AI for Healthcare. Big Data for Improved Health Outcomes*. Ed. Apress, Berkeley, 2021, pág.77. Disponible en: https://link.springer.com/book/10.1007/978-1-4842-6537-6#toc

[51] CHAPELLE,O , SCHÖLKOPF,B Y ZIEN,A: *Semi-Supervised Learning*, Ed.The MIT Press, Londres, 2006, pág.17.

puede procesar una amplia gama de recursos de datos pero requiere de menor intervención humana en el procesamiento previo.

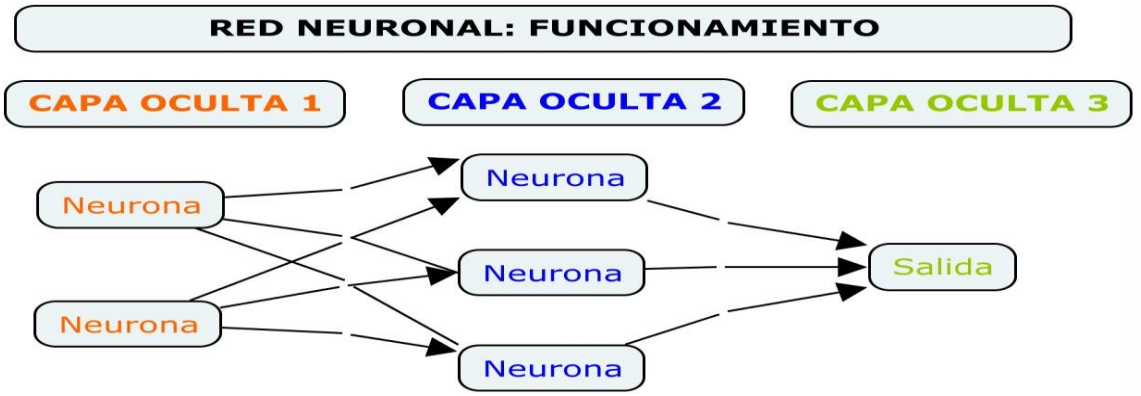
Por *aprendizaje profundo* o *deep learning* aludimos a aquellas técnicas basadas en redes neuronales que cuentan con varias capas de entrada y salida, las cuales, permiten al algoritmo aprender de la relación general que existe entre ellas[52].Se trata de un enfoque que en los últimos años ha recibido una atención especial fruto de la precisión que muestra a la hora de adoptar decisiones[53]. El *deep learning* se diferencia de los enfoques de aprendizaje automático antes mencionados por la menor intervención humana que se requiere durante el procesamiento de los datos. Así, estos algoritmos funcionan de la siguiente manera. Cada neurona -también conocida como nodo- se le ingresa un conjunto de entradas y esta emite una única salida. Es decir, las entradas están formadas por distintas variables independientes de una misma muestra y la neurona analiza y procesa ese conjunto de variables de una muestra concreta y las convierte en uno o varios valores de salida. El resultado puede ser uno o varios valores, según el número de valores que pretendemos predecir. En definitiva, lo que se busca es que haya un resultado concreto para una muestra que está conformada por un conjunto de variables independientes de esa muestra, es decir, para una fila.



Elaboración propia

[52] Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea.*Una definición de la inteligencia artificial: Principales capacidades y disciplinas científicas*, 2019, pág.6.

[53] Destaca el trabajo de GOODFELLOW, I, POUGET-ABADIE,J , MIRZA,M, XU,B , WARDE-FARLEY,D OZAIRZ,S COURVILLE,A , BENGI,Y., "Generative Adversarial Nets", 10 de junio de 2014.
Disponible en: https://arxiv.org/abs/1406.2661



Elaboración propia

Cada salida de una neurona es a su vez una entrada o input de otras neuronas que están en otras capas. Por tanto, los valores de salida de una neurona se convierten en la siguiente capa en el valor de entrada de la siguiente neurona y así sucesivamente. A más capas se establezcan mayor número de neuronas y mayor número de interrelaciones. Se genera así la llamada red neuronal. Cada neurona va otorgando un peso distinto a las variables adoptando un resultado final.

Finalmente encontramos el *aprendizaje federado o distributivo*. Se trata de una técnica recientemente desarrollada donde el análisis y entrenamiento de los datos se realiza de forma descentralizada[54]. Es un método que favorece el cumplimiento de las normas en materia de protección de datos ya que las distintas organizaciones que entrenan sus datos no las comparten con el resto. Así, el entrenamiento y análisis de los datos se realiza en los distintos *data center* donde se encuentren los datos. Una vez procesados esas bases de datos de forma separada, estas generan distintos modelos. Cada modelo es enviado a un servidor central o maestro conformado por todos los modelos que se han creado por los diversos *data center* o bases de datos. Posteriormente, ese modelo maestro es utilizado por las distintas organizaciones que en su caso analizaron los datos de forma separada. Se comparte por tanto el modelo algorítmico y no los datos. Ejemplo: Varios hospitales deciden realizar una investigación científica ayudándose de la IA para prevenir una determinada enfermedad. Para ello, cada hospital analiza de forma separada su base de datos generando distintos modelos algorítmicos. Seguidamente, cada modelo generado por las distintas bases de

[54] Autoritat catalana de protecció de dades. *Intel·ligència Artificial. Decisions Automatitzades a Catalunya*, op.cit., pág.129.

datos se envía al servidor central y se conforma el modelo maestro. Este último modelo es utilizado por los distintos hospitales[55].

b.4 Big data

Aunque no existe un concepto uniforme sobre el *Big data*. Podemos definirlo como el conjunto de técnicas que permiten el procesamiento y recopilación a gran velocidad de importantes masas de datos en distintos formatos y estructuras[56]. Hasta no hace mucho tiempo, existían límites tecnológicos que no permitían aprovechar al máximo la información presente en las grandes bases de datos. El big data permite al mismo tiempo analizar gran cantidad de datos (volumen), con diferentes formatos y estructuras (variedad) y de forma muy rápida (velocidad). Se consigue además filtrar y detectar la información más relevante (veracidad), permitiendo con ello la obtención de conocimiento escondido en las bases de datos (valor)[57]. Gracias al *Big Data* se puede conseguir un análisis y procesamiento prácticamente omnipresente de todos los datos que pueden estar circulando en un determinado entorno. Ello permite que se pueda extraer conocimiento nuevo de aquellos datos que se estudian en el momento que se generan. Ha sido en los últimos años cuando este concepto ha adquirido importancia fruto esencialmente del abaratamiento de los sistemas computacionales y de almacenamiento de datos.

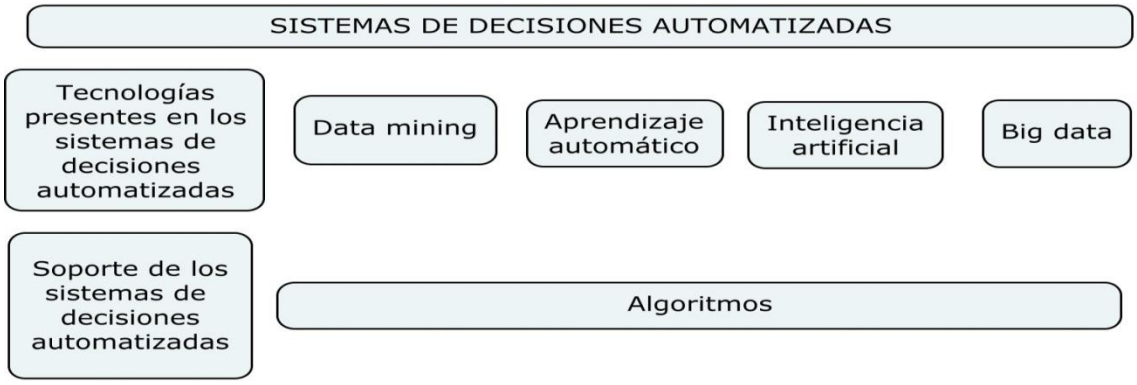
El *Big data* se compone de distintas fases a la hora de obtener conocimiento. Así, en primer lugar se recopilan y se obtienen los datos. Esta obtención se realiza a través de multitud de formas. Destaca la obtención de los datos a través de sistemas o aparatos interconectados, así como los datos generados en internet. En segundo lugar esos datos pueden o no pre procesarse. Así, es habitual que las técnicas del *Big data* permitan el análisis de los datos en bruto sin necesidad de realizar técnicas de limpieza

[55] Fuente de la noticia: GALINDO, F, J,C: "Aprendizaje federado, la técnica de IA para proteger la privacidad". *Muy interesante.* 26/05/2020. Información disponible en: https://www.muyinteresante.es/tecnologia/inteligencia-artificial/articulo/aprendizaje-federado-la-tecnica-de-ia-para-proteger-la-privacidad-361590439434

[56] Australian Human Rights Commission. *Human rights and Technology. Discussion Pape*r, op., pág.61.

[57] Las conocidas como las 5 uves, estas son: volumen, variedad, velocidad, veracidad y valor son las características o atributos que están presentes en las bases de datos sobre las que se procesan las técnicas basadas en big data. En: GIL GONZÁLEZ,E : *Big data, privacidad y protección de datos*, Ed. BOE, Madrid, 2016, págs 21 a 23. Disponible en: https://www.aepd.es/sites/default/files/2019-10/big-data.pdf . También en: ATHMAJA,S; , HANUMANTHAPPA,M ; KAVITHA,V: "A survey of machine learning algorithms for big data analytics," *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, 2017, pág.1. Texto disponible en: https://ieeexplore.ieee.org/abstract/document/8276028

de datos[58]. En último lugar se procede al procesamiento de los datos a través de potentes algoritmos, los cuales, arrojan determinados resultados que en muchas ocasiones pueden alumbrar conocimiento que hasta ese momento no se había tenido en cuenta o estaba oculto.



C) Distintas tecnologías, elemento común: la toma de decisiones automatizadas

Las técnicas que hemos descrito en el apartado anterior se han presentado de forma separada y muestran diferencias en cuanto a sus objetivos y finalidades principales. Sin embargo, existe cierto grado de interrelación entre ellas ya que presentan muchos elementos comunes. Así, todas ellas se valen de las mismas herramientas para poder operar, esto es, datos y algoritmos para procesarlos. De esta manera, a pesar de que cada una de estas tecnologías responda a distintas finalidades, parte de sus objetivos han convergido en uno común, esto es, el diseño y creación de sistemas de decisiones automatizadas. Decisiones que como ya sabemos implica en muchas ocasiones el uso de algunas de estas tecnologías y sobre las cuales se aplicará la normativa de protección de datos cuando entren dentro de su ámbito de aplicación.

Es turno de analizar algunas de las mencionadas interrelaciones.

c.1) Aprendizaje automático, data mining e inteligencia artificial

Como hemos comprobado, el surgimiento del *data mining* y el de la IA responden a intereses y objetivos distintos. El primero dijimos que trata de obtener

[58] Existen multitud de herramientas para analizar datos masivos de datos, indicamos entre otras Apache Hadoop y Apache Spark. En: SANTOS, P: "Apache Spark VS Hadoop Map Reduce". 2019. Disponible en: https://openwebinars.net/blog/apache-spark-vshadoop-map-reduce/

conocimiento de los datos presentes en un conjunto de datos. El segundo busca crear sistemas inteligentes capaces de adoptar decisiones racionales o adecuadas. Pues bien, a pesar de estas diferencias, ambas técnicas utilizan el *machine learning* para cumplir parte de los fines en los que se apoyan. Así, aunque el aprendizaje automático es un sub-campo de la IA que surgió del esfuerzo para construir máquinas que aprendieran y fueran capaces de adaptarse a situaciones imprevistas. El *data mining* se ha valido y se vale de dicha técnica para desarrollar de forma más incisiva la búsqueda de patrones ocultos de los datos.

Además, *data mining* e IA se entrecruzan en la medida que el desarrollo de una tecnología se apoya en la otra. En este sentido, la inteligencia artificial apoyada en los algoritmos de aprendizaje automático también se vale de algunos elementos fuertemente desarrollados en la minería de datos como pudiera ser los enfoques basados en la resolución de problemas o las técnicas de clasificación o agrupamiento[59]. A su vez, el *data mining* también ha aprovechado otras técnicas propias de la IA para desarrollar aún más sus objetivos[60].

Es decir, aunque los fines de estas tecnologías son distintos, ambas se valen del aprendizaje automático para lograr parte de los mencionados objetivos.

c.2) Data mining y big data

La minería de datos y el *Big data* tienen un elemento en común, ambas tratan de obtener conocimiento de los datos. Ahora bien, se podría decir que en el *Big data* supera a las técnicas del *data mining* tradicionales en la medida que permite el procesamiento de una gran volumen y variedad de datos que la minería de datos tradicional no está capacitada. Así, mientras que el *data mining* está sobre todo configurado para analizar grandes bases de datos presentes en una organización de forma estática. Las herramientas de *Big data* permiten un análisis dinámico de los datos en distintos formatos y variables.

[59] En el ámbito científico también está patente esa interrelación, siendo habitual que expertos de estas técnicas acudan a las principales conferencias internacionales de estas tecnologías.. IEEE ICDM International Conference on Data Mining, ICML International Conference on Machine Learning, NIPS Neural Information Processing Systems.

[60] OLSON,D Y DESHENG DASH,W: *Predictive Data Mining Models*. Ed. Springer, 2017, Singapur, pág. 71. También en: KOTU, V Y DESHPANDE,B:.*Predictive Analytics and Data Mining : Concepts and Practice with RapidMiner*, Ed. Elsevier Science & Technology, 2015, pág 4.

Dicho esto, a día de hoy ambas definiciones se entremezclan y en muchos casos *Big data* y *data mining* se utilizan de forma indiferenciada. Es por ello que las fronteras entre estas técnicas sean en determinadas ocasiones un tanto difusas. Aunque es un poco osado señalarlo por alguien que no es un experto en la materia. Se podría decir que el *Big data* es una evolución del *data mining* debido a las posibilidades que a día de hoy brinda el procesamiento masivo de datos que no estaba disponible en la época en la que se desarrolló la minería de datos.

c.3) Big data e inteligencia artificial

Curiosamente, *Big data* e IA son dos tecnologías que van unidas de la mano y se necesitan la una a la otra. Así, para la IA, el *Big data* se convierte en un aliado perfecto al habilitarle una mejor disponibilidad y pre procesamiento de ingentes cantidades de datos que muchas herramientas basadas en IA necesitan para operar en entornos altamente variables e impredecibles. A su vez, el propio *Big data* se ve necesitado de los avances que se van generando en el propio progreso de la IA al desarrollarse cada vez mejores técnicas que facilitan el propio procesamiento de las ingentes cantidades de datos que se recopilan a través del *Big data*.

c.4) Resumen: Algoritmos, datos y tecnologías

Los algoritmos necesitan de inputs para poder operar, concretamente necesitan datos para poder arrojar determinados resultados. Gracias a las técnicas del *Big data* o la minería de datos estos datos pueden ser recopilados y posteriormente procesados por los algoritmos. Tras el procesamiento de estos datos los sistemas basados en inteligencia artificial pueden arrojar resultados altamente fiables que permiten a las organizaciones incorporarlos a los procesos decisorios de toma de decisiones.

2. Datos

Todas las tecnologías a las que hemos hecho previamente referencia se ven necesitados inexcusablemente de un elemento básico para que puedan operar, esto es, los datos. Estos se convierten en el *combustible* necesario para que dichos sistemas puedan adoptar decisiones eficientes y adecuadas. Como ahora veremos, la necesidad

del dato no sólo está presente durante la fase del diseño de estos sistemas sino que su protagonismo también es palmario cuando el sistema se implanta en el entorno donde adoptará decisiones.

El ansia por recopilar y obtener todo tipo de información no es un elemento característico de la sociedad de hoy día. La información siempre ha sido poder. En este sentido, tanto organizaciones públicas como privadas siempre se han afanado en recopilar y utilizar la información para todo tipo de fines que pudieran ser más o menos legítimos. Precisamente, en parte, el germen del derecho fundamental a la protección de datos tanto en Europa como la *privacy* en EEUU surgió como respuesta a los peligros que se empezaban a detectar en los usos que podían realizar las grandes organizaciones eminentemente públicas de los datos de los que disponían[61].Producto de esa preocupación destacan la *Privacy Act* de 1974 en EEUU o el Convenio 108 del Consejo de Europa sobre tratamiento de datos personales[62].

Así, aunque la información y los datos siempre han sido relevantes. Lo que ocurre ahora es que esa ambición por obtener más datos se ve acrecentada por el hecho de que las nuevas técnicas a las que previamente hemos aludido permiten descubrir y obtener conocimiento no detectado a través de un simple análisis. Es precisamente esa la gran diferencia que hace que los datos hoy en día se hayan convertido en el nuevo petróleo o energía que mueve parte de la sociedad. Insistimos, no por los datos en sí, que también, sino por lo que puede extraerse de esos datos una vez se aplican las técnicas idóneas.

Proceso de Knowledge Discovery in Databases (KDD)

Fuente. Evaluando Sofwfare.com[63]

[61] En 1973 el Departamento de Salud de los Estados Unidos presentó un informe sobre las consecuencias perjudiciales que podían resultar del uso de sistemas automatizados de datos personales por parte de autoridades públicas. En este informe se propusieron una serie de recomendaciones sobre cómo debían tratarse esos datos. Visto en: Secretary's Advisory Committee on Automated Personal Data Systems. *Records, Computers and the Rights of Citizens*. July, 1973.
Disponible en: https://epic.org/privacy/hew1973report/
Este informe fue utilizado posteriormente para el desarrollo de la *Privacy Act* de 1974.

[62] Convenio 108 para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal. 1981. Consejo de Europa.

[63] Imagen obtenida en la web: https://www.evaluandosoftware.com/tecnicas-data-mining/

Este cambio de paradigma en la búsqueda de información oculta en los datos ha empujado a toda una serie de procesos y alteraciones en la sociedad por obtener y recopilar datos de toda índole.

Así, un fenómeno reciente que se deriva de este proceso es la llamada *datificación de la sociedad*. Esta última trata de cuantificar y estudiar cualquier aspecto de la vida para transformarlo en dato[64]. Es decir, cualquier acontecimiento, movimiento o hecho puede quedar registrado y convertirse en un dato perfectamente medible que ayude a entender mejor el entorno que se trata de observar. De esta manera, hoy en día es perfectamente medible y por tanto cuantificable todo tipo de situaciones hasta ahora impensables. Entre otros ejemplos: el estado de ánimo a través del número de clics del ratón el ordenador[65], la probabilidad de sufrir un accidente a través de la forma de conducción, la posibilidad de sufrir un apagón en base a los consumos eléctricos de una vivienda[66], los flujos de movimiento y seguimiento en las redes sociales, número de pasos realizados en un día, la temperatura del cuerpo, etc.

Por otro lado, esta obsesión por obtener todo tipo de datos y cuantificar cualquier elemento de la sociedad ha empujado a un número ingente de empresas a tratar de obtener datos a cualquier pecio. En los últimos años han proliferado los llamados *data bróker* o vampiros de datos. Estas organizaciones se dedican a recopilar todo tipo de datos de personas para después venderlas a terceras empresas que pueden estar interesadas en dichos datos para distintos motivos[67]. Estos mismos *data bróker* pueden realizar elaboraciones de perfiles sobre todo tipo de personas y vender las inferencias que resultan de esas aplicaciones algorítmicas. En otros casos la obtención

---

[64] Sobre el concepto de datificación véase: GIL GONZÁLEZ,E : *Big data, privacidad y protección de datos*, op.cit., pág.18. También en HERRERO SUÁREZ,C: "Big Data y derecho de la competencia". En PIÑAR MAÑAS,J,L Y DE LA QUADRA SALCEDEO,T (dir): *Sociedad digital y derecho.* Ed. BOE, Madrid, 2018, págs. 659-681, pág. 662. Disponible en: https://www.boe.es/biblioteca_juridica/publicacion.php?id=PUB-NT-2018-97&tipo=L&modo=2

[65] Resulta de interés para los titulares de un sitio web valorar si una persona muestra cabreo ante determinado contenido que visualiza. Ello puede ser medible a través de los llamados clics de la ira, es decir, repeticiones fuertes y constantes del ratón. https://onezero.medium.com/almost-every-website-you-visit-records-exactly-how-your-mouse-moves-4134cb1cc7a0

[66] Los datos de consumo eléctrico relacionadas con las mediciones referidas al consumo individual de energía eléctrica asociadas a cada punto de suministro y su código se consideran datos personales. STS Sala de lo Contencioso-Administrativo, Sección 3ª), Sentencia núm. 1062/2019 de 12 julio, FJ 3º *in fine*. Misma interpretación la encontramos en Estados Unidos. Naperville Smart Meter Awareness v. City of Naperville, No. 16-3766 (7th Cir. 2018). Disponible en: https://cases.justia.com/federal/appellate-courts/ca7/16-3766/16-3766-2018-08-16.pdf?ts=1534456820

[67] En Estados Unidos la figura de los *Data Brokers* está regulada en algunos Estados. En: LLANEZA,P: *Datanomics: Todos los datos personales que das sin darte cuenta y todo lo que las empresas hacen con ellos*. Ed. Deusto, 2019, pág.41.

de datos responde al objetivo esencial de mejorar el entrenamiento de los sistemas de inteligencia artificial.

A su vez, existen cada vez más organizaciones del sector privado que centran sus actividades en el diseño de productos que tienen como objetivo principal la recopilación de datos personales para posteriormente ofrecer publicidad personalizada. Así, en el pasado, el dato no era el objeto del negocio sino parte de los elementos necesarios para facturar los servicios[68]. Esta situación se altera cuando el valor secundario de los datos resulta más rentable que el propio servicio que se ofrece. El proceso es muy interesante. Dado que los datos se convierten en el negocio que sustenta a esas organizaciones, estas deben ofrecer un producto atractivo y "gratuito" que permita a los usuarios facilitar el mayor número de datos durante el mayor tiempo posible[69]. Gran parte de esos datos por tanto se obtiene inicialmente con el objetivo de brindar un servicio, pero en su esencia, lo que se pretende es obtener esos datos para todo tipo de finalidades más o menos legítimas y que en gran parte tienen como objetivo el de utilizar los datos para fines secundarios para los cuales fueron recopilados. Este modelo está fuertemente implantado en el comercio electrónico donde se ofrece multitud de contenidos y servicios digitales a cambio de datos. Ello ha llevado al legislador europeo a regular en parte la relación jurídica creada por este fenómeno[70].

Los datos personales y no personales se convierten en un activo de las organizaciones desde el momento en el que dichos datos dejan de ser únicamente utilizados para prestar el servicio principal para el que fueron recopilados. Se otorga por tanto una ventaja competitiva a aquellas empresas que los utilizan para obtener nuevo conocimiento y además, aplicar ese conocimiento al despliegue de sistemas que adoptarán decisiones automatizadas en dicha organización. No es casualidad que las multinacionales que más datos tanto personales como no personales albergan en su poder sean a su vez aquellas que más capacidad de negocio ostentan en la actualidad[71]. En el sector público, casos como el que destapó Eduard Snowden[72] o el sistema de

---

[68] LLANEZA,P: *Datanomics: Todos los datos personales que das sin darte cuenta y todo lo que las empresas hacen con ellos,* op.cit., pág.15.

[69] Las grandes plataformas y redes sociales basan parte de su negocio en esta operatoria. A su vez, multitud de Apps de todo tipo también presentan este modelo de negocio.

[70] Véase el Considerando 24 y artículo 3 de la DIRECTIVA (UE) 2019/770 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 20 de mayo de 2019 relativa a determinados aspectos de los contratos de suministro de contenidos y servicios digitales

[71] Las 4 big: Google, Microsoft, Amazon, Apple y Facebook. En esta ecuación se suma tecnología y uso de datos.

[72] Eduard Snowden denunció el control masivo de la información al que se someten los ciudadanos de Estados Unidos y Reino Unido a cargo de la NSA. Visto en: https://citizenfourfilm.com/ Concretamente,

crédito social utilizado en China también reflejan la relevancia del uso secundario de la información que pueden hacer las Administraciones Públicas[73].

Recopilador del dato	Finalidad Inicial	Finalidad secundaria
Policía	Plantear una denuncia por robo	Detectar denuncias falsas[74]
Servicio de salud	Atención sanitaria	Investigación sanitaria
Red social	Alta en la red social	Evaluar la solvencia financiera del usuario[75]

De lo dicho anteriormente se desprende la dependencia e importancia del dato en la sociedad de hoy en día y en concreto en todas las organizaciones que pretenden desarrollar o implantar sistemas de decisiones automatizadas en sus procesos internos. Es por ello que la recopilación y tratamiento de dichos datos deberán realizarse con las mayores garantías posibles para evitar posteriormente decisiones injustas e inadecuadas[76]. En los siguientes capítulos se tratarán los elementos jurídicos aplicables a estas situaciones.

---

en Reino Unido, la interceptación masiva de comunicaciones e intercambio de inteligencia entre estos países se ha declarado contraria al Convenio Europeo de Derechos Humanos por parte del Tribunal Europeo de Derechos Humanos. Véase la sentencia CASE OF BIG BROTHER WATCH AND OTHERS v. THE UNITED KINGDOM *(Applications nos. 58170/13, 62322/14 and 24960/15*) de 25 de mayo de 2021. Disponible en: https://hudoc.echr.coe.int/eng#{%22documentcollectionid2%22:[%22GRANDCHAMBER%22,%22CHAMBER%22],%22itemid%22:[%22001-210077%22]}

[73] El gobierno chino está implantando un sistema de crédito social en varias ciudades. Cada ciudadano tendrá un puntaje que se verá afectado por la evaluación y el seguimiento constante de la vida diaria de las personas a través de las redes digitales. El objetivo principal del sistema de crédito social es fomentar la honestidad entre los ciudadanos. En: LI XAN WONG, K Y SHIELDS DOBSON,A: "We're Just Data: Exploring China's Social Credit System in Relation to Digital Platform Ratings Cultures in Westernised Democracies." *Global Media and China* 4, no. 2, June 2019, págs. 220–32.
Disponible en: https://doi.org/10.1177/2059436419856090.

[74] En 2018 la el cuerpo de policía nacional de España puso en marcha VeriPol. Esta aplicación detecta las denuncias falsas interpuestas en casos de robos con violencia e intimidación. Dicha herramienta analiza el texto de la denuncia e indica si dicha denuncia es o no falsa. Para desarrollar este sistema la policía nacional entrenó un modelo basado en miles de denuncias. Unas falsas y otras verdaderas. Visto en: http://www.interior.gob.es/ca/prensa/noticias/-/asset_publisher/GHU8Ap6ztgsg/content/id/9496864

[75] En un informe reciente del FMI se destaca la idea de que las redes sociales se pueden convertir en el futuro en los nuevos protagonistas del mundo en la intermediación financiera gracias a la recopilación masiva de datos que obtienen de sus plataformas. Así, través de la aplicación de técnicas de aprendizaje automático las plataformas digitales pueden llegar a predecir la solvencia financiera de una persona con mayor fiabilidad que una entidad bancaria. En: BOOT A; HOFFMANN,P ; LAEVEN,L ; RATNOVSKI,L: "Financial Intermediation and Technology: What's Old, What's New?", *Fondo Monetario Internacional*, WP/20/161, 2020, pág.11.

[76] Uno de los principales problemas es que no sabemos que está sucediendo realmente con la recogida de datos y su procesamiento. En: AI 360 | COPENHAGEN workshop, *Human Brain Project***,** 2019, pág.21 Disponible en: https://tekno.dk/article/ai-360-steering-ai-for-societal-benefit/?lang=en

3. El papel de las personas

Junto a la tecnología y los datos, el tercer factor que ha de estar presente en el uso de los sistemas automatizados es el elemento humano. Es decir, independientemente de la cantidad de los datos que se ostenten y de las tecnologías con las que cuenten las organizaciones para procesarlos, el equipo humano encargado del desarrollo de estos sistemas es sumamente relevante. Ello es así porque su impacto durante las distintas fases que engloban el diseño y despliegue de estos sistemas influirá decisivamente en cómo estos posteriormente funcionarán, es decir, cómo adoptarán unas u otras decisiones.

Por citar algunas improntas del equipo de trabajo en estas fases. En lo que respecta a la fase de desarrollo o diseño, las decisiones más relevantes de los programadores se centrarán en marcar los objetivos o problemas que se pretende resolver con la implantación del sistema, también resulta importante la elección de las variables que se utilizarán y el algoritmo o algoritmos que procesarán tales datos. A su vez, y en lo concerniente a la fase de despliegue, también se debe valorar el papel que tendrán las personas encargadas de interpretar los resultados dictados por la máquina, tanto si la decisión es total como parcialmente automatizada.

De esta manera, aunque el proceso decisorio en la toma de decisiones tienda a automatizarse en las organizaciones, el grado de implicación de las personas durante su diseño denota que tales sistemas en ningún momento pueden considerase neutrales ya que las improntas del equipo humano que está detrás de su creación incorpora a su programación aquellas suposiciones, metas, valores y objetivos presentes en los mismos[77]. Ello obliga a contar con un grupo de personas altamente cualificadas que no sólo ostenten habilidades técnicas adecuadas[78] sino además un conocimiento profundo

---

[77] LEMLEY, A,M Y CASEY,B: "Remedies for Robots." *The University of Chicago Law Review*, vol. 86, no. 5, 2019, págs. 1311–1396, pág. 1339. Disponible en: www.jstor.org/stable/26747441 En el mismo sentido, ARELLANO TOLEDO,W: "El derecho a la transparencia algorítmica en Big Data e inteligencia artificial". *Revista General de Derecho Administrativo*. 50, 2019, pág.8.

[78] Entre los perfiles técnicos más demandados encontramos entre otros a los *ingenieros informáticos* que se centran en la construcción y desarrollo de bases de datos y plataformas que generan y capturan los mismos. También es relevante el papel de los *ingenieros de telecomunicaciones* que se encargan de implantar la transmisión de los datos. Los *gestores de proyectos*, centrados en fijar los objetivos de los sistemas que adoptarán las decisiones automatizadas. Finalmente los *analistas o científicos de datos*, estos últimos velan por la calidad del ciclo de vida de los datos, analizan los datos y faciliten a los directivos de las organizaciones la generación de conocimiento. Más información en: https://datos.gob.es/es/blog/los-profesionales-de-los-equipos-de-ciencia-de-datos

del contexto del problema que se pretenda resolver[79]. Además, resulta sumamente necesario que este equipo cuente con un cierto bagaje formativo de las principales cuestiones éticas y jurídicas que pueden verse implicadas en el desarrollo de estos sistemas. En este sentido, no es que se apueste por formar a estas personas como expertos juristas. La idea más bien gira en torno a crear perfiles profesionales que no sólo se preocupen por exprimir y analizar datos en busca del hallazgos o correlaciones sino que además, también se pregunten y cuestionen en las distintas fases del diseño de estas herramientas por los principales riesgos que pueden acarrear los sistemas que en un futuro adoptarán decisiones relevantes para los individuos.

Precisamente, la demanda de perfiles y carreras profesionales que ostentan este cúmulo de cualidades cada vez es mayor por parte de los sectores que desarrollan estas tecnologías. Como respuesta a esta necesidad, en los últimos años han aumentado las titulaciones universitarias que apuestan por la formación de analistas o científicos de datos[80].

4. ¿Por qué ahora y no antes? La tormenta perfecta

Ya hemos analizado los factores básicos que están presentes en todos los sistemas automatizados. Es turno de analizar las razones por las que ahora y no antes el uso de sistemas automatizados basados esencialmente en tecnologías de inteligencia artificial está alcanzando una fuerte implantación en los distintos sectores a los que aludíamos anteriormente.

Pues bien, si bien existen distintas causas. Resumidamente podemos hacer mención a las tres que han sido esenciales. Estas son: i) mayor disponibilidad y recopilación de datos, ii) desarrollo exponencial del procesamiento de los datos y abaratamiento de almacenamiento de los mismos y, por último, iii) mejora del rendimiento de los sistemas computacionales[81].

---

[79] KUHN,M, JOHNSON,K: *Applied Predictive Modeling*. Ed. Springer Science + Business Media, New York, 2013, pág. 6.

[80] Varias universidades españolas ya cuentan con un grado de ciencia de datos entre su oferta formativa. Otras tantas ofrecen títulos de posgrado específicos.

[81] MONDAL, B: "Artificial Intelligence: State of the Art". En: BALAS V; KUMAR R; SRIVASTAVA, R; (eds): *Recent Trends and Advances in Artificial Intelligence and Internet of Things*.Ed. Springer, Cham, vol 172, 2020, págs. 389-425. Disponible en: https://doi.org/10.1007/978-3-030-32644-9_32

En primer lugar y por lo que se refiere al aumento en la *disponibilidad de datos*. Ello ya ha sido puesto de manifiesto en párrafos anteriores cuando hablábamos del poder de los datos y su importancia en nuestros días. Entre otros factores ya comentados, esta disponibilidad se ha visto favorecida en los últimos años por el desarrollo de toda una serie tecnologías y fenómenos muy relevantes. Así, la revolución digital en la que estamos envueltos ha potenciado tecnologías como la Internet de las cosas. Estos sistemas están diseñados para recibir y tratar enormes volúmenes de datos procedentes de distintas fuentes[82]. Ello ha permitido que innumerables objetos e instrumentos que incorporan estas herramientas emitan en todo momento todo tipo de datos perfectamente cuantificables gracias a la infinidad de sensores que llevan integrados[83]. Dichos datos, al ser generados en línea, permiten a las organizaciones recopilarlos y generar una imagen fiable de la realidad que pretenden analizar o moldear. Por otro lado, y aunque el origen de internet data de finales de los 90, ha sido en los inicios de este siglo cuando el acceso universal a internet ha eclosionado. Ello ha permitido a todo tipo de organizaciones ofrecer sus servicios a través de este ecosistema. Destacan en este ámbito sin lugar a dudas las redes sociales y grandes plataformas al desarrollar entornos cerrados donde gran parte de la población permanece horas y horas generando ingentes cantidades de datos dispuestos a ser procesado, analizados y posteriormente convertidos en conocimiento para ofrecer servicios y publicidad personalizada.

En segundo lugar, esta disponibilidad masiva de datos requiere de herramientas y técnicas capaces *de procesar el mayor número de datos y el almacenamiento de estos*. Así, podemos destacar la figura de los *data center* o centros de procesamiento de datos. Este tipo de espacios han proliferado en los últimos años tanto en el sector público como sobre todo en el sector privado. Muchas organizaciones cuentan con sus propio *data center* lo que les permite una gestión de los sistemas de información y datos actualizada a los servicios que ofrecen. A su vez, y gracias a la computación en nube o *cloud computing*, todo tipo de organizaciones pueden contar hoy día con una gestión y

[82]Comisión Europea. *Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica.* Publicado el 19 de febrero de 2020. Pág.3.

[83] Una definición de internet de las cosas podemos encontrarla en: GARRIGA DOMÍNGUEZ, A: "La elaboración de perfiles y su impacto en los derechos fundamentales. Una primera aproximación a su regulación en el reglamento general de protección de datos de la Unión Europea". Derechos y libertades, Número 38, Época II, enero 2018, págs.113 y 114.
Disponible en: https://e-archivo.uc3m.es/handle/10016/28000

procesamiento de sus datos sin la necesidad de contar con una infraestructura física[84]. Es decir, la gestión de sus datos se realiza en entornos virtuales. Consecuencia de ello, los costes por almacenar y gestionar los datos que posteriormente se utilizarán para desarrollar sistemas inteligentes y de toma de decisiones automatizadas se han visto drásticamente abaratados.

En tercer lugar y muy relacionado con lo señalado anteriormente, en la actualidad se cuenta con computadoras capaces de procesar la inmensa cantidad de datos que ahora están disponibles. Ello permite que dichas computadoras puedan analizar entornos altamente cambiantes y excesivamente complejos con la suficiente precisión para obtener conocimiento adecuado[85].

La conjunción de estos factores ha permitido que muchas de las tecnologías en las que descansan los sistemas de toma de decisiones automatizadas eclosionen. Técnicas como el *Big data* o la IA han encontrado su momento de esplendor ahora que cuentan con los elementos necesarios para su correcto desarrollo. Así, muchos de los enfoques teóricos sobre los que descansan algunas técnicas de la IA datan de mediados de los años 50. Sin embargo, no ha sido hasta épocas recientes cuando se está descubriendo el gran potencial que pueden ofrecer a la sociedad. Sin ir más lejos, el *machine learning* se hizo popular a partir de 2000, En cambio, sus primeras aproximaciones teóricas aparecieron mucho antes[86]. A su vez, el concepto de red neuronal aplicado a sistemas automatizados fue teorizado por vez primera en 1943[87]. No ha sido hasta la irrupción del deep learning en 2014 cuando su uso ha alcanzado

[84] La escalabilidad o elasticidad del *cloud computing* es una de las principales ventajas que presenta esta nueva tecnología ya que permite que los clientes o usuarios solo usen aquellos recursos que necesitan en el momento o periodo concreto. En: CORRALES COMPAGNUCCI, M: *Big Data, Databases and "Ownership" Rights in the Cloud*. Ed. Springer, Singapore, 2020, pág. 60. Disponible en: https://doi.org/10.1007/978-981-15-0349-8_3

[85]"Desde mediados del siglo pasado, cada generación de dispositivos hardware ha conllevado un aumento en la velocidad de proceso y en la capacidad de almacenamiento, así como una reducción de precios. La potencia de los computadores se dobla cada 18 meses aproximadamente y seguirá a este ritmo durante una o dos décadas más". En: RUSSELL, S Y NORVING, P: *Inteligencia Artificial. Un enfoque moderno*, op.cit., págs. 16 y 17.

[86] Aunque se hizo popular a mediados de la década de 2000, sus primeras definiciones aparecieron mucho antes. En 1959, Arthur Samuel definió el *machine learning* como la "habilidad de aprender sin ser programado explícitamente". En SAMUEL, A,L: "Some Studies in Machine Learning Using the Game of Checkers". *IBM Journal of Research and Development* , Volume: 3, Issue: 3, July 1959.

[87] MCCULLOCH,W Y PITTS,W., "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics,* Vol.5, 1943, págs. 115 a 133.

efectos prácticos[88]. Se ha pasado por tanto de la mera experimentación de dichas teorías a la puesta efectiva de estas con resultados cada vez más prácticos para la vida diaria en multitud de espacios. Ejemplos como la traducción de textos, el reconocimiento facial o los asistentes de voz han dejado de ser meras conjeturas de laboratorio a ser auténticas herramientas con una utilidad y funcionalidades extraordinarias. Ello ha empujado a que organizaciones privadas en una primera fase, como públicas en un momento posterior, se lancen en el desarrollo e inversión de estos sistemas dadas las funcionalidades que pueden ofrecer en distintos ámbitos[89], entre ellos, la toma de decisiones automatizadas.

A) La carrera por el desarrollo de estas tecnologías

Por tanto, aquellas organizaciones que ostenten los elementos necesarios para desarrollar las tecnologías bajo las que descansan los sistemas de decisiones automatizadas se encuentran mejor posicionadas para aprovechar sus bondades, ya sea para usarlas internamente o venderlas a terceros. En este sentido, no todas las organizaciones se encuentran en esa posición privilegiada. Como ocurre desde hace ya bastantes décadas, el sector privado se encuentra a la cabeza del desarrollo de productos tecnológicos. Los sistemas de decisiones automatizados inteligentes tampoco son una excepción.

Así, en esta nueva revolución digital se encuentran a la cabeza las grandes tecnológicas o plataformas que precisamente disponen de los factores básicos antes mencionadas. Estos son: capital humano, datos y técnicas avanzadas de procesamiento de los mismos. Grandes empresas como *Apple*, *Google*, *Microsoft*, *Amazon*, *Alibaba*, *Facebook* son a día de hoy las que canalizan gran parte de las interrelaciones que suceden en el mundo virtual entre las personas físicas y jurídicas. En ese mundo, la circulación incesante de grandes cantidades de datos personales y no personales que requieren de un procesamiento y una gestión es inimaginable. Es precisamente ahí donde radica la principal ventaja de las grandes plataformas tecnológicas en la actualidad. Son además estas últimas las que en muchos casos promocionan las historias

[88] El *deep learning* se trata de uno de los últimos avances en inteligencia artificial. Los avances en esta técnica fueron publicados en 2014 En: GOODFELLOW, I, POUGET-ABADIE,J , MIRZA,M, XU,B , WARDE-FARLEY,D OZAIRZ,S COURVILLE,A , BENGI,Y., "Generative Adversarial Nets", 2014.

[89] Desde el año 2016 la inversión por parte de empresas especializadas en IA se ha ido incrementando de forma exponencial. OCDE, *Artificial Intelligence in Society*, op., 2019, págs. 17 y 39.

de éxito de las aplicaciones de IA que ellas mismas desarrollan generando en muchos casos unas expectativas de éxito y bondades que no siempre son reales[90].

Junto a las *big tech*, tampoco se encuentran en una mala posición todas aquellas empresas privadas cuyo negocio se centra en ofrecer productos basados en IA[91]. Además, empresas dedicadas a sectores como el bancario, seguros, sanitario o eléctrico también se sitúan en un lugar privilegiado ya que desde hace tiempo dedican parte de sus inversiones al estudio y análisis de los datos a través de las técnicas del *data mining*[92].

Por otro lado, en la otra posición encontramos todo tipo de organizaciones privadas y sobre todo Administraciones Públicas que por diversas razones no disponen de los elementos básicos para desarrollar estas estructuras tecnológicas. Son precisamente las Administraciones Públicas las que debido a su incapacidad para desarrollar tecnologías punteras de procesamiento de datos y desarrollo de algoritmos potentes se ven abocadas a la adquisición de los productos ofrecidos por el sector privado. En otros supuestos, la gestión de estos algoritmos en el sector público se lleva a cabo por los propios proveedores de estos productos, los cuales, se han nutrido de los datos de los ciudadanos sobre los que las administraciones ofrecen sus servicios.

Organización	Datos	Tecnología	Capital Humano
Grandes plataformas	Ostentan un potencial de datos incalculable	Punteras en el desarrollo tecnológico	Alta especialización de su personal
Organizaciones privadas especializadas en IA	Su potencial de datos es más reducido.	Punteras en el desarrollo tecnológico	Alta especialización de su personal
Entidades que históricamente han tratado de obtener conocimiento de los datos.	Disponen de una disponibilidad alta de datos del sector al que pertenecen.	Desarrollan sistemas altamente punteros en su sector.	Cuentan con equipos especializados.
Administraciones	Ostentan un	Desarrollo muy	Déficit de

[90] Muchos de los nuevos avances en materia de inteligencia artificial son propiciados por estas tecnológicas. Además, los grandes gurús son en muchos casos trabajadores de las mismas.

[91] Junto a las grandes tecnológicas existen una serie de empresas altamente especializadas en IA como Tempus, Salesforce, Sift, Nauto, Siemens, Sherpa, Deepmind (adquirida por Alphabet Inc.) etc. Fuente: https://www.datamation.com/artificial-intelligence/top-artificial-intelligence-companies.html

[92] Estos sectores simplemente dan un paso más en la evolución del manejo de los datos, los cuales, son esenciales para prestar sus servicios esenciales. Sin perjuicio eso sí, de los nuevos riesgos que se pueden derivar del uso de estas nueva técnicas. En relación con el mercado de los seguros Véase: MUÑOZ PAREDES, M,L: "Big data y contratos de seguro: Los datos generados por los asegurados y su utilización por los aseguradores". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed. Aranzadi, Navarra, 2020, pág. 129 y ss.

Públicas	potencial de datos incalculable.	pobre de la tecnología.	personal especializado.
Otras organizaciones privadas que demandan servicios de IA	Puede o no disponer de masas de datos relevantes.	Carecen de la tecnología adecuada para integrarla en sus procesos internos.	Pueden o no tener personal especializado.

Factores que ostentan las organizaciones para desarrollar sistemas de inteligencia artificial. Elaboración propia.

Precisamente y con el objetivo de poder hacer frente a esas limitaciones en las que se encuentran muchas organizaciones del sector privado para implantar sistemas de IA adecuados, las grandes plataformas y empresas especializadas en estas tecnologías ofrecen a día de hoy un elenco de servicios excepcionales. Así, herramientas como *TensorFlow* de Google, *Microsoft Azure* de Microsoft, *Amazon web series* de Amazon o *Deep learning as a service* de IBM permiten a todo tipo de empresas contar con toda una serie de tecnologías puntera en IA y obtener rendimiento de los mismas[93]. Así, las limitaciones de computación o procesamiento se salvan a través de estos servicios. De esta manera, aunque se facilita el acceso de estas tecnologías a un número mayor de empresas y personas, a su vez se va generando una dependencia gradual en favor de las grandes plataformas. Dicho lo anterior, cabe destacar que actualmente existen toda una serie de programas y software con código fuente en abierto que permite el acceso gratuito a gran cantidad de productos que facilitan el desarrollo de sistemas IA[94].

En el plano público este déficit acaba normalmente cubriéndose con la adquisición de estos productos. Herramientas predictivas en el ámbito penal o sistemas de reconocimiento facial son ejemplos claros donde las Administraciones Públicas dependen altamente de los productos elaborados por las grandes compañías. Dicho esto, en los últimos tiempos se denota una tendencia por tratar de potenciar la colaboración público-privada en estos sectores, esto es, Administraciones Públicas y

[93]No todas las organizaciones pueden permitirse elaborar una red neuronal adecuada a sus necesidades. Ello puede deberse a un déficit de datos o limitaciones computacionales. Las grandes tecnológicas hoy ofrecen esos servicios tanto para el sector público como para el privado.
Google: https://cloud.google.com/products/ai/
Microsoft: https://azure.microsoft.com/en-gb/services/cognitive-services/
Amazon: https://aws.amazon.com/machine-learning/
IBM: https://www.ibm.com/analytics/data-science

[94] En el ámbito del almacenamiento de datos destacan las versiones gratuitas de Oracle, MongoDB y MySQL.

empresas especializadas en IA.[95]. De esta manera, cada uno de estos actores se beneficia de los principales activos que ostenta la otra y compensa el déficit o hándicap de la que carece la otra. En esta interacción, la transacción es clara. El sector público facilita datos mientras que la organización privada ofrece las herramientas para procesarlos y tratarlos. Esta colaboración es muy relevante dada la carencia actual que presenta el sector público para elaborar sus propios sistemas automatizados de decisiones[96]. En este sentido, es importante destacar que los datos que se producen en las Administraciones Públicas resultan de sumo valor para la sociedad. Es por ello que en los últimos años las autoridades europeas estén fomentando distintas iniciativas que favorecen la disponibilidad de datos públicos. Entre otros ejemplos podemos destacar la creación de espacios de datos europeos comunes en sectores claves o la actualización de las disposiciones normativas que potencia la reutilización de datos[97]. Concretamente la Directiva de datos abiertos y de reutilización de la información del sector público de 2019[98]y la Propuesta de Reglamento relativo a la gobernanza europea de datos van por este camino[99]. No obstante, la cantidad de datos que se ofrecen en los portales abiertos sigue siendo residual respecto del conjunto de datos que a día de hoy se utilizan y se analizan. En este sentido, la mayor parte de los datos siguen generándose y analizándose en el seno interno de las empresas.

---

[95] Comisión Europea. *Libro Blanco sobre la inteligencia artificial. Un enfoque europeo orientado a la excelencia y la confianza*, Bruselas, 19.2.2020, pág. 9. También en: Comisión Europea. *Una Estrategia Europea de Datos*, 19.2.2020, pág.8.

[96] Un ejemplo de la colaboración público-privado la podemos encontrar en el sistema implantado por el Gobierno de Colombia para facilitar la resolución de casos judiciales en la Corte Constitucional de Colombia. Disponible en:
https://www.corteconstitucional.gov.co/noticia.php?PRETORIA,-un-ejemplo-de-incorporacion-de-tecnologias-de-punta-en-el-sector-justicia-8970

[97] Entre los sectores claves que la Comisión Europea ha considerado claves destacan: industria, movilidad, salud, agricultura, energía. Comisión Europea. *Una Estrategia Europea de Datos*, op.cit., págs. 7 y 23.

[98] La reutilización de datos en España se encuentra hoy regulada en la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. Esta ley pronto deberá ser adaptada a las nuevas exigencias que se derivan la mencionada directiva europea. Esto es: 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. Un resumen de sus principales novedades puede encontrarse en:
https://datos.gob.es/es/noticia/aprobada-la-reforma-de-la-normativa-europea-sobre-datos-abiertos-y-reutilizacion-de-la

[99] Esta propuesta viene a complementar la Directiva de reutilización de datos previamente comentada. Se centra sobre todo en potenciar la reutilización de datos que por sus especiales características no están sometidos a dicha directiva. Concretamente son datos relativos a secretos comerciales, propiedad intelectual y datos personales.
Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos), Bruselas, 25.11.2020.

II. LA ELABORACIÓN Y APLICACIÓN DE LOS SISTEMAS DE DECISIONES AUTOMATIZADAS EN LAS ORGANIZACIONES

1. Introducción. Fases relevantes

Uno de los objetivos principales que tienen las organizaciones que desarrollan sistemas de decisiones automatizadas es que estos acaben aplicándose en la vida real. Ahora bien, antes de que un sistema de este tipo acabe implantándose en un proceso de toma de decisiones de una organización pública o privada es necesario que dicho sistema supere distintas etapas. Así, en términos generales, cabe señalar esencialmente dos fases bien diferenciadas.

Por un lado encontramos la *fase del diseño* del sistema. La cual engloba a su vez una serie de sub-etapas como la planificación del proyecto, la recogida de datos, la preparación de dichos datos, el entrenamiento de los mismos, la elaboración del sistema y en su caso la validación y conformación de este.

Por otro lado, una segunda fase se refiere al *despliegue y aplicación* de dicho sistema en la vida real. Este periodo a su vez se divide en otra serie de etapas como son; la adquisición del sistema, la incorporación de los datos, las inferencias generadas por el sistema, la decisión final del sistema y la monitorización del mismo.

2. La fase de diseño de los sistemas automatizados

El diseño de los sistemas de decisiones automatizadas engloba un conjunto de actividades o fases que pueden variar según el tipo y finalidad del proyecto. Muchas de las actividades que vamos a describir no son realizadas por todos los desarrolladores y además en muchos supuestos varían en función del tipo de problema que se pretenda resolver, el tipo de enfoque empleado o el objetivo pretendido con el proyecto. Así, no serán los mismos los pasos a seguir en proyectos de aprendizaje supervisado o no supervisado o en aquellos casos donde entra en juego el aprendizaje profundo. Tampoco es similar un proyecto que tiene como objetivo la investigación sanitaria de una determinada dolencia a un proyecto que pretende desarrollar un sistema capaz de detectar a posibles defraudadores. Además, en cada proyecto, la posibilidad de avanzar

en fases y retroceder es bastante habitual. Como veremos, cada una de las etapas analizadas se encuentra estrechamente interrelacionadas entre sí conformando un proceso circular en el cual, la superación de una de estas, no supone la imposibilidad de retornar a fases anteriores cuando se detecten errores o se requiera de evaluaciones.

El objetivo de esta fase no es otro que diseñar un modelo que posteriormente se implantará en un sistema de toma decisiones. Tal sistema tratará de representar de la forma más aproximada la realidad donde este desplegará sus efectos. Se reduce así la posibilidad de decisiones automatizadas posteriores incorrectas o poco precisas. Es por ello que la elección de las variables y datos más adecuados o el tipo de algoritmo y desarrollo del modelo resultan esenciales para que tal realidad que se pretende moldear sea la más adecuada. En este sentido, desde el punto de vista de los efectos que pueden generar estos sistemas, la fase del diseño de los mismos se muestra como la más esencial ya que las deficiencias presentes durante la misma conllevarán errores posteriores en su despliegue. En la práctica ello se traducirá en decisiones inadecuadas y por tanto, en posibles vulneraciones de derechos, incorrectas asignaciones de recursos públicos o privados, daños para las personas, etc. . En respuesta a esto, la necesidad de analizar las principales actividades englobadas en esta fase se hace sumamente necesaria para poder valorar posteriormente las respuestas y soluciones jurídicas previstas por la normativa derivada del derecho fundamental a la protección de datos.

Es turno por tanto de ir analizando una a una este conjunto de actividades que ahora mencionamos telegráficamente, estas son: planificación del proyecto, recogida y preparación de los datos, construcción de los modelos, evaluación de los modelos y validación de los mismos.

A) La planificación del proyecto

Planificar el proyecto supone el primer paso que han de encarar todas aquellas organizaciones que potencialmente pretendan desarrollar un sistema de decisiones automatizadas. Así, en esta primera fase los desarrolladores deben plantearse y fijar toda una serie de elementos que desde este momento es recomendable que se establezcan. Estos son; objetivos iniciales, papel del sistema, contexto e hipótesis:

Objetivo: resulta elemental determinar los objetivos que se pretenden con el desarrollo del sistema[100]. Como regla general, aquella organización que decide diseñar un sistema de decisiones automatizadas tiene en mente la resolución de un problema que afecta a dicha organización o que puede afectar a terceras organizaciones. En el caso de las empresas privadas estos objetivos generalmente responderán a una idea de negocio que busca la venta de ese producto a terceros compradores o también la posible incorporación del sistema a procesos internos de las mismas. Por lo que se refiere a las organizaciones públicas, como regla general, dichos sistemas se incorporarán directamente a los procesos internos de su funcionamiento sin perjuicio de su posible desarrollo y facilitación de estos sistemas a terceros. En este sentido y como ya indicamos anteriormente, cada vez será más frecuente el desarrollo de sistema de decisiones automatizadas en colaboración público-privada donde las distintas organizaciones aportan diversos elementos que ayuden a conformar el sistema y donde ambas se beneficiarán del producto creado[101].

Papel del sistema: Otro aspecto elemental es valorar la hipotética función que cumplirán los sistemas de decisiones automatizadas en la organización que los utilice. Así, no es lo mismo que un sistema se incorpore a un proceso general donde se tienen en cuenta multitud de factores para adoptar una u otra decisión a que un sistema sustituya todo un procedimiento y se convierta en el agente que adopta las decisiones. Las repercusiones y exigencias jurídicas pueden ser muy distintas y ello se debería ya de tener en cuenta desde las fases iniciales del desarrollo del sistema.

Contexto: es sumamente relevante valorar y precisar el hipotético sector o ámbito donde el sistema potencialmente irradiará sus efectos. El objetivo es que el sistema encaje adecuadamente en el entorno donde se pretende que este tome decisiones. Ello obliga a considerar las distintas situaciones generales a las que se enfrentará dicho sistema. Tal contexto estará estrechamente relacionado con el

[100] OCDE, *Artificial Intelligence in Society*, op., 2019, págs. 27 y 28.

[101] Un ejemplo de esa colaboración público privada lo encontramos en la aplicación que desarrolló la empresa especializada DeepMind y la Royal Free London NHS Foundation Trust sobre diagnóstico de lesión renal aguda. De esta manera, la Royal Free London NHS facilitó a deep mind alrededor de 1,6 millones de historiales de pacientes para crear el sistema.
En: Select Committee on Artificial Intelligence: House of Lord. *AI in the UK: ready, willing and able?* Report of Session 2017–19, HL Paper 100, pág. 90. Disponible en:
https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/10002.htm

ordenamiento jurídico donde este finalmente se implantará y los potenciales destinatarios sobre los que se tomarán las decisiones.

Hipótesis previa: Un proyecto basado en inteligencia artificial puede o no responder a una hipótesis previa. En muchas ocasiones, tales proyectos pueden estar basados en meras conjeturas o teorías no contrastadas donde un conjunto de datos es procesado en la búsqueda de correlaciones y patrones ocultos. Asimismo, existen otros supuestos donde existe una sólida teoría y el procesamiento de los datos busca corroborarla. Esto último resulta frecuente en el ámbito de la investigación donde la búsqueda de hallazgos, patrones y correlaciones ocultas en los datos presentes es consustancial a la propia idea de este sector. Creemos que tanto la hipótesis previa como en su caso la corroboración de la misma deberían quedar documentadas y registradas a efectos de facilitar futuras auditorías o rendición de cuentas del funcionamiento de los sistemas.

Pues bien, fijar desde fases tempranas claramente los objetivos, funcionalidades y contexto donde desplegará los efectos el sistema se muestra sumamente relevante por varias razones. En primer lugar ayudará a vislumbrar que conjunto de normas jurídicas serán hipotéticamente aplicables y por tanto que exigencias se requerirán a dicho sistema una vez este se implemente[102]. Ello es elemental porque se consigue que, desde el diseño del proyecto, los desarrolladores valoren en cada una de las fases que componen el desarrollo del sistema los elementos técnicos que deben tener en cuenta para que el sistema pueda amoldarse a dichas exigencias jurídicas y se evite o se reduzcan posteriormente efectos indeseados. En este sentido, decisiones como la categoría de datos que se utilizarán para el entrenamiento de los sistemas, el modelo que se elegirá para implantarlo en el sistema, el tipo de algoritmos o la presencia mayor del humano en la toma de la decisión final están estrechamente relacionados y tendrán como referencia los objetivos y contextos fijados previamente. De manera que, fijados estos, las posteriores decisiones indicadas se acompasarán a los mismos.

[102] Ya se ha resaltado en varias ocasiones la necesidad de valorar desde las primeras fases del diseño de estos sistemas los propósitos y efectos secundarios no deseados de los mismos. Concretamente si la incorporación del sistema puede afectar los derechos fundamentales o las normas básicas del Estado constitucional. Comité de ética alemán. Gutachten der Datenethikkommission, 2019,pág.164. Disponible en: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.html

A su vez, la fijación clara del contexto y los objetivos evita los costes de oportunidad que se pueden derivar de un incorrecto diseño del sistema. No porque el mismo adolezca internamente de fallos sino a causa de las exigencias jurídicas que en el contexto donde potencialmente se implante no permitan su encaje. A modo de ejemplo, un sistema puede adoptar decisiones totalmente adecuadas y precisas, pero estas, al ser totalmente inexplicables, no tienen encaje en un sector cuyo ordenamiento jurídico exija taxativamente la motivación de la decisión. En otros casos la imposibilidad podría sobrevenir porque el fin que se pretende con el despliegue del sistema es contrario al ordenamiento jurídico en sentido estricto. Por ejemplo, crear un sistema que use técnicas de reconocimiento facial para detectar la orientación sexual de una persona puede tener fuertes implicaciones jurídicas que no permitan la implementación del mismo[103]. Es decir, en estos supuestos, el obstáculo para implantar estos sistemas no derivaría de los posibles defectos que incorpore internamente el sistema sino de la imposibilidad para adaptarse al contexto jurídico en el que desplegará sus efectos.

Dicho lo anterior, no siempre será fácil marcar en fases tempranas todos los elementos indicados previamente. Sin embargo, antes se tengan claros, más probable es que la organización pueda adaptar el diseño a dichos elementos y evitar las consecuencias nefastas posteriores por un encaje inadecuado de los mismos.

B) La recopilación y obtención de datos

Los datos de los que disponen las organizaciones y los métodos que tienen para recopilarlos pueden ser muy diversos. Hasta no hace mucho tiempo la regla general era que las organizaciones disponían de un *dataset* o conjunto de datos donde quedaban almacenados grandes cantidades de datos históricos que posteriormente eran analizados en la búsqueda de información útil. Esa ha sido la práctica habitual utilizada con las técnicas del *data mining*. No obstante, en los últimos años y gracias al desarrollo de tecnologías como el *Big data*, muchos de los nuevos proyectos que se basan en técnicas de aprendizaje automático no sólo obtienen información escondida de los datos históricos de las organizaciones sino que a ello hay que sumarle el valor que se deriva del estudio y análisis de los flujos de datos que estos generan en línea. Las fuentes de

[103] La pregunta sería, ¿puede la tecnología usarse con el fin de determinar la orientación sexual de una persona? Consultado en: https://www.theguardian.com/world/2017/sep/08/ai-gay-gaydar-algorithm-facial-recognition-criticism-stanford

estos últimos datos provienen esencialmente de los llamados *ficheros logs* presentes en un gran número de aparatos digitales, los sensores que se incorporan a distintas aplicaciones e instrumentos y la extracción de datos obtenidos directamente de la red[104]. Todos estos datos, a diferencia de los datos históricos, no reflejan las pautas del pasado sino del presente. Contar con una base de datos, *data lake* o *data warehouse* adecuado resulta elemental debido a que serán dichos datos la base que sustentará el procesamiento de los mismos y en definitiva la conformación del sistema que posteriormente adoptará las decisiones automatizadas[105]. En este sentido, uno de los grandes desafíos a los que se enfrenta cualquier organización que pretende llevar a cabo proyectos de diseño de sistemas es el tipo de dato que recopila y cómo los recopila. La calidad de los datos es elemental para que posteriormente los sistemas no operen inadecuadamente. Los "datos sucios", como son frecuente llamados en los procesos de *data mining*[106], deben eliminarse ya que dichas incorreciones iniciales se traducirán posteriormente en decisiones inadecuadas. Y es que, si lo que se pretende es crear un sistema que sea capaz de moldear un determinado entorno en el cual posteriormente adopte decisiones adecuadas, los datos que tratarán de moldear el entorno deben ser los más adecuados y representativos del mismo. Basar el funcionamiento de un sistema y las decisiones que este adopte en un conjunto de datos incompletos conformará un sistema alejado de la realidad que pretende moldear, generando un sistema errático.

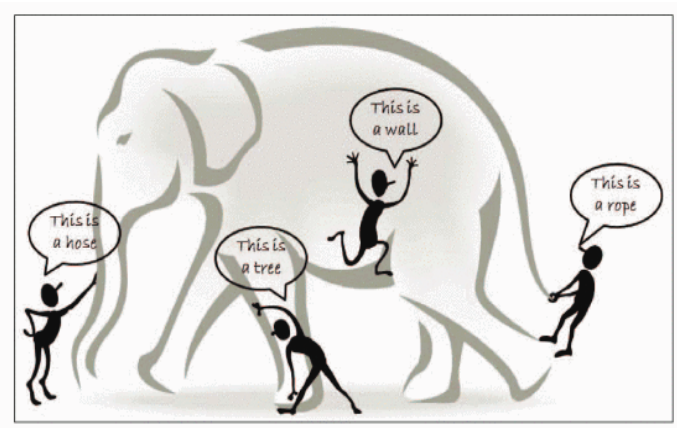
---

[104] "Los *ficheros logs* son generados automáticamente por las aplicaciones y los aparatos digitales y graban datos de las actividades que se hacen en ellos. Por ejemplo, los servidores web almacenan el número de ellos, las visitas y otras características de los usuarios". En: GARCÍA ALSINA, M: *Big data: gestión y explotación de grandes volúmenes de datos.* Barcelona: Ed. UOC, Barcelona, 2017, pág.19.

[105] Por base de datos entendemos aquella base de datos que contiene información organizada en columnas, filas y tablas que se indexa periódicamente para hacer más accesible el acceso a la información relevante. Por otro lado, por *data warehouse* o almacén de datos hemos de considerar aquella base de datos diseñada para almacenar, filtrar, extraer y analizar grandes cantidades de datos. Estos almacenes de datos están desarrollados para trabajar eficazmente con las técnicas de big data permitiendo visualizar y analizar la información de forma simultánea, sin tener que mezclar y consolidar resultados de diferentes fuentes de datos. Leído en: Bismart. https://blog.bismart.com/en/when-choose-data-warehouse-instead-of-database. Por otro lado, en los *data lakes*, la información que se almacena no está preparada y lista para el consumo, sino que se recoge en estado natural. Esto permite que los usuarios puedan dar a los datos un uso más creativo que no queda marcado por el fin para el que se han definido al momento de su carga, tal y como sucede con el data warehouse. En: Powerdata. https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/data-lake-vs-data-warehouse.-veamos-sus-principales-diferencias

[106] Por datos sucios o dirty data se engloba una serie de datos que se caracterizan por ser incorrectos, mal recopilados y en muchos casos intencionadamente manipulados. En: RICHARDSON, R; SCHULTZ, J Y CRAWFORD, K.: "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice". *New York University Law Review*, 2019, pág. 4.

Como ejemplo para describir lo señalado utilizamos prestada la imagen de abajo.



En dicha imagen podemos ver a un conjunto de ciegos analizando distintas partes de un gran elefante[107]. Cada persona ciega obtendrá distintas conclusiones del entorno que está moldeando. Esto es, las diversas partes del elefante. Sin embargo, sino cuentan con un análisis global de dicho entorno, es decir, el cuerpo completo del animal, el modelo que diseñarán puede resultar totalmente ineficiente y las conclusiones serán erróneas. Así, si lo que pretendemos es desarrollar un sistema para que irradie sus efectos en una determinada población, los datos que se deben recopilar han de ser lo más representativos de esa concreta población. De otra manera, el modelo diseñado estará lejos de aproximarse a la realidad que pretende moldear y en la cual se adoptarán decisiones. En este sentido, un modelo, por muy representativo que sea difícilmente será similar a la realidad que pretende moldear[108]. Por tanto, es necesario que los datos que se obtengan se aproximen en la mayor medida posible a ese entorno.

En definitiva, procedencia y obtención de datos se convierten en elementos relevantes a la hora de influir en las fases de desarrollo previas ya que los datos que se obtienen y recopilan son los que posteriormente se utilizarán para entrenar a los algoritmos y por consiguiente, diseñar unos u otros modelos más o menos precisos.

C) El pre procesamiento de los datos

La fase de pre procesamiento de datos comprende un conjunto de procesos que tiene como objetivo principal transformar los datos a un formato que ayude o permita al algoritmo poder trabajar adecuadamente en el objetivo previsto[109]. Es decir, una vez que

[107] Imagen obtenida en: XINDONG,W ; XINGQUAN, ZHU; GONG-QING, WU, Y WEI, D: "Data mining with big data", *Transactions on knowledge and data engineering*, vol. 26, no. 1, january 2014, pág. 2. Disponible en: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6547630

[108] SMART,A: *Más allá de ceros y unos: Robots, psicodelia y conciencia*. Ed. Clave intelectual. Madrid, 2018, pág.132.

[109] AGGARWAL,CH.: *Data Mining. The Textbook*. Ed.Springer, Nueva York, 2015, pág.5.

las organizaciones cuentan con una cantidad de datos suficiente, el siguiente paso consistirá en refinarlos para el procesamiento ulterior de los mismos.

La fase de pre procesamiento de los datos se convierte en una de las más relevantes y a la que en muchas ocasiones más tiempo le dedican los diseñadores de este tipo de sistemas presentando la misma un importante desafío[110]. Ello es así porque los datos que alberga una base de datos, *data lake o data warehouse* pueden presentar distintos formatos, resultar incompletos o incorporar incorreciones. Por tanto, la necesidad de normalizarlos y convertirlos a una dimensión que sea reconocible para los algoritmos se presenta como una tarea esencial. Ello mejorará el rendimiento de los algoritmos a la hora de procesar los datos y además reducirá la posibilidad de que dicho programa sea incorrectamente creado. En los últimos años las técnicas de pre procesamiento manual han sido sustituidas o mejoradas por técnicas basadas en el *Big data*. Estas son capaces en muchos casos de procesar datos sin necesidad de normalizarlos u homogenizarlos a un concreto formato.

Desde un punto de vista técnico resulta habitual distinguir entre datos estructurados, semi estructurados y no estructurados[111]. Así, por datos estructurados hemos de hacer referencia a aquellos datos que están organizados e incorporados en una base de datos con filas y columnas y que son fácilmente relacionables entre sí. Sin duda, estos datos son los más simples de procesar. En cambio, los datos no estructurados se refieren a toda aquella información que no está organizada de una manera predefinida. Ejemplos: documentos que incorporan fechas, frases, números, imágenes, registros webs, así como audios, videos, etc[112]. La dificultad para su procesamiento es mayor que los estructurados. También se ha hecho referencia a una tercera clase, esto es, los

[110] La variabilidad de los datos y sus formatos se han convertido en muchas situaciones en importantes escollos a la hora de desarrollar proyectos de este tipo. Tales problemas fueron puestos de manifiesto en el proyecto llevado a cabo entre DeepMind (empresa privada dedicada a desarrollar sistemas de inteligencia artificial) y la Royal Free London NHS Foundation en el despliegue de una App que ayudará en el diagnostico decisiones renales agudas.
Visto en: https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/10011.htm#_idTextAnchor111

[111] ORTEGA JIMÉNEZ, A: *Las aplicaciones del Big Data en el ámbito asegurador y el tratamiento legal de sus datos. Una perspectiva desde el derecho internacional privado*. Ed. Fundación Mapfre, Madrid, 2019, págs.. 31 y 32.

[112] Unos investigadores utilizaron las sentencias del Tribunal Europeo de Derechos Humanos para desarrollar un modelo algorítmico que tiene como objetivo predecir pronunciamientos futuros de este mismo tribunal. Los textos de estas resoluciones judiciales son datos no estructurados. En: MEDVEDEVA, M; VOLS, M. Y WIELING, M: "Using machine learning to predict decisions of the European Court of Human Rights". *Artificial Intelligence and Law, volume 28,* 2020. Disponible en: https://doi.org/10.1007/s10506-019-09255-y.

llamados datos semi estructurados. En principio estos no están en una base de datos, si bien, su inclusión puede resultar fácil a través de códigos alfanuméricos que representan un determinado hecho o situación[113].

	Sexo	**Nivel de retribución**	**Modelo de coche**
Cliente 1	Hombre	1000 €	Renault Clio
Cliente 2	Mujer	1200 €	Renault Megane
Cliente 3	Mujer	1850 €	Seat Ibiza
Cliente 4	Hombre	1300 €	Volkswagen Golf

Datos estructurados en una base de datos

Ahora bien, incluso cuando los datos que se pretendan utilizar se presenten en un formato estructurado, es necesario hacer una selección adecuada de aquellos que puedan ser los más relevantes para el correcto desempeño de las fases posteriores. En este sentido, la selección de los datos de capacitación del sistema debe ser lo suficientemente representativa de la realidad que se pretende moldear con el sistema, esto es, el contexto en el que irradiará sus efectos. De ahí la importancia de haber establecido en fases previas el contexto y los objetivos del proyecto. Así, en muchos casos el volumen mayor o menor de datos será irrelevante si estos sólo cubren una fracción de la realidad en la que trabajará el sistema tal y como previamente indicábamos[114].

Es turno de analizar los principales procesos y técnicas que aplican los diseñadores cuando llevan a cabo el pre procesamiento de datos.

c.1) Selección de los datos

Hasta ahora, al aludir a las entradas que alimentan los sistemas hemos hecho referencia al concepto general de dato. Sin embargo, es necesario realizar una aproximación más profunda del conjunto de elementos que engloba el concepto de dato como insumo que alimentará el algoritmo. En este sentido y por lo que se refiere a las técnicas de aprendizaje supervisado, a efectos de este trabajo es necesario explicar los conceptos de muestra, característica/variables y etiqueta.

Muestra: cuando hablamos de muestra nos referimos al conjunto de ejemplos con los que cuenta la organización para llevar a cabo su análisis. Así, cada muestra

[113] COTINO HUESO,L: "El alcance e interactuación del régimen jurídico de los datos personales y *big data* relacionado con salud y la investigación biomédica". *Revista de derecho y genoma humano: genética, biotecnología y medicina avanzada*, Nº 52, 2020, págs.74 y 75.
[114] The Norwegian Data Protection Authority. *Artificial intelligence and privacy.* 2018, pág.11.

presenta un conjunto de características que se procesarán posteriormente por los algoritmos elegidos. Ejemplo: Cliente 1 o Pepito Pérez.

Característica/Variable independiente: conjunto de atributos que conforman cada muestra y que se utilizarán para predecir posteriormente las etiquetas. Cada característica puede presentar distintos o similares valores. En estadística, los atributos o característica son denominados variables independientes. Las variables pueden ser categóricas o numéricas. Ejemplo: Variable categórica; sexo, edad, modelo de coche. Variable numérica: 1500 €. Todas esas características pertenecen al cliente 1 o a Pepito Pérez.

Valor: es el número o definición que presenta cada atributo en una muestra. Ejemplo: el valor del atributo sexo en el cliente 1 es hombre o el valor del atributo edad del cliente 1 es 22 años.

Etiqueta o clase/Variable dependiente: la etiqueta o clase es la asignación que se establece a cada una de las muestras que conforman el conjunto de características. Dicha etiqueta es fijada por aquellos que han elaborado la base de datos y que tiene como referencia normalmente una serie de datos históricos. En estadística, la etiqueta o clase se conoce como la variable dependiente debido a que la misma depende del conjunto de variables independientes. Ejemplo: Buen pagador, *spam*, tumor maligno, discurso de odio. Ejemplo: El paciente 3 (muestra) que presenta problemas cerebrales, antecedentes familiares cancerígenos (variable independiente/característica) presenta un tumor maligno. (Clase o etiqueta).

El etiquetado de las muestras es un proceso lento y costoso debido a que es necesario marcar uno a uno los datos que posteriormente alimentarán el sistema. En este sentido, no es infrecuente que a día de hoy miles de personas se dediquen en exclusiva al etiquetado de imágenes, textos, sonidos o documentos con el fin de que el sistema sea lo más preciso posible[115].En otros supuesto, los problemas de etiquetado no se derivan del número de etiquetas que se pretenden marcar sino de la complejidad a la hora de

[115] Etiquetar imágenes de posibles tumores, analizar fotografías pornográficas o escuchar tipos de estornudos se convierten en las tareas diarias de multitud de personas. Dichos datos posteriormente se utilizarán por empresas para entrenar sus sistemas, los cuales, ofrecerán servicios médicos de detección de enfermedades, controlarán el contenido ilícito vertido en la red o desplegarán asistentes virtuales. Fuente de la noticia: METZ,C: "A.I Is learning form Humans. Many Humans". The New York Times. 16/08/2019. Información disponible en: https://www.nytimes.com/2019/08/16/technology/ai-humans.html?action=click&module=Discovery&pgtype=Homepage

concretar la variable dependiente Por ejemplo: Para una aplicación que pretende detectar el discurso del odio se requiere que previamente este sistema sea alimentado con contenido que refleje dicha variable. Ello no siempre será sencillo. Más teniendo en cuenta que en muchas situaciones serán los propios desarrolladores de los sistemas los que en su caso tomarán esta decisión[116]. Precisamente, aquellos que se dediquen al etiquetado de los datos han de tener unas pautas muy claras sobre dicho proceso y en muchos casos cierta experiencia sobre la etiqueta que se pretende asignar. Así, la complejidad presente para etiquetar una imagen de un perro no será la misma a la de una imagen de un tumor.

Muestras /Ejemplos	**Características/Variables independientes**			**Etiqueta Variable dependiente**
	Sexo (Variable 1)	Nivel de retribución (Variable 2)	Modelo de coche (Variable 3)	
Cliente 1	Hombre (Valor)	1000 €	Renault Clio (Valor)	Mal pagador
Cliente 2	Mujer (Valor)	1200 € (Valor)	Renault Megane	Mala pagadora
Cliente 3	Mujer (Valor)	1850 € (Valor)	Seta Ibiza	Buena pagadora
Cliente 4	Hombre (Valor)	1300 € (Valor)	Volkswagen Golf	Buen pagador

En sentido estricto, el concepto de dato solo englobaría a las características o atributos que se utilizan para diseñar el sistema, ya que la muestra simplemente engloba todas las características y la etiqueta es la inferencia que se deriva de tales características. No obstante, en nuestra opinión, dicha inferencia también habría de considerarse dato de la persona debido a que también se relaciona con esta. La diferencia entre dato y dato inferido será posteriormente explicada en los capítulos siguientes ya que resulta de suma importancia para nuestro estudio[117].

Por lo que se refiere a la elección de las características o atributos, es recomendable que dicha selección sea realizada por personas expertas en la materia que

[116] DUARTE,N Y LLANSÓ,E: "Mixed Messages? The Limits of Automated Social Media Content Analysis", *Center for Democracy and technology*, 2017, pág.16. Disponible en: https://cdt.org/insights/mixed-messages-the-limits-of-automated-social-media-content-analysis/

[117] Véase el Capítulo II, aparatado II, punto 4, epígrafe B) de esta tesis.

tengan un gran conocimiento del comportamiento de las variables y la posible eficacia de estas en el entorno donde el sistema irradiará sus efectos[118]. Ello es así porque estas personas son las que se encuentran en la mejor posición para perfilar y comprender el problema que se pretende resolver con la creación del sistema mencionado. De esta manera, el primer paso elemental será el de justificar la relevancia y adecuación de las variables seleccionadas poniéndolas en relación con la etiqueta o clase que se pretende predecir cuando el sistema se aplique en la práctica. Por ejemplo; resulta justificado elegir el salario como atributo para predecir el riesgo de impago de un préstamo bancario. Sin embargo, a priori, no resultaría pertinente o debería justificarse en mayor medida la selección del atributo del color de los ojos para evaluar dicha solvencia financiera. Lo mismo ocurriría si queremos crear un sistema que sea capaz de distinguir entre perros y gatos y utilizamos como variable el número de patas. La justificación de las variables elegidas debería quedar registrada. Ello permitirá posteriormente acreditar la idoneidad de dichas variables ante la puesta en entredicho de tal elección por parte de terceros afectados por las decisiones adoptadas por el sistema.

Muestra/ Ejemplos	**Atributos/característica**			**Etiqueta**
	Edad	Color de ojos	Salario	
Cliente 3	45	Marrones ¿?	1500 €/mes	Buen pagador/a

Elección y justificación del atributo

Una incorrecta elección de las características o variables independientes puede acarrear unas consecuencias muy negativas sino se tienen en cuenta adecuadamente las repercusiones que puede tener la elección de tal variable en el contexto donde el sistema irradiará sus efectos. A modo de ejemplo, en EEUU existe un algoritmo que incluye a las personas que mayores costos generan para el sistema sanitario en un programa específico con mayores necesidades de atención[119]. Dicho algoritmo utiliza como variable dependiente o etiqueta el costo que supone una persona para el sistema potenciando variables independientes relacionadas con dicho costo. De manera que a mayor costo, aumentan las opciones de que esa persona se pueda beneficiar de dicho programa. Sin embargo, aunque la premisa del sistema es que las personas con mayores

[118] The Norwegian Data Protection Authority. *Artificial intelligence and privacy.* 2018, op.cit., pág.11.
[119] OBERMEYER, Z; POWERS, B; VOGELI,CH Y MULLAINATHAN,S: "Dissecting racial bias in an algorithm used to manage the health of populations". *Science*, 2019. Disponible en: https://science.sciencemag.org/content/366/6464/447

necesidades de atención se beneficien del sistema, lo cierto es que, al centrarse sólo en los costos y no también en la salud de los pacientes, existe la posibilidad de que pacientes que presentan la misma gravedad de un enfermedad o incluso superior no sean asignados a dicho programa por no presentar los mismos costes para el sistema. En este sentido, ha quedado probado que por diversas razones en EEUU los pacientes negros suponen un menor coste para el sistema que los pacientes blancos[120]. Al valorar únicamente características relacionadas con el costo. Una persona de raza negra que presenta la misma enfermedad que una persona de raza blanca podría no quedar incluida en ese programa precisamente porque sólo se tienen en cuenta los costes que supone para el sistema. Al no tenerse en cuenta el sesgo estructural existente en dicho contexto, esto es, los pacientes negros cuestan menos al sistema porque por ejemplo tienen mayores dificultades para acceder al mismo, las características elegidas son altamente perjudiciales para ese colectivo.

Por otro lado, las características elegidas deben ser concretas y cuantificables[121]. Así, por ejemplo, establecer como atributo la conducción temeraria podría ser un atributo muy impreciso y subjetivo. Tal atributo podría sustituirse por otras características más objetivas como el número de denuncias por estado de embriaguez que ostentan los sujetos observados.

Muestra/ Ejemplos	**Atributos/característica**				**Etiqueta**
	Edad	Color de ojos	Salario	Conducción temeraria	
Cliente 3	45	Marrones	1500/mes	Grave ¿?	Buen pagador/a

Precisamente, la cuantificación o la frecuencia de una variable pueden ser muy importante a la hora de valorarlas en mayor o menor grado. Así, si mantenemos el ejemplo anterior, no es lo mismo que una persona haya sido detenida cuatro veces por

[120] Aunque las persona de raza negra acuden más a los servicios de asistencia sanitaria primaria, son las personas de raza blanca las que finalmente utilizan en mayor medida los servicios más costosos como pueden ser la realización de operaciones quirúrgicas. En: OBERMEYER,Z; POWERS,B; VOGELI,CH Y MULLAINATHAN,S: "Dissecting racial bias in an algorithm used to manage the health of populations", op.cit., pág. 4.

[121] Puede haber atributos que sólo tienen sentido si otros atributos tienen algunos valores especiales, por ejemplo, la característica "nombre del marido" sólo es útil si la persona está casada. Estos valores pueden ser problemáticos a la hora de tratar de modelar una determinada realidad. En: ANRIG,B; BROWNE,W; GASSON,M: "The Role of Algorithms in Profiling". En: HILDEBRANDT, M; GUTWIRTH,S (eds.): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008, pág.69.

circular en estado de embriaguez a que solo haya sido sancionado una sola vez por esa conducta. Reducir por ejemplo esa variable a una simple afirmación o negación perjudicaría al desarrollo posterior del modelo ya que representaría inadecuadamente esa variable que se ingresa y además comprimiría las posibles influencias de esta en la variable dependiente que se pretende predecir[122].

Muestra / Ejemplos	**Atributos**		**Etiqueta**
	Edad	Delitos de seguridad vial	
Cliente 3	45	1,2,3, etc.	Buen pagador

Muestra / Ejemplos	**Atributos**		**Etiqueta**
	Edad	Delitos de seguridad vial	
Cliente 3	45	Si/No	Buen pagador

c.2) Limpieza y reducción de los datos

Reducir el número de características a elegir y centrarse en aquellas que son más pertinentes ayuda a desarrollar modelos más sencillos y fáciles de entrenar. Ello facilita la comprensión del sistema y reduce el tiempo de computación de los mismos. En este sentido, en muchos casos, varios atributos elegidos pueden presentar una fuerte correlación entre sí. En estos supuestos, ambos podrían estar midiendo la misma información subyacente[123]. Por tanto, la eliminación de una de esas características no debería comprometer el rendimiento del sistema. Se reducen características de recopilación y se gana en interpretabilidad.

En la búsqueda de reducción de datos también es posible eliminar aquellas muestras que presenten atributos atípicos respecto del conjunto de características utilizadas. Es frecuente que estas irregularidades obedezcan a una incorrecta incorporación de dichos atributos a la base de datos. No obstante, tal supresión debe realizarse con cautela ya que muchos de esos atributos atípicos pueden representar a una parte minoritaria del conjunto de datos generales que presentan características más o menos similares. Por ejemplo: en la tabla anterior, un atributo atípico puede considerarse el importe de 100 €/mes en concepto de salario. Se puede llegar a entender

[122] GILLINGHAM, P: "Can Predictive Algorithms Assist Decision-Making in Social Work with Children and Families?" *Child Abuse Review*, 28, 2019, pág.122. Disponible en: https://doi.org/10.1002/car.2547
[123] KUHN,M, JOHNSON,K: *Applied Predictive Modelling*. Ed. Springer Science + Business Media, New York, 2013, op.cit., págs. 43 y 44.

que cuando se introdujeron esos datos pudo existir un error y en vez de marcar la cifra 1000€ se introdujo un cero menos. Sin embargo, ese mismo atributo, esto es, 100€, podría reflejar a un colectivo específico de personas que, por diversas razones[124], se separan de la regla general pero aun así, presentan esos atributos que reflejan una realidad que debe de ser observable por el modelo[125]. Y es que, tal realidad minoritaria podría representar un grupo desfavorecido o poco representado que no debería quedar fuera de la realidad que se pretende moldear con el sistema. Esta cautela a la hora de valorar la eliminación o no de los atributos atípicos debe aumentar cuando el conjunto de datos que se presente para el análisis sea pequeño[126].

A veces, también se pueden suprimir muestras que estén duplicadas, ya sea porque presentan los mismos valores en los distintos atributos o bien porque al ingresar los datos se duplicaron las mismas[127].

A su vez, también resulta conveniente comprobar si el conjunto de muestras que se van a utilizar presenta un número de atributos con todos los valores completos. En la medida que todos los valores de un atributo estén presentes en todas las muestras o en la mayoría de estas, el modelo tendrá mayor facilidad para aprender cómo se relaciona el atributo con la etiqueta. Es decir, el mayor número de muestras completas que presentan la misma característica permite al modelo visualizar dicha característica en diferentes escenarios y por tanto, determinar con mejor precisión la predicción que se pretende de la etiqueta. A la inversa, si los valores de la característica seleccionada aparecen en pocas muestras, el modelo no realizará predicciones adecuadas basadas en dicho

[124] Las técnicas como el Big Data "pueden operar omitiendo sistemáticamente a personas que viven en los "márgenes tecnológicos, ya sea a causa de la pobreza, su situación geográfica, su estilo de vida, o por cualquier otra causa que los saque del "data field". En: MORENTE,PARRA,V: "Big data o el arte de analizar datos masivos. una reflexión crítica desde los derechos fundamentales". *Derechos y libertades*, número 41, Época II, junio 2019, pág.229.

[125] Las técnicas como el Big Data "pueden operar omitiendo sistemáticamente a personas que viven en los márgenes tecnológicos, ya sea a causa de la pobreza, su situación geográfica, su estilo de vida, o por cualquier otra causa que los saque del "data field". En: : MORENTE,PARRA,V "Big data o el arte de analizar datos masivos. una reflexión crítica desde los derechos fundamentales". *Derechos y libertades*, número 41, Época II, junio 2019, pág. 229. Otros autores muestran también su preocupación por aquellos colectivos que por diversas razones no son tenidos en cuenta en estos sistemas. COTINO HUESO,L: "Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales". *Dilemata*, nº 24, 2017, pág.138.
Disponible en: https://www.dilemata.net/revista/index.php/dilemata/article/view/412000104/494

[126] KUHN,M, JOHNSON,K: *Applied Predictive Modeling*. Ed. Springer Science + Business Media, New York, 2013, op.cit., pág.33.

[127] En: ANRIG,B; BROWNE,W; GASSON,M: "The Role of Algorithms in Profiling". En: HILDEBRANDT, M; GUTWIRTH,S (eds.): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. op.cit.,, pág.68

atributo. Para resolver este problema, es posible eliminar las muestras incompletas por presentar valores faltantes en un concreto atributo.

Muestra	**Atributos/característica/v.independiente**			**Etiqueta**
	Edad	Color de ojos	Salario	
Cliente 1	45	Marrones	1500 €/mes	Buen pagador
Cliente 2	34	Marrones	1200 €/mes	Buen pagador
Cliente 3	¿?	Azules	1100€/mes	Mal pagador
Cliente 4	28	¿?	1400E/mes	Buen pagador

Conjunto de datos con filas de atributos incompletas

No obstante, los desarrolladores, para evitar la eliminación de muestras que pueden esconder información valiosa, aplican distintas técnicas en función del tipo de valor que se pretende reemplazar. Así, para el caso de que el valor que se desconoce pertenezca a una variable numérica, existe la posibilidad de reemplazar esos valores desconocidos por otros valores que resulten de la aplicación de distintas técnicas estadísticas como la mediana o la media al conjunto de valores de los que sí se disponen. La aplicación de una u otra técnica dependerán de la distribución más o menos asimétrica de los valores[128]. Ejemplo de la tabla anterior: el valor del atributo edad del cliente 3 aparece incompleto. Antes de proceder a eliminar la muestra del cliente 3, el valor incompleto podría sustituirse por la media del resto de valores "edad" del conjunto de características de ese atributo. En este caso el valor que se incorporaría sería 36. Por otro lado, cuando el valor sea categórico normalmente se utiliza la técnica estadística de la moda del conjunto de valores, esto es, el valor que tiene mayor frecuencia absoluta.

Muestra	**Atributos/característica**			**Etiqueta**
	Edad	Color de ojos	Salario	
Cliente 1	45	Marrones	1500 €/mes	Buen pagador
Cliente 2	34	Marrones	1200 €/mes	Buen pagador
Cliente 3	**36**	Azules	1100€/mes	Mal pagador
Cliente 4	28	**Marrones**	1400€/mes	Buen pagador

Conjunto de datos con filas aplicando a los atributos faltantes la técnica de la media y la moda del resto de variables.

[128] Cuando el conjunto de valores sean muy asimétricos la mediana es una medida mucho más robusta debido a que no se ve afectada por los valores extremos. A sensu contrario, cuando el conjunto de valores de un atributo no presenten una gran asimetría la media es la medida ideal. Pudiendo comprobar tal asimetría de datos a través de técnicas como el diagrama de cajas o el diagrama de dispersión. .
Visto en: https://www.youtube.com/watch?v=XKshQsg8TGw (minuto 30:21)

En otras ocasiones, a los valores desconocidos del atributo se le imputa un valor aleatorio del conjunto de valores pertenecientes al atributo observado[129]. Esta técnica resulta válida tanto para atributos categóricos como numéricos. Ejemplo: el valor desconocido del atributo edad se le asigna un valor aleatorio del resto, esto es, 45. También es frecuente que si el *dataset* cuenta esencialmente con variables numéricas, las categóricas sean transformadas en las primeras.

Muestra	**Atributos/característica**			**Etiqueta**
	Edad	Color de ojos	Salario	
Cliente 1	45	Marrones	1500 €/mes	Buen pagador
Cliente 2	34	Marrones	1200 €/mes	Buen pagador
Cliente 3	**45**	Azules	1100€/mes	Mal pagador
Cliente 4	28	**Azules**	1400€/mes	Buen pagador

Conjunto de datos con filas aplicando a los atributos faltantes valores aleatorios obtenidos del resto de variables.

Por otro lado, y al igual que ocurre con los atributos que se seleccionan, es posible que en la base de datos exista una desproporción importante entre las etiquetas o clases asignadas. Ejemplo: Supongamos que tenemos una base de datos con 400 muestras a las que, 375 hemos asignado la etiqueta de malos pagadores y a 25 la de buenos pagadores. En estos casos se dice que las etiquetas utilizadas son desproporcionadas o están desbalanceadas lo que también exigirá la previsión de una serie de técnicas para evitar que el modelo solo sea preciso respecto de una determinada etiqueta, esto es, aquella que aparece desproporcionada en la base de datos[130].

En definitiva, la selección adecuada de atributos se muestra esencial para que los sistemas presenten toda una serie de requisitos mínimos que eviten el menor número de errores y riesgos en fases posteriores[131]. Así, en muchos supuestos no será relevante la cantidad de variables elegidas sino más bien la precisión a la hora de elegirlas. Es decir, es necesario optar por aquellas variables que realmente tengan relación con la variable dependiente u objetivo que se pretende con el modelo a desarrollar y no simplemente utilizar aquellas de las que se dispone en el *dataset*.

[129] En estadística esta técnica se conoce como muestreo aleatorio simple.
[130] Entre otras técnicas cabe destacar el sub muestreo, la creación de muestras sintéticas o la penalización de las etiquetas más representativas.
En: https://www.aprendemachinelearning.com/clasificacion-con-datos-desbalanceados/
[131] Las características que deben estar presentes en los datos que alimentarán los sistemas son exactitud, completos, únicos, oportunos, válidos, suficientes, relevantes, representativos, consistentes. En: Office for artificial intelligence. Government Digital Service of UK. *A guide to using artificial intelligence in the public sector*. 2020, pág. 16. Disponible en: https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector

c.3) Riesgos derivados de una incorrecta selección y limpieza de los datos

La selección adecuada de las características trata de evitar toda una serie de efectos negativos que se puede derivar de un incorrecto pre procesamiento de los datos. Aunque a lo largo de este trabajo se harán referencia a potenciales riesgos, vamos a tratar de sistematizar aquellos que están estrechamente relacionados con la fase del pre procesamiento de datos, estos son: dimensionalidad de los datos, sobre ajuste del modelo y sesgos en las decisiones.

Por lo que se refiere a la *dimensionalidad* de los datos. Esta sucede frecuentemente cuando aumentamos el número de atributos a analizar y estos realmente no están relacionados con la etiqueta o, estando, el número de muestras es insuficiente para que el modelo pueda llegar a aprender de ese conjunto de atributos. Así, conforme el número de atributos aumente, mayor número de muestras se requerirán para que el modelo sea capaz de aprender de todas las combinaciones presentes en las características seleccionadas respecto de las etiquetas asignadas.

Muestra	**Atributos/característica**			**Etiqueta**
	Nº de puertas	Nº de Ruedas	Altura	Avería
Coche 1	4	4	1.64	Defectuoso
Coche 2	3	4	1,43	No defectuoso
Coche 3	4	4	1.56	No defectuoso
…	…	…	…	
Coche 50	4	4	1.60	No defectuoso

Muestra con sólo 3 atributos para predecir la etiqueta de posible avería.

Supongamos que la tabla anterior engloba un conjunto de 50 muestras y en cada muestra se contienen 3 atributos. Con esos tres atributos pretendemos que el algoritmo sea capaz de generalizar y aprender las características principales de la etiqueta deseada, esto es, coche que presentará averías o no. Por otro lado, presupongamos que en la tabla posterior existe un número menor de muestras, esto es 25, y además, en vez de 3 atributos, se utilizan 6. Pues bien, es muy probable que en el primer ejemplo el algoritmo sea capaz de aprender las principales características presentes en ese conjunto de datos respecto de la etiqueta deseada. Sin embargo, en el segundo supuesto, ello puede resultar más complicado debido a la falta de muestras y el aumento de variables

al quedar limitada la capacidad de generalización del algoritmo y consecuentemente aumentando la posibilidad de que se desarrolle un modelo inadecuado[132].

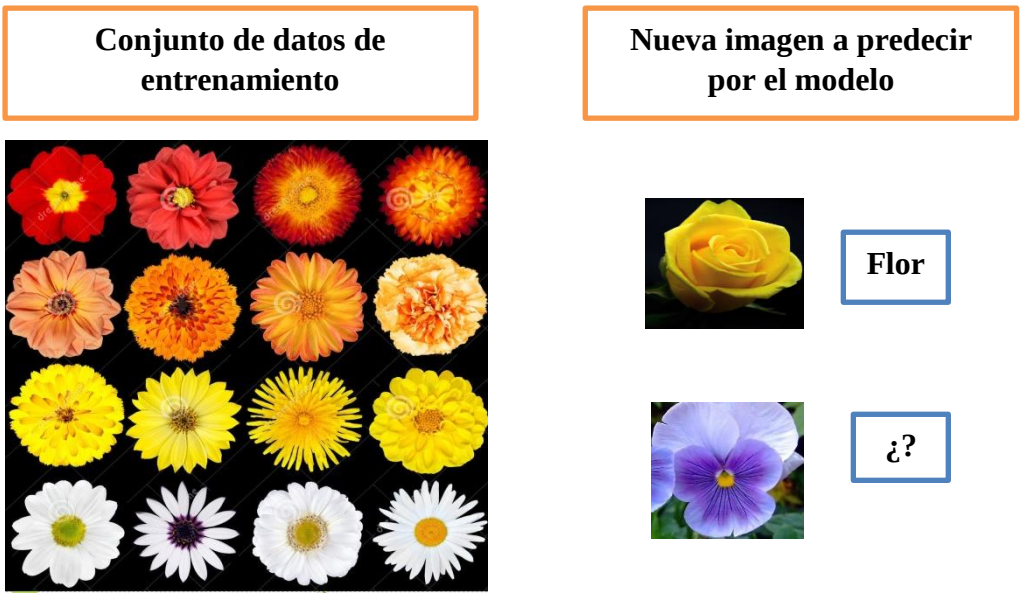
	Atributos/Características						**Etiqueta**
Muestra	Nº de puertas	Altura	Anchura	Color	Nº de ruedas	Nº de pasajeros máx.	Avería
Coche 1	4	1.64	1.83	Rojo	4	5	Defectuoso
Coche 2	4	1.76	1.76	Azul	4	7	No defectuoso
Coche 3	4	1.58	1.79	Negro	4	5	Defectuoso
…	…	…	…	…	…	…	
Coche 25	3	1.76	1.80	Negro	4	5	No defectuoso

Muestra con 6 atributos para predecir la etiqueta "avería"

Junto a la dimensionalidad, la insuficiencia adecuada de determinadas características puede llevar al *sobreajuste* del sistema. Se dice que un sistema está sobre ajustado cuando aprende excesivamente de los datos con los que ha sido entrenado pero no es capaz de adoptar decisiones correctas cuando se implanta en la realidad[133]. Es decir, el sistema es incapaz de reconocer las nuevas instancias u ofrece errores porque los datos con los que se entrenó al sistema no eran lo suficientemente representativos. La falta de representatividad puede tener su origen en una desproporción en las etiquetas o atributos presentes en el *dataset*. A su vez, el sobreajuste se pude deber al hecho de que el sistema se haya centrado en aprender correlaciones que son irrelevantes y que estas realmente no están presentes en la realidad que se pretende modelar. En ambos supuestos, el sistema generará estimaciones poco precisas y por consiguiente se adoptarán decisiones erróneas[134]. En este sentido, es frecuente que conforme aumente el número de variables, mayor probabilidad de que se genere tal sobre ajuste.

A modo de ejemplo presentamos la siguiente imagen que muestra un modelo sobre ajustado.

[132] The Norwegian Data Protection Authority. *Artificial intelligence and privacy.* 2018, op.cit., pág.11.
[133] BRAMER,M: *Principles of Data Mining*. Ed.Springer, Londres,2016, pág.121.
[134] El sobreajuste de los datos en determinados contextos puede llevar a situaciones muy complejas donde pueden quedar afectados colectivos especialmente sensibles. Ejemplo: probabilidad de que un menor sufra maltrato infantil. En: GILLINGHAM, P: "Can Predictive Algorithms Assist Decision-Making in Social Work with Children and Families?", op.cit., pág.119.



Sistema sobre ajustado. Elaboración propia.

Un sistema podría estar sobre ajustado cuando en la fase de entrenamiento solamente le mostramos imágenes de flores de varias tonalidades concretas que no representan todos los colores que pueden estar presentes en el resto de flora existente. Así, en la fase de entrenamiento este modelo ha podido aprender muy bien a identificar las flores cuando las imágenes comprenden las tonalidades rojas, amarillas, naranjas y blancas pero puede volverse impreciso cuando lo que trate de predecir sea la imagen de una flor que no presentan las tonalidades presentadas durante el entrenamiento, tal y como puede ocurrir con una flor de color violeta. Es decir, este modelo ha generalizado adecuadamente los datos presentes en el *data set* o conjunto de datos pero sin embargo, al no entrenarse con otros datos que también pueden estar presentes en el entorno al se enfrentará el sistema cuando adopte decisiones, dicho sistema será impreciso por estar sobre ajustado. El sistema probablemente acertará y predecirá fácilmente las flores que presenten las tonalidades rojas, naranjas, amarillas y blancas, al no haber visto otra realidad más allá de esa, cuando se presenten flores con otras tonalidades nunca vistas por el sistema, este producirá errores. Para él, en el mundo que ha moldeado sólo existen las flores presentadas durante el entrenamiento.

Finalmente, y no menos importante, otro de los efectos negativos de una incorrecta selección de las características será el de los *sesgos* que puedan estar presentes al elegir determinadas variables. En este sentido, la elección de unas variables o atributos en defectos de otros pueden esconder unos sesgos que posteriormente se pueden manifestar cuando el sistema adopte decisiones. Así, es muy habitual que a la

hora de elegir entre unas u otras variables los analistas utilicen criterios propios basados en su experiencia personal o profesional que de forma no intencionada pueden reflejar sesgos. En otras ocasiones, los atributos elegidos son intencionadamente elegidos[135]. Tal elección no tiene que ser directa sino que en muchas ocasiones una característica que puede entenderse innocua realmente busca sesgar a las personas. Ejemplo: Para evitar que se tilde como sesgado un sistema que utiliza el atributo sexo los desarrolladores pueden utilizar otros atributos *proxy* que intencionadamente se asocian con el sexo. Tal *proxy* podría ser el modelo de coche y la etiqueta el riesgo de sufrir un accidente. Queremos creer no obstante, que en la mayoría de los casos, los sesgos que se generan se derivan de la elección incorrecta de las características de forma no intencionada. Y es que no podemos olvidar que los sesgos que estén presentes en el *dataset* elegido se reproducirán con toda seguridad en el funcionamiento posterior del sistema[136]. De esta manera, en muchas ocasiones, el problema no deviene del sistema en sí, sino de los datos que lo alimentan y lo entrenan.

c.4) Análisis de los datos

Una vez seleccionados y perfilados los datos, resulta habitual que los analistas procedan a su observación a través de una serie de técnicas antes de que estos mismos sean procesados por los algoritmos. Dicha observación permite a los analistas hacerse una idea general de ese conjunto de datos y obtener una serie de conclusiones que pueden ser prácticas para el procesamiento ulterior.

Dichas técnicas de observación proceden en su mayoría de la estadística y facilitan el estudio de los datos seleccionados. Entre otras, podemos destacar las siguientes: la media, la mediana, la moda, desviación típica, varianza o las tablas de contingencia.

Por ejemplo, si al aplicar alguna de estas técnicas se obtiene la conclusión de que el 80% del *dataset* está formado por muestras del género masculino. Es muy probable que posteriormente el sistema tenga en cuenta esta variable de forma más incisiva que la del género femenino. No estamos diciendo que ello pueda generar un sesgo favor de

[135] European Parliamentary Research Service.*A governance framework for algorithmic accountability and transparency*. 2019, pág.20.

[136]"Dada la naturaleza de la predicción, un pasado racialmente desigual producirá necesariamente resultados racialmente desiguales". En: MAYSON, S,G: "Bias In, Bias Out". *Yale Law Journal*, vol 108, 2019, pág. 2224. Disponible en: https://www.yalelawjournal.org/article/bias-in-bias-out

uno u otro género pero es muy probable que ello suceda. Es decir, este análisis de los datos seleccionados permite observarlos y por tanto avisar a los propios analistas de incoherencias o descompensaciones no observadas hasta la fecha. Ello puede llevar a estos últimos a plantearse si es necesario realizar alteraciones en el conjunto de muestras que componen el *dataset.* Y es que, esos datos serán los que alimentarán y conformarán al sistema que posteriormente adoptará decisiones. Por tanto, detectadas esas incoherencias, lo lógico es que se corrijan en este momento.

A su vez, también puede servir de ayuda valerse de todo tipo de instrumentos que faciliten la representación del conjunto de datos que se tienen. La visualización de los datos pude reflejarse a través de herramientas como la escala de valores, los histogramas o los mapas de calor entre otros[137].

c.5) Separación de los datos

Una técnica habitual que se realiza en algoritmos de aprendizaje supervisado consiste en la separación del conjunto de datos en dos grupos.

Por un lado, los primeros son utilizados para entrenar al algoritmo que los procesará y que desarrollará el modelo. El resto se destina a la evaluación posterior de dicho modelo. Generalmente el conjunto de datos destinados al entrenamiento es muy superior al de pruebas o test. Alrededor del 70 u 80 por ciento de la *dataset* se destina a la fase del entrenamiento y el resto para la fase de prueba. La idea es sencilla, con el conjunto de datos de entrenamiento el modelo aprende y con el conjunto de datos de prueba el modelo es evaluado. La separación del conjunto de datos debe ser aleatoria[138]. Como es lógico, cada uno de los grupos en los que se divide el *dataset* debe estar conformado por muestras distintas[139]. Lo que se pretende con ello es que el sistema en la fase de entrenamiento sea capaz de generalizar adecuadamente la realidad que pretende perfilar a través de ese primer grupo de datos. Posteriormente, a través de los datos de prueba, se puede evaluar si el sistema se ha entrenado y ha generalizado

[137] Para más información sobre la visualización de datos véase: CLAUS, O,W: *Fundamentals of Data Visualization*. Ed. O'Reilly Media, Inc, 2019.
Disponible en: https://serialmentor.com/dataviz/index.html
[138] Government Digital Service of UK. *A guide to using artificial intelligence in the public sector*, op., p-29. También en: RUSSELL, S Y NORVING, P: *Inteligencia Artificial. Un enfoque moderno*, op.cit., pág.761.
[139] Esas muestras deben ser distintas pero todas han de formar parte de la *data set*. Ello es así porque en ese conjunto de datos está la realidad que pretendemos moldear. Simplemente la separamos con el objetivo de evaluar la fase de aprendizaje del sistema.

adecuadamente el entorno. Ese objetivo no se lograría si se utilizan los mismos datos que se utilizaron durante el entrenamiento ya que el modelo arrojaría unos resultados óptimos de funcionamiento cuando realmente no lo son. Esta técnica de evaluación es conocida comúnmente como *cross validation*, la cual, posteriormente analizaremos.

D) El desarrollo del modelo. La importancia del entrenamiento del modelo

Cuando hablamos de modelo nos estamos refiriendo al resultado derivado del procesamiento de los datos previamente seleccionados a través de distintos algoritmos. Por tanto, de una misma base de datos pueden generase tantos modelos como procesamientos de datos se realicen y algoritmos y técnicas se apliquen a dicha base de datos. De esta manera, una vez que hemos seleccionado los datos que se procesarán, el siguiente paso será la creación del modelo a través del entrenamiento de los mismos. Para ello, es necesario seleccionar primeramente el algoritmo y las técnicas que se utilizarán. Es habitual que en esta fase los programadores opten por el desarrollo de diversos modelos. Con esta forma de proceder los diseñadores posteriormente pueden comparar cómo trabaja cada uno de los modelos y elegir aquel que más se adecúe a los objetivos iniciales por los que se decidió iniciar el proyecto de desarrollo de un sistema de toma de decisiones automatizadas.

Dicho lo anterior, el paso inicial será elegir el algoritmo o algoritmos que se pretendan utilizar para desarrollar el modelo o modelos que se quieren crear. La elección de los algoritmos a utilizar puede variar en función de muchas causas. Sin embargo, en este momento mencionaremos algunas razones técnicas que llevan a los desarrolladores a optar por unos u otros.

d.1.) Objetivo y contexto del sistema

El objetivo que se pretende con el desarrollo del sistema y el contexto potencial donde se implantará resultan esenciales a la hora de optar entre unos u otros algoritmos.

Cuando una organización se embarca en un proyecto de inteligencia artificial esta última tiene en mente el cumplimiento de un propósito o resolución de un problema. En este sentido, al hacer referencia al concepto de *data mining* y de aprendizaje automático indicábamos los principales problemas/objetivos que se pretendían resolver con estas técnicas, poniendo el acento en cuatro de ellos, esto es:

clasificación, regresión, asociación y agrupamiento. Pues bien, la elección de uno u otro algoritmo va a estar estrechamente vinculada al tipo de problema que pretenda resolver la organización con el proyecto que despliega. Además, en función del propósito que se busque también será relevante valorar el modelo que pretendamos crear, esto es, un modelo predictivo o un modelo descriptivo. Finalmente, en función de la técnica de aprendizaje que se pretenda utilizar, esto es, supervisado, no supervisado y de refuerzo. El tipo de algoritmo también puede variar. Es decir, el objetivo inicialmente marcado en el proyecto vinculará necesariamente al tipo de algoritmo u algoritmos a elegir. A modo de ejemplo: si nuestro objetivo es crear un sistema que sea capaz de distinguir entre personas altas y no altas para limitar el acceso a determinados espacios, lo habitual es que la organización trate de resolver un problema de clasificación, esto es, es o no alto. En este caso, lo ideal es acudir a algoritmos de aprendizaje supervisado como podría ser el árbol de decisión.

Por lo que se refiere al *contexto* donde se implantará el sistema hemos de considerar que en esta fase del desarrollo del mismo ya debe estar en gran medida definido. Pues bien, en función de las exigencias jurídicas que se deriven de la toma de decisiones automatizadas en uno u otro sector, las organizaciones en muchos casos deberán optar entre aquellos algoritmos que más se adecúen a tales ordenamientos jurídicos. Por ejemplo, algunos algoritmos son más interpretables que otros y por tanto, si el contexto donde se pretende implantar el sistema exige la explicación de la decisión, las organizaciones deberían optar por aquellos que de forma más adecuada cumplan con la garantía jurídica exigida por ese contexto, en este caso, la motivación de la decisión. Es decir, las organizaciones deben elegir el algoritmo que tenga un mejor encaje en el ordenamiento jurídico donde el sistema desplegará sus efectos. Esto supone que en muchos casos determinados algoritmos o modelos de aprendizaje no puedan ser utilizados en determinados contextos. Siguiendo con el ejemplo anterior, a priori, un sistema basado en un algoritmo que arroja resultados totalmente indescifrables no podría tener cabida en contextos jurídicos donde se requiere necesariamente de cierta explicación de la decisión.

Junto al contexto y el objetivo del sistema, aspectos como la tipología o el número de datos con los que se cuenta en el *dataset* serán también elementos relevantes a la hora de optar por uno u otro algoritmo. Así, si se cuentan con pocos datos

etiquetados lo normal es que se acuda a algoritmos basados en técnicas de aprendizaje no supervisados o semi supervisado.

d.2) Algoritmos presentes

Aunque la variedad de algoritmos es muy heterogénea y en muchos casos algunos de estos pueden aplicarse tanto en aprendizaje supervisado como no supervisado, en la siguiente tabla se hace mención a algunos de los principales que a día de hoy son generalmente utilizados por las organizaciones que se embarcan es estos proyectos[140]. Cada uno de los algoritmos que se presentan contienen diversas característica técnicas y resultan más o menos adecuados para los distintos objetivos que se pretende con cada proyecto de inteligencia artificial.

Algoritmo	**Tipo de aprendizaje**	**Problema o tarea**	**Modelo**	**Principales características**
Árbol de decisión	Supervisado	Clasificación	Predictivo	Muy interpretable
Bosque aleatorio	Supervisado	Clasificación	Predictivo	Poco interpretable
Clasificador Bayesiano ingenuo (Naive Bayes)	Supervisado	Clasificación	Predictivo	Muy interpretable
Máquinas de soporte vectorial	Supervisado	Clasificación	Predictivo	Poco interpretable
K vecino más cercano	Supervisado	Clasificación	Predictivo	Medianamente interpretable
Regresión Lineal	Supervisado	Clasificación /Regresión	Predictivo	Muy interpretable
Regresión Logística	Supervisado	Clasificación	Predictivo	Muy interpretable
K-means	No supervisado	Agrupamiento	Descriptivo	Poco interpretable
Red neuronal artificial	Deep Learning Supervisado/No supervisado	Clasificación /Regresión	Predictivo	Poco interpretable

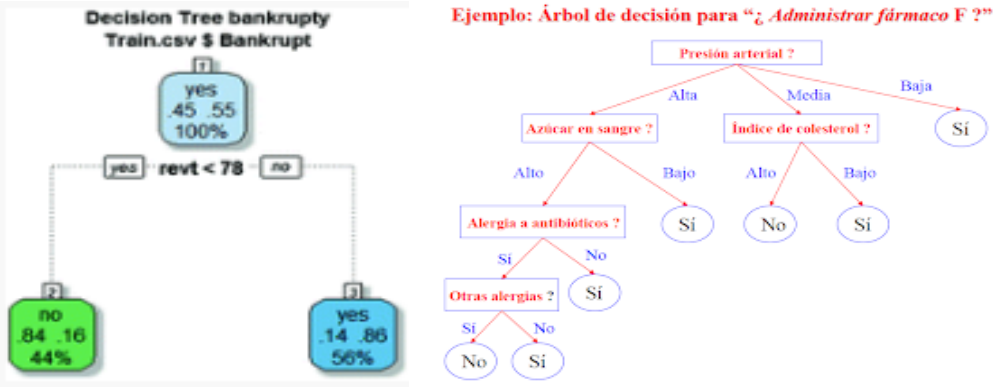
[140] Cuadro extraído de: Information Commissioner's Office. *Explaining decisions made with AI. Part 1. The basics of explaining AI*, mayo de 2020. Accesible en: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence/annexe-2-algorithmic-techniques/

No podemos olvidar que en muchas ocasiones ni siquiera será necesario el uso de estos algoritmos sino que se puede acudir a los modelos estadísticos tradicionales para resolver el problema que se pretende resolver.

d.3) Decisiones técnicas relevantes

En función del algoritmo elegido los desarrolladores también deben adoptar toda una serie de decisiones relacionadas con el comportamiento que se desea que este desarrolle al generar el modelo. Pues bien, pese a que cada modelo presentará una fuerte personalización en función del tipo de algoritmo y la experiencia de los desarrolladores, mencionaremos algunas de las decisiones más relevantes que se suelen tomar y que en muchos casos tendrán consecuencias jurídicas cuando el sistema se implante. Entre otras:

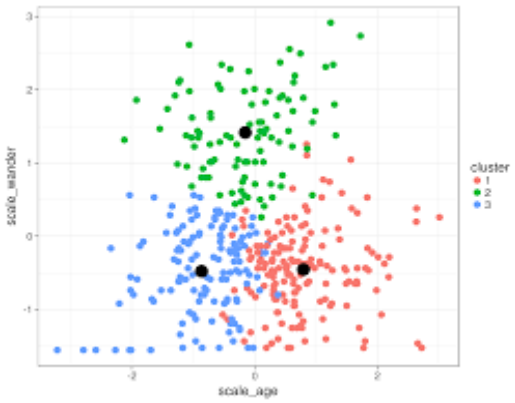
- El algoritmo conocido como *árbol de decisión* se caracteriza por establecer métodos de ramificaciones en los cuales divide los datos en nodos de decisión interrelacionados[141]. Estas ramificaciones terminan en clasificaciones o predicciones. El algoritmo determina automáticamente qué variables son las más importantes, basándose en su capacidad para clasificar los datos en la categoría de salida correcta.[142]. Pues bien, corresponde a los desarrolladores decidir el número de "ramas" o ramificaciones que contendrá el árbol de decisión. Así, un mayor número de ramificaciones dará como resultado un algoritmo más complejo y menos interpretable.



Árbol de decisión simple (izquierda). Árbol de decisión con varias ramificaciones(derecha)

[141] Imágenes obtenidas en:
https://sites.google.com/site/extraccionyrecuperaciondeinfo/home/arbolesdedecision
[142] OLSON,D Y DESHENG DASH,W: *Predictive Data Mining Models*, op.cit., pág.82.

- El algoritmo de *random forest* o bosque aleatorio construye un modelo predictivo combinando y promedia los resultados de múltiples árboles de decisión que son entrenados en subconjuntos aleatorios de características compartidas y datos de entrenamiento. Los diseñadores del modelo deben decidir cuantos números de árboles de decisión quieren que tenga su modelo. A mayor número de árboles, más complejo será el sistema.
- Por su parte, el algoritmo *K-means* es utilizado en métodos de aprendizaje no supervisado y su función esencial es la de agrupar el conjunto de datos de los que se dispone en distintos grupos o *cluster*. Así, los diseñadores deben decidir en cuantos grupos se debe dividir el conjunto de datos. En función de ello, se pueden obtener unos u otros resultados.



Como se puede observar en la imagen de la izquierda[143]. En este caso se ha optado por establecer tres grupos de diferentes colores. Cada punto negro es el centro de cada uno de los grupos del total de datos que comprende el *data set.*

- Anteriormente ya explicamos en qué consistían las redes neuronales. Ahora hemos de mencionar algunas de las decisiones técnicas que han de tomar los diseñadores de estas redes. En concreto, durante el procesamiento de los datos deben indicar el número de capas que pretenden que contenga la red neuronal, el conjunto de neuronas por capa y el número de épocas que debe realizar el algoritmo para optimizar el modelo, es decir, el número de repeticiones.
- A su vez, muchos algoritmos de clasificación indican como resultado una puntuación específica de predicción. Dicha puntuación establece la certeza del modelo de que la muestra que se aporta al sistema pertenece a la etiqueta positiva, esto es, la que pretendemos predecir. Para tomar la decisión sobre si la observación debe clasificarse como positiva o negativa, los diseñadores normalmente establecen un umbral de clasificación o corte para que el propio algoritmo al arrojar el resultado y en función de este decida cómo asignar la

[143] Imagen obtenida en: https://rpubs.com/cyobero/k-means

etiqueta. Ejemplo: Supongamos que se quiere implantar un sistema que al introducir una imagen de un gato sea capaz de distinguirla de otras imágenes de no gatos. De manera que la etiqueta positiva sería gato y la negativa no gato. Así, normalmente, cuando a un modelo se le incorpora la muestra, en nuestro caso una determinada fotografía, expulsará un resultado en forma de porcentaje. Ese porcentaje puede ser 0,7, 0,3, 0,99 sobre 1. A los diseñadores le corresponde establecer en qué concreto porcentaje, esto es, en qué umbral, una determinada muestra que se incorpora a un sistema debemos considerarla que pertenece a la clase deseada, en este caso gato. Si se establece que el umbral de corte es de 0,7, todas las imágenes ingresadas en el modelo que superen ese umbral serán clasificadas por el algoritmo como pertenecientes a la etiqueta de salida gato. A la inversa, todas las fotografías o muestras que se incorporen al algoritmo y este emita un porcentaje inferior al señalado lo clasificará en la etiqueta negativa o no deseada, en este caso, no gato.

d.4) La fase del entrenamiento

Elegido el algoritmo, el último paso para desarrollar el modelo será el procesamiento de los datos en dicho algoritmo. Ese procesamiento es conocido como la fase de entrenamiento. Es decir, durante esta fase el sistema es entrenado a través del procesamiento de datos para conseguir el objetivo previamente marcado. En función del problema que se pretenda resolver en el entrenamiento se tratará de clasificar, agrupar, asociar, etc. Además, dicho entrenamiento podrá ser tanto supervisado como no supervisado. Una vez que el modelo ha sido entrenado este ya se habrá creado. El siguiente paso será su evaluación para comprobar el nivel de bondad del mismo, es decir, el nivel de precisión del mismo.

E) Evaluación del modelo. Prueba del modelo

Aunque el proceso de evaluación de resultados que arroja el modelo debe ser permanente, es en la fase de su diseño donde la evaluación tiene más relevancia. En este sentido, las pruebas iniciales de evaluación van a reflejar los primeros indicios de que un determinado modelo puede estar tomando decisiones poco precisas o no adecuadas

ya que lo que se pretende en esta fase es evaluar el error del modelo a la hora de predecir la salida ante nuevos datos de entrada. Esto permitirá a los desarrolladores realizar alteraciones en el diseño del modelo que hasta ese momento no se hubieran tenido en cuenta para revertir los potenciales errores. De esta manera, para el desarrollador del modelo el coste de oportunidad de lanzar un producto al mercado que toma decisiones incorrectas se reduce ya que antes de su despliegue se valora su precisión. Además, los particulares que posteriormente se verán sometidos a estas decisiones también estarán más protegidos dado que el sistema ha sido testado y se han corregido las principales consecuencias negativas que el modelo puede efectuar.

Al igual que ocurría con el desarrollo de unos modelos u otros, el tipo de algoritmo y la técnica de aprendizaje que se utilicen hacen que los métodos de evaluación a realizar también sean distintos. En los siguientes párrafos mencionamos y explicamos algunos de los que frecuentemente se utilizan por los desarrolladores. Sin embargo, somos conscientes que existen muchas otras técnicas que no se mencionan en este trabajo.

e.1) Testar el sistema

Uno de los métodos esenciales utilizados para evaluar el sistema ya ha sido mencionado en páginas anteriores cuando hacíamos referencia a la fase de pre procesamiento de los datos. Esto es, la llamada validación cruzada o *cross validation*[144]. Así, mediante esta técnica, el conjunto de datos con los que se contaba eran separados en dos grupos de forma aleatoria[145]. Es decir, datos de capacitación o entrenamiento, utilizados para el desarrollo del modelo, y datos de prueba o validación. Pues bien, es en este momento cuando esos datos de prueba que se reservaron se utilizarán para chequear el funcionamiento del modelo o modelos elegidos. Con ello se trata de valorar si el modelo ha sido capaz de generalizar las principales correlaciones existentes en los datos de entrenamiento para que, una vez se introducen nuevos datos no vistos anteriormente, el sistema los pueda catalogar adecuadamente en base a las variables que presentan esas nuevas muestras. En este sentido, cabe mencionar que la partición de datos es un

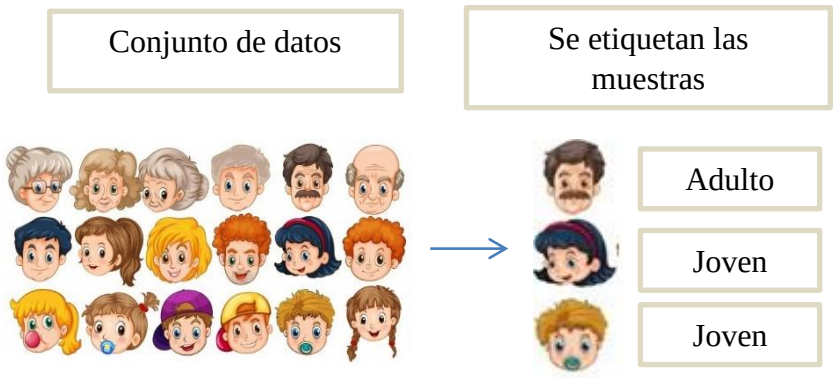
[144] ELKAN, C: "Evaluating classifiers". *University of San Diego, California*, 2012, pág.4. Disponible en: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.233.2746

[145] Dentro de la validación cruzada existen a su vez distintos enfoques o tipos como *hold-out*- que es el más simple-, el *k-fold cross-validation, el exhaustive cross validation, el leave-one-out cros-validation, etc.* En: ARLOT,S Y CELISSE,A: "A survey of cross-validation procedures for model selection". *Statistics Surveys*, Vol. 4, 2010, págs.. 40–79

proceso que se realiza generalmente cuando se utilizan algoritmos de aprendizaje supervisado. Ello es así debido a que, dado que en este tipo de algoritmos los datos utilizados están etiquetados, es decir, se sabe de antemano el valor de la etiqueta que se busca predecir, las técnicas de evaluación de partición permiten comparar el valor real con el valor predicho por el modelo ya creado.

Como ya se ha señalado anteriormente. Los datos que se utilizan en la fase de entrenamiento han de ser distintos a los de prueba[146]. Dado que lo que se pretende es evaluar cómo el sistema ha aprendido, no es posible utilizar los mismos ya que dicha valoración no sería real debido a que arrojaría una alta precisión por el simple hecho de que la muestra coincide, no por la generalización que el sistema haya podido realizar.

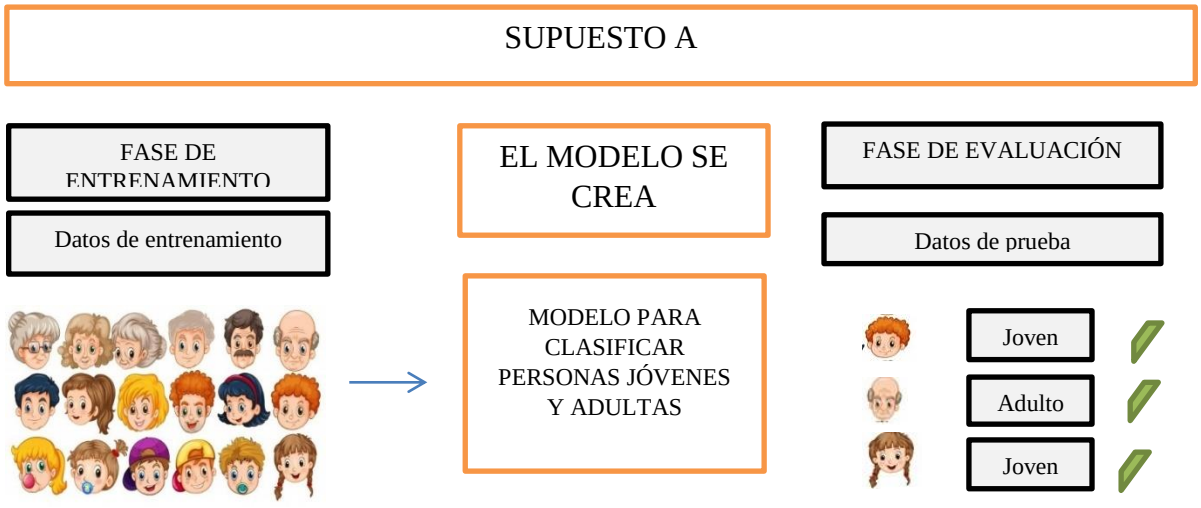
Para reflejar lo señalado mostramos dos ejemplos de separación de los datos de un conjunto de datos o *data set*. Imaginemos que queremos crear un sistema que sea capaz de diferenciar personas adultas de jóvenes. Para ello contamos con un *data set* conformado por imágenes de personas adultas y jóvenes. El primer paso consiste en etiquetar los datos de los que disponemos. En este caso, cada imagen se etiqueta como joven o adulta.



Elaboración propia

[146] Aunque es un error importante, a veces se suelen utilizar parte de las muestras que se utilizaron en la fase de entrenamiento para la fase de prueba. En: CHOULDECHOVA, A; BENAVIDES-PRADO, D; FIALKO, O Y VAITHIANATHAN, R: "A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions". *Proceedings of machine learning research*, 2018, pág.16. Disponible en: http://proceedings.mlr.press/v81/chouldechova18a.html

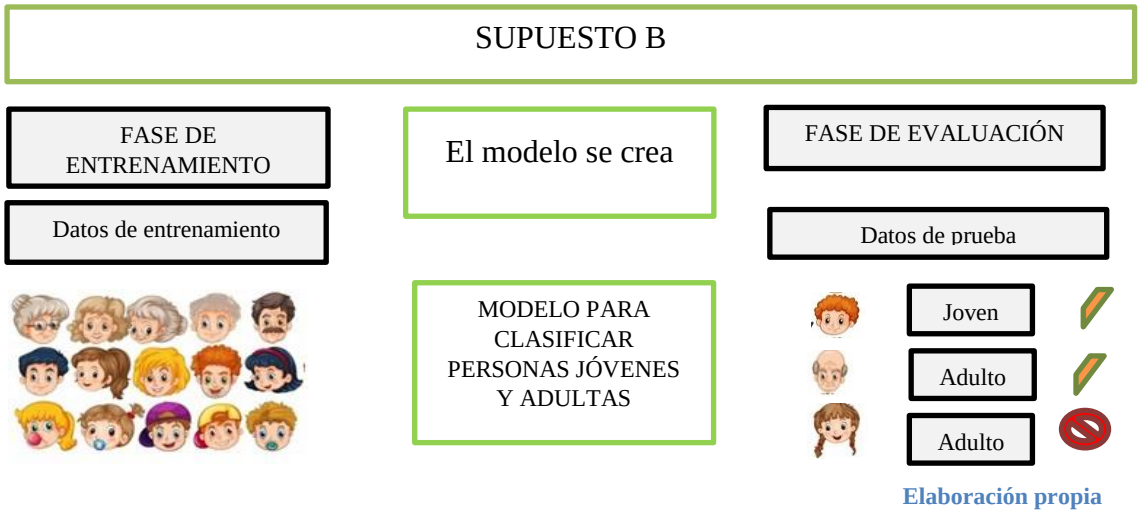
Supuesto A. Los datos se separan pero se utilizan los mismos tanto para la fase de entrenamiento como la de prueba.



Elaboración propia

Como utilizamos los mismos datos de prueba y entrenamiento. Al utilizar los datos de prueba el sistema predecirá correctamente todas las imágenes. Es decir, podríamos entender que tiene un nivel de precisión del 100%. Sin embargo ,ello no es fiable, ya que realmente, las imágenes que se han utilizado en la fase de prueba son idénticas a las que se utilizaron durante la fase de entrenamiento, por tanto, el sistema ha predicho las etiquetas por el mero hecho de que en los datos de entrenamiento esa persona también había sido catalogada como adulta. Aunque este modelo puede que haya generalizado bien, a través de esta validación cruzada no podremos conocerlo a ciencia cierta.

Supuesto B: Los datos se separan. Parte se destinan a la fase de entrenamiento y parte a la fase de prueba.



Elaboración propia

En este supuesto los datos de prueba son distintos a los de entrenamiento. Al incorporar en esa fase de prueba datos que el modelo nunca había visto, el nivel de acierto de las predicciones será bastante fiable en cuanto a la precisión del modelo. En este caso, el sistema ha predicho correctamente 2 etiquetas y se ha equivocado en 1. Su nivel de precisión es del 66%. En este caso, los desarrolladores cuentan con una información fiable de la generalización del sistema. A partir de ahí deben valorar si siguen con el desarrollo del modelo o en su caso prefieren mejorar esa precisión probando con nuevos algoritmos o datos.

Con la validación cruzada no se consigue que el sistema mejore el entrenamiento o la capacidad de precisión sino que se utiliza como método para valorar si el sistema es realmente preciso o no. Esa es la clave, medida de testeo, no de mejora del sistema. Insistimos, obviar la validación cruzada no significará que el sistema no aprenda adecuadamente durante la fase de entrenamiento. Simplemente no sabremos si el sistema ha generalizado adecuadamente o no. Se previene que, una vez el modelo se ingrese en la vida real, este adopte decisiones con un grado de acierto relativamente bajo o no asumible por la organización que lo desarrolló por falta de evaluaciones previas.

e.2) La matriz de confusión

La matriz de confusión es una herramienta que permite la visualización del rendimiento de un modelo algorítmico. Tal herramienta aparece representada normalmente en una tabla. De manera que cada fila representa las etiquetas de la clase predicha por el modelo mientras que cada columna representa las etiquetas de la clase real[147].

		Valor predicho		
		Buen pagador	**Mal pagador**	
Valor Real	**Buen Pagador**	13	4	17
	Mal pagador	7	70	77
		20	74	94

Matriz de confusión de modelo que pretende predecir la solvencia patrimonial

[147] https://en.wikipedia.org/wiki/Confusion_matrix

Como resultado lógico, esta técnica es utilizada generalmente en la evaluación de algoritmos de aprendizaje supervisado debido a que estos se entrenan con datos etiquetados. Esta primera matriz de confusión representa el conjunto de etiquetas que el modelo algorítmico ha predicho adecuadamente respecto del total de etiquetas que se han utilizado para chequear el sistema. Concretamente en este modelo lo que se pretendía predecir es si una persona es solvente o no etiquetando las salidas en buen y mal pagador. Etiqueta positiva y negativa respectivamente. Pues bien, de las 20 personas que estaban etiquetadas como buenas pagadoras el modelo ha predicho bien a 13 de ellas, mientras que ha errado en 7 al catalogarlas como malas pagadores cuando realmente eran buenas. A su vez, de las 74 personas catalogadas como malas pagadoras, el sistema ha acertado en 70 ocasiones, errando únicamente en 4 al considerarlas buenas pagadoras cuando realmente eran malas.

		Valor predicho		
		Tumor maligno	**Tumor benigno**	
Valor Real	**Tumor maligno**	13	4	17
	Tumor benigno	7	70	77
		20	74	94

Matriz de confusión para un modelo que valora el riesgo de sufrir un tumor maligno

A través de estos datos se pueden llegar a conclusiones importantes sobre el rendimiento del modelo pudiendo valorar los distintos resultados. Para realizar un correcto análisis del modelo es necesario indicar cuál es la clase positiva y negativa que se pretende ya que nos permitirá valorar el número de verdaderos y falsos positivos y el número de falsos y verdaderos negativos. En este caso, la clase positiva es detectar tumores malignos y la negativa detectar tumores benignos en un total de 94 muestras.

	Verdadero positivo: 13	**Falso positivo: 4**
El sistema predice	Tumor maligno	Tumor Maligno
La realidad es	Tumor benigno	Tumor benigno
	Falso negativo: 7	**Verdadero negativo: 70**
El sistema predice	Tumor benigno	Tumor Benigno
La realidad es	Tumor Maligno	Tumor Benigno

Número de verdaderos positivos y negativos, así como de falsos positivos y negativos.

De esta manera, los *verdaderos positivos* (VP) quedan abarcados por todas las salidas que el modelo predice como clase positiva correctamente. Es decir, el modelo predice que una persona es buena pagadora y realmente lo es. A su vez, los *verdaderos negativos* (VN) serán los resultados que el modelo predice correctamente la clase negativa. Esto es, el modelo pronostica que una persona es mala pagadora y realmente lo es. Por otro lado, el conjunto de resultados que el modelo catalogue incorrectamente la clase positiva se identifican con los *falsos positivos* (FP). A saber, el modelo indica que una persona es buena pagadora pero en realidad no lo es. Finalmente, cuando catalogue incorrectamente la clase negativa, estos resultados se conocerán como *falsos negativos (FN)*. Es decir, el modelo indica que una persona es buena pagadora pero en realidad no lo es.

Insistimos, el evaluador del modelo conoce el valor real de la muestra analizada porque dichas muestras o ejemplos están etiquetados. Esa es la gran virtud de este tipo de evaluaciones.

e.2.1) Métricas de evaluación del rendimiento del modelo

La matriz de confusión puede arrojar mucha información del desempeño de un modelo. Concretamente, con los resultados que se proyectan se puede analizar distintas métricas y extraer conclusiones muy relevantes para los desarrolladores de los modelos. Nos centraremos en las principales métricas que se suelen utilizar en estos contextos. Estas son: exactitud, precisión, sensibilidad y especificidad[148].

Tasa de acierto /accuracy: La exactitud es una métrica que nos indica el porcentaje de acierto del sistema. Es decir, la fracción de predicciones que el modelo realiza correctamente. Desde un punto de vista técnico. La exactitud se representa a través de la siguiente ecuación:

$$Exactitud = VP+VN\ /VP\ +VN\ +FP\ +FN$$

Uno de los principales inconvenientes de la exactitud es que, dado que esta métrica únicamente se centra en captar los aciertos del sistema, al no medir los falsos negativos y los falsos positivos, los resultados podrían ser muy engañosos. De manera que un sistema puede ser altamente exacto porque tiene un porcentaje de exactitud del

[148] PANESAR, A: *Machine Learning and AI for Healthcare. Big Data for Improved Health Outcomes*, op cit., págs.95 y ss y 196 y ss.

92%, pero en cambio, puede que el otro 8 % pertenezca a un conjunto de falsos negativos. En determinados ámbitos un número de falsos negativos relativamente bajo puede no ser tolerable[149].

Ejemplo: Etiqueta positiva: Presenta riesgo de enfermedad

	Verdadero positivo 13	**Falso positivo 4**
El sistema dice	Presenta enfermedad	Presenta enfermedad
La realidad es	Presenta enfermedad	No presenta enfermedad
	Falso negativo 7	**Verdadero negativo 70**
El sistema dice	No presenta enfermedad	No presenta enfermedad
La realidad es	Presenta enfermedad	No presenta enfermedad

Precisión o valor predictivo positivo: Esta métrica trata de evaluar el desempeño del modelo en relación con las predicciones positivas. Es decir, en estos casos lo que se pretende resolver es la siguiente pregunta. ¿Qué proporción de predicciones indicadas por el modelo como positivas son realmente correctas? La precisión se define de la siguiente manera:

$$Precisión = VP / VP + FP$$

Como puede observarse, en esta métrica se tienen en cuenta los verdaderos positivos y los falsos positivos debido a que lo que se pretende valorar es el acierto de las predicciones positivas realizadas por el modelo.

Sensibilidad/ /Tasa de verdaderos positivos[150]*:* En este supuesto lo que se valora es la proporción de positivos reales que se identificó por el sistema correctamente. La sensibilidad se representa de la siguiente manera:

$$Sensibilidad = VP / VP + FN$$

[149] En muchos entornos médicos, un falso positivo no es tan grave como un falso negativo. Por ejemplo, los riesgos de no detectar una enfermedad normalmente serán mayores que diagnosticar erróneamente una enfermedad que realmente esa persona no ostenta. Un falso negativo significaría que no se diagnosticó a alguien, lo que quizás sea más preocupante. En: PANESAR, A: *Machine Learning and AI for Healthcare. Big Data for Improved Health Outcomes*, op cit., pág.155.

[150] Scikit-Learn: Metrics and scoring: quantifying the quality of predictions. Visto en: https://scikit-learn.org/stable/modules/model_evaluation.html

A diferencia de la métrica de la precisión, en este caso se tienen en cuenta los verdaderos y los falsos negativos ya que son el conjunto de datos que realmente son positivos. Aunque el modelo en algunos casos lo haya catalogado como negativos.

Especificidad/ *Tasa de verdaderos negativos:* En la especificidad lo que se trata de valorar es la proporción de negativos reales que se identificaron correctamente por el sistema[151]. Esta métrica se representa de la siguiente manera:

Especificidad = VN/VN+FP

Como puede observarse, la especificidad es lo contrario a la sensibilidad. En estos casos se tienen en cuenta el conjunto de datos que realmente son negativos independientemente de que el modelo haya catalogado algunos de ellos como positivos.

Buen pagador/mal pagador

Etiqueta positiva: Buen pagador
Etiqueta negativa: Mal pagador

	Verdadero positivo 13	**Falso positivo 4**
El sistema dice	Buen pagador	Buen pagador
La realidad es	Buen pagador	Mal pagador
	Falso negativo 7	**Verdadero negativo 70**
El sistema dice	Mal pagador	Mal pagador
La realidad es	Buen pagador	Mal pagador

Métricas.

- Tasa de acierto *accuracy*: VP+VN/ VP+FP+FN+VN = 83/94 = 0,88 → **88% de exactitud**.

El sistema tiene un acierto general en sus predicciones de un 88%. El 88% de las veces que el sistema indica que una persona es buena pagadora o mala pagadora, acierta.

-Precisión/ valor predictivo positivo: VP / VP+FP = 13/17 = 0,76 → **76 % de precisión.**

[151] Towards Data Science: Various ways to evaluate a machine learning model's performance. Visto en: https://towardsdatascience.com/various-ways-to-evaluate-a-machine-learning-models-performance-230449055f15

Cuando el sistema indica que alguien es un buen pagador, acierta un 76 % de las veces. A la inversa, el 24% de las ocasiones que el sistema indica que una persona es buena pagadora no lo es.

-Sensibilidad/ *Tasa de verdaderos Positivos* VP/VP+FN= 13/20= 0,65 → **65% de sensibilidad.**

El sistema detecta realmente al 65% de los buenos pagadores. El 35% de las ocasiones el sistema no detecta a los buenos pagadores.

-Especificidad/ *Tasa de verdaderos negativos*: VN/ VN+FP = 70/74 = 0,95 → **95% especificidad**. El sistema detecta realmente al 95% de los malos pagadores. El 5 % de las ocasiones el sistema no detecta a los malos pagadores.

Otro ejemplo:

Tumor maligno /tumor benigno

Etiqueta positiva: Tumor maligno
Etiqueta negativa: Tumor benigno

	Verdadero positivo 13	**Falso positivo 4**
El sistema dice	Tumor maligno	Tumor maligno
La realidad es	Tumor maligno	Tumor benigno
	Falso negativo 7	**Verdadero negativo 70**
El sistema dice	Tumor benigno	Tumor benigno
La realidad es	Tumor maligno	Tumor benigno

Métricas.

- Tasa de acierto */accuracy*: VP+VN/ VP+FP+FN+VN = 83/94 = 0,88 → **88% de exactitud**.

El sistema tiene un acierto general en sus predicciones de un 88%. El 88% de las veces que el sistema indica que una persona tiene un tumor maligno o benigno acierta.

-Precisión/ valor predictivo positivo: VP / VP+FP = 13/17 = 0,76 → **76 % de precisión.**

Cuando el sistema indica que alguien tiene un tumor maligno, acierta un 76 % de las veces. A la inversa, el 24 % de las ocasiones que el sistema indica que una persona tiene un tumor maligno, no lo tiene.

-Sensibilidad/Tasa de verdaderos positivos: VP/VP+FN= 13/20= 0,65 → **65% de sensibilidad.**

El sistema detecta realmente en un 65% de las ocasiones que alguien tiene un tumor maligno. El 35% de las ocasiones el sistema no detecta los tumores malignos.

-Especificidad/ *Tasa de verdaderos negativos*: VN/ VN+FP = 70/74 = 0,95 → **95% especificidad**.

El sistema detecta realmente en un 95% a las personas que tienen un tumor benigno El 5 % de las ocasiones el sistema no detecta los tumores benignos.

	Etiqueta a predecir	
	Buen pagador/Mal pagador	Tumor maligno/Tumor benigno
Tasa de acierto /accuracy	88% de las ocasiones el sistema detecta a los buenos y malos pagadores. Existe un 12 % de veces que el sistema no acierta ni los buenos pagadores ni los malos	88% de las veces el sistema detecta a los tumores malignos y benignos. Existe un 12 % de veces que el sistema no acierta ni los tumores malignos ni los benignos.
Precisión/ valor predictivo positivo	El 76% de las veces que el sistema indica un ejemplo como buen pagador acierta. El 24 % de las veces que el sistema indica que una persona es buena pagadora no lo es. (FP)	El 76 % de las veces que el sistema indica que una persona presenta un tumor maligno, lo tiene. El 24% de las veces que el sistema indica que una persona tiene un tumor maligno, no lo tiene en realidad.
Sensibilidad/ Tasa de verdaderos positivos	El Sistema detecta el 65% de las muestras que son buenas pagadoras. El 35 % de las veces, el sistema no detecta a los buenos pagadores. (FN)	El sistema detecta el 65 % de los ejemplos que presentan tumores malignos. El 35 % de las veces, el sistema no detecta los tumores malignos.
Especificidad/	El sistema detecta en un 91% de	El sistema detecta el 95% de las muestras que presentan tumores

Tasa de verdaderos negativos	ocasiones a los malos pagadores. El sistema no detecta al 9% de malos pagadores.	benignos. El sistema no detecta el 5% de muestras que presentan tumores benignos.

Elaboración propia

Una vez tenemos todos los resultados que se han derivada del conjunto de métricas es turno de analizar dichos resultados para valorar si estos son adecuados y permisibles para el contexto donde se pretende implantar el sistema. Así, los efectos de los falsos negativos en un sistema que pretende predecir el riesgo de padecer cáncer son totalmente distintos a los efectos de los falsos negativos en un sistema que pretende predecir si una persona es solvente económicamente. Para la primera situación un paciente no será adecuadamente diagnosticado al considerarse por el sistema que no padece cáncer mientras que en el segundo supuesto una persona no recibirá un crédito aunque realmente cumplía con los requisitos.

Los evaluadores del sistema deben realizar un análisis de este conjunto de métricas y valorar si, basados en los resultados que se ofrecen, consideran adecuado que este modelo pueda ser trasladado a un contexto donde el sistema irradiará sus efectos en un futuro. Es decir, lo que deben preguntarse en este momento aquellos que analizarán estos resultados es si los porcentajes que se han derivado de la prueba pueden o no ser tolerables en esa realidad en la que se pretende implantar el sistema. En este sentido, podría ser perfectamente asumible que un sistema tenga una sensibilidad muy baja en unos contextos o una exactitud muy alta en otros o a la inversa. Todo dependerá del contexto y las consecuencias que se derivan de esos errores. Veamos varios ejemplos de lo indicado previamente:

		Valor predicho		
		Tumor maligno/Buen pagador	**Tumor benigno/Mal pagador**	
Valor Real	**Tumor maligno/Buen pagador**	13	4	17
	Tumor benigno/Mal pagador	7	70	77
		20	74	94

Tasa de acierto /*accuracy*: a priori, si sólo atendemos a la exactitud, es decir, precisión global y error global. Podría entenderse que el sistema es bastante adecuado en los dos ámbitos analizados. Quizás, en el caso de los tumores sería arriesgado establecer un sistema que el 12% de las veces no sabe diferenciar entre tumores benignos y malignos. Aunque podría compensarse con ayuda profesional humana.

Precisión/ valor predictivo positivo: si analizamos la precisión, tiene un acierto relativamente alto en relación con las etiquetas positivas que buscamos, esto es, tumor maligno y buen pagador. Por otro lado tiene un porcentaje medio-bajo de errores en la asignación de la etiqueta positiva.

Así, para valorar el coste de los falsos positivos:

- Sector bancario: en un 76% de las solicitudes de crédito la concesión es adecuada. Sin embargo, en un 24% de ocasiones se concede un crédito a un mal pagador. Para una entidad bancaria asignar el 24% de las ocasiones crédito a personas inadecuadas puede ser un coste muy alto.
- Sector sanitario: el sistema indica un número alto de tumores malignos (76 %). Sin embargo, existe un número de supuestos relativamente alto en el cual, el sistema considera que una persona presenta un tumor maligno, cuando realmente dicho tumor es benigno. La consecuencia directa en estos casos será que pacientes sanos reciban un tratamiento inadecuado. Lo ideal es que se compense con presencia humana que corrobore ello.

Sensibilidad/Tasa de verdaderos positivos: si analizamos la sensibilidad el sistema no tiene un porcentaje muy alto para detectar la etiqueta positiva, además, este presenta un porcentaje relativamente alto en no detectar las etiquetas positivas.

Así, para valorar el coste de los falsos negativos:

- Sector bancario: aunque el 65 % de los buenos pagadores son detectados por el sistema existe un porcentaje relativamente alto de personas que siendo buenas pagadoras no recibirán un crédito porque el sistema no las ha tenido en cuenta.
  A una entidad bancaria puede que no le interese que tantas personas no reciban crédito cuando realmente eran solventes.
- Sector sanitario: aunque el sistema detecta un porcentaje relativamente alto de tumores malignos (65 %), existe un porcentaje inasumible de casos en los que el

sistema no detectará los tumores malignos (35 %). Tanto las Administraciones Públicas como una clínica privada deben plantearse si pueden asumir tal índice de errores teniendo en cuenta las consecuencias que se pueden derivar.

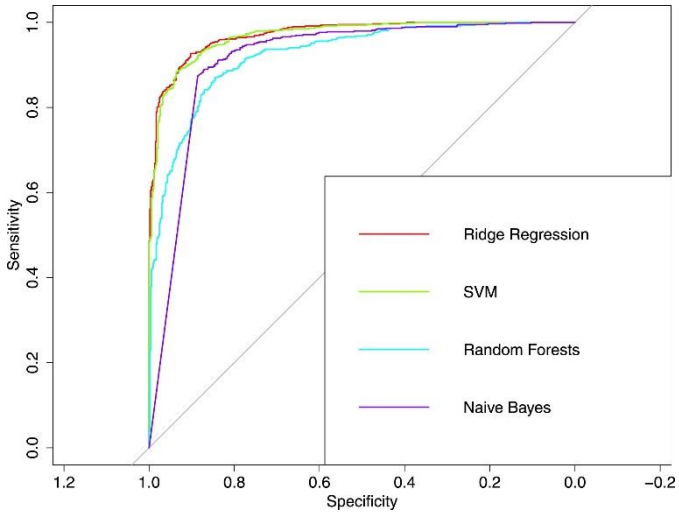
Especificidad/ Tasa de verdaderos negativos: el modelo presenta unos resultados muy adecuados para detectar la etiqueta negativa, esto es, malos pagadores y tumores benignos.

Así:

- Sector bancario: el 95 % de malos pagadores son detectados por el sistema. Una entidad bancaria que esté centrada en conceder créditos a pagadores de dudoso reintegro puede estar muy interesada en este modelo.
- Sector sanitario: el porcentaje de tumores benignos que detecta el sistema es muy alto. Tanto una Administración como una clínica pueden estar interesadas en adquirir este modelo si el objetivo esencial es ese.

Los resultados que arrojen estas métricas pueden llevar a los diseñadores del sistema a plantearse toda una serie de cuestiones. Por ejemplo, si es necesario volver a entrenar al modelo o incluso diseñar uno nuevo. También es posible que el papel inicial que se pretendía con el modelo se altere. Por ejemplo: si las métricas arrojan que un modelo presenta un número relativamente alto de falsos negativos esa organización podría replantearse la plena automatización o no de la decisión que adoptará posteriormente el sistema.

e.2.2) Técnicas para mostrar los resultados



Junto a las métricas señaladas anteriormente, también resulta habitual utilizar otra serie de herramientas que permitan mostrar gráfica e intuitivamente los resultados que se derivan de la matriz de confusión. Nos estamos refiriendo a las curvas ROC.

La curva ROC es una técnica utilizada para visualizar, organizar y seleccionar modelos de clasificación[152]. A través de dicho gráfico se pueden obtener conclusiones que permiten la representación de los resultados que arrojan distintos modelos cuando se están evaluando. Ello facilita la elección de aquel modelo que presente unos resultados óptimos para el objetivo que se persigue. A través de estas gráficas se puede representar la tasa de falsos positivos y la tasa de falsos negativos. [153]
En la imagen se puede visualizar una Curva ROC con diversos modelos. Cada color representa los resultados de distintos modelos basados en los algoritmos elegidos, los cuales, han procesado los mismos datos para diseñar tales modelos.

Además de las Curvas ROC, podemos indicar otras técnicas de representación de los resultados como los histogramas u otro tipo de herramientas que persiguen dicho fin.

e.3) Correlación/Causalidad

Muchos proyectos de inteligencia artificial parten de una o varias hipótesis basadas en que determinado conjunto de variables puede reflejar determinados comportamientos y patrones ocultos. Por ejemplo, las variables capacidad económica y temporalidad en el trabajo pueden ser características que reflejen la solvencia financiera. En otros supuestos, simplemente, y sin previa hipótesis, un conjunto de datos y variables son procesados en la búsqueda de las mentadas correlaciones[154]. Ambas combinaciones son posibles en los algoritmos tanto de aprendizaje supervisado como no supervisado. Así, resulta probable que los modelos detecten correlaciones esperadas y no esperadas en las variables observadas. Sin embargo, no todas las correlaciones observadas están relacionadas con un hecho causal sino que en muchos casos tal relación puede derivarse de simples casualidades. En este sentido, cabe destacar los

[152] FAWCET,T: "An introduction to ROC analysis". *Pattern Recognition Letters*, 27, 2006, pág. 861. Disponible en: https://www.sciencedirect.com/science/article/pii/S016786550500303X
[153] Imagen obtenida de: QUIJANO-SÁNCHEZ,L; LIBERATORE,F; CAMACHO-COLLADOS,J, CAMACHO-COLLADOS,M: "Applying automatic text-based detection of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the police". *Knowledge-Based Systems*,Volume 149, 2018,pág.159.
[154] En el big data "la causalidad de los hechos ya no es el objetivo del análisis de los datos, sino que lo que interesa es saber qué está pasando mediante el descubrimiento de perspectivas nuevas que han pasado desapercibidas a primera vista, gracias al análisis de las conexiones entre datos, pautas y correlaciones". GARCÍA ALSINA, M: *Big data: gestión y explotación de grandes volúmenes de datos*, op.cit., pág. 11.

principales errores que se derivan de las llamadas correlaciones espurias, estos son: el error por azar y el error por confusión[155]. El error por azar aparece sobre todo cuando se utilizan un número alto de variables para estudiar un fenómeno. Así, a mayor número de variables mayor es la probabilidad de que se detecten correlaciones entre estas variables que no tienen por qué obedecer a causas razonables[156]. Debido a la gran cantidad de datos que se pueden procesar a través de las técnicas de aprendizaje automático, la probabilidad de que se detecten errores por azar aumenta[157]. Junto al error por azar también es habitual que aparezca el llamado error por confusión. En este caso a primera vista puede considerarse que existe una correlación entre dos variables pero ciertamente, existe una tercera que es la que realmente está correlacionada con ambas[158].

En ambos errores los resultados arrojados por los modelos deben ser analizados detenidamente para evitar que aquellas correlaciones espurias acaben convirtiéndose en la regla general de funcionamiento del sistema. Es decir, la gravedad de diseñar un sistema que se basa en correlaciones espurias tendrá consecuencias muy negativas cuando el sistema se ponga en funcionamiento en la vida real y acabe adoptando decisiones automatizadas que afecten a los derechos de los particulares. En muchos casos las correlaciones espurias serán fácilmente constatables, sobre todo, si quien las analiza es un experto en el contexto en el que se estudia el patrón[159]. En cambio, conforme aumenta el número de datos a analizar y el tipo de correlaciones observadas la posibilidad de detectar aquellas que obedecen a la simple casualidad se reduce. Así, existen multitud de teorías descabelladas que pueden ser perfectamente puestas en práctica a través de estas nuevas tecnologías[160]. Basta con disponer de un número importante de datos y buenas herramientas tecnológicas. Por tanto, incluso en aquellos casos en los que la hipótesis inicial sea corroborada por las técnicas algorítmicas tal

[155] Elena Gil González explica de forma muy nítida los tipos de errores clásicos que existen cuando se analizan correlaciones espurias. GIL GONZÁLEZ,E : *Big data, privacidad y protección de datos*, op.cit., págs. 28 y 37. En el mismo sentido: POLITOU,E; ALEPIS,E: PATSAKIS,C: "Profiling tax and financial behaviour with big data under the GDPR". *Computer Law & Security Review*, Volume 35, Issue 3, 2019, pág.4. Disponible en: https://www.sciencedirect.com/science/article/pii/S026736491830133X

[156] GIL GONZÁLEZ,E : *Big data, privacidad y protección de datos*, op.cit., pág. 30.

[157] HERSCHEL,R Y MIORI,V,M: "Ethics & Big Data". *Technology in Society*,Volume 49,2017,pág.33.

[158] GIL GONZÁLEZ,E : *Big data, privacidad y protección de datos*, op.cit., pág.33

[159] En el ámbito médico, cuando un modelo arroja o indica una correlación que contradice todos los experimentos médicos realizados hasta el momento se ha de cuestionar.. Ahora bien, han de ser expertos médicos los que deben cuestionar tal correlación. En: En: LEMLEY, M,A Y CASEY,B: "Remedies for Robots.", op.cit., pág. 1336.

[160] Hay estudios poco fiables que indican que de acuerdo a los rasgos faciales es posible averiguar si una persona es o no homosexual o si una persona es o no conflictiva. Aunque las teorías indicadas pueden ser puestas en entredicho, a través del aprendizaje automático estas pueden ponerse en práctica. Visto en: https://www.nytimes.com/2019/07/10/opinion/facial-recognition-race.html?smid=tw-share

hipótesis deberá ponerse en entredicho y cuestionarse. Obligando a los diseñadores a ser críticos con los resultados y correlaciones que se obtienen, todo ello, como paso previo a la puesta en funcionamiento del sistema.

Contrastar la correlación o patrón observado que se deriva del modelo no sólo restringirá la posibilidad de que el sistema arroje posteriormente decisiones injustas sino que además ello podría llegar a limitar determinadas responsabilidades futuras de los diseñadores del sistema al demostrar que fueron debidamente diligentes en el estudio de dichos patrones[161]. Tales correlaciones y su correspondiente justificación deberían quedar registradas a efectos de posteriores explicaciones y fundamentaciones de las decisiones adoptadas.

F) La elección del modelo: Conformación del sistema

Una vez que se ha evaluado el rendimiento de los modelos el siguiente paso es elegir aquel modelo que más se adecúe al contexto y objetivos previamente fijados durante la planificación del proyecto. Este modelo será el que se integrará posteriormente en el sistema que adoptará las decisiones automatizadas y que desplegará sus efectos en la vida real. Ahora bien, la elección de un modelo u otro dependerá de multitud de factores que deben ser tenidos en cuenta en este momento dado los efectos potenciales que puede generar este sistema en el entorno al que se enfrentará.

Como regla general, en la mente de cualquier desarrollador de este tipo de sistemas reside el objetivo principal de crear y elegir aquel modelo que sea capaz de generar las predicciones o resultados más precisos. Siendo secundarias otras características como la interpretabilidad o la simplicidad del sistema. En este sentido, y aunque cada vez existe un mayor interés por crear sistemas interpretables, lo cierto es que a día de hoy los sistemas que suelen ser más precisos son a su vez lo más complejos y menos interpretables[162]. La disyuntiva entre elegir uno u otro modelo no es baladí

[161] Así se ha indicado expresamente en: Australian Human Rights Commission. *Human rights and Technology. Discussion Paper*, op.cit., pág.87.

[162] Interpretabilidad y exactitud son dos parámetros que están en continuo conflicto a la hora de optar por un modelo u otro. EN: CHOULDECHOVA, A; BENAVIDES-PRADO, D; FIALKO, O Y VAITHIANATHAN, R: "A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions". *Proceedings of machine learning research*, op.cit., pág.10.

cuando se pretenda implantar en un entorno en el que se requiere necesariamente de la explicación de los resultados emitidos. En estos supuestos la elección por un modelo más interpretable en defecto de otro menos interpretable y más preciso puede ser una obligación.

Junto a la interpretabilidad y simplicidad del sistema, factores como la facilidad de adaptación del modelo a distintos contextos, el número de falso positivos y negativos o la frecuencia con la que es necesaria volver a reentrenar al sistema también resultarán elementos esenciales a la hora de optar por un modelo u otro.

Tras la elección del modelo sigue resultando recomendable volver a realizar chequeos previos y simulaciones antes de que el mismo tome decisiones automatizadas que afecten a los particulares[163]. Las evaluaciones previas a la puesta en marcha del sistema pueden realizarse por la propia organización que lo ha diseñado o por aquellas terceras organizaciones que lo adquiere[164]. Estas pruebas se clasifican en estáticas y dinámicas.

Por lo que se refiere a las pruebas estáticas estas consisten habitualmente en la lectura del código fuente del sistema elegido. Ello permite a los analistas valorar las tecnologías involucradas en la implementación del mismo. Aunque esta información ya puede haber sido facilitada por el desarrollador del sistema, no está de más que el nuevo adquiriente valore estos elementos.

Junto a los métodos estáticos resulta habitual que también se realicen pruebas piloto donde se ponga en práctica el sistema. Son las llamadas pruebas o métodos dinámicos. A diferencia de los métodos estáticos donde sólo se analizan las fallas internas a través del código fuente, en las dinámicas el sistema es probado en un entorno parecido o similar al que en el futuro previsiblemente operará. Ello permitirá a la organización hacerse una idea de cómo el sistema funcionará y adoptará las decisiones en el futuro. No obstante, incluso en estos casos, tales pruebas siempre serán limitadas

---

PROCEEDINGS OF THE 1ST CONFERENCE ON FAIRNESS, ACCOUNTABILITY AND TRANSPARENCY, IN PMLR</I> 81:134-148. También en: BIBAL, A; LOGNOUL, M ; DE STREEL, A:. "Legal requirements on explainability in machine learning". *Artif Intell Law*, 2020,pág.16. Disponible en: https://doi.org/10.1007/s10506-020-09270-4

[163] Consejo de Europa. *Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data*, 2017, pág. 4.

[164] KROLL, A,J;. HUEY, J; BAROCAS, S; FELTEN, E, W; REIDENBERG, J,R;.ROBINSON, D,G;. Y YU, H:, "Accountable Algorithms". *University of Pennsylvania Law Review*, Vol. 165, 2017, págs .646 y ss. Disponible en: https://scholarship.law.upenn.edu/penn_law_review/vol165/iss3/3/

ya que resulta cuanto menos imposible testar cada una de las situaciones a las que puede enfrentarse el sistema o prever cualquier hecho improbable que puede llevar a la adopción de decisiones no previsibles. Se debe ser consciente de que no siempre el sistema funcionará bien y por tanto, la fiabilidad y robustez no será plena[165].

Por tanto, elegido y revisado el modelo, este ya se encuentra en condiciones de ser implantado en el sistema que se utilizará para la toma de decisiones. Para finalizar, resulta pertinente indicar que todas las fases del diseño de los sistemas queden registradas. La elaboración de un manual técnico que incluya información de cómo trabaja el sistema también resultará sumamente útil y ayudará a la trazabilidad del mismo.

FASE	Etapas	Principales actuaciones que se realizan
DISEÑO	**Planificación del proyecto**	-Establecer hipótesis. - Fijar los objetivos pretendidos por el sistema. - Situar el marco normativo.
	Recopilación de los datos	-Bases de datos donde se obtienen los datos. -Fuente de procedencia de los datos.
	El pre-procesamiento de datos	-Selección de los datos a procesar. -Limpieza de datos. -Análisis de los conjuntos de datos elegidos. -Separación del conjunto de datos.
	El desarrollo de los modelos. El entrenamiento.	-Elección de algoritmos. -Elección de modelo de aprendizaje. -Fase de entrenamiento de los datos. -Toma de decisiones técnicas: Umbral de decisión, nº de árboles de decisión, nº de capas de la red neuronal, etc.
	Evaluación de los modelos	-Testar y probar los distintos modelos. -Aplicar métricas de evaluación del rendimiento.
	Elección de los modelos	-Elección de los modelos. -Verificación y validación del modelo elegido.

3. La fase de despliegue o toma de decisiones de los sistemas automatizados

Cuando una organización, ya sea pública o privada, se embarca en un proyecto de *data mining* o inteligencia artificial siempre tendrá como objetivo poder obtener resultados óptimos de él. Predecir determinados comportamientos, integrar un sistema

[165] Andrew Smart ilustra con varios ejemplos cómo los sistemas inteligentes no siempre son infalibles y que una confianza ciega en estos puede llevar consigo consecuencias muy negativas. En: SMART,A: *Más allá de ceros y unos*. Ed.Clave intelectual, Madrid, 2018, págs.105 y ss.

en un artefacto robótico o meramente observar a través de los datos la tendencia de un entorno real son algunas de las virtualidades que ofrecen estas tecnologías. Concretamente, y por lo que nuestro estudio nos incumbe, en el plano de la toma de decisiones automatizadas los proyectos basados en inteligencia artificial tienen como objetivo principal tratar de crear sistemas que puedan implantarse en los procesos internos de las organizaciones o en su caso ofrecer esos sistemas a terceros para que puedan utilizarlos en la toma de decisiones. Es decir, el resultado final perseguido por estos proyectos es el de crear un sistema lo suficientemente óptimo para que pueda desplegar sus efectos en un entorno real. Así, y una vez desgranada la fase de diseño de estos sistemas, es turno de analizar la fase de despliegue de los mismos, esto es, el conjunto de etapas que abarcan la integración e implementación de los sistemas en el proceso organizativo y la puesta en marcha de dicho sistema en el contexto real en el cual adoptará decisiones.

A) La Adquisición, implantación y puesta en marcha del sistema

Aunque resulta habitual que la organización que diseñó el sistema acabe utilizándolo para sí misma, es frecuente también que, una vez diseñado y desarrollado el sistema, este acabe vendiéndose como producto a terceras organizaciones. Los roles que a partir de esto momento pueden presentar cada una de estas organizaciones pueden ser muy variados. Destacamos los siguientes:

- El sistema es adquirido por una organización distinta a la que lo desarrolló y adquiere los plenos derechos sobre el producto. Como es lógico, el coste de esta adquisición será elevado debido a que desde ese momento la organización que adquiere el sistema es la plena propietaria y tiene plena libertad para utilizar dicho sistema a su antojo. Por ejemplo: las autoridades públicas que obtienen un sistema para la toma de decisiones automatizadas pueden adquirir productos del mercado privado. Dado que en muchos casos las obligaciones de transparencia exigirán que dichos sistemas sean abiertos, las empresas que vendan estos productos deben ser consencientes que las entrañas de estos podrán ser conocidas[166].

[166] La Comisión de Garantía del derecho de acceso e información púbica de Cataluña (GAIP) consideró que el código fuente de un algoritmo cuya función básica era elegir automáticamente a los profesores que evaluarían en la PAU era información a los efectos de la Ley de Acceso a la transparencia y por tanto, y

- El sistema es adquirido por una organización distinta a la que desarrolló el producto sin embargo esta última establece cláusulas de confidencialidad sobre determinados aspectos técnicos y funcionales del producto. Se trata de una modalidad que puede permitir que organizaciones que se ven limitadas para el desarrollo de sistemas automatizados y que además tienen dificultados para adquirir dichos sistemas por sus excesivos costes puedan beneficiarse de las funcionalidades que puede ofrecer los mismos. Problemas relacionados con la distribución de responsabilidades en casos de decisiones inadecuadas o controles de funcionamiento del sistema pueden limitar esta práctica.

- El sistema es gestionado por la organización que lo desarrolló pero es otra tercera la que en su caso lo utiliza para la adopción de decisiones automatizadas. Las fórmulas para instrumentalizar esta operatoria pueden ser muy variopintas. Bajo este paraguas cabe destacar una forma de cogestión que en nuestra opinión acabará implantándose en muchos casos. Se trata efectivamente de la gestión público-privada de los sistemas automatizados donde el sector privado facilita las herramientas y técnicas algorítmicas mientras que el sector público proporciona los datos para alimentar a los sistemas y observar los entornos. Ambas organizaciones se benefician. En este sentido, ejemplos de esta práctica empiezan a observarse en iniciativas como las *smart city* o en proyectos de sectores específicos como podría ser el sanitario o la investigación.

En los casos señalados previamente el desarrollador del sistema debería ofrecer al adquirente toda aquella información que pueda resultar necesaria para la posterior implementación del programa. Es por ello tan relevante que en la fase del diseño del sistema se registren todos los procesos llevados a cabo y en su caso sea elaborado un manual de funcionamiento del prototipo[167]. En este sentido, la propuesta de Reglamento europeo sobre la regulación de los sistemas de inteligencia artificial conocida como *Artificial Intelligence Act* –en adelante PRAI- obliga a los desarrolladores/proveedores

dado que no afectaba a otros intereses la transparencia de dicho código fuente, este se podía poner a disposición del solicitante. Resolución 200/2017, de 21 de junio de 2017. Resolución disponible en: http://www.gaip.cat/ca/detall/normativa/2017-0200

[167] European Parliamentary Research Service. *A governance framework for algorithmic accountability and transparency*, op.cit., pág.63.

de los sistemas de inteligencia artificial de alto riesgo a facilitar a las organizaciones que los utilizarán toda una serie de información técnica relacionada con el sistema[168].

Por otro lado, elementos como la interpretabilidad de los resultados, el número de falsos positivos o negativos generados durante las fases de evaluación o el tipo de algoritmo elegido se presentan también como un conjunto de elementos esenciales a conocer. Dicha comprensión le ayudará al adquirente a valorar si tal sistema tiene cabida en el entorno donde pretende implantarlo, reduciendo la posibilidad de resultados posteriores inadecuados. Con ello se puede evitar una posible adquisición de un sistema que no cumpla con los requisitos mínimos exigidos en ese contexto valorando en su caso la compra de otro. En este sentido, no es la primera ni será la última vez que un sistema es diseñado para unos fines específicos y acaba desplegando sus efectos en un entorno totalmente diferente. Aunque esto último no debería ser problemático para la toma de decisiones, en muchos casos puede ser la principal causa de un comportamiento inadecuado del sistema. Así, una aplicación puede haberse diseñado inicialmente para etiquetar a personas en una red social y posteriormente acaba desplegando sus efectos en un sistema que se encarga de buscar y comparar personas en una base de datos de personas que han cometido delitos graves[169]. En otros casos sin embargo, el desarrollo inicial puede buscar un enfoque general -sistemas de reconocimiento facial- y posteriormente dicho sistema es adquirido por terceros que lo usan para distintos propósitos –como pueden ser los señalados previamente- Las consecuencias de los errores que se pueden derivar del uso de estos sistemas son muy diversas en cada uno de estos contextos. De esta forma, dado que en la fase del diseño se ha podido establecer un nivel de error razonablemente adecuado para el entorno inicial previsto, puede que ese nivel de error no pueda ser asumible en el entorno posterior en el que pretenda integrar el sistema el adquirente.

Indicado lo anterior, tanto si el sistema es adquirido por terceros, como si dicho sistema es implantado por la misma organización que lo desarrolló, varios elementos deben quedar fijados en este momento. Entre ellos, resulta elemental valorar el grado de participación o nivel de autonomía que tendrá el sistema a la hora de tomar decisiones.

[168] Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

[169] LEARNED-MILLER,E; ORDÓÑEZ,V; MORGENSTERN,J Y BUOLAMWINI,J**: "**Facial recognition technologies in the wild: a call for a federal office". Algorithmic Justice League, 2020, págs.24 y 25.

También es relevante que se fije el papel que se le asignará a los resultados que el sistema arroje o la fase del proceso decisorio donde se implementará el sistema.

Junto a lo indicado, resulta nuevamente recomendable realizar test del sistema. Estas pruebas sí que podrían ser ya experimentos piloto ya que en esta fase se ha de conocer el contexto donde el sistema desplegará sus efectos y los posibles colectivos que pueden quedar afectados por las decisiones que se adoptarán. Lo ideal es que dichos test se realicen de forma local y limitada a un grupo de personas y que en su caso dichas decisiones no tengan aún implicaciones en sus esferas. Así, es posible que los resultados mostrados durante la fase del diseño y testeo previos se alteren una vez el sistema se ponga en funcionamiento[170].

También se ha señalado la posibilidad de que los datos que se utilicen para testar los sistemas sean sintéticos[171]. De manera que no queden afectados en esas fases iniciales directamente personas. Estas pruebas piloto pueden ser una buena toma de contacto inicial previa de cómo se comportará el sistema cuando esté en pleno funcionamiento. Mostrándose imprescindibles cuando la organización que diseñó e implantó el sistema no sea la misma. Las auditorías previas o la participación de terceros ajenos a la organización que traten de sabotear o poner contra las cuerdas al sistema también se han puesto como ejemplo de testo previo a los mismos[172].

B) Se introducen los datos en el sistema

Una vez que el sistema se ha implantado en el proceso interno de toma de decisiones y se han realizado los chequeos previos al mismo. Este último se hallaría en perfectas condiciones para ser operable. Es decir, ya puede procesar los datos que se transformarán en salidas, las cuales, se utilizarán para adoptar decisiones o arrojar

[170] Sobre todo sucederá en entornos donde el sistema continuamente está aprendiendo de dicho contexto en el que interactúa. COVINGTON,P; ADAMS,J Y SARGIN,E: "Deep Neural Networks for YouTube Recommendations". *Google*, 2016,p-2. Disponible en: http://dx.doi.org/10.1145/2959100.2959190

[171] European Parliamentary Research Service. *A governance framework for algorithmic accountability and transparency*, op.cit., pág.32.

[172] Se habla por ejemplo de establecer equipos rojos que intenten destruir deliberadamente el sistema para encontrar posibles vulnerabilidades. En: Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Directrices éticas para una IA fiable*, 2019, pág.27.

resultados que pueden ser útiles para la toma de decisiones posteriores. Es turno de analizar algunos de los aspectos más relevantes de esta fase.

b.1) ¿Quién y cómo se ingresan los datos?

Existen multitud de formas a través de las cuales un sistema de decisiones automatizadas inteligente recibe los inputs que alimentarán al sistema. Centrándonos en el sujeto u objeto a través del cual se producirá el ingreso podemos mencionar algunas situaciones.

En primer lugar es frecuente que sea una persona de la organización la que ingrese manualmente los datos al sistema. De manera que una vez los ha recopilado del individuo sobre el que se tomará la decisión, esta persona lo ingresará en el sistema para que los procese. Aunque se trata de un procedimiento cada vez menos practicado sigue resultando habitual en muchos sectores y contextos. En segundo lugar, es el propio sistema el que capta esos datos y procesa los datos directamente. Este funcionamiento es frecuente en aquellos supuestos en los que el sistema se integra en un entorno dinámico e interactúa con él. Así, todos los servicios que utilizan algoritmos en plataformas en línea operan de esta manera. Procesan datos de forma automática y generalizada sin reparar en ninguna presencia humana durante el ingreso de los mismos. Finalmente y en tercer lugar, la forma de ingresar los datos puede ser mixta. Es decir, parte de los datos que se recopilan e ingresan se realiza de forma manual mientras que el resto son directamente captados por el sistema fruto de la interacción del individuo que se está analizando con el entorno en el que está integrado el sistema.

En los casos indicados previamente donde es el propio sistema el que capta los datos, resulta habitual que dicho sistema lleve incorporado una serie de sensores que les permitan al mismo poder procesarlos adecuadamente. De esta manera, al igual que los seres humanos disponemos de sentidos para captar todo tipo de sensaciones y a partir de ahí adoptar una u otra decisión, los sistemas automatizados tienen integrados una serie de sensores que permiten a los mismos captar y procesar los datos. Así, aunque existen multitud de sensores, alguno de ellos puede ser: cámaras para sistemas de video vigilancia o reconocimiento facial, micrófonos para interactuar con *chat bots*, teclados

para medir los patrones de escritura, sensores térmicos o táctiles para permitir el acceso a determinados lugares, marcas y señales viarias[173], etc.

b.2) La obtención de los datos

Junto al quién y cómo se captan e ingresan los datos en el sistema, también es relevante analizar la forma en que estos datos se obtienen y si los individuos son o no consciente de que tales datos están siendo utilizados para adoptar decisiones sobre ellos. Así, en función de ese grado de conocimiento podemos distinguir datos obtenidos directamente del individuo y datos obtenidos indirectamente del individuo.

La obtención directa de los datos facilitados por el particular que se ve sometido a la toma de decisiones ha sido el cauce habitual hasta no hace muchos años. Por datos directamente obtenidos del individuo hay que entender todos aquellos que este último expresamente conoce y es consciente que se están recopilando y utilizando para adoptar las decisiones automatizadas. Así, cuando una persona se dirige a una entidad aseguradora y esta le solicita que aporte determinados datos para evaluar el riesgo y en función de ello se le asigna una prima u otra, hemos de entender que esos datos son obtenidos directamente por el individuo. No sólo porque el individuo lo ha facilitado sino porque además este es consciente de que dichos datos se utilizarán con el fin de que el sistema adopte una u otra decisión. Es decir, el individuo conoce que dichos datos se integrarán en el procesado de datos que conformarán la decisión que posteriormente le afectará.

Sin embargo, y fruto de la capacidad actual que tienen las organizaciones para obtener todo tipo de datos de nuevas fuentes. Es cada vez más frecuente que a la hora de adoptar decisiones automatizadas no sólo se tengan en cuenta los datos obtenidos directamente del afectado por la decisión sino que además se incorporen toda otra serie de datos de los que el particular nunca ha sido consciente. Esta obtención indirecta de datos abarca una malgama de técnicas de todo tipo que facilitan la captación de datos del afectado por la decisión. Así, resulta habitual que terceros ajenos a las

[173] Para un vehículo autónomo resulta esencial que las marcas y señales viarias de las carreteras estén adecuadamente indicadas ya que favorecerá una correcta conducción autónoma del vehículo al poder procesar adecuadamente los datos que va percibiendo a través de sus distintos sensores. En este sentido pueden verse los considerandos 20, 22 y 30 y artículo 6.quarter de la DIRECTIVA (UE) 2019/1936 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 23 de octubre de 2019 por la que se modifica la directiva 2008/96/CE sobre gestión de la seguridad de las infraestructuras viarias.

organizaciones que adoptan la decisión faciliten datos de los individuos[174]. En otros supuestos, aunque el particular tiene una relación mercantil/laboral/pública con la organización, este desconoce que determinados datos o comportamiento que realiza están sometidos a un estudio que posteriormente arrojará una determinada decisión. Ejemplo: Una aseguradora, a la hora de conceder un seguro no sólo tiene en cuenta los datos aportados por el particular –directos-, sino que puede valorar otros datos como el estilo de conducción que desarrolla ese mismo usuario u otra serie de información aportada por terceros (datos observados o indirectos). En estos supuestos, si el usuario no es consciente de que dichos datos son utilizados para conformar la decisión hemos de entender que se obtienen indirectamente. En este sentido, el Grupo del artículo 29, en adelante GT29, ha establecido distintas categorías de datos en función de su origen, estos son: i) datos facilitados activamente por el interesado, por ejemplo, dirección postal, nombre de usuario, edad. ii) Datos observados facilitados por el interesados, por ejemplo, el historial de búsqueda, los datos de tráfico y los datos de ubicación de una persona y, iii) datos inferidos, Por ejemplo, el resultado de una evaluación de la salud de un usuario o el perfil creado[175].

C) Se genera un resultado

Una vez que los datos son incorporados al sistema este último ha de estar preparado para interpretarlos, procesarlos y transformarlos en un determinado resultado. El siguiente paso por tanto será el de fijar y determinar el valor que se le otorgará a esa salida.

Es turno de analizar algunos elementos necesarios relacionados con la interpretación y valorización de los resultados.

En primer lugar, lo primero que debe valorar toda organización es si el resultado que arrojará el sistema se convertirá inmediatamente en una decisión que afecte a los individuos o dicho resultado formará parte de otra serie de factores que fundamentarán la decisión final. Es decir, si la decisión será o no plenamente automatizada.

[174] Cabe destacar aquí la figura de los *data bróker*, los cuales recopilan todo tipo de datos de personas y elaboran perfiles sobre las mismas, posteriormente, los datos que se infieren de esos perfiles son vendidos a terceros. Es una práctica habitual en el sector bancario de Estados Unidos. En: HERTZA,V,A: "Fighting Unfair Classifications in Credit Reporting: Should the United States Adopt GDPR-Inspired Rights in Regulating Consumer Credit?". *New york university law review*, volume 93, num 6, 2018, pág.1735.

[175] Grupo del Artículo 29. *Directrices sobre el derecho a la portabilidad de los datos. Adoptadas el 13 de diciembre de 2016*. Revisadas por última vez y adoptadas el 5 de abril de 2017. Págs. 11 y 12.

En segundo lugar también es relevante conocer la forma a través de la cual dicha salida se mostrará a los miembros de la organización. En este sentido, los tipos representación de los resultados serán infinitos y ello variará en función de factores como el tipo de algoritmo elegido, el problema que se pretenda resolver con la incorporación del sistema o el modelo finalmente implantado. Por ejemplo: habrá sistemas en los que el resultado arrojado se manifieste a través de gráficas indicando los puntos calientes donde es más probable que se cometa un delito. En otros supuestos los resultados se pueden mostrar a través de determinados porcentajes en los cuales se representa una probabilidad de acierto, de coincidencia, etc. En otras ocasiones un resultado indicará que el objeto o persona que se buscaba ha sido identificado. A su vez, el resultado puede manifestarse en forma de resultado numérico[176].

En tercer lugar, una vez que tengamos ese resultado en los distintos formatos indicados, otro elemento interesante es que desde la organización esté claramente definido cómo debe ser interpretado ese resultado por parte de aquellos que en su caso lo observan. En este orden de cosas, cabe hacer especial mención a aquellos supuestos en los que el sistema arroja un determinado porcentaje. Así, aunque es habitual que en la fase de diseño del sistema ese porcentaje se le haya asignado directamente una salida específica a través del umbral de decisión. No es infrecuente que dicho umbral no se haya establecido y por tanto se deba interpretar ese porcentaje por parte de los miembros de la organización.

Resultado indicado por el sistema	**Umbral de decisión:** **Resultado >= 0,5 → Se deniega. (clase positiva)** **Resultado < 0,5 → Se concede. (clase negativa)**
0,655	No se concede el crédito
0,995	No se concede el crédito
0,728.	No se concede el crédito
0,428	Se concede el crédito

Asignación del resultado con umbral de decisión. Elaboración propia.

[176] En Estados Unidos se ha implantado un sistema que trata de prevenir el riesgo de que una menor sufra violencia en el ámbito doméstico. Una vez que se ingresan los datos en el sistema, este expulsa un resultado numérico que oscila entre el número 1 y el 20. Tal puntuación ha sido considerada la más adecuada por los desarrolladores de este modelo. En: CHOULDECHOVA, A; BENAVIDES-PRADO, D; FIALKO, O Y VAITHIANATHAN, R: "A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions". *Proceedings of machine learning research*, op.cit., pág.5.

Resultado indicado por el sistema	**Interpretación del resultado por parte del personal de la organización**
0,70 % → Mal pagador	No se concede el crédito
0,52 %→ Mal pagador	Se concede el crédito con un interés alto.
0,25 % → Buen pagador.	Se concede el crédito con interés bajo

Asignación del resultado basado en instrucciones. Elaboración propia.

En las tablas anteriores se describe cómo una organización puede establecer distintas consecuencias jurídicas para cada uno de los resultados emitidos por un sistema. De esta manera, en la primera tabla, si el sistema arroja un resultado de 0,71 esa persona se le deniega el crédito, al no superar la puntuación de corte previamente establecida[177]. En cambio, si el resultado es de 0,69 recibirá el crédito. La diferencia entre un porcentaje y otro puede ser nimia. Sin embargo, al cuantificar ese umbral de certeza y asignar a dicho umbral unas determinadas consecuencias tal porcentaje de corte se vuelve elemental ya que lo que se está fijando es el sentido otorgado a cada una de las decisiones que arrojará el sistema. Es decir, el porcentaje probabilístico indicado por el sistema. Fijar por tanto ese umbral de decisión resulta esencial al permitir que la incertidumbre que se deriva de los resultados probabilísticos emitidos por el sistema se concretice en un valor concreto. Ahora bien, ello supone aceptar y ser consciente que probabilistamente el sistema adoptará decisiones erróneas[178].

D) Se adopta la decisión. Plena automatización o no

Toda organización que integra en sus procesos de toma de decisiones un sistema basado en las tecnologías que venimos explicando en este trabajo deberá valorar el grado de autonomía que tiene dicho programa en la decisión final que se adoptará. En este sentido y aunque se pueden establecer distintas clasificaciones en función de una mayor o menor presencia humana en la decisión que se adopta, en este momento del trabajo vamos a distinguir dos grandes grupos. Estas son, las llamadas decisiones automatizadas y las decisiones parcialmente automatizadas.

[177] En: MEIKE K, KÖRFFER B & MEINTS,M: "Profiling of Customers and Consumers - Customer Loyalty Programmes and Scoring Practices". En: HILDEBRANDT, M & GUTWIRTH, S (eds): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008, pág.206.

[178] BOIX PALOP,A : "Los algoritmos son reglamentos: La necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones". *Revista de Derecho Público: Teoría y Método*, Vol. 1,Madrid, 2020, pág.233.

Por lo que se refiere a las primeras, estas se caracterizan por la plena automatización del resultado adoptado. Es decir, una vez que se ingresan los datos en el sistema y este los procesa el resultado que arroja se transforma automáticamente en una decisión que afecta a determinados individuos. Ejemplo: Al acceder a una página web de contenidos multimedia recibes publicidad personalizada de acuerdo al historial de tu motor de búsqueda. En estas situaciones no existe intervención del personal de la organización que adopta la decisión ya que el sistema por si solo recopila los datos, los procesa conforme a la instrucciones más o menos establecidas por los programadores y finalmente vierte un resultado en función de esos datos previamente recopilados[179].

Por otro lado, las decisiones parcialmente automatizadas se caracterizan por el hecho de que la salida que vierte el sistema no es el único elemento a tener en cuenta a la hora de adoptar la decisión final sino que se valoran otra serie de factores en la formación de dicha decisión. Ejemplo: Un sistema de *scoring* evalúa la solvencia financiera de una persona. Este sistema puede arrojar un determinado porcentaje que indica un alto grado de impago futuro de su préstamo. Una entidad bancaria puede depositar toda la confianza en el sistema y otorgar directamente consecuencias a dicho resultado vertido, en este caso, conceder o no el préstamo solicitado. Sin embargo, también puede utilizar ese resultado junto con otros factores como la opinión del trabajador de la entidad bancaria u otros documentos que facilite el solicitante del préstamo. Así, la decisión ya no será totalmente automatizada sino parcialmente automatizada ya que si bien ha existido un grado de automatización en la formación de la decisión -rating emitido por el sistema-, en la decisión final participan o se han tenido en cuenta otros elementos. Es decir, en este tipo de decisiones el *output* arrojado por el sistema se utiliza por los miembros de la organización como soporte, el cual, y junto con otras variables o factores, conforman la decisión final.

La plena o parcial automatización de las decisiones es un elemento muy relevante que además puede presentar distintas consecuencias jurídicas para aquellas organizaciones que deciden implantar en sus procesos este tipo de sistemas. Como

[179] En otros casos incluso, esa salida puede servir como entrada para uno o más algoritmos. En: MCGREGOR, L; MURRAY, D Y NG, V: "International human rights law as a framework for algorithmic accountability". *International and Comparative Law Quarterly*, 68(2), 2019, pág.318. Disponible en: https://www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6

ahora se verá, en materia de protección de datos esta característica presenta algunas diferencias relevantes.

E) Evolución del modelo

Desde el momento que un sistema emprende la toma de decisiones este último comienza a interactuar con el entorno para el cual fue diseñado. Es decir, obtiene los datos, los procesa y finalmente arroja determinadas salidas. Durante ese proceso existen sistemas que continuamente aprovechan los datos de entrada para mejorar y ajustarse a los cambios que acontecen en el entorno. A su vez, encontramos otros sistemas que no cambian con el uso y las nuevas entradas que se incorporan al mismo. De manera que los resultados siempre son los mismos ante el ingreso de nuevos datos. Precisamente y fruto de esa adaptación o no al entorno, los primeros sistemas a los que aludimos son conocidos como dinámicos o adaptativos mientras que los segundos son tildados de estáticos o no adaptativos[180].

Los sistemas dinámicos son mucho más prácticos en entornos excesivamente cambiantes donde se requiere de una continua adaptación a las múltiples alteraciones del contexto donde irradia sus efectos. Por ejemplo, el uso de algoritmos en entornos bursátiles o en todo tipo de plataformas de internet muestra que dichos programas son realmente efectivos. La capacidad adaptativa y de aprendizaje de cada nueva entrada del sistema le permite al mismo adecuarse y moldearse al contexto específico donde toma decisiones refinando así las relaciones ya establecidas en la fase del diseño del mismo[181]. Ahora bien, debido al alto grado de alteración al que se ve sometido el sistema las decisiones que puede adoptar en determinados momentos pueden volverse erróneas e imprevisibles cuando el entorno sea hostil o impredecible. Ejemplo paradigmático y archiconocido lo encontramos con el *bot* que fue desarrollado por Microsoft. Este se integró a través de un perfil en la red social *Twitter*[182]. En pocas horas el sistema comenzó a escribir *tuits* de contenido misógino y nazi. Tal actitud la

[180] The Norwegian Data Protection Authority. *Artificial intelligence and privacy*. 2018, op.cit., pág.10. En el mismo sentido y haciendo referencia a la inteligencia artificial adaptativa y no adaptativa. En: Autoritat Catalana de Protecció de dades. *Intel·ligència Artificial. Decisions Automatitzades a Catalunya*, op.cit., pág.110.

[181] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, 2020, pág.6.

[182] El Algoritmo en pocas horas se volvió racistas y sexista. Intencionadamente, un grupo de usuarios interactúo con el *bot* y este adoptó dicha actitud. En: https://www.xataka.com/robotica-e-ia/microsoft-retira-su-bot-de-ia-despues-de-que-este-aprendiera-y-publicara-mensajes-racistas

adoptó en gran parte porque muchos usuarios interactuaron intencionadamente con el *bot* para que aprendiera ese comportamiento. Este tipo de errores resultan sumamente problemáticos debido a que se pueden llegar a diluir posibles responsabilidades legales en cuanto al sujeto al que habría de asignarles las responsabilidades de las decisiones adoptadas. Así, tanto desarrollador del sistema como aplicador pueden echarse en cara los problemas iniciales o posteriores generados, ya que muchos de los errores generados en la fase de despliegue puede que no se hayan podido prever o anticiparse durante la fase del diseño. Para mitigar algunos de estos problemas los propios desarrolladores pueden programar que el sistema vaya registrando los principales eventos o alteraciones de este y ello puede permitir posteriormente la revisión del momento en el que se han adoptado las decisiones más sensibles[183]. La transparencia ligada al funcionamiento del sistema también se ve limitada desde el momento que las alteraciones del sistema no permiten explicar las decisiones adoptadas. Para estos sistemas resulta sumamente relevante la realización de seguimientos y evaluaciones meticulosas de su funcionamiento fruto de las alteraciones que posiblemente sufrirá el algoritmo desde el momento que este despliegue sus efectos[184]. En este sentido, se trataría de evitar el conocido como *concept o model drift*, esto es, alteraciones o cambios en el sistema derivados de los nuevos datos de entrada que este recibe al interactuar con el entorno[185].

Por su parte, los *modelos estáticos* resultan ideales en entornos que no son excesivamente cambiantes. Estos modelos dejan de aprender una vez han sido diseñados de manera que cuando comienzan a desplegar sus efectos en un entorno real su programación no se altera. Ello facilita la interpretación de los resultados objeto de análisis ya que son altamente predecibles. Ahora bien, dado que su capacidad de adaptación al contexto es limitada, un cambio brusco del mismo exigirá la sustitución del modelo por otro debido a que el mismo se volverá obsoleto.

Pues bien, aunque los sistemas estáticos conservan un papel relevante en la adopción de decisiones automatizadas[186]. En los últimos tiempos los sistemas dinámicos

---

[183] KROLL, A,J;. HUEY, J; BAROCAS, S; FELTEN, E, W; REIDENBERG, J,R;.ROBINSON, D,G;. Y YU, H:, "Accountable Algorithms". *University of Pennsylvania Law Review*, op.cit., pág.651.

[184] Comisión Europea. *Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica*, 2020, pág.8.

[185] Este fenómeno se conoce como "la deriva de concepto" o *drift in machine learning*. Lo que exige realizar continuos testeos para comprobar el nivel de errores. En: https://machinelearningmastery.com/gentle-introduction-concept-drift-machine-learning/

[186] The International Committee of the Red Cross (ICRC). *Autonomy, artificial intelligence and robotics: Technical aspects of human control*, agosto 2019, pág. 15. Disponible en:

están ganando protagonismo. Es precisamente el grado de adaptación a dichos entornos altamente cambiantes el que ha empujado a multitud de organizaciones a incorporar en su proceso decisorio estos sistemas donde antes era imposible debido a que el ingreso de los mismos no generaba un grado de acierto lo suficientemente permisible o adecuado.

Implantar unos u otros sistemas y valorar los elementos señalados anteriormente también resultarán relevantes para las organizaciones que pretendan desplegar sistemas de decisiones automatizadas en sus procesos.

FASE	Etapas	Principales actuaciones que se realizan
DESPLIEGUE	**Adquisición del sistema/Puesta en marcha**	-Fijar el grado de autonomía que tendrá el sistema. -Automatización o no plena de las decisiones. -Pruebas piloto del sistema en el entorno donde desplegará sus efectos.
	Se introducen/ procesan los datos	-Establecer las personas/máquinas que ingresan los datos. -Fuente de procedencia de los datos.
	Se emite el resultado	-Interpretación del resultado por parte de los miembros de la organización. -Automatización o no del resultado emitido por el sistema.
	Se adopta la decisión	-Decidir si la decisión es total o parcialmente automatizada. -Papel del personal de la organización a la hora de interpretar/alterar/no seguir la decisión adoptada por el sistema.
	Evolución del sistema	-Monitorear el sistema a lo largo de su interrelación con el entorno y el resto de particulares que interactúan con dicho sistema.

4. Consideraciones finales

En definitiva, el desarrollo y despliegue de estos sistemas está formado por una serie de fases fuertemente interrelacionadas entre sí. Las decisiones que se adopten en una pueden afectar al resto. Además, multitud de factores y variables se han de tener en cuenta. En este sentido, distintas legislaciones y ordenamientos jurídicos pueden resultar aplicables. En los capítulos posteriores trataremos de presentar y analizar aquellas garantías que desde el derecho a la protección de datos personales son

https://www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control

requeridas y en su caso necesarias implantar para entender que los particulares están protegidos en esta esfera.

CAPÍTULO II. INTRODUCCIÓN AL RÉGIMEN NORMATIVO DE LOS SISTEMAS DE TOMA DE DECISIONES AUTOMATIZADAS EN EL DERECHO FUNDAMENTAL A LA PROTECCIÓN DE DATOS

En este capítulo se realiza una aproximación a un conjunto de elementos y nociones que resultan relevantes a la hora de afrontar las implicaciones que presenta el uso de sistemas de toma de decisiones automatizadas en el Reglamento General de Protección de Datos, en adelante RGPD[187]. Para ello, en primer lugar se analizan los conceptos de decisión y elaboración de perfiles. En segundo lugar se realiza una acercamiento a las distintas tipologías de datos que están presentes durante el ciclo de vida de estos sistemas y cuáles de estos se consideran datos personales. Finalmente, en tercer lugar se estudia el encaje legal de los principales agentes y sujetos presentes durante las fases que engloban dicho ciclo de vida de los sistemas de toma de decisiones automatizadas. Las nociones extraídas de este capítulo se utilizarán como base para el desarrollo de los siguientes capítulos y marcarán el contenido legal del objeto de esta tesis.

I. LAS DECISIONES AUTOMATIZADAS Y LA ELABORACIÓN DE PERFILES

A lo largo del capítulo anterior hemos descrito el conjunto de fases que engloba el diseño y despliegue de los sistemas que adoptan decisiones automatizadas. En estas etapas ha quedado más que patente la importancia que tienen los datos durante todo el proceso. Es decir, desde que se eligen las variables hasta que finalmente el modelo algorítmico se integra en un sistema que acaba adoptando decisiones automatizadas basadas en los datos del particular. A su vez, también se ha podido vislumbrar que en cada una de estas etapas los datos sufren toda una serie de operaciones para alcanzar el objetivo marcado.

Pues bien, la relevancia que ostentan los datos en todas estas fases junto con el conjunto de tratamientos a los que se someten los mismos induce a pensar que, cuando en dichas operaciones se traten datos personales, la normativa encargada de regular el

[187] 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.

derecho fundamental a la protección de datos personales resultará plenamente aplicable. En este sentido, el legislador europeo, al igual que con otros tratamientos, no ha optado por regular específicamente todo el ciclo de vida que acompaña a estos sistemas sino que únicamente se ha centrado en determinadas fases o procesos que los ha considerado más relevantes. Concretamente, el RGPD ha regulado la llamada elaboración de perfiles y una categoría específica de decisiones adoptadas por estos sistemas, esto es, aquellas que son totalmente automatizadas y generan efectos relevantes en la esfera de los particulares[188]. El hecho de que se haya optado por regular únicamente de forma expresa estos dos tratamientos no supone que el resto de fases queden fuera del ámbito de aplicación del RGPD. Esta norma es perfectamente aplicable a todas las etapas que abarcan este fenómeno debido a que entran dentro de su ámbito de aplicación. Únicamente, el legislador, al mencionar expresamente tales tratamientos[189], hace hincapié en los mismos por entender que por sus características requieren de un tratamiento normativo particularizado dados los riesgos que presentan para los derechos y libertades de las personas.

1. El concepto de decisión: Tipos de decisiones

Como ya conocemos, el funcionamiento básico de los sistemas de inteligencia artificial a la hora de adoptar decisiones es el siguiente: 1) se introduce el input o dato, 2) dicho dato se procesa por el algoritmo o algoritmos correspondientes y, 3) tras ese proceso, se expulsa un output o resultado. En función de si la salida automáticamente se convierte o no en una decisión que genera efectos en los particulares hablamos de decisiones automatizadas o decisiones parcialmente automatizadas respectivamente.

Ahora bien, ¿a qué nos estamos refiriendo cuando hacemos alusión al concepto de decisión? Según la RAE, por decisión hemos de entender la *resolución que se toma o se da en una cosa dudosa*[190]. Trasladado a nuestro estudio, el concepto de decisión abarcará todas aquellas resoluciones que adopte una organización basadas en el

[188] Artículo 22 del RGPD.
[189] Desde 1990 se palpa un especial interés por parte del legislador europeo por reguilar este tipo de tratamientos de forma específica. Véase la Comunicación de la Comisión sobre la protección de las personas en lo referente al tratamiento de datos personales en la Comunidad y a la seguridad de los sistemas de Información. COM(90) 314 final - SYN 288, 24 de septiembre de 1990. (pág.16)
https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:51990DC0314&from=EN
[190] Real Academia Española. Se ha utilizado la primera acepción. Disponible en:
https://dle.rae.es/decisi%C3%B3n

resultado que emite un sistema algorítmico tras el procesamiento de datos. Son varios los elementos que hemos de destacar a la hora de acotar el término analizado.

Así, en *primer lugar* hay tener en cuenta el peso que tiene el resultado que emite el algoritmo en el proceso de formación de la decisión que se adopta y que afecta a los particulares. Por un lado, no existirán dudas del carácter de decisión cuando el resultado expulsado por el sistema se convierta automáticamente en una decisión que afecta a un particular. En estos supuestos, el proceso decisorio se automatiza por completo. Nos encontramos ante las llamadas decisiones automatizadas. Por otro lado, también formarán parte de la noción de decisión todas aquellas resoluciones donde el resultado del sistema tenga un peso alto o guarde relación significativa con la adopción de la decisión final. De esta manera, aunque el resultado no está completamente automatizado y durante el proceso decisorio intervienen personas, el resultado emitido por el sistema es de tal calibre que el mismo presenta una impronta real y material en la formación de la decisión finalmente adoptada. Nos encontramos ante las decisiones parcialmente automatizadas. A sensu contrario, cuando la salida emitida por el algoritmo tenga una relevancia residual en todo el proceso decisorio, dado que no existe una ayuda material y real del programa algorítmico durante la formación de la decisión, esta no la podemos considerar incluida en el objeto de nuestro estudio[191]. Quedan fuera por ejemplo todas aquellas situaciones donde se utilizan todo tipo de programas algorítmicos que, si bien están presentes en el proceso de formación de la decisión, su participación no se relaciona directamente con la misma. Por ejemplo, aquí encuadraríamos el uso de procesadores de textos inteligentes o traductores automáticos[192]. Dichos programas son necesarios y sirven de apoyo a la persona que adopta la decisión pero tales elementos no han tenido influencia en la misma.

[191] Imaginemos que una persona que toma decisiones debe sopesar tres factores a la hora de adoptar una decisión. Si esa persona se basa en una herramienta de IA para formarse una opinión sobre uno de estos tres factores es probable que el sistema de IA haya sido material o significativo en el proceso de toma de decisiones. En cambio, imaginemos que esta misma persona teclea la decisión usando una sofisticada aplicación de procesamiento de textos que fue desarrollada usando un sistema de IA. Si la función de la aplicación de procesamiento de textos es simplemente para registrar la decisión y no para ayudar a sopesar los tres factores que influyen en la decisión, el uso de esta herramienta de procesamiento de textos no sería un uso material o significativo de la IA en la formación de la decisión. En: Australian Human Rights Commission. *Human rights and Technology. Discussion Pape*r, pág. 64.

[192] La tarea asignada al programa algorítmico durante la fase de formación de la decisión puede ser clave para valorar si ese sistema realmente ha tenido una presencia significativa en la formación de la decisión. New York City Automated Decision Systems (ADS). *Automated Decision Systems Task Force Report*, November, 2019, pág.20.
El uso de una hoja de Excel por parte de un trabajador no conforma la decisión que finalmente se adopta En: https://algorithmwatch.org/en/story/kees-verhoeven-algorithm-registry/

En *segundo lugar*, y tomando como base el concepto de decisión definido en el párrafo anterior, cabe destacar que el objeto de esta tesis no queda circunscrito a las decisiones que se adopten con una tecnología específica. Así, el RGPD adopta un enfoque neutral a la hora de regular las tecnologías que tratan datos personales[193]. De esta manera, pese a que un número importante de decisiones tendrán como base el uso de sistemas de inteligencia artificial, ello no quiere decir que el objeto de este estudio quede reducido a las decisiones que se deriven de estas tecnologías.

En definitiva, una decisión a los efectos de esta tesis no deja de ser una resolución que adopta una organización basada en el resultado que emite un sistema algorítmico tras analizar los datos personales del individuo sobre el que se adoptará la mentada decisión. El tipo de tecnología sobre la que descansa la salida emitida por el algoritmo no es un elemento relevante a efectos de la consideración de decisión.

Pues bien, definido el concepto de decisión, cabe señalar que el RGPD establece un régimen específico para aquellos tratamientos donde los sistemas algorítmicos acaban adoptando decisiones plenamente automatizadas y dichas decisiones generan efectos relevantes en la esfera de los particulares. Así, en concreto, se prevé toda una serie de garantías en favor de los particulares sometidos a dichas decisiones a lo largo de esta norma[194].Para el resto de decisiones, como hemos indicado antes, el régimen general de la normativa de protección de datos se seguirá aplicando enteramente[195].

Es turno de analizar los distintos tipos de decisiones.

---

[193] El considerando 15 del RGPD establece que *a fin de evitar que haya un grave riesgo de elusión, la protección de las personas físicas debe ser tecnológicamente neutra y no debe depender de las técnicas utilizadas.*

[194] Como mera introducción, las garantías específicas se encuentran reconocidas en los siguientes artículos. Derechos de información específicos: artículos .13.2.f) y14.2.g). Derecho de acceso específico: artículo.15.h). Prohibiciones específicas y bases legales reforzadas: artículo.22.1 y 22.2. Derecho a la participación humana, revisión y expresar punto de vista: artículo 22.3. Evaluación de impacto: artículo 35. 3.a)

[195] Como ha indicado la doctrina, para el resto de decisiones siempre quedará el París del régimen general de protección de datos aplicable. COTINO HUESO, L: "Derecho y garantías ante el uso público y privado de inteligencia artificial, robótica y big data". En: BAUZÁ REILLY,M (eds): *El Derecho de las TIC en Iberoamérica.* Ed. La Ley Uruguay, Montevideo, 2019, pág.936.

	Decisiones plenamente automatizadas que generan efectos relevantes tomadas por algoritmos	**Resto de decisiones tomadas por algoritmos.**
Garantías	-Derecho de información sobre la lógica del tratamiento. -Derecho de acceso específico. -Prohibiciones específicas. -Bases legales de legitimación reforzadas. -Derecho a solicitar la participación humana. -Derecho a expresar el punto de vista. -Derecho de revisión de la decisión. -Evaluación de impacto obligatoria. -Normativa general del RGPD.	-Normativa general del RGPD.

A) Las decisiones plenamente automatizadas que generan efectos relevantes

El tratamiento de datos al que hacemos referencia aparece definido en el artículo 22 del RGPD, concretamente, este precepto indica que:

> *Todo interesado tendrá derecho a no ser objeto de u**na 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.* (La negrita y cursiva son nuestras)

Del contenido de este precepto pueden extraerse a primera vista dos elementos esenciales que configuran el tratamiento de datos específico que estamos analizando. *En primer lugar*, la decisión que afecta al particular ha de ser adoptada por medios enteramente automatizados. *En segundo lugar*, esta decisión debe producir efectos jurídicos en los particulares o generar efectos significativamente similares a los jurídicos. Es decir, ha de producir efectos relevantes para los particulares afectados por esta decisión.

Son por tanto dos elementos acumulativos los que deben estar presentes en las decisiones algorítmicas para considerarlas que entran dentro del concepto de decisión automatizada que estamos analizando. Pasamos a su estudio.

a.1.Decisión basada únicamente en un tratamiento automatizado

El primer factor que establece el RGPD para que entendamos que se aplican las especiales garantías previstas para este tipo de tratamientos es que la decisión que se derive del mismo sea *únicamente* automatizada. Por tanto, todos aquellos sistemas

algorítmicos en los que durante el proceso decisorio se incorpore la presencia humana quedarían fuera de su ámbito de aplicación. El carácter automatizado como elemento relevante para prever mayores garantías cuando las organizaciones hacen uso de estas decisiones pone de manifiesto que el legislador europeo consideró que estos tratamientos conllevan importantes riesgos para los particulares. Además, se vislumbra un recelo al hecho de que durante todo el proceso decisorio que abarca la adopción de la decisión no exista una mínima intervención humana[196]. En este sentido, en los últimos años, distintas normas han vetado la automatización de diversos procesos decisorios en concretos ámbitos o sectores por la especial sensibilidad que comporta la incorporación de la plena automatización en esos concretos espacios obligando a incorporar el elemento humano previó a la adopción de la decisión[197].

Debido a las especiales garantías que se exigen cuando se llevan a cabo este tipo de decisiones, las organizaciones, con el objetivo de eludirlas, pueden introducir el factor humano antes de la toma de la decisión. La incorporación de la persona en el proceso decisorio llevaría consigo considerar que tal decisión es parcialmente automatizada y por tanto, ya no estaríamos en el concepto de decisiones que estamos analizando. Pues bien, con el objetivo de evitar ello, el grupo del artículo 29 -en adelante GT29- ya ha advertido que el papel del humano durante el proceso decisorio no puede resultar simbólico sino que ha de ser significativo[198]. Es decir, la persona que analiza el resultado expulsado por el sistema algorítmico debe participar activamente en la decisión que finalmente se adopte, ya que si no, aunque se haya establecido que tal

[196] La desconfianza hacia las decisiones que puede tomar una máquina en relación con una persona fueron reseñadas en los documentos previos a la elaboración de la Directiva 95/46. En: Commission of the European Communities, COM (92) 422 final - SYN 287 Brussels, 15 October 1992. (Págs. 26 y 27). Disponible en: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:51992PC0422&from=DE

[197] *Las autoridades competentes no adoptarán ninguna decisión que produzca efectos jurídicos adversos para una persona o que afecte negativamente a una persona únicamente en razón del tratamiento automatizado de datos PNR.* Artículo 14.6 de la Ley Orgánica 1/2020, de 16 de septiembre, sobre la utilización de los datos del Registro de Nombres de Pasajeros para la prevención, detección, investigación y enjuiciamiento de delitos de terrorismo y delitos graves. Menos reciente pero en la misma sintonía encontramos el artículo 6 del Real Decreto 190/1996, de 9 de febrero, por el que se aprueba el Reglamento Penitenciario indique que *ninguna decisión de la Administración penitenciaria que implique la apreciación del comportamiento humano de los reclusos podrá fundamentarse, exclusivamente, en un tratamiento automatizado de datos o informaciones que ofrezcan una definición del perfil o de la personalidad del interno.* Por último, el artículo 14.5 de la (PRAI) Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la unión también prevé la verificación del resultado del algoritmo antes de la toma de la decisión por al menos 2 personas de la organización cuando los sistemas de IA estén destinados a la identificación biométrica a distancia "en tiempo real" y "a posteriori" de personas físicas. Resolución de 21 de abril de 2021

[198] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018. Pág. 23

decisión es parcialmente automatizada, de facto, hemos de entender que es automatizada y por tanto la misma entraría dentro del ámbito de aplicación del artículo 22 del RGPD. En este sentido, por decisión automatizada fáctica debemos considerar aquellas decisiones que, si bien a primera vista pueden considerarse que son realizadas por un humano, tal intervención no es significativa o real ya que de los hechos se desprende que realmente es la máquina, y no el humano, la que decide el resultado final. Son muchas las situaciones donde se puede enmascarar una decisión automatizada fáctica, describimos algunas de ellas.

En *primer lugar* nos encontramos con aquellos supuestos en los que el humano, aunque tiene cierto margen de decisión, queda restringido a las distintas posibilidades que establece el propio algoritmo. Es decir, se ofrecen varios resultados y el personal valora cuál de los resultados elegir[199]. Aunque la persona ha intervenido en la decisión[200], al elegir una de las salidas emitidas por el sistema, ha sido la máquina la que realmente la ha adoptado ya que la organización no permite que el humano pueda tener en cuenta otros elementos distintos a los expulsados por el sistema. También es frecuente que la incorporación del humano se utilice meramente para comprobar que el sistema ha actuado correctamente y no ha emitido un falso positivo. Así, por ejemplo, imaginemos que el revisor únicamente se encarga de constatar que la persona indicada por el sistema como de alto riesgo forma parte de un grupo de individuos a los que se pretende realizar un especial seguimiento. Si el sistema arroja un resultado positivo y esas personas no forman realmente parte de ese grupo, el revisor procedería al descarte de las mismas (falso positivo). Para los casos en los que esa persona detectada por el sistema sí forme parte de ese grupo (verdadero positivo), el revisor procede a realizar las actuaciones investigadoras correspondientes. En estos casos, si el papel del humano es de simple revisor y no tiene en cuenta otros elementos, también habrá de considerarse que nos encontramos ante una decisión automatizada.

---

[199] ALMADA, M,. "Human Intervention in Automated Decision-Making: Toward the Construction of Contestable Systems", *17th International Conference on Artificial Intelligence and Law* (ICAIL 2019), p.2. En sentido contrario, ROIG I BATALLA,A: Las garantías frente a las decisiones automatizadas del Reglamento general de Protección de Datos a la gobernanza algorítmica. Ed. Bosch. Barcelona, 2020, pág., 66.

[200] Algunas autoridades consideran que este tipo de decisiones si bien limitan ampliamente la labor del humano, no se consideran plenamente automatizadas ya que la persona tiene la facultad de optar por uno u otro resultado que expulsa el sistema. En: Comité de ética alemán. Gutachten der Datenethikkommission, 2019, págs. 25 y 161.
https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.html

Sistema para detectar fraude fiscal en determinados grupos	**Papel del revisor del resultado del algoritmo**	**Tipo de decisión**
El sistema indica riesgo alto de fraude	Falso positivo → Se descarta	Decisión plenamente automatizada
	Verdadero positivo → -Se inicia investigación -Se incorpora a una base de datos de futuros investigados	

En *segundo lugar* también serán frecuentes este tipo de decisiones en aquellos entornos donde la celeridad exigida a la hora de adoptar la decisión por parte del humano le empuja a este último a acudir reiteradamente al resultado que arroja el algoritmo. Es decir, dado que la persona designada para evaluar el resultado del algoritmo carece de tiempo para realizar un análisis adecuado, esta se ve irremediablemente destinada a confiar en lo que indica la máquina. Por ejemplo, algunas grandes plataformas indican que las denuncias presentadas por los particulares que ven retirado el contenido que suben a las mismas son estudiadas por personas[201]. Dada la celeridad de algunas de las contestaciones que reciben estos usuarios y la cantidad de denuncias que pueden presentarse de retirada de contenidos, es cuanto menos dudoso que dicho análisis de tal denuncia sea realizada por una persona. Otro ejemplo lo podemos encontrar en el personal presente en los controles fronterizos de los aeropuertos que se ven obligados a evaluar en un periodo relativamente corto a posibles personas sospechosas[202].

En *tercer lugar*, también debemos considerar que nos encontramos ante una decisión automatizada fáctica en aquellos supuestos en los que la propia organización haya previsto distintos modos de proceder a la hora de interpretar los resultados que expulsa el algoritmo y en algunas de estas actuaciones dicho resultado se convierta automáticamente en una decisión que afecta al particular.

[201] En este caso la plataforma YouTube retiró un video a unos usuarios de esa red. Tras impugnar la retirada de contenido, los usuarios recibieron la respuesta a tal denuncia dos minutos después de haberla planteado. Según esta plataforma, en el momento que sucedieron los hechos dicha revisión era realizad por humanos y no por máquinas. Dado el poco tiempo de respuesta que medió entre la presentación de la denuncia y la respuesta recibida, existen dudas de que efectivamente dicha decisión fuese tomada por un humano y no por un algoritmo. Fuente de la noticia: GARCÍA,J: "YouTube elimina la parodia de Pantomima Full sobre los negacionistas de la COVID-19 por cuestionar las directrices de la OMS", *Xataka*. 2/10/2020. Disponible en: https://www.xataka.com/aplicaciones/youtube-elimina-parodia-pantomima-full-negacionistas-covid-19-cuestionar-directrices-oms

[202] European Union Agency for Fundamental Rights, *Preventing unlawful profiling today and in the future:a guide*, op.cit., pág.56.

Resultado indicado por el sistema Probabilidad de impago	**Interpretación del resultado por parte del personal de la organización**	**Tipo de decisión**
≥0,7	No se concede el crédito	Decisión plenamente automatizada
0,69>0,31	Se analiza el resultado junto con otros elementos a la hora de valorar la concesión del crédito	Decisión parcialmente automatizada
≤0,3	Se concede el crédito	Decisión automatizada

Atendiendo a esta tabla, sólo cuando el porcentaje estimado por el algoritmo oscile entre 0,69 y 0,31 como probabilidad de impago de un determinado crédito el personal de la entidad bancaria tendrá capacidad real para alterar y valorar la decisión mostrada por el sistema. De manera que únicamente en este tipo de decisiones debemos considerar que nos encontramos ante una decisión parcialmente automatizada, en el resto de casos, la automatización es plena y no cabe alegar que un humano ha analizado el resultado por el mero hecho de visualizar dicho resultado y automáticamente adoptar la decisión que se le atribuye a ese resultado, es decir, denegación o concesión del crédito[203].

En *cuarto lugar*, tampoco podemos considerar una decisión como parcialmente automatizada cuando por ejemplo, la persona encargada de analizar el resultado del sistema no tiene la suficiente formación para saber interpretar dicho resultado debido a que el sistema es poco interpretable[204]. En este sentido, la autoridad de protección de datos portuguesa consideró que una decisión era plenamente automatizada pese a que durante el tratamiento de datos existía intervención humana ya que la persona encargada

[203] Normalmente, la puntuación que vierte el algoritmo suele utilizarse para estimar la calificación de un préstamo, establecer la duración de un crédito, asignar un límite de crédito o fijar un tipo de interés u otro. En: MEIKE K, KÖRFFER B & MEINTS,M: "Profiling of Customers and Consumers - Customer Loyalty Programmes and Scoring Practices". En: HILDEBRANDT, M & GUTWIRTH,S (eds): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008, pág.206.

[204] LAZCOZ MORATINOS:G: "Análisis jurídico de la toma de decisiones algorítmicas en la asistencia sanitaria". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed.Aranzadi, Navarra, 2020, pág.295.También en: Privacy international. Data is power: *Towards additional guidance on profiling and automated decision-making in the GDPR*, 2017, pág.14. A su vez también en este mismo sentido se ha pronunciado la Information Commissioner's Office, disponible en: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/how-do-we-ensure-individual-rights-in-our-ai-systems/

de analizar los resultados del algoritmo carecía de instrucciones y criterios claros a la hora de interpretar dichos resultados[205].

Finalmente, en *quinto lugar*, también estará presente la automatización de la decisión cuando las personas encargadas de analizar los resultados pese a que ostentan la formación necesaria para interpretarlos, acaban siempre inclinándose por los salidas indicadas por el sistema ya que dejan de ser críticos con tales resultados a causa de la plena confianza que deposita en el algoritmo[206], es decir, sobrestima sus resultados sin cuestionarlos[207].

En definitiva, a efectos del artículo 22 del RGPD, una decisión será plenamente automatizada cuando el resultado que emite el modelo algorítmico esté automatizado por completo y el mismo afecte directamente al particular. La decisión seguirá considerándose automatizada cuando el responsable haya indicado que existe participación human pero la misma no es real, o resultando real, la decisión quede circunscrita a lo indicado por el sistema.

a.2 Decisión que genera efectos jurídicos o significativamente similares. Los efectos relevantes

Junto a la plena automatización, el tratamiento descrito en el artículo 22 del RGPD exige que dicha decisión genere efectos jurídicos o significativamente similares a los mismos en la esfera del particular afectado. Es decir, no basta con que la decisión sea totalmente ejercida a través de medios mecanizados sino que además, dicha decisión debe generar o producir efectos relevantes para la persona afectada por la decisión que

[205] Comissão Nacional de Proteção de Dados. Deliberação 622/2021. Apartado 55. Resolución disponible en: https://www.cnpd.pt/decisoes/deliberacoes/

[206] En el sector crediticio, si el agente de crédito adopta sistemáticamente la estimación indicada por la puntuación de crédito sin reconsiderar y verificar la decisión con las circunstancias individuales se debe considerar que la decisión es totalmente automatizada. En: MEIKE K, KÖRFFER B & MEINTS,M: "Profiling of Customers and Consumers - Customer Loyalty Programmes and Scoring Practices". op.cit., pág.210.

[207] Se ha de tratar de evitar el sesgo del humano a la hora de interpretar el resultado. Este sesgo se genera en los casos en los que el personal encargado de interpretar los resultados los acepta sin espíritu crítico asumiendo un "principio de autoridad" del sistema derivado de las expectativas creadas por dichos sistemas. En: Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.8. En este mismo sentido, se ha indicado que confiar en la máquina puede llevarnos a considerar que la realidad ha de obedecer a los sistemas y no a la inversa. GORUNESCU,F: *Data Mining: Concepts, Models and Techniques*…. op.cit., pág.36.

ha adoptado previamente el algoritmo[208]. Así, el GT29, a la hora de interpretar que se entiende por *efectos jurídicos* ha indicado que dicha decisión debe afectar a los derechos o al estatuto jurídico del titular que se ve sometido a dicho tratamiento[209]. Algunos ejemplos que se encuadran en este concepto pueden ser la denegación de una prestación económica[210], la no obtención de un contrato o el despido de un trabajador[211]. Por otro lado, y en relación con las decisiones que *afectan significativamente de modo similar* a los efectos jurídicos. Hay que entender que dentro de este tipo de decisiones quedan aglutinadas todas aquellas en las que, sin que exista un cambio en las obligaciones o derechos jurídicos del particular afectado por la decisión algorítmica, esta última resulte ser de una importancia cuanto menos similar al de una decisión que produzca un efecto jurídico[212]. Por ejemplo, los precios o anuncios personalizados que se ofrecen en una web no generan en la esfera de los particulares efectos jurídicos o cambios en las obligaciones per se, sino que más bien, dicha decisión tomada por el algoritmo, esto es, el precio o el anuncio publicitario, puede potencialmente alterar la conducta o el comportamiento del particular afectado por la decisión.

Aunque no es fácil establecer una lista de las principales decisiones que pueden considerarse relevantes para los particulares, sí que podemos tomar como referencia una serie de criterios que nos orienten sobre cuándo un responsable puede considerar que su

---

[208] El carácter relevante de las decisiones automatizadas como elemento que obliga a imponer mayores garantías quedó patente desde las primeras propuestas que trataron de regular este tipo de tratamientos específicos. Así, ya en 1990 se señaló que estas especiales salvaguardias derivaban del objetivo esencial de *proteger e1 interés de1 interesado en participar en aquellas decisiones que sean importantes para él*. En: Comisión de las Comunidades Europeas COM(90) 314 final - SYN 288 Bruselas, 24 de septiembre de 1990, pág.22. Disponible en: :https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:51990DC0314&from=EN

[209] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*., op cit., pág.23.

[210] Por ejemplo, en 2018, el gobierno británico deportó a más de 7.000 estudiantes extranjeros tras acusarles falsamente de hacer trampas en sus exámenes de equivalencia de inglés. El gobierno utilizó un software de reconocimiento de voz para determinar si el propio estudiante estaba haciendo realmente el examen o, en cambio, había enviado a otra persona para que realizara el examen en su nombre. Cuando el análisis automatizado de la voz se cotejó con el análisis humano se comprobó que era erróneo en más del 20% de los casos. En: International Human Rights Program (Faculty of Law, University of Toronto) and the Citizen Lab. *Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System*.2018, Págs. 52 y 53. Disponible en: https://citizenlab.ca/2018/09/bots-at-the-gate-human-rights-analysis-automated-decision-making-in-canadas-immigration-refugee-system/

[211] No se duda de la producción de efectos jurídicos cuando lo que se pretende es automatizar la ejecución del contrato. En: VILLALBA SÁNCHEZ,A: "El principio de transparencia en la ejecución automatizada del contrato de trabajo: una aproximación jurídica a la tecnología blockchain y a la inteligencia artificial". *Revista Española de Derecho del Trabajo*. num.224/2019, pág.30.

[212] Grupo del artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*., op cit., pág.24.

tratamiento se encuadra en este tipo de decisiones, entre otros elementos podemos indicar:

Grado de impacto de la decisión en los particulares: Un primer factor es valorar el grado de injerencia que supone la decisión para los derechos e intereses de los particulares en relación con la resolución que se toma. Así, no presentarán las mismas consecuencias un sistema que se encargue de recomendar películas a otro que se encargue de conceder o denegar una hipoteca. El error en el primer sistema supondrá que el particular pierda dos horas de su vida viendo una película mala. En cambio, un error en el segundo de los sistemas puede suponer un impago futuro de la hipoteca o la no concesión de la misma cuando esa persona se lo merecía. Por ejemplo, la Autoridad Catalana de Protección de Datos ha considerado que la elaboración de perfiles automatizados con el fin de estudiar la conducta de los ciudadanos a la hora gestionar sus residuos orgánicos tienen efectos relevante desde el momento que a través de ese análisis se pueden conceder bonificaciones o reducciones en la tasa del servicio de basura[213]. A modo orientativo, la PRAI establece un listado de sectores y finalidades donde se ha de considerar que un sistema de inteligencia artificial genera un alto riesgo para los derechos y libertades de las personas. Esta lista puede servir como referencia para los responsables de tratamiento a la hora de considerar que un sistemas algorítmicos que están utilizando o pretende utilizar adopta decisiones que generan efectos relevantes[214]. Esta lista es orientativa ya que existen sectores que a día de hoy no se han incorporado a la misma y entendemos que deberían estarlo, tal y como ocurre con el sector seguros[215].

[213] Autoritat Catalana de Protecció de dades. Consulta nº CNS 36/2020, págs.15 y 16. Texto disponible en: https://apdcat.gencat.cat/web/.content/Resolucio/Resolucions_Cercador/Dictamens/2020/Documents/ca_cns_2020_036.pdf

[214] Artículo 6.3 y anexo III. Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021. El CEPD ha considerado que esta lista indicada por esta norma hace presumir que los tratamientos de datos llevados a cabo bajo esas finalidades han de ser considerados de alto riesgo a los efectos de la normativa de protección de datos. En: Comité Europeo de Protección de Datos. EDPB-EDPS. Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act). Resolución de 18 de junio de 2021. Apartado 21, pág.9.

[215] De las finalidades designadas como de alto riesgo en la PRAI, no se hace alusión al uso de sistema de toma de decisiones automatizadas a la hora de calcular el pecio de la póliza de seguros. A pesar de que no se considere de alto riesgo, ello no es óbice para entender que desde el punto de vista del RGPD no pueda considerarse que ese tipo de decisiones puedan generar efectos relevantes en la esfera de los individuos a los efectos del artículo 22.

Grupos o colectivos a los que se dirigen las decisiones: Existen determinados colectivos que por sus características pueden resultar especialmente afectados por determinadas decisiones que en términos generales no son consideradas relevantes. Ello puede ir desde colectivos como los menores hasta grupos minoritarios o adultos vulnerables en determinados contextos[216]. Es decir, la relevancia de la decisión en la esfera del particular no deviene en muchos casos del tipo de decisión sino de cómo esa decisión afecta a ese concreto colectivo. El GT29 específica por ejemplo que mostrar publicidad en línea personalizada no supone a priori que tal tratamiento sea relevante. Sin embargo, cuando esta publicidad se muestra por ejemplo a personas que se conoce que tenga o que pueda tener dificultades financieras y reciban por ello anuncios de préstamos a tipos de interés elevados, dado que esa persona es potencial comprador de ese producto, es posible que acabe suscribiéndose a dicha oferta y posiblemente aumentar su deuda. El efecto para esos colectivos sobre los que se conoce que presentan esa característica resulta muy relevante[217]. Otro ejemplo sería cuando un sistema automatizado recomiende a personas que sufren un trastorno alimentario anuncios o contenidos para bajar de peso o practicar el ayuno cuando el responsable es consciente de esos problemas[218]. En estos supuestos consideramos que dicha decisión en formato anuncio o recomendación personalizada se debe considerar relevante a los efectos del artículo 22 del RGPD. Y ello es así, no porque el anuncio o la recomendación resulten más o menos persuasibles, algo que se presupone en la publicidad, sino por el grado de conocimiento que se ostenta de esas personas y las implicaciones negativas que dicha decisión comporta para las mismas.

Decisiones tomadas en un determinado periodo de tiempo: para el GT29 una decisión se considera relevante si tiene un impacto prolongado en la esfera del particular. Así, por ejemplo, un tribunal holandés consideró que el bloqueo de una cuenta de un trabajador de la plataforma por parte de un algoritmo no puede considerarse que genere dichos efectos relevantes ya que dicho trabajador podía reactivar la cuenta en un periodo relativamente corto de tiempo, y por tanto, dicho

[216] El considerando 38 del RGPD pone el acento en las garantías adicionales que han de preverse cuando se lleven a cabo tratamientos de datos personales con fines de mercadotecnia o elaboración de perfiles de personalidad.

[217] "Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679". Grupo del Artículo 29. op.cit., pág25. En el mismo sentido también: Center for Democracy and Technology. *Digital Decisions,2017*,pág.8. Disponible en: https://cdt.org/insights/digital-decisions-tool/

[218] Fuente de la noticia: CRIDDLE,C: "Instagram fixes mistake promoting harmful diet content", *BBC News*. 15/04/2021. Disponible en: https://www.bbc.com/news/technology-56750088

bloqueo no generaba un efecto a largo plazo o permanente[219]. Es decir, aunque a priori pudiera considerarse relevante esa decisión, bloqueo de la cuenta, dado que el efecto que provocó es temporal y no permanente, no puede considerarse que genere efectos relevantes en los términos descritos en el artículo 22 del RGPD. En muchos casos, una decisión no logra dicho efecto sino que más bien, la conjunción en un determinado espacio temporal de multitud de estas puede originar repercusiones importantes en la esfera de los particulares. Es decir, una decisión por sí sola no es relevante pero en conjunto con otras, sí que puede adquirir tal relevancia[220]. Por otro lado, también puede ocurrir que en condiciones normales las decisiones que adopta un sistema algorítmico no se consideren relevantes, pero, en concretos periodos dichas decisiones alcancen tal grado de relevancia. Así, por ejemplo, la función de autocompletado implementada en los motores de búsqueda no deja de ser un algoritmo que emite decisiones en formato de recomendación. Estas decisiones en términos generales no pueden considerarse relevantes. Sin embargo, pueden surgir dudas cuando dicho autocompletado opera en periodos electorales recomendando determinadas búsquedas en las que podrían quedar afectados algunos derechos como el de participación política[221]. Y es que, existe la posibilidad de que dicha información brindada a través de los resultados ofrecidos por estas búsquedas pueda cambiar las opiniones de los votantes indecisos[222].

La personalización de la decisión: otro de los elementos esenciales para valorar la relevancia de la decisión es el grado de personalización previo que haya existido

[219] Este tribunal señaló que: *de la explicación de Uber de su proceso antifraude que dio a la audiencia se desprende que después de una señal de fraude, el acceso del conductor a la aplicación Driver se bloquea temporalmente hasta que el conductor se ha puesto en contacto con un empleado de Uber. El acceso a la aplicación Driver se reactiva tan pronto como el conductor se ha puesto en contacto. En vista de esta explicación de Uber, el tribunal asume que la decisión de bloquear temporalmente el acceso a la aplicación Driver después de una señal de fraude se tomará automáticamente, sin intervención humana. Sin embargo, este bloqueo temporal no tiene un efecto a largo plazo o permanente, por lo que la decisión automatizada no tiene consecuencias legales o afecta significativamente al conductor.*
Tribunal de Ámsterdam. Resolución de 11 de marzo de 2021. Apartado 4.25.
Disponible en: https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBAMS:2021:1018

[220] EDPS Ethics Advisory Group | Report, *Towards a digital ethics*, 2018 págs. 12 y 13 y Parlamento Europeo. *The ethics of artificial intelligence: Issues and initiatives,*2020, pág.20.

[221] La función de autocompletado del motor de búsqueda de Google recomienda regularmente consultas completas una vez que los usuarios comienzan a escribir palabras .Google anunció que actualizaría sus políticas de autocompletado en las búsquedas relacionadas en las elecciones presidenciales de los Estados Unidos de 2020 durante el periodo electoral, eliminando aquellos autocompletados engañosos que pudieran afectar a estos comicios electorales. Fuente de la noticia: DAVALOS,J: "Google Blocks Search Suggestions to Stop Election Misinformation", *Bloomberg,* 10/10/2020. Disponible en: https://www.bloomberg.com/news/articles/2020-09-10/google-blocks-search-suggestions-to-stop-election-misinformation

[222] URMAN,A ; MAKHORTYKH, M; ULLOA,R : "The Matter of Chance: Auditing Web Search Results Related to the 2020 U.S. Presidential Primary Elections Across Six Search Engines". *Social Science Computer Review.* April 28, 2021. Texto disponible en:
https://journals.sagepub.com/doi/10.1177/08944393211006863

antes de que se adopte la decisión[223]. De esta manera, a mayor personalización exista en la decisión, más posibilidades de que esta se considere relevante. En este sentido, la personalización de las decisiones va unida de la mano de un mayor conocimiento que tienen las organizaciones que hacen uso de estos sistemas automatizados sobre los particulares. De manera que la afectación en su esfera es más patente. Sentado lo anterior, lo cierto es que, la personalización de las decisiones presenta distintos grados de relevancia en función de los sectores o el tipo de decisión que se adopte.

- Así, por lo que se refiere a la *publicidad personalizada*, como ya hemos señalado antes, como regla general, esta no se entenderá que genera efectos relevantes[224]. No obstante, desde el punto de vista normativo, cuando dicha publicidad compartimental se ofrece en determinados espacios, como pueden ser las plataformas digitales, esta puede presentar importantes riesgos en la esfera de los particulares y por lo tanto, se han de prever especiales garantías cuando la misma se muestre en dichos entornos. Ello es así porque los sistemas publicitarios utilizados por las plataformas en línea de muy gran tamaño tienen más capacidad para dirigirse y llegar a los destinatarios del servicio en función de su comportamiento dentro y fuera de la interfaz de ese contexto[225].
- Algo parecido ha ocurrido cuando lo que se *personaliza es el contenido informativo* que recibe el particular. Así, a la hora de abordar si la personalización de noticias se considera que genera un efecto relevante, la doctrina ha distinguido cómo se produce la previa personalización que lleva a la toma de la decisión. Si la personalización es establecida por el usuario a través de una serie de filtros que él mismo elige en función de las opciones que establece la propia organización que ofrece la noticia, habría que entender que las decisiones que se basen en dicha personalización no se consideran relevantes. Sin embargo, en aquellos casos en los

[223] La Autoritat Catalana de Protecció de dades ha considerado que en principio, un servicio público personalizado que elabora perfiles con el fin de informar a las personas sobre los servicios públicos que pueden afectar a estas de manera más previsible no generaría efectos significativos en la esfera de los particulares. En: Autoritat Catalana de Protecció de Dades. Informe nº PD 9/2019, pág.10. Texto disponible en: https://apdcat.gencat.cat/web/.content/Resolucio/Resolucions_Cercador/Dictamens/2019/Documents/ca_pd_2019_009.pdf

[224] Aunque el Grupo del Artículo 29 así lo ha señalado, no todas las organizaciones o colectivos piensan lo mismo. Consideran que la publicidad compartimental en la mayoría de los supuestos va más allá de las exceptivas que esperan los consumidores que la reciben. En: Privacy international. Data is power: *Towards additional guidance on profiling and automated decision-making in the GDPR*. op.cit., pág.13.

[225] Véase los considerandos 53,54,55,56,63 y 66 y artículo 30 de la 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.

que la personalización se derive esencialmente en cómo el proveedor de los noticias las ofrece en función de los datos que ostenta del particular, esa afectación se vuelve más relevante[226]. Es decir, en función de la participación más o menos activa del particular, los efectos relevantes podrán ser mayores o menores. En este mismo orden de cosas, se ha señalado que, la actitud más o menos pasiva del particular deja de ser un factor a tener en cuenta cuando dicha personalización de contenido informativo se genera en las grandes plataformas digitales. En estos contextos se indica que cualquier tipo de personalización puede generar efectos significativos en la esfera de esos individuos[227]. En este sentido, y al igual que ocurre con la publicidad compartimental, la propuesta de reglamento de servicios digitales propone también una serie de garantías específicas en favor de los particulares que ven personalizado el contenido informativo en el seno de estas grandes plataformas[228].

- Finalmente, cabe mencionar la llamada *personalización de precios*. Para el GT29, el artículo 22 del RGPD puede ser aplicado cuando el precio mostrado sea prohibitivamente elevado[229], de manera que impida a una persona acceder a determinados bienes o servicios fruto de la personalización derivada de los datos de ese particular. Se pone el acento esencialmente en que dicho precio coarte o limite las posibilidades de obtención de ese producto por el coste del mismo. Para valorar ese elemento prohibitivo, la doctrina propone que se tome como referencia el precio que ofrece el sistema y se compare con el precio de mercado del producto en cuestión o el precio promedio cobrado por el vendedor por ese producto[230].Así, si el precio ofrecido es desproporcionado habría de considerarse relevante. En nuestra opinión, además, también se debe valorar el grado de intrusión o conocimiento que se pueda llegar a lograr por parte de la organización respecto del

[226] ESKENS,S: "A right to reset your user profile and more: GDPR-rights for personalized news consumers". *International Data Privacy Law*, Volume 9, Issue 3, August 2019, pág.15. Disponible en: https://doi.org/10.1093/idpl/ipz007
Los medios de comunicación denominados intermediarios pueden ser éticamente problemáticos en su diseño concreto. Sus sistemas proporcionan a los usuarios una selección personalizada de información. Ello origina una selección sobre los contenidos mostrados. Sin embargo, dado que la gran mayoría del contenido no se muestra o sólo se muestra de forma subordinada, el espectro de percepción del individuo se reduce. En: Comité de ética alemán. Gutachten der Datenethikkommission, 2019, op.cit., pág.176.
Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a un mercado único de servicios digitales (Ley de servicios digitales). Artículo 29.
Grupo del Artículo 29. "Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679",op.cit., pág.25.
WONG,B: "Online personalised pricing as prohibited automated decision-making under Article 22 GDPR: a sceptical view". Information & Communications Technology Law, 30:2, 193-207, 2021, pág.200. Disponible en: https://doi.org/10.1080/13600834.2020.1860460

particular sobre el que se muestra el precio personalizado. Y es que no podemos olvidar que el precio es uno de los elementos que más influye y que más tienen en cuenta los potenciales compradores a la hora de optar por la adquisición de un producto u otro, por ello, una personalización más agresiva presentará normalmente ese carácter significativo[231]. Distinto es, sin embargo, aquellos sistemas automatizados en los que se muestran precios dinámicos. En estos casos, dado que no existe personalización y el precio mostrado afecta a la totalidad de los usuarios que acceden a esa oferta, el impacto no sería tan significativo. En estos supuestos, los bienes jurídicos afectados se focalizan más en las posibles implicaciones que este tipo de precios pueden generar en materia del derecho de competencia.

Por tanto, a la hora de valorar la relevancia de la decisión, el responsable ha de tener en cuenta los efectos que esta puede generar en los particulares. Por tanto, factores como; la alteración del estatuto jurídico, el acceso a bienes o servicios, el impacto que generen las decisiones, los colectivos a los que vaya destinadas las mismas o el grado de personalización de la actuación favorecerán a la consideración de dicho carácter relevante o significativo.

a.3. Características de las decisiones. Positivas y negativas

Para terminar con el análisis del concepto de decisiones automatizadas del artículo 22. Cabe indicar que el término "*afectar de manera significativa*" no debe entenderse como que el mismo únicamente engloba las decisiones que generan efectos negativos sobre el particular. También forman parte del ámbito de aplicación de este precepto las decisiones que reporten beneficios o sean positivas para los individuos afectados por las mismas. En este sentido, si bien será frecuente que los particulares sobre todo se preocupen por las decisiones que le afecten negativamente, no hay justificación para considerar que aquellas que puedan resultar favorables no reciban el mismo tratamiento. En este sentido, un responsable en muchas ocasiones no será consciente de si su algoritmo a priori tomará decisiones negativas o positivas sobre los

[231] Hay quien incluso propone que la personalización de precios debería estar prohibida. En: TENA ARREGUI, R: "¿Son justos los precios personalizados mediante algoritmos?". *El notario del siglo XXI: revista del Colegio Notarial de Madrid*, Nº. 87, 2019, págs. 14 a 19.

particulares específicos. Es más, la regla general es que adopte tanto decisiones positivas como negativas. Así, un particular puede haber recibido un crédito de forma automatizada (decisión positiva a priori), pero dicho crédito puede haberlo obtenido con un interés más alto que el establecido para otra persona (decisión negativa) por el mero de que el sistema, a la hora de procesar los datos, ha considerado distintos tipos de interés e función de esos datos personales. Además, por otro lado, si el legislador europeo hubiera pretendido incorporar el carácter negativo o perjudicial como nota específica de las decisiones descritas en el artículo 22 así lo habría indicado[232], tal y como lo establecen expresamente otras normas que regulan este mismo fenómeno[233].

a.4.Evaluación ex ante y ex post del tipo de decisiones que adopta el sistema

Como es lógico, corresponde al responsable del tratamiento valorar si las decisiones que adoptará su sistema son encajables en la definición de tratamiento reconocida en el artículo 22 del RGPD. Esta valoración se ha de realizar antes de que el sistema comience a adoptar decisiones. Ello es sumamente relevante teniendo en cuenta las garantías que en su caso se prevén para este tipo de decisiones. Así, el responsable se encuentra en la mejor posición para valorar si las decisiones que adoptará su sistema se automatizarán plenamente y si además dichas decisiones pueden generar efectos relevantes en la esfera de los particulares. En este sentido, tanto el análisis de riesgos como la evaluación de impacto servirán para realizar esa evaluación ex ante del tipo de

[232] Durante la tramitación de la de la Directiva de Protección de Datos que regulaba las decisiones automatizadas, en algunas propuestas se indicó que estas decisiones solo deberán referirse a aquellas decisiones que generarán efectos perjudiciales. Sin embargo, esta propuesta fue descartada en fases posteriores. Véase, la propuesta modificada de la Directiva del Consejo relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (l) 192 / C 311 / 04 COM(92) 422 final — SYN 287 de 27 de noviembre de 1992. Pág.21. Texto disponible en:
https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:JOC_1992_311_R_0030_01&from=EN

[233] *Los Estados miembros dispondrán la prohibición de las decisiones basadas únicamente en un tratamiento automatizado, incluida la elaboración de perfiles, que produzcan efectos jurídicos negativos para el interesado o le afecten significativamente, salvo que estén autorizadas por el Derecho de la Unión o del Estado miembro a la que esté sujeto el responsable del tratamiento y que establezca medidas adecuadas para salvaguardar los derechos y libertades del interesado, al menos el derecho a obtener la intervención humana por parte del responsable del tratamiento*. Artículo 11.1 de la 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. En el mismo sentido, véase también el dictamen del Grupo del artículo 29 sobre algunas cuestiones fundamentales de la Directiva sobre protección de datos en el ámbito penal (UE 2016/680). Aprobado el 29 de noviembre de 2017. (pág.13).

decisiones que se pretenden adoptar[234]. En aquellos supuestos en los que el responsable del tratamiento no considere que las decisiones que se adoptarán en su organización se encuadran en la definición del 22 habrá de justificarlo adecuadamente y deberá dejar constancia de ello por escrito. Esta justificación se adecuará a las razones específicas que han llevado a considerar que su sistema de adopción de decisiones no cumple alguno de los dos requisitos acumulativos que prevé el artículo 22, es decir, el carácter plenamente automatizado y la relevancia significativa de las decisiones. Para los supuestos en los que se aluda que no existe una plena automatización, el responsable deberá justificar el papel del humano en el proceso decisorio tratando de evidenciar que la presencia de la persona es relevante y no existe una decisión fáctica automatizada. Por otro lado, para justificar la falta de relevancia de las decisiones adoptadas, las organizaciones pueden utilizar todo tipo de documentos, informes o guías oficiales que apoyen su argumento adaptadas al contexto específico de su proceso decisorio.

Junto a la justificación que acredita que un sistema algorítmico no adopta las decisiones definidas en el artículo 22. El responsable del tratamiento también deberá establecer mecanismos que le permita valorar de forma periódica si dicho sistema por diversas circunstancias comienza a adoptar decisiones que pudieran encuadrase en tal definición. Como dijimos en páginas anteriores, los sistemas pueden alterar su comportamiento por su interacción con el entorno, ello puede llevar a que el algoritmo comience a adoptar decisiones no esperadas. Además, cualquier cambio en el colectivo al que vayan dirigidas las decisiones o el momento temporal en el que se adoptan estas también puede suponer un replanteamiento inicial del tipo de decisiones que se adoptan[235]. Por tanto, es recomendable que las organizaciones realicen testeos periódicos del sistema, tanto internos como externos. A su vez, conviene que establezcan cauces para que los particulares afectados por las decisiones puedan presentar quejas del funcionamiento del sistema.

En definitiva, encuadrar adecuadamente el tipo de decisión que adopta el sistema es una tarea esencial que han de establecer los responsables. El legislador europeo ha considerado que los particulares sometidos a las decisiones que se encuadran en el artículo 22 del RGPD han de recibir un nivel de garantías superiores al resto de

[234] Véase el capítulo III, apartados I y II de esta tesis.

[235] Ya hemos hecho referencia anteriormente a las decisiones que pueden destinarse a colectivos especialmente vulnerables. A su vez, también hemos indicado como un sistemas que en principio no generan efectos significativos como puede ser el autocompletado de un buscador puede comenzar a afectar a la esfera de los particulares de forma relevante cuando por ejemplo propone determinadas búsquedas en el periodo electoral.

decisiones. Ello es así porque considera que estas pueden afectar en mayor grado a los derechos e intereses de estos particulares y por tanto se requiere de una protección especial. El responsable no puede eludir esta definición creando falsas intervenciones de personas durante el proceso decisorio u obviar la relevancia de las decisiones que adopta el sistema. Con ello contribuirá a una mayor protección de estos particulares y además mostrará un cumplimiento efectivo de la norma.

a.5 Los requisitos acumulativos de la plena automatización y la relevancia de las decisiones. ¿Divergencias con otros textos normativos?

Ya ha quedado claro en las páginas anteriores que el RGPD establece una serie de garantías específicas en favor de aquellos particulares que se ven sometidos a decisiones plenamente automatizadas relevantes. Sin embargo, en muchas ocasiones, las decisiones que en mayor grado afectan a la esfera de los ciudadanos no provienen de sistemas plenamente automatizados sino de aquellos donde la decisión es parcialmente automatizada. Conscientes de esta realidad, la mayoría de los textos normativos que han afrontado la regulación de los sistemas de decisiones automatizadas han focalizado su atención en la gravedad de las decisiones que adoptan los sistemas, dejando en un segundo plano la plena o no automatización de los procesos decisorio. En este sentido, países como Canadá[236], Nueva Zelanda[237] o la propia Unión Europea[238] tienen en cuenta el potencial riesgo que un sistema puede generar para un particular en sus derechos como parámetro para asignar mayores o menores exigencias jurídicas. Pues bien, aunque a priori podría pensarse que el RGPD genera una desprotección de las decisiones parcialmente automatizadas relevantes, lo cierto es que una interpretación sistemática del texto lleva irremediablemente a otra conclusión. En este sentido, este texto normativo prevé todo un conjunto de medidas de responsabilidad activa. Estas medidas se han de ponderar entre otros factores basándose en el riesgo potencial que dichos tratamientos pueden generar en los derechos y libertades de los particulares[239]. Por tanto, en muchas ocasiones resultará habitual que tratamientos de datos donde se

[236] Directive on Automated Decision-Making de 1 de abril de 2019. Véase el texto en: https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592

[237] Algorithm Charter for Aotearoa New Zealand de Julio de 2020. Véase el texto en: https://data.govt.nz/assets/data-ethics/algorithm/Algorithm-Charter-2020_Final-English-1.pdf

[238] Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

[239] Véase el artículo 24 del RGPD sobre responsabilidad del responsable del tratamiento.

adopten decisiones parcialmente automatizadas generarán mayores riesgos que tratamientos donde se haya optado por la plena automatización. Como consecuencia de ello, es probable que los responsables de los tratamientos acaben implementando en muchas ocasiones las mismas medidas de cumplimento normativo independientemente del tipo de sistema que se utilice, esto es, que se total o parcialmente automatizado el proceso decisorio.

B) Decisiones parcialmente automatizadas que generan efectos relevantes

En este tipo de decisiones, los resultados que emite el algoritmo no afectan automáticamente al particular sino que estos se utilizan como apoyo a la decisión final que se adoptará[240]. Es decir, la salida que vierte el sistema es uno de los elementos, entre otros, que tiene en cuenta la organización para configurar la decisión que posteriormente afectará a las personas. Este tipo de decisiones resultan ser hasta la fecha las que más atención han causado a la doctrina debido a que son las que habitualmente acaban implantando las organizaciones tanto públicas como privadas en sus procesos de toma de decisiones[241]. Son diversas razones las que empujan a empresas y poderes públicos a desplegar este tipo de decisiones que no alcanza la plena automatización, algunas son:

En primer lugar, hay que indicar que a día de hoy, los sistemas basados en inteligencia artificial todavía carecen de toda una serie de elementos que están presentes en las decisiones adoptadas por los seres humanos y que resultan muy relevantes en los procesos decisorios. Así, el factor sentimental[242], emocional o el conocimiento amplio del contexto social o económico presente sobre el que subyace la decisión que se pretende tomar siguen resultando muy difíciles de incorporar al código del algorítmico.

---

[240] HOGAN-DORAN,D: "Computer says "no": automation, algorithms and artificial intelligence in Government decision-making". *The Judicial Review*, 2017, pág.2 y 3.

[241] OLSEN, H.P., SLOSSER, J., HILDEBRANDT, T., & WIESENER, C: "What's in the Box? The Legal Requirement of Explainability in Computationally Aided Decision-Making in Public Administration". *Political Economy: Structure & Scope of Government eJournal*.. 2019, p-9. En el sector público véase distintos usos en de este tipo de sistemas en: COTINO HUESO,L: Hacia la transparencia 4.0: el uso de la inteligencia artificial y big data para la lucha contra el fraude y la corrupción y las (muchas) exigencias constitucionales. En RAMIÓ,C (coord): "Repensando la Administración Pública. Administración digital e innovación pública". Ed. INAP, Madrid, 2021. Págs.169 a 196. Véase también las resoluciones de la Red de Derecho Administrativo e Inteligencia Artificial (DAIA). Más información en: http://reddaia.org/

MARTÍNEZ MARTÍNEZ,R: "Cuestiones de ética jurídica al abordar proyectos de Big Data. El contexto del Reglamento general de protección de datos". *Dilemata*, año 9, nº 24, 2017, pág.154.Texto disponible en: https://www.dilemata.net/revista/index.php/dilemata/article/view/412000105/495

Por tanto, en aquellos sectores o ámbitos donde ese grado de subjetividad este presente y sea necesario, por ahora, y dadas las limitaciones tecnológicas, los humanos deberán mantener ese papel relevante en la decisión final. A su vez, en muchos casos, por mucho que se pretenda moldear un determinado entorno y tratar de representar esa realidad física, la imposibilidad para traducirlo en variables que midan cada aspecto de dicha realidad resulta materialmente imposible[243]. Ello empuja a las organizaciones a incorporar el elemento humano con el objetivo de compensar ese déficit técnico.

En segundo lugar, los sistemas algorítmicos que se utilizan como apoyo a la decisión final suelen ser idóneos en sectores donde la formación previa de la persona junto con el resultado emitido por el sistema dan lugar a decisiones más precisas que si las mismas fueran enteramente adoptadas por una máquina o por un humano. Ello suele ocurrir en ámbitos donde el algoritmo realiza la tarea más laboriosa o tediosa y la persona con el resultado emitido por el sistema aplica su conocimiento para adoptar una u otra decisión final. Por ejemplo, en el ámbito sanitario, cada vez se están diseñando más sistemas que tienen como objetivo tratar de identificar diversas enfermedades presentes y futuras basándose únicamente en los historiales de un paciente. Pues bien, este sistema al detectar esa posible enfermedad, permite al facultativo médico valorar el resultado dictado por el algorítmico y a partir de ahí, decidir qué actuaciones acometer[244]. De esta manera, la experiencia del humano junto con la labor del algoritmo permiten obtener una combinación más precisa que en aquellos supuestos en los que tales decisiones fueran adoptadas directamente por la máquina, o solamente por la persona[245].

En tercer lugar, la incorporación del humano durante el proceso decisorio suele justificarse en aquellos ámbitos donde un error en el sistema que vierte decisiones

[243]"*La matemática sólo conecta con la realidad cuando las entidades matemáticas, abstractas, representan adecuadamente una realidad física numéricamente mensurable. Sólo si se atribuyesen valores matemáticos a realidades físicas que pueden ser así medidas, las fórmulas matemáticas servirían para una predicción fiable"*. En: DE LA OLIVA SANTOS, A: "Justicia predictiva, interpretación matemática de las normas, sentencias robóticas y la vieja historia del justizklavier". *El cronista del estado social y democrático de derecho*, Nº 80, 2019, pág.34.

[244] Es el caso del sistema PathAI, este es mucho más preciso cuando el mismo se conjuga con la supervisión de una persona autorizada que entiende adecuadamente el sistema. Véase en: https://www.pathai.com/

[245] El uso del VAR en el ámbito deportivo es un ejemplo de ello. La máquina ayuda a los árbitros en las jugadas deportivas más polémicas. Fuente de la noticia: MALIK,K: "Technology will never replace human judgment. Look at football…". The Guardian. 16/11/2019. Noticia disponible: https://www.theguardian.com/commentisfree/2019/nov/16/technology-will-never-replace-human-judgment-look-at-football

automatizadas pudiera causar graves perjuicios para los particulares. De manera que, la presencia del humano sirve como filtro para valorar dicho resultado algorítmico y por tanto, probado que el sistema ha vertido un resultado erróneo, la decisión no se adopta. Con ello se evitan o reducen las consecuencias negativas que se derivarían de esa decisión si esta fuera totalmente automatizada y por tanto generará efectos inmediatos a los particulares, con independencia de que posteriormente fueran o no reparados los daños causados. Así, no es de extrañar que diversas normativas y resoluciones judiciales, previendo los riesgos que se pueden derivar de esa plena automatización, impongan la necesidad de establecer la presencia previa del humano antes de que se adopte una decisión altamente relevante para los particulares afectados por la misma[246].

De lo dicho hasta ahora en este apartado se desprende la importancia que tienen las personas encargadas de analizar los resultados emitidos por los sistemas algorítmicos. De esta manera y como ya comentábamos anteriormente, se hace necesario que el responsable del tratamiento justifique las razones que le han llevado a decidir la presencia del humano en dicho tratamiento. Deben quedar claras las funciones que tiene asignadas y las facultades que ostenta para poder alterar el resultado emitido por el algoritmo. Además, ha de contar con la suficiente formación para entender entre otras cosas: los resultados del sistema, cómo dicho sistema interactúa con la realidad y por supuesto, un conocimiento amplio del entorno donde opera el mismo[247]. En muchos casos, dado que estos sistemas se alterarán continuamente, estas personas deberán recibir actualizaciones periódicas de los avances técnicos que puedan sufrir estos

[246] Artículos 12.4 y 16.1 de la Ley Orgánica 1/2020, de 16 de septiembre, sobre la utilización de los datos del Registro de Nombres de Pasajeros para la prevención, detección, investigación y enjuiciamiento de delitos de terrorismo y delitos graves. Estos preceptos obligan a las Unidades de información sobre Pasajeros (UIP) a realizar una evaluación individualizada y humana de aquellos resultados positivos en los que un sistema algorítmico detecta una coincidencia en determinadas bases de datos. De esta manera, y una vez se analice ese resultado positivo, la Unidad de información sobre Pasajeros (UIP) puede enviar dicho resultado a las autoridades policiales para que lleven a cabo las actuaciones pertinentes. Por otro lado, el análisis previo del resultado por parte de una persona antes de que se adopte una decisión que afecta negativamente al particular también fue puesto de manifiesto en el Dictamen 1/15 del TJUE (Gran Sala) de 26 de julio de 2017 sobre el Proyecto de Acuerdo entre Canadá y la Unión Europea con relación a la transferencia de los datos del registro de nombres de los pasajeros aéreos desde la Unión a Canadá. Véase el considerando 173.
Un análisis doctrinal de este dictamen puede encontrarse en: VEDASCHI,A: "Privacy and data protection versus national security in transnational flights: the EU–Canada PNR agreement", *International Data Privacy Law*, Volume 8, Issue 2, May 2018, págs. 124–139.
Texto disponible en: https://doi.org/10.1093/idpl/ipy004

[247] ICO. Guidance on AI and data protection. How do we ensure individual rights in our AI systems? Información disponible en:
https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/how-do-we-ensure-individual-rights-in-our-ai-systems/

sistemas algorítmicos. El objetivo que se pretende con todas estas exigencias no es otro que el de permitir que la tarea asignada a las personas que revisan los resultados vertidos por los algoritmos sea la más adecuada posible.

C) Decisiones plenamente automatizadas que no generan efectos relevantes

El uso de sistemas automatizados en la toma de decisiones que no generan efectos relevantes está ampliamente implantado en multitud de sectores y ámbitos. Son sin duda las que predominan a día de hoy. Es decir, todo tipo de decisiones totalmente mecanizadas pero que no generen ningún efecto significativo en los particulares o los efectos producidos son insignificantes. Ejemplos de estas decisiones los encontramos en los autocompletados de los buscadores, recomendación de anuncios personalizados o contenidos audiovisuales basados en perfilados poco intrusivos, etc. La normativa de protección de datos también es aplicable enteramente a este tipo de decisiones. No obstante, teniendo en cuenta que el único elemento destacable de estos sistemas es la plena automatización de la decisión derivada de los posibles datos personales que se obtengan del particular afectado por la misma[248], el responsable del tratamiento sobre todo deberá tener especiales cautelas en aquellos casos en los que dichos sistemas puedan en determinados momentos comenzar a adoptar decisiones que sí pudieran generar efectos significativos en la esfera de los particulares. Estas cautelas ya han sido señaladas previamente cuando hacíamos mención a las decisiones descritas en el artículo 22.

En definitiva, el RPGD aplica a cualquier tratamiento de datos que involucre la toma de decisiones, ya sean o no plenamente automatizadas. Este texto ha previsto una regulación específica para aquellos tratamientos donde se adopten decisiones totalmente automatizadas que generan efectos relevantes. Sin embargo, para el resto de decisiones, toda la normativa de protección de datos se aplica enteramente. Gracias al enfoque del riesgo sobre el cual, el RGPD configura el establecimiento de mayores o menores medidas de cumplimiento normativo[249], el responsable en muchos supuestos deberá

[248] Ya se ha puesto de relieve que la publicidad compartimental a priori no se considera que genere efectos significativos en la esfera de los particulares. En: MORENTE PARRA, V: "Big data o el arte de analizar datos masivos. una reflexión crítica desde los derechos fundamentales". *Derechos y libertades*, Número 41, Época II, junio 2019, p-242. En cuanto al uso de las cookies, también parece existir consenso en que las decisiones automatizadas que se tomen basadas en ellas tampoco hay que entender que generan efectos significativos. En: Agencia Española de protección de datos. *Guía sobre el uso de las cookies*. pág.17.

[249] El enfoque basado en el riesgo se analiza en el capítulo III de esta tesis.

implantar prácticamente las mismas garantías en favor de los particulares que se vean sometidos a decisiones total o parcialmente automatizadas relevantes.

2. La elaboración de perfiles

El segundo concepto básico que está estrechamente relacionado con el objeto de esta tesis y para el cual el RGPD no sólo prevé ciertas garantías como ocurría con las decisiones automatizadas sino que además establece una definición específica es la elaboración de perfiles. Es turno de analizar dicha definición y la interpretación que se ha realizado de este concepto. En capítulos posteriores se estudiarán las reglas específicas aplicables a este tratamiento de datos.

A) Definición

De acuerdo al artículo 4 apartado 4 del RGPD la elaboración de perfiles es:

> *toda forma de tratamiento automatizado de datos personales consistente en utilizar datos personales para evaluar determinados aspectos personales de una persona física, en particular para analizar o predecir aspectos relativos al rendimiento profesional, situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, ubicación o movimientos de dicha persona física*[250]*;*

De esta manera, para el GT29 son tres los elementos básicos que caracterizan a este tratamiento, en primer lugar, este ha de ser automatizado, si bien, puede existir cierta presencia del humano, en dicho tratamiento han de tratarse datos personales y, el objetivo de dicha elaboración de perfiles perseguirá la evaluación de aspectos personales sobre una persona física[251].

---

[250] El Comité Económico Social Europeo propuso durante la fase de tramitación del RGPD que se estableciera una definición específica de que se entendía por elaboración de perfiles, ya que si bien, dicho concepto aparecía mencionado en distintos lugares del RGPD, no se establecía una definición específica al uso. Véase al apartado 4.6.5 de esa propuesta. En:
https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex%3A52012AE1303
Seguidamente, el Parlamente Europeo propuso una definición específica a través de la enmienda 98. Enmienda aportada por la Comisión de Libertades Civiles, Justicia y Asuntos de Interior y que incorporó el Parlamento Europeo en su propuesta.
https://www.europarl.europa.eu/doceo/document/A-7-2013-0402_ES.html?redirect#title3

[251] "Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679". Grupo del Artículo 29, op,.p-10 y 18. Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, 2020, pág.10.

Así, por *perfil* hemos de entender al conjunto de datos que caracterizan a una categoría de personas[252]. De manera que dichos atributos presentes en esas personas se le asignan un determinado perfil. Por ejemplo, un perfil sería: mujeres blancas, de edad comprendida entre los 25 y 40 años con titulación universitaria. Para crear un perfil y en su caso aplicar dichos perfiles a personas físicas con el objetivo de evaluar su comportamiento normalmente se llevan a cabo tres fases;

En *primer lugar* se han de recopilar los datos que se utilizarán para la creación de los perfiles. Estos datos se podrán obtener de multitud de fuentes aunque normalmente estarán disponibles en los *data center* o *data lakes* de los que disponen aquellos que pretendan realizar la elaboración de perfiles. Ejemplo: Conjunto de datos disponibles de los afiliados de un partido político o historial clínico y de compras de una serie de clientes. También se pueden incorporar otros datos provenientes de los portales de datos abiertos publicados por las Administraciones Públicas.

En *segundo lugar*, esos datos serán procesados por distintos algoritmos que se basarán en técnicas de *data mining* y *machine learning*. A través de la aplicación de estas herramientas se extraerán determinados patrones y correlaciones, los cuales, se analizarán y en su caso se utilizarán para crear los perfiles de dichos conjuntos de datos. Indicamos varios ejemplos de lo indicado hasta ahora.

Conjunto de datos formado por datos médicos y hábitos de compra					
Muestras	**Atributos**				
	Peso	Enfermedad grave	Fumador/a	Litros/Día	Color del pelo
Asegurado 1	Entre 55-60 kg	Si	No	0,25	Castaño
Asegurado 2	Entre 55-60 kg	No	Si	0,5	Rubio
Asegurado 3	Más de 100 kg	Si	No	0	Moreno
Asegurado 4	Entre 70 y 85 kg	No	No	0	Moreno
Asegurado 5	Menos de 50 kg	No	Si	1	Moreno

Por ejemplo, en la tabla anterior encontramos un conjunto de datos de una compañía de seguros compuesto por toda una serie de datos de salud, hábitos de comportamiento e historiales de compra pertenecientes a sus clientes. Esta compañía pretende crear toda una serie de perfiles para posteriormente aplicarlos a futuros

[252] Véase la Recomendación CM/Rec(2010)13 del Comité de Ministros a los Estados miembros sobre la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal en el contexto de la creación de perfiles.
https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=09000016805cdd2a

potenciales clientes que pretenda contratar con ellos. Para ello, dichos datos son procesados por diferentes algoritmos en buscas de patrones. Imaginemos que tras el procesamiento de dichos datos se detectan una fuerte correlación entre el peso del cliente y el consumo de más de medio litro de leche al día por un lado, y el padecimiento de una enfermedad grave. De esta manera y aplicadas estas correlaciones de forma probabilística se llega a la conclusión de que las personas que ingieran más de medio litro al día de leche y que además tengan un peso comprendido entre 70 y 85 kilos tienen un riesgo alto de presentar una enfermedad grave.

Conjunto de datos o dataset de un partido político					
Muestras	**Atributos**				
	Género	Afinidad líder	Código Postal	Asunto de relevancia	Edad
Simpatizante 1	Hombre	Si	1235	Economía	35
Simpatizante 2	Mujer	No	3456	Salud	24
Simpatizante 3	Mujer	Si	7658	Salud	28
Simpatizante 4	Hombre	Si	7658	Trabajo	75
Simpatizante 5	Hombre	No	1235	Trabajo	50

Otro ejemplo, un partido político tiene como objetivo la mejora de sus resultados electorales en las próximas elecciones. Esta organización ha realizado un cuestionario a un conjunto de simpatizantes y afiliados del partido. Dicha información es procesada a través de distintos algoritmos y se obtienen varias correlaciones. Una primera correlación indica que las personas que habitan en ciudades de menos de 10.000 habitantes tienen una afinidad baja al líder político que pretende presentarse en las elecciones. A su vez, también se detecta una fuerte correlación entre las mujeres mayores de 35 años que residen en ciudades de más de 15.000 habitantes y su interés por asuntos económicos. De esta manera, la agrupación política extrae dos grupos a los que se les asocia un perfil. Por un lado, personas que habitan en pueblos de menos de 10.000 habitantes y por otro, las mujeres mayores de 35 años que residen en ciudades de más de 15.000 habitantes.

Con estos dos ejemplos, se puede ver cómo, durante esta fase, las organizaciones extraen aquellos perfiles que potencialmente presentan toda una serie

de correlaciones. Estos perfiles determinan distintos grupos de personas y las organizaciones pueden utilizarlos para todo tipo de actuaciones.

En *tercer lugar*, y una vez creados los perfiles, las organizaciones puede aprovecharlos para aplicarlos a una persona física concreta que presente las mismas características del mencionado perfil con el fin de prever e inferir sus preferencias, comportamiento o actitudes personales. Encuadradas en dicho perfil[253], estas personas se le asignará la inferencia atribuida a los miembros de ese grupo y las consecuencias que se hayan previsto para aquellas personas que se amolden al mencionado perfil.

De esta manera y volviendo a los supuestos anteriores, ante un potencial cliente que pese 74 kilos y consuma un litro de leche al día, la aseguradora, teniendo en cuenta el perfil al que se incorpora, entenderá que esta persona presenta un riesgo alto de sufrir una enfermedad y por tanto, podrá denegar la póliza o establecer una más alta. Por lo que se refiere al partido político, teniendo en cuenta que se detectaron dos correlaciones entre los datos, el partido político puede aplicar esos perfiles a distintos grupos de personas. Para el supuesto de los ciudadanos que habitan en ciudades de menos de 10.000 habitantes, la agrupación política puede desarrollar publicidad personalizada que ensalce la labor del líder. Para ello, cada vez que acceda una persona a la web del partido y se conecte a través de una dirección IP correspondiente a un código postal de una ciudad de menos de 10.00 habitantes, estas personas recibirán esa publicidad personalizada. Por otro lado, para el segundo perfil, esto es, mujeres mayores que residan en ciudades de más de 15.000 habitantes, la agrupación política podría facilitar esos perfiles a una plataforma digital para que esta ofreciera publicidad personalizada a las mujeres que presenten ese perfil en la que se potencie propaganda relacionada con las propuestas económicas que pretende implantar esta agrupación política en las siguientes elecciones.

Perfil creado por la compañía de seguros	**Elaboración del perfil a una persona específica**	**Objetivo de la elaboración del perfil**
Personas entre 70 y 85 kilos que consumen más de medio litro →	Potencial cliente que pesa 74 kilos y 1 litro de leche/día.	Predecir el costo de la póliza del potencial cliente.

[253] GARRIGA DOMÍNGUEZ,A:. *Nuevos retos para la protección de datos personales: en la Era del Big Data y de la computación ubicua*. Ed. Dykinson, Madrid, 2016, pág.67.

tienen un riesgo alto de padecer cáncer.	→ Riesgo alto de padecer una enfermedad →Se deniega la póliza o se establece un precio mayor	

Perfil creado por el partido político (1)	**Elaboración del perfil y consecuencia prevista**	**Objetivo de la elaboración del perfil**
Personas que viven en ciudades de menos de 10.000 → presentan una baja afinidad con el líder del partido.	Persona que accede a través del ordenador de su casa a la web del partido en una ciudad de menos de 10.000 →Baja afinidad con el líder →Publicidad personalizada presentando una buena imagen del líder.	Predecir su afinidad política y tratar de alterar el comportamiento para que estas personas tengan una imagen más afín al líder político.

Perfil creado por el partido político (2)	**Elaboración del perfil aplicado por la plataforma digital y consecuencia prevista**	**Objetivo de la elaboración del perfil**
Mujeres mayores de 35 años que residen en ciudades de más de 15.000 habitantes → interés por asuntos económicos	Mujer que tiene 38 años y reside en una gran ciudad y es usuaria de una plataforma digital →Están interesadas en asuntos económicos. →Publicidad personalizada sobre asuntos económicos.	Predecir su afinidad política y tratar de alterar el comportamiento para potenciar que este colectivo sea más afín a este partido político.

Como puede observarse, a través del proceso de elaboración de perfiles, personas físicas que ostentan determinados atributos presentes en un perfil previamente creado, se le asignan determinadas consecuencias, las cuales, se han desarrollado o configurado tras las correlaciones previamente detectadas en los atributos que son relevantes para construir el perfil.

B) Tipos de perfiles

En función de diversos factores la doctrina ha establecido distintas clasificaciones de perfiles, nosotros vamos a hacer mención a aquellos que consideramos más relevantes para el objeto del estudio de esta tesis[254]. Estos es, los perfiles individuales y grupales, y, dentro de estos últimos, los distributivos y no distributivos.

En *primer lugar*, por lo que se refiere a los perfiles individuales, estos comprenden aquellos en los que el perfil que se elabora es específico para la persona sobre la que se aplica el mencionado perfil. En estos supuestos, el perfil individual agrega información sobre un individuo y la usa para derivar o inferir características desconocidas. Así, se ha señalado que si bien, a día de hoy los perfiles grupales siguen presentando su preponderancia[255], conforme aumente la disponibilidad de datos masivos, estos cada vez presentarán predicciones más personalizadas e individualizadas[256].

En *segundo lugar*, por lo que se refiere a los perfiles grupales, estos categorizan a las personas en un grupo que comparte todas o algunas características, las cuales, estando presentes, permiten encuadrar a esa persona en ese perfil. Por lo tanto, son la conjunción de varios atributos presentes en un sujeto los que permiten incorporarlo a dicho grupo. Dentro de los perfiles grupales la doctrina además ha distinguido entre grupos distributivos y no distributivos[257]. Los *perfiles distributivos* están conformados por toda una serie de características que todos los miembros de esos grupos las

[254] Por todos, HILDEBRANDT, M & GUTWIRTH,S: *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed. Springer, 2008. Véase también: POLITOU, E, ALEPIS,E AND ATSAKIS,C: "Profiling tax and financial behaviour with big data under the GDPR", *Computer Law & Security Review*, Volume 35, Issue 3, May 2019, pág.3. También GIL GONZÁLEZ, E., DE HERT, P.: "Understanding the legal provisions that allow processing and profiling of personal data—an analysis of GDPR provisions and principles". *ERA Forum* 19**,** 597–621, 2019-. Págs. 608 y ss.

[255] *Más que individualización, debemos hablar de segmentación del riesgo y agrupación de asegurados por tribus. Estas tribus o pools cada vez son más pequeñas, gracias al uso de grandes masas de datos, pero no hay una individualización plena. El asegurado, con los datos que tiene de cada solicitante de póliza, elabora un perfil de riesgo y lo incluye en el pool de asegurados que presentan un perfil similar*. MUÑOZ PAREDES, M,L: "Big data y contrato de seguro: los datos generados por los asegurados y su utilización por los aseguradores". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed.Aranzadi, Navarra, 2020, pág.147.

[256] GARRIGA DOMÍNGUEZ, A: "La elaboración de perfiles y su impacto en los derechos fundamentales. Una primera aproximación a su regulación en el reglamento general de protección de datos de la Unión Europea". op.cit., pág.137.

[257] Entre otros, En: HILDEBRANDT, M: "What is profiling? Defining Profiling: A New Type of Knowledge?" En: HILDEBRANDT, M & GUTWIRTH,S (eds): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008, p-21. y GONZÁLEZ, E., DE HERT, P :"Understanding the legal provisions that allow processing and profiling of personal data—an analysis of GDPR provisions and principles". *ERA Forum* 19**,** 597–621, 2019-. Págs. 610 y ss.

presentan. A su vez, en los *no distributivos*, no todos los miembros individuales del grupo presentan todas las características En este segundo supuesto, la personas que se le asigna en dicho perfil está "condenada" a recibir el mismo tratamiento que se haya previsto para el resto de personas que forman parte del grupo, y ello, a pesar de que no presenten todas las características similares.

C) Ventajas y riesgos de la elaboración de perfiles

Señalado lo anterior, cabe destacar que la elaboración de perfiles no es un fenómeno novedoso. Es más, la atribución de ciertas consecuencias a ciertos grupos que presentan ciertas características más o menos definidas lleva realizándose mucho tiempo. Las personas, basadas en la experiencia, prejuicios o conocimientos propios agrupamos y perfilamos a otras personas en grupos y además les atribuimos una serie de inferencias o consecuencias por pertenecer a ese grupo. Dicha categorización además lleva realizándose también por parte de las organizaciones mucho tiempo, las cuales, intentan dar sentido a la compleja realidad en la que estas operan, clasificando las interacciones presentes entre la organización y sus usuarios[258]. En este sentido, la categorización no hay que entenderla como algo negativo, sino más bien positiva cuando se basa en una sólida estadística[259].

Sin embargo, ahora, y fruto de los avances en las técnicas de *big data* y los sistemas de inteligencia artificial capaces de procesar grandes masas de datos, las inferencias y variables utilizadas permiten crear todo tipo de perfiles muchos más

[258] En: HILDEBRANDT, M: "What is profiling? Defining Profiling: A New Type of Knowledge?" En: HILDEBRANDT, M & GUTWIRTH,S (eds): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008, op.cit., pág.22.

[259] Hildebrandt, apoyándose en Schauer, viene a indicar que las personas y las organizaciones aplican de forma rutinaria la generalización, esta generalización es, sobre todo, algo bueno, especialmente si se basa en una "sólida base estadística", En opinión de Schauer , las evaluaciones rutinarias basadas en la generalización no sólo son necesarias para hacer frente a la complejidad y la multiplicidad, sino que también proporcionan decisiones *justas* en lugar de *arbitrarias*, debido a la apelación a una norma general, que crea un tipo de previsibilidad (esencial, por ejemplo, para la seguridad jurídica). De esta manera, según Hildebrandt, -inconscientemente,- agrupamos diferentes acontecimientos, cosas o personas en categorías para evaluar lo que cabe esperar y poder decidir cómo actuar. Los estereotipos así permiten la anticipación. En: HILDEBRANDT, M: "What is profiling? Defining Profiling: A New Type of Knowledge?" En: HILDEBRANDT, M & GUTWIRTH,S (eds): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008, op.cit., pág.,24. Hildebrandt cita SCHAUER, F.: *Profiles Probabilities and Stereotypes*, Belknap Press of Harvard University Press, Ed, Cambridge, Mass. London, 2003.

precisos y a la vez más incisivos en la vida de las particulares[260]. Ello ha llevado consecuentemente a un mayor uso de los mismos y un aumento de los riesgos que estos pueden generar[261].

Precisamente y por lo que se refiere a su uso, la elaboración de perfiles encuentra su espacio en una gran cantidad de organizaciones dadas las ventajas que puede ofrecer a las mismas, entre otras: permite una mejor asignación de los recursos, se pueden segmentar adecuadamente los mercados, se desarrollan productos que se adaptan a la oferta y demanda de los usuarios, etc.[262]Así, por ejemplo, una entidad bancaria a través de una aplicación adecuada de *scoring* puede reducir la probabilidad de impago y los riesgos de morosidad de los productos que ofrece a sus potenciales clientes. Consecuencia de ello, se obtiene un mayor beneficio por parte de la entidad bancaria y además, se evita la concesión de créditos de dudoso cobro, lo que redunde en un sistema bancario saneado[263].

Pues bien, pese a las ventajas indicadas y la justificación del uso de la elaboración de perfiles, no podemos pasar por altos algunos de los principales inconvenientes que se han señalado por parte de la doctrina[264].

Así, en el caso de los perfiles grupales no distributivos, dado que estos tienden a prever las mismas consecuencias o derivar las mismas inferencias a grupos de personas que reúnen ciertas características comunes, en muchos supuestos, los miembros

---

[260] ORTEGA GIMÉNEZ, A: *Las aplicaciones del Big Data en el ámbito asegurador y el tratamiento legal de sus datos una perspectiva desde el derecho internacional privado*. Ed. Fundación MAPFRE ,2019 pág.37.

[261] MANTELERO,A: "Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection", *Computer Law & Security Review*, Volume 32, Issue 2, 2016, pág.239.

[262] Consejo de Europa. *Recomendación CM/Rec(2010)13 del Comité de Ministros a los Estados miembros sobre la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal en el contexto de la creación de perfiles*, pág.2. Disponible en: https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=09000016805cdd2a

[263] GILLIS, TALIA B; SPIESS, JANN L: "Big Data and Discrimination," *University of Chicago Law Review*:Vol.86, iss.2.Article 4, 2019, pág.459.

[264] Hace ya algún tiempo Murillo de la Cueva hizo hincapié en señalar los riesgos que presentaban la elaboración de perfiles. Así, este venía a decir que: el *peligro potencial y real al que nos enfrentamos radica, por una parte, en el volumen de información, a menudo aparentemente irrelevante, que sobre nosotros se maneja. Por la otra, en la posibilidad cierta de obtener —mediante el tratamiento de esos datos— nuevos elementos informativos que nos afectan….., tales procedimientos permiten lograr el conocimiento de aspectos de nuestra vida que, además de personales, merecen ya el calificativo de íntimos. Por último, existe el riesgo de que, a partir de ese cúmulo informativo, se elaboren o construyan perfiles de nuestra personalidad en función de los cuales se tomen decisiones sobre nuestros derechos y expectativas, por ejemplo, a la hora de conseguir una vivienda en alquiler, obtener un crédito bancario o una simple tarjeta de crédito o aspirar a un puesto de trabajo*. En: MURILLO DE LA CUEVA,P,L: "La construcción del derecho a la autodeterminación informativa". *Revista de Estudios Políticos* (Nueva Época) Núm. 104. Abril-Junio1999, pág.38. Texto disponible en: http://www.cepc.gob.es/publicaciones/revistas/revistaselectronicas?IDR=3&IDN=261&IDA=17224

incorporados a estos grupos pueden ser tratados injustamente pese a que no presentan todos los atributos presentes en el grupo[265]. Por ejemplo, una persona puede catalogarse como mala pagadora por adquirir productos en determinados establecimientos que la entidad bancaria, o mejor dicho, el algoritmo, ha correlacionado con personas que son malas pagadoras. De esta manera, no es necesario que la persona que compra en la tienda sea culpable de tener un mal crédito, sino que, al adquirir productos en dichas tiendas, se infiere que esa persona es mala pagadora y ello independientemente de que esta persona tenga un historial crediticio impoluto. No podemos olvidar que a través de la elaboración de perfiles se trata de generalizar y agrupar a distintas personas en grupos. Sin embargo, cada persona presenta variables específicas propias que las pueden diferenciar del resto de miembros que forman parte del grupo al que se incorpora. Por tanto, la generalización que pretenden estos algoritmos difícilmente será completa y podrá envolver todos los factores individuales que están presentes en esas personas[266]. Consecuencia de ello, siempre quedará un recoveco para considerar que las inferencias derivadas de estos sistemas arrojan resultados inadecuados.

Por otro lado, e independientemente del tipo de perfil que se elabore, tenemos que partir de que las correlaciones extraídas por estos algoritmos no dejan de ser eso, es decir, una serie de patrones que en la mayoría de los casos no responden a causas sino a determinadas casualidades en las cuales se representan una probabilidad de que las cosas salgan igual en el futuro. Además, normalmente, estas correlaciones no son contrastadas careciendo por tanto de rigor científico. Por tanto, a pesar de que la persona que se incorpore al perfil presente las mismas características del grupo, aún seguirán existiendo riesgos de inferencias inadecuadas por un incorrecto diseño del sistema a la hora de presentar las correlaciones. En este sentido, a través del perfilado se crean grupos o categorías completamente nuevos que se agrupan según características más o menos aleatorias y que normalmente[267], los miembros que forman parte de ese grupo desconocen su integración en él y las consecuencias que ello comporta[268].

---

[265] GONZÁLEZ, E., DE HERT, P :"Understanding the legal provisions that allow processing and profiling of personal data—an analysis of GDPR provisions and principles", op.cit., pág.610.

[266] MORENTE,PARRA,V: "Big data o el arte de analizar datos masivos. una reflexión crítica desde los derechos fundamentales", op.cit., págs..251 y ss.

[267] Comité de ética alemán. Gutachten der Datenethikkommission, 2019, pág.168. Disponible en: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.html

[268]Informe del Relator Especial sobre la extrema pobreza y los derechos humanos, Philip Alston, presentado de conformidad con la resolución 35/19 del Consejo de Derechos Humanos. 11 de octubre de 2019, p-9 y ss. Disponible en: http://undocs.org/es/A/74/493

3. Las decisiones plenamente automatizadas y la elaboración de perfiles en el artículo 22: Los límites de su aplicación

Hemos analizado los dos tratamientos de datos que explícitamente el RGPD hace alusión y que tienen especial relevancia con el uso de sistemas algorítmicos en la toma de decisiones. Es turno de valorar cuándo estos tratamientos quedan englobados en el ámbito de aplicación del artículo 22. Así, podemos distinguir distintas situaciones:

En primer lugar, es posible que tanto la elaboración de perfiles como la toma de decisiones automatizada se realicen de forma separada y sin ningún nexo de unión entre ambos tratamientos. Por lo que respecta a la elaboración de perfiles, un *data bróker* podría recoger información de distintas fuentes públicas y privadas para elaborar perfiles sobre personas[269]. Esos perfiles de personas podría venderlos a terceros que posteriormente podrán hacer uso de los mismos para distintas finalidades[270]. Por otro lado, en relación con las decisiones automatizadas, por ejemplo, una determinada administración pública podría incorporar un sistema de reconocimiento de matrículas automático por el cual, al detectar que dicha matrícula coincide con los registros de datos de bases de datos de matrículas sospechosa el sistema arroja una alerta y automáticamente la policía lleva a cabo la posible búsqueda del vehículo[271]. En este caso, se introduce el dato, el sistema lo procesa y, cuando detecta un *match* o coincidencia, se adopta una decisión que afecta al particular. Sin embargo, no existe un perfilado previo sobre dicha persona sobre la que se adopta la decisión arrojada por el sistema.

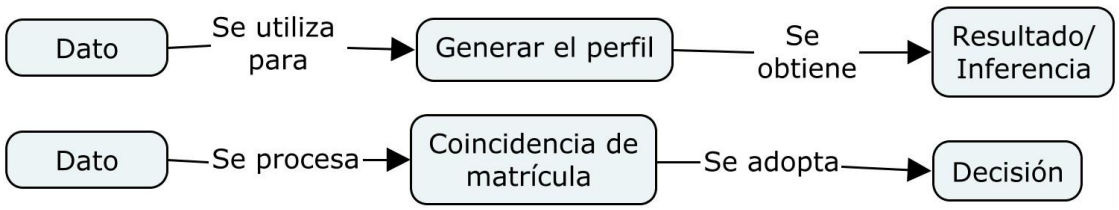
---

[269] En Estados Unidos muchas empresas se dedican a recoger información y realizar perfiles sobre consumidores. Posteriormente, los resultados arrojados por esos perfiles son enviado a terceras organizaciones, las cuales, utilizan esos resultados de la forma que consideran más adecuada. Empresas como Sift, Zeta Global, Retail Equation, Riskfield o Kustomer son ejemplo de ello. Visto en: https://www.nytimes.com/2019/11/04/business/secret-consumer-score-access.html

[270] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679.* Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág. 8.

[271] El uso de sistemas de reconocimiento automático de matrículas está ampliamente extendido en distintos países. Así, en el Reino Unido estos sistemas se utilizan por parte de la policía para prevenir delitos. Visto en: https://www.police.uk/pu/advice-crime-prevention/automatic-number-plate-recognition-anpr/

En Alemania también se implantó otro sistema parecido. Sin embargo, el Tribunal Constitucional Alemán declaró que un reconocimiento masivo automático de matrículas era contrario a la Constitución por no ser una medida proporcional. *Bundesverfas-sungsgericht (1 BvR 142/15), 18 de diciembre* 2018. Disponible en:https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/2018/12/rs20181218_1bvr014215en.html



Elaboración de perfiles y decisión automatizada separados.

Por un lado, la mera elaboración de perfiles, aunque pueda generar efectos relevantes o significativos en el particular, quedará fuera del ámbito de aplicación del precepto si el resultado inferido del perfil no deriva en una decisión plenamente automatizada. En este sentido, es importante reseñar que en los casos en los que el resultado del perfilado sea la única base sobre la cual un responsable adopta una decisión, el hecho de que haya transcurrido un determinado tiempo entre la conversión del resultado y la decisión automatizada no debe ser óbice para entender inaplicado el artículo 22. Así, en la sentencia holandesa que analizaba el sistema SyRI cuya finalidad era detectar potenciales defraudadores del sistema de seguridad social. El tribunal de la Haya consideró que los informes de riesgo que se derivaban de este algoritmo y que en muchos casos permanecían hasta 2 años en manos de los responsables de este tratamiento generaban efectos relevantes por el mero hecho de apuntar dicho riesgo[272]. El carácter relevante estaba presente ya que las autoridades podían iniciar actuaciones inspectoras en cualquier momento[273]. Aunque este tribunal no se pronunció expresamente si dicho tratamiento le resultaba aplicable el artículo 22, en nuestra opinión, tal tratamiento sí que se encuadraría dentro del ámbito del mencionado precepto. De esta manera, si una organización realiza un perfil de riesgo de una persona y lo mantiene durante un periodo de tiempo archivado, si ese perfil de riesgo posteriormente es utilizado por dicha organización para iniciar una investigación sin

[272] Un análisis de esta resolución judicial puede encontrarse en: COTINO HUESO,L: "«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, Nº. 4 (Abril-junio 2020),pág.2.. Véase sobre este mismo autor, COTINO HUESO,L: *Hacia la transparencia 4.0: el uso de la inteligencia artificial y big data para la lucha contra el fraude y la corrupción y las (muchas) exigencias constitucionales*. En RAMIÓ,C (coord.): "Repensando la Administración Pública. Administración digital e innovación pública", op.cit,. Págs.169 a 196.

[273] El Tribunal de la Haya consideró que un informe de riesgo puede tener un impacto significativo en la esfera de los particulares basándose en la definición que establece el Grupo del artículo 29, concretamente, *estos informes de riesgo pueden almacenarse durante dos años y los participantes en el proyecto SyRI pueden utilizarlo durante un máximo de 20 meses. Además, el Ministerio Público y la policía pueden ser notificados del informe de riesgo cuando lo soliciten. El hecho de que un informe de riesgo no siempre dé lugar a una investigación más a fondo, o a una sanción administrativa o penal o tampoco pueda utilizarse como la única base para una decisión de ejecución no altera el efecto significativo en la vida privada del interesado*. Véase el apartado 6.59 de la mentada sentencia. Disponible en: https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878

tener en cuenta otros elementos[274], hay que entender que se aplica el artículo 22, y ello, independientemente del periodo que haya mediado entre el momento en el que se generó el resultado inferido del perfil y la toma de la decisión posterior. Y es que no podemos olvidar que el artículo 22 trata de establecer un plus de garantías para aquellos procesos decisorio donde el resultado que vierte el algoritmo sea la única base sobre la cual se adopta la decisión que afecta a los particulares. El elemento clave en estos casos para considerar aplicado el artículo 22 será valorar el uso que se haga del resultado de ese perfilado, es decir, que el mismo se automatice y genere efectos relevantes en el particular.

Por otro lado, sí que se encuadrarán en el ámbito de aplicación del artículo 22 todas aquellas decisiones que sean plenamente automatizadas y generen efectos relevantes[275], independientemente de que haya habido o no una previa elaboración de perfiles sobre la que se base dicha decisión automatizada[276]. Esta conclusión podría estar abierta a debate ya que en el origen histórico -allá por 1978- de este precepto se puso el acento esencialmente en los riesgos que se podían derivar de la elaboración de perfiles plenamente automatizada[277]. Así, las primeras propuestas del artículo 15 de la Directiva de protección de datos –precepto que fue sustituido por el actual artículo 22 del RGPD- centraban su atención en las decisiones que tuvieran como objetivo la definición de un perfil o la personalidad del particular afectado por dichas decisiones[278].

---

[274] Aquí es donde el Tribunal de la Haya no se pronuncia, podría llegar a interpretarse que el artículo 22 también se aplica a perfiles significativos aunque no sean totalmente automatizados. Sin embargo, como hemos indicado, creemos que ello no es posible dada la literalidad del precepto analizado.

[275] The Norwegian Data Protection Authority. *Artificial intelligence and privacy.* 2018, págs. 19 y 20

[276] MURGA FERNÁNDEZ,J,P: "Derechos de los individuos". En: MURGA FERNÁNDEZ, J,P; FERNÁNDEZ SCAGLIUSI, M,A ; ESPEJO LERDO DE TEJADA,M: (dirs.): *Protección de datos, responsabilidad activa y técnicas de garantía*. *Curso de delegado de protección de datos.* Ed.Reus, Madrid, 2018, pág.103. En el mismo sentido véase: ROIG I BATALLA,A: *Las garantías frente a las decisiones automatizadas del Reglamento general de Protección de Datos a la gobernanza algorítmica*, op.,cit,, págs.34 y 36.

[277] El artículo 22, antiguo artículo 15 de la Directiva 95/46 de Protección de datos, tiene su origen en la Ley Francesa n ° 78-17 del 6 de enero de 1978 relativa al procesamiento de datos, archivos y libertades (Versión vigente el 23 de julio de 1978). En esta norma se palpa como el legislador estaba pensando esencialmente en las decisiones que se basan en la elaboración de perfiles. (Artículo 2). Así, este precepto indicaba, *Ninguna decisión judicial que implique una evaluación del comportamiento humano puede basarse en el procesamiento automatizado de información que dé una definición del perfil o la personalidad de la persona en cuestión. Ninguna decisión administrativa o privada que implique una evaluación del comportamiento humano puede tener como única base un tratamiento automatizado de la información que dé una definición del perfil o personalidad del interesado*. Disponible en: https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000886460/1978-07-23/

[278] Las primeras propuestas legislativas de la Directiva 95/46 a la hora de analizar este fenómeno continuaron con la senda marcada por la ley francesa Así, esta propuesta de precepto establecía que los interesados tenían derecho *a no tener que someterse a ninguna decisión administrativa o privada que implique una apreciación de su comportamiento fundada únicamente en un tratamiento automatizado de datos personales que dé una definición de su perfil o de su personalidad*; Véase: Comunicación de la Comisión sobre la protección de las personas en lo referente al tratamiento de datos personales en la

No obstante, las sucesivas reformas y propuestas del artículo 22 han mostrado una intención por ampliar al ámbito de aplicación del mencionado precepto a las decisiones plenamente automatizadas aunque no exista como base dicha elaboración de perfiles[279]. En este sentido, el propio artículo 22 del RGPD hace mención a la elaboración de perfiles como una base más de tratamiento que puede generar decisiones automatizadas, pero no la única[280]. Por otro lado, descartar aquellas decisiones que se deriven de sistemas algorítmicos inteligentes que no se basen en elaboración de perfiles limitaría la esfera de protección de los interesados que se ven sometidos a estas decisiones en un ámbito donde la variedad de las tecnologías que están presentes pueden permitir la adopción de decisiones automatizadas relevantes en las que no hay un perfilado previo de la persona.

En *segundo lugar* puede suceder que la elaboración de perfiles realizada se utilice como base para la adopción de la decisión posterior. Estos supuestos coincidirían

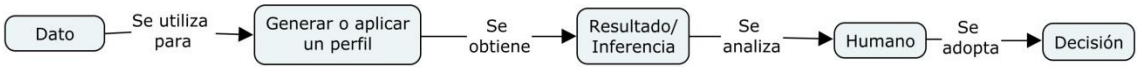
---

Comunidad y a la seguridad de los sistemas de información. COM (90) 314 final - SYN 288 Bruselas, 24 de septiembre de 1990**.** Pág.48. Disponible en:
https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:51990DC0314&from=EN
Más expresivo fue posteriormente el Consejo en algunos de sus informes sobre la Directiva al señalar que las decisiones del mentado precepto *sólo se referían a aquellos tratamientos que atribuyen a los datos relativos al interesado variables que permitan determinar un perfil de personalidad tipo (considerado bueno o malo), lo que excluye todos los casos en que el sistema no define el perfil de personalidad: por ejemplo, el hecho de que una persona no pueda obtener el importe solicitado en una cajero automático, porque ya haya superado su crédito, no se incluye en esta definición*. Véase. Propuesta modificada de DIRECTIVA DEL CONSEJO relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación**.** COM(92) 422 final- SYN 287 Bruselas, 15 de octubre de 1992. Págs. 27 y 28.
https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:51992PC0422&from=EN

[279] Por ejemplo, el Consejo, durante la tramitación del RGPD indicaba que las decisiones automatizadas a las que se refiere el artículo 22 puede incluir la elaboración de perfiles, es decir, es o no posible dicha inclusión. Véase: Posición del Consejo en primera lectura con vistas a la adopción de un Reglamento del Parlamento Europeo y del Consejo 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). 5419/16 ADD 1 REV 1. 31 de marzo de 2016. (pág.18). Disponible en:
https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CONSIL:ST_5419_2016_ADD_1_REV_1&from=ES
Más expresivamente, el propio Grupo del artículo 29 indicó que el artículo 22 se aplica a las *decisiones individuales automatizadas (incluyan o no la elaboración de perfiles)*. En: "Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679". Grupo del Artículo 29. Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.6

[280] El término "incluida" al que hace referencia el artículo 22 pensamos que hay que interpretarlo como un elemento más que puede o no estar presente en este tipo de decisiones. Es decir, puede estar o no incluida la elaboración de perfiles en la formación de la decisión automatizada. Así también lo ha considerado la ICO al indicar que :"*The UK GDPR restricts you from making solely automated decisions, including those based on profiling, that have a legal or similarly significant effect on individuals*". Disponible en:
https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/rights-related-to-automated-decision-making-including-profiling/

en su mayor parte con las decisiones parcialmente automatizadas. Es decir, el sistema en este caso aplicaría un perfil determinado y arrojaría un resultado o inferencia, dicho resultado derivado del perfilado se utilizaría por el responsable del tratamiento para adoptar finalmente una decisión u otra. Por ejemplo, la policía nacional, utiliza un sistema cuyo objetivo principal es detectar denuncias falsas presentadas por particulares que alegan haber sufrido determinados ilícitos penales. Pues bien, el sistema analiza dicha denuncia y pronostica una determinada probabilidad de que dicha denuncia sea falsa[281]. Ese resultado es finalmente analizado por un policía, el cual, valora si procede al archivo de la mentada denuncia o se continúa con las diligencias policiales correspondientes para indagar el posible hecho denunciado. En este supuesto, la decisión parcialmente automatizada se basa en gran parte en la elaboración del perfil que realiza el sistema sobre esta persona, en este caso, evaluar la probabilidad de que haya podido mentir a la hora de presentar la denuncia, si bien, no es el único elemento a tener en cuenta. En estos supuestos, a pesar de que existe un perfilado y la decisión puede generar efectos relevantes, dado que todo el proceso no es plenamente automatizado, el artículo 22 no resultaría aplicable.



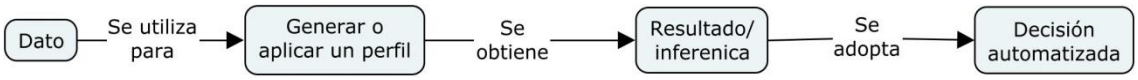
Elaboración de perfiles junto a decisión parcialmente automatizada

En *tercer lugar*, la elaboración del perfil puede utilizarse como base para la adopción de una decisión plenamente automatizada. En estos supuestos, el resultado derivado de la elaboración de perfiles se convierte automáticamente en una decisión que afecta al particular sin que medie intervención humana entre el resultado que se deriva de aplicar el perfil y la decisión que afecta directamente al particular sobre el que se aplica dicho perfil. En estos supuestos estaríamos ante elaboración de perfiles totalmente automatizados, sobre los cuales, el artículo 22 resulta plenamente aplicable[282]. El legislador europeo por tanto, menciona estos dos tratamientos de forma expresa y combinada con el objetivo de hacer ver la importancia de los mismos cuando

[281] Sistema Veripol. Más información en: http://www.interior.gob.es/prensa/noticias/-/asset_publisher/GHU8Ap6ztgsg/content/id/9496864

[282] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, 2020, pág.10.

se conjugan, pero haciendo hincapié en que el elemento clave de este precepto deriva de la plena automatización del resultado del algoritmo.



Elaboración de perfiles junto a decisión totalmente automatizada

En definitiva, como puede comprobarse, la incorporación de estos tratamientos dentro del ámbito de aplicación del artículo 22 dependerá de factores como la finalidad que se pretenda con los mismos, el papel que se pretende con el sistema, la intervención o no humana en el proceso decisorio y la relevancia o no de dichas decisiones en la esfera de los particulares.

Tratamientos incluidos en el artículo 22	**Tratamientos no incluidos en el artículo 22**
Elaboración de perfiles+ Decisiones relevantes totalmente automatizadas Decisiones relevantes totalmente automatizadas	Elaboración de perfiles sean o no relevantes Decisiones parcialmente automatizadas sean o no relevantes
APLICA EL RGPD	

II. TIPOS DE DATOS

Aunque la normativa de protección de datos se centra en los datos personales, lo cierto es que existen otra serie de datos que, pudiendo ser o no datos personales, requieren de nuestra atención en este momento de la tesis. Ello es así porque muchas de las variantes de datos a las que haremos referencia tienen una especial relevancia en el contexto del uso de los algoritmos para la toma de decisiones automatizadas. Por tanto, se hace necesario realizar una aproximación a los mismos ya que en las páginas posteriores se aludirán a ellos.

1. Dato personal

El concepto de dato personal resulta basilar en la normativa de protección de datos personales. Aunque no es el objeto de esta tesis estudiar esta definición. Sí que vamos a hacer referencia a aquellos elementos más relevantes de la misma que presentan una especial incidencia cuando se llevan a cabo tratamientos de datos a través del uso de sistemas para la toma de decisiones automatizadas y la elaboración de perfiles.

Así, el apartado primero del artículo 4 del RGPD define los datos personales como:

> *toda información sobre una persona física identificada o identificable («el interesado»); se considerará 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;*

Esta definición se ha mantenido más o menos uniforme desde las primeras legislaciones europeas que regularon esta materia[283]. Este concepto, tal y como ha indicado el TJUE[284], ha de recibir una interpretación amplia que le permita abarcar multitud de información que de alguna manera pueda vincular a una persona. Así, para el GT29, el concepto de dato personal incluye todo tipo de información. Esto es, tanto aquella que sea objetiva como puede ser la edad de una persona, como aquella que sea subjetiva, es decir, las evaluaciones realizadas sobre una persona o las inferencias realizadas por un sistema algorítmico sobre una persona. En ambos casos, resulta

[283] Véase el artículo 2 del CONVENIO 108 del Consejo de Europa para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, hecho en Estrasburgo el 28 de enero de 1981. La Directiva 95/46 añadió que se entendía por persona identificable (en 1990, Propuesta de Directiva del Consejo relativa a la protección de las personas en lo referente al tratamiento de datos personales COM(90) 314 final — SYN 288, artículo 2.1. a)
https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:51990PC0314(01)&from=EN
Finalmente, el RGPD aclara aún más el concepto fijando además ejemplos para clarificar que se considera identificable ya sea de forma directa o indirecta. (Considerando 26 y artículo 4 RGPD)

[284] Por ejemplo, el TJUE ha llegado a considerar como dato personal la IP dinámica de un usuario. En: Tribunal de Justicia de la Unión Europea de 19 de octubre de 2016. Asunto C-582/14, caso Patrick Breyer**.** ECLI:EU:C:2016:779. FJº 49. Resolución disponible en:
https://eur-lex.europa.eu/legal-content/es/TXT/?uri=CELEX:62014CJ0582

irrelevante que dicha información sea o no verídica[285]. Ejemplo: Manuel tiene 40 años (información objetiva) o Manuel sufrirá una enfermedad grave (información subjetiva).

Por otro lado, y desde el punto de vista del formato sobre el que la información está contenida, el concepto de datos personales abarca la información disponible en cualquier forma, ya sea alfabética, numérica, gráfica, fotográfica o sonora[286]. Trasladado al objeto de nuestro estudio, uno de los campos de la inteligencia artificial que más expansión está teniendo y que más implicaciones jurídicas puede desempeñar a la hora de la toma de decisiones automatizadas está presente en los sub campos del lenguaje natural o reconocimiento de textos. Así, por ejemplo, el cuerpo de la denuncia que se analiza como inputs en un sistema de inteligencia artificial que se utiliza con el fin de evaluar si una denuncia es o no falsa entraría dentro del ámbito de la normativa de datos cuando dicho cuerpo de textos, conjunto de datos no estructurados, sea referido a esa persona identificable[287].

A su vez, en muchos casos, la elaboración de perfiles y la toma de decisiones automatizadas se realizan en base a datos que se obtienen de un particular sin llegar a identificarlo como tal. Es decir, se sabe que hay un particular detrás de esos datos pero no se logra asignar o vincularle a un nombre o apellidos. Pues bien, como ha señalado el GT29 , la consideración de datos personal no requiere la necesidad de identificar a una persona por sus nombres y apellidos, en determinadas ocasiones, bastará con que el identificador o identificadores que se utilicen consigan singularizar a alguien[288]. Ello ocurre por ejemplo en el uso de las cookies personalizadas para mostrar anuncios o las huellas digitales que se derivan de los dispositivos electrónicos que utilizamos cuando accedemos a la red[289]. En estos casos, los datos que se van recopilando permiten en un

[285].Grupo del Artículo 29. *Dictamen 4/2007 sobre el concepto de datos personales.* Adoptado el 20 de junio de 2007, pág.4

[286] Vid, pág.8.

[287] Como ya se ha comentado, la policía nacional de España puso en marcha VeriPol. Esta aplicación detecta las denuncias falsas interpuestas en casos de robos con violencia e intimidación. Dicha herramienta analiza el texto de la denuncia e indica si dicha denuncia es o no falsa. Visto en: http://www.interior.gob.es/ca/prensa/noticias/-/asset_publisher/GHU8Ap6ztgsg/content/id/9496864

[288] "Dictamen 4/2007 sobre el concepto de datos personales". Grupo del Artículo 29. Adoptado el 20 de junio de 2007, pág.15. Así, la APED ha considerado dato personal la dirección mac de un ordenador o móvil, ya que se trata de un identificador único que cada fabricante le asigna a la tarjeta de red de estos dispositivos conectados. En: AEPD, *informe N/REF: 0017/2019*, pág.6.

[289] El concepto de huella digital del dispositivo o *fingerprints* ha sido definido por la AEPD como la recopilación sistemática de información sobre un determinado dispositivo remoto con el objetivo de identificarlo, singularizarlo y, de esa forma, poder hacer un seguimiento de la actividad del usuario del mismo con el propósito de perfilarlo. Dado que lo habitual es que las personas no compartan sus equipos, ya sea este un teléfono móvil, tableta, portátil u ordenador de trabajo, individualizar el terminal supone

momento óptimo de recopilación llegar a identificar de forma inequívoca y única a un determinado dispositivo, el cual, tras el mismo, estará una persona.

2. Los datos de categoría especial

El considerando 51 del RGPD establece que determinados datos personales, por su naturaleza, son particularmente sensibles ya que los mismos en el contexto de su tratamiento podrían entrañar importantes riesgos para los derechos y las libertades fundamentales de las personas. Es por ello que el RGPD le otorga una protección específica a los mismos imponiendo a aquellos que pretendan tratarlos unas exigencias mayores. Así, de acuerdo al artículo 9.1 del RGPD, son considerados datos de categoría especial aquellos datos personales que *revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o las orientación sexuales de una persona física*.

Como puede apreciarse, las categorías de datos que son considerados especiales aglutinan dos grupos de información sensible; por un lado aparecen referenciados gran parte de los rasgos o características consideradas como categorías sospechosas de discriminación[290]. En este sentido, la doctrina ha señalado que a pesar de que no se incluya la variable sexo o género como un dato de categoría especial en el artículo 9 de forma explícita, la misma quedaría integrada dentro de los datos referidos a la vida sexual de las personas o los datos genéticos[291]. Por otro lado, también se incluyen otra serie de información que al estar estrechamente relacionados con el individuo, el tratamiento de la misma puede generar importantes riesgos para dichas personas.

---

individualizar a la persona que lo utiliza y por tanto, se puede perfilar individualmente. Véase: "Estudio Fingerprinting o Huella digital del dispositivo". Agencia Española de Protección de Datos. Adoptado en Febrero 2019. Disponible en: https://www.aepd.es/sites/default/files/2019-09/estudio-fingerprinting-huella-digital.pdf

[290] Estas son, los datos que *revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, datos relativos a la vida sexual o las orientación sexuales de una persona física.* Artículo 9 RGPD. Tanto el Convenio Europeo de Derecho Humanos (artículo 14), la Carta de Derechos Fundamentales de la Unión Europa (artículo 21) como la Constitución Española. (artículo 14) prohíben la discriminación por estos motivos.
Sobre la discriminación y su relación con la protección de datos véase el Capítulo IV, apartado VII, punto 2 de esta tesis.

[291] SORIANO ARNANZ, A: *Data protection for the prevention of algorithmic discrimination.* Ed. Thomson Reuters Aranzadi. Pamplona. 2021, pág.130.

Tratamiento de categorías especiales de datos personales	
Rasgos o características protegidas por el principio de no discriminación	Otros datos de categoría especial estrechamente relacionados con el interesado
-Origen étnico o racial. -Opiniones políticas y afiliación sindical. -Convicciones religiosas o filosóficas. -Datos relativos a la vida sexual o la orientación sexual. (sexo/género)	-Datos relativos a la salud[292]. -Datos biométricos destinados a la identificación. -Datos genéticos.

Ahora bien, una lectura global del artículo 9 del RGPD lleva a la conclusión que bajo este precepto no sólo se contemplan los datos que son estrictamente de categoría especial sino también aquellos datos que "revelen" o sean "relativos" a esos datos de categoría especial. Esta apreciación resulta muy relevante en el contexto del uso de sistemas automatizados. Así, puede darse la posibilidad de que los datos iniciales que se incorporen al sistema no sean considerados especiales, tal y como define el artículo 9 del RGPD, pero, tras el procesamiento algorítmico de los mismos, dichos datos inferidos sí que se conviertan o se refieran a datos especialmente sensibles. Ante esta situación, la pregunta que habría que hacerse es cómo han de considerarse dichos datos iniciales, esto es, de categoría especial o no. La pregunta no es baladí ya que el régimen previsto por el RGPD es más protector para los interesados cuando los responsables tratan datos personales especialmente sensibles respecto del resto de datos personales convencionales[293].

[292] De acuerdo al considerando 35 del RGPD: *entre los datos personales relativos a la salud se deben incluir todos los datos relativos al estado de salud del interesado que dan información sobre su estado de salud física o mental pasado, presente o futuro* (….) *todo número, símbolo o dato asignado a una persona física que la identifique de manera unívoca a efectos sanitarios; la información obtenida de pruebas o exámenes de una parte del cuerpo o de una sustancia corporal, incluida la procedente de datos genéticos y muestras biológicas, y cualquier información relativa, a título de ejemplo, a una enfermedad, una discapacidad, el riesgo de padecer enfermedades, el historial médico, el tratamiento clínico o el estado fisiológico o biomédico del interesado, independientemente de su fuente, por ejemplo un médico u otro profesional sanitario, un hospital, un dispositivo médico, o una prueba diagnóstica in vitro*.

[293] La normativa de protección de datos no es la única que prevé medidas más garantistas para el tratamiento de datos personales sensibles. Así, en el Código Penal, para que se aplique el tipo penal descrito en el artículo 197.2 de descubrimiento y revelación de secretos, se exige que se demuestre que dicho acceso se ha realizado en perjuicio de tercero. Pues bien, cuando se accede a datos sensibles en los términos descritos por el artículo 9 del RGPD, dicho perjuicio queda demostrado con el mero conocimiento de esos datos, sin que sea necesario probar un perjuicio añadido. Perjuicio añadido que sí se deberá demostrar cuando el acceso de esos datos no sean sensibles. STS (Sala de lo Penal) 1153/2021 - ECLI:ES:TS:2021:1153 de 22/03/2021, FJº 1º.

Especiales garantías cuando se tratan datos de categoría especial
-Prohibición general de tratar estos datos. Bases legales específicas. Artículo 9 -Elemento a tener en cuenta a la hora de valorar la compatibilidad entre el fin inicial y el posterior. Artículo 6.4.c) -Garantías específicas cuando se adopten decisiones automatizadas relevantes basadas en estos datos. Artículo 22.4. -Factor a tener en cuenta a la hora de implementar o no determinadas medidas de responsabilidad activa: designar un representante (art.27.2.a), llevar acabo el registro de actividades de tratamiento (art.30.5), evaluación de impacto (art.35.3.b), designar delegado de protección de datos (art.37.1.c)

Es turno de analizar qué datos quedan englobados en el artículo 9 del RGPD y por tanto son considerado de categoría especial en el contexto de los tratamientos presentes durante el ciclo de vida de los sistemas automatizados.

En *primer lugar*, resulta claro que cualquier dato que forme parte de alguna de las categorías indicadas por el artículo 9 se entenderá incorporada dentro del ámbito de aplicación del mismo. Así, si un responsable utiliza el historial clínico de un paciente para incorporarlo a un sistema automatizado que predice el riesgo que tiene ese individuo de sufrir una enfermedad, ese dato, dado que es relativo a la salud, el tratamiento del mismo se integra en este precepto. Es decir, desde el inicio del tratamiento el dato tratado es considerado de categoría especial.

En *segundo lugar*, del contenido del artículo 9 se desprende que su ámbito de aplicación no queda reducido a los datos que estrictamente pudieran considerarse más sensibles o especiales sino también a las finalidades sensibles que se pretendan con esos datos[294]. Así, por ejemplo, una aseguradora, con el fin de estimar el riesgo que tiene una persona de sufrir una determinada enfermedad y por tanto asignar una u otra prima de riesgo, utiliza un sistema automatizado que recopila datos como el lugar de residencia del interesado, la actividad deportiva que realiza a través de una App y las enfermedades actuales que padece. Pues bien, a priori, tanto el dato de la residencia como el dato de la actividad deportiva no son considerados datos relativos a la salud[295].

[294] El artículo 9 del RGPD hace alusión a los datos que revelen o sean relativos a esas categorías especiales. Además, algunas de las excepciones previstas a la prohibición general de este tipo de datos previstas en el artículo 9.2 del RGPD se refieren expresamente a finalidades específicas. Es decir, el tratamiento de esos datos de categoría especial sólo se autoriza si se pretenden esas finalidades.

[295] Por ejemplo, el dato personal de la residencia puede ser utilizado por ejemplo a la hora de conceder una ayuda pública. En este supuesto, dicho dato no puede considerarse especialmente protegido. A su vez, el dato de la actividad deportiva que realiza una persona como puede ser el número de pedaladas, velocidad en la bicicleta o kilómetros realizados por día pueden ser útiles para una empresa que pretenda contratar a un ciclista. Al no utilizarse datos de categoría especial ni pretender finalidades especiales o

Sin embargo, dado que la finalidad pretendida es precisamente evaluar la salud de esa persona, todos los datos que se recopilen para esa finalidad serán considerados para ese concreto tratamiento como datos especialmente sensibles o de categoría especial. Por tanto, el artículo 9 abarca tanto los datos personales que en sentido estricto son especiales como los datos personales convencionales que se utilicen para tratamientos que tengan como finalidad obtener información o datos especiales o sensibles. En el caso anterior, el dato estrictamente sensible sería el dato correspondiente a las enfermedades que esa persona padece y el resto de datos convencionales también se considerarían sensibles debido a la finalidad pretendida. En este sentido, el Tribunal Constitucional consideró que los datos personales que puede recopilar un partido político en las redes sociales con el fin de evaluar las orientaciones políticas de los interesados son datos especialmente sensibles o de categoría especial[296]. Así, la actividad que realiza esa persona en las plataformas digitales a través de por ejemplo; los "me gustas"[297], el contenido que comparte o las aficiones deportivas más relevantes se consideraran datos de categoría especial cuando la finalidad del tratamiento sea precisamente la de revelar u obtener la opinión política de esa persona. El GT29 ha indicado que, en el caso de que se infieran datos de categoría especial derivados de la elaboración de perfiles, hemos de entender que los datos iniciales que se recopilaron y se trataron para esos perfiles sean considerados especiales[298]. En este sentido, el Comité Europeo de Protección de Datos, en adelante CEPD, ha señalado que el tratamiento de

---

sensibles, dichos datos iniciales no pueden quedar englobados en el ámbito de aplicación del artículo 9 del RGPD.

[296] El TC declaró inconstitucional la Disposición final tercera de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales que modificaba la Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General. A través de esta modificación se habilitaba a los partidos políticos a recopilar información a través de redes sociales para inferir orientaciones políticas. Esta disposición se consideró inconstitucional entre otras razones por que el TC consideró que el tratamiento que se estaba legitimando a través de esta norma no había previsto las suficientes medidas de garantía. Véase la STC Sentencia 76/2019, de 22 de mayo de 2019. FJº 5 Y 8. Texto disponible en: https://hj.tribunalconstitucional.es/HJ/docs/BOE/BOE-A-2019-9548.pdf
Una aproximación al caso en cuestión puede verse en: ADSUARA VARELA, B: "El perfilado ideológico de los ciudadanos por los partidos políticos". *El Consultor de los Ayuntamientos*, Nº III, Sección Crónica, Julio 2019, págs. 77 y ss. También sobre la misma cuestión véase: ARENAS RAMIRO,M: "Partidos políticos, opiniones políticas e internet: la lesión del derecho a la protección de datos personales". *Teoría y Realidad Constitucional*, núm. 44, 2019, págs.341 a 372

[297] Un estudio ha establecido que a través de la aplicación "Me gusta" se pueden inferir atributos como la orientación sexual, el origen étnico o las opiniones políticas de los usuarios. KOSINSKI,M ;STILLWELL,D; THORE,G: "Private traits and attributes are predictable from digital records of human behavior", *PNAS*, 9/4/2013, Información disponible en: https://doi.org/10.1073/pnas.1218772110

[298] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*. Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.16. Así también lo ha indicado la Autoridad de protección de datos Noruega. The Norwegian Data Protection Authority. *Artificial intelligence and privacy.* 2018, op.,cit, págs.20 y 21.

una mera declaración, o de un solo dato de localización o similar que revele que un particular ha visitado (una o varias veces) un lugar típicamente visitado por personas con determinadas creencias religiosas no se considerará, por lo general, tratamiento de datos de categorías especiales. Sin embargo, puede llegar a considerarse tratamiento de categorías especiales de datos debido al contexto en el que se tratan o a los fines para los que se utilizan[299]. Así, para la ICO, el elemento esencial es que exista esa intención a la hora de utilizar ese dato con el fin de inferir un dato de categoría especial. De esta manera, y aunque a priori ese dato no sea formalmente especial, si se pretende inferir datos de categoría especial hay que considerarlos como tal desde el momento inicial. Por otro lado, cuando se infieran datos sensibles a partir de datos convencionales y no exista interés por parte del responsable para inferir tales datos, la ICO ha señalado que, para estos casos, a priori, no se puede considerar que se esté tratando con datos especialmente sensibles. Ahora bien, dado que puede existir ese riesgo, el responsable debe estar atento para valorar si en cualquier momento se pueden llegar a estar infiriendo datos especialmente protegidos[300]. Por ello, si el responsable obtiene esas inferencias de forma no intencionada este último no podrá utilizarlas para las finalidades previstas en el artículo 9, ya que si las utiliza, habrá que entender que está tratando ese tipo de datos.

En *tercer lugar,* y muy relacionado con el apartado anterior, también se ha de considerar dato de categoría especial aquel que, siendo inicialmente convencional, resulte ser un *proxy* de otro dato de categoría especial[301]. Así, según la AEPD, una variable *proxy* es aquella que se usa en lugar de la variable de interés cuando esa variable de interés no se pueda o no se quiera medir directamente[302]. Por ejemplo, una aseguradora podría utilizar como elemento a la hora de valorar el precio de una póliza

[299] Comité Europeo de Protección de Datos. *Guidelines 8/2020 on the targeting of social media users.* Versión 1.0. Directrices adoptadas el 2 de septiembre de 2020. Apartado 115, pág.30.
[300] Information Commissioner's Office, disponible en: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/
[301] En palabras del GT29, la expresión "datos *que revelen el* origen racial o étnico, las opiniones políticas, las convicciones religiosas o filosóficas, la pertenencia a un sindicato" debe entenderse en el sentido de que no sólo los datos que, por su naturaleza, contengan información sensible están cubiertos por esta disposición, sino también los datos de los que pueda deducirse información sensible con respecto a una persona. En: Grupo del Artículo 29. *Advice paper on special categories of data ("sensitive data"),* 2011, pág.6 *ab initio.* Texto disponible en: https://ec.europa.eu/justice/article-29/documentation/other-document/files/2011/2011_04_20_letter_artwp_mme_le_bail_directive_9546ec_annex1_en.pdf
[302] Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág39.

de seguros la marca del coche del potencial asegurado. A priori, el modelo del coche no es un dato de categoría especial. Sin embargo, es posible que ese dato sea un *proxy* del género o sexo de la persona. De esta manera, no sería descabellado pensar que las aseguradoras identifiquen estadísticamente que determinados modelos o tipos de coche suelen ser comprados por mujeres. Por tanto, ese dato aparentemente inocuo puede estar escondiendo una variable *proxy* considerada de categoría especial. Por ello, en este contexto, las variables *proxy* son aquellas que muestran una correlación lo suficientemente estrecha entre la variable/dato de categoría especial y la variable/dato elegida que se utiliza para la toma decisiones.

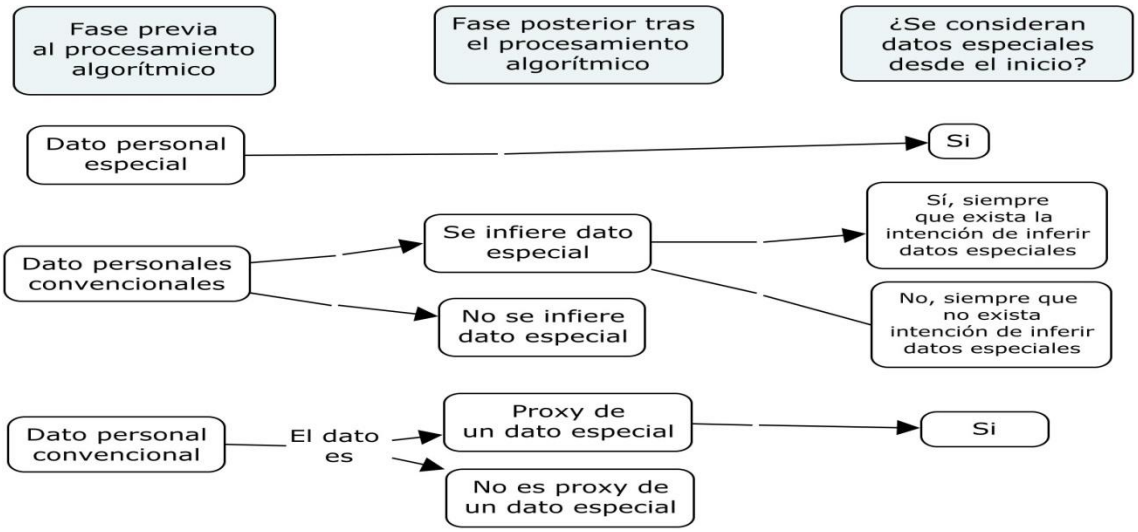
La tormenta perfecta a la que aludíamos en el primer capítulo de esta tesis, la cual permite la recopilación y el análisis masivo de todo tipo de datos personales, permite que las organizaciones traten todo tipo de datos aparentemente inocuos con finalidades que resultan especialmente protegidas. Esto ha llevado a algún sector de la doctrina a considera que los responsables, cuando pretendan utilizar sistemas de toma de decisiones automatizadas, deberán considerar todos estos datos como especiales ya que de una manera u otra estos siempre estarán unidos a otros datos que pueden revelar información sensible o de categoría especial[303]. En nuestra opinión, pese a que la irrupción de estas nuevas tecnologías permite el uso de todo tipo de datos convencionales y por tanto se produce una vis expansiva de los datos que pueden llegar a considerarse de categoría especial[304], siempre habrá que estar a la finalidad que se pretende con el tratamiento de esos datos o al posible *proxy* que se utilice para considerar que los mismos son datos de categoría de especial de acuerdo a lo indicado en el artículo 9 del RGPD. Como se puede apreciar, en muchos casos el problema ya no dependerá del carácter sensible del dato sino de la inferencia que se derive de su tratamiento ya que estas nuevas técnicas permiten extraer conocimiento íntimo a partir de datos insignificantes[305].

---

[303] L,JANSSEN,H:, "An approach for a fundamental rights impact assessment to automated decision-making", op.cit., pág.98.

[304] Así, por ejemplo, el concepto y ámbito de los datos relacionados con la salud debe englobar todos aquellos datos que pueden utilizarse con esa finalidad. COTINO HUESO,L: "El alcance e interactuación del régimen jurídico de los datos personales y *big data* relacionado con salud y la investigación biomédica". *Revista de derecho y genoma humano: genética, biotecnología y medicina avanzada*, Nº 52, 2020, págs. 70 y 71. En este mismo sentido, HILDEBRANDT, M: "Profiling and the Identity of the European Citizen" En: HILDEBRANDT, M & GUTWIRTH,S (eds): *Profiling the European Citizen. Cross-Disciplinary Perspectives*, op.cit., pág. 3014.

[305] HILDEBRANDT, M: "Profiling and the Identity of the European Citizen" En: HILDEBRANDT, M & GUTWIRTH,S (eds): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008, op.cit., pág. 320.



Elaboración propia

3. Dato alternativo

Cualquier información referente a una persona puede resultar ser un dato alternativo. Es decir, por dato alternativo hemos de considerar toda aquella información que comúnmente no ha sido utilizada por parte de una organización para un específico proceso de toma de decisiones pero ahora, y gracias al análisis masivo de datos, dicha información adquiere una fuerte relevancia. Por ejemplo, hasta no hace muchos años, una entidad bancaria siempre acudía a datos tradicionales como pudiera ser la nómina del sujeto a la hora de valorar la solvencia financiera del mismo. En los últimos años, gracias a las correlaciones que se pueden detectar en los modelos de aprendizaje automático, esta misma organización puede decidir dicho *scoring* basado en otro tipo de datos totalmente ajenos a la actividad financiera tradicional tal y como pueden ser el tipo de webs que visita, la actividad en las redes sociales o las calificaciones académicas[306]. Precisamente, y gracias a estas tecnologías, todo tipo de datos que hasta

[306] Entre los datos alternativos que actualmente se utilizan a la hora de evaluar la solvencia financiera de un particular encontramos: los pagos de alquiler, pagos de telefonía móvil, pagos de televisión por cable, actividades en redes sociales, nivel de estudios de la persona o su ocupación. Visto en: Experian. *The State of Alternative Credit Data. How the financial services industry is adopting and benefiting from these new data sources*. 2018, p-5. También se han considerado datos alternativos la información de las compras en línea de mercados como Amazon, los datos de envío de los servicios postales, los patrones de navegación, el tipo de teléfono o navegador utilizado. En: OCDE, "Artificial Intelligence in Society", 2019, pág.58.Disponible en: *https://doi.org/10.1787/eedfee77-en*. En el caso del sector crediticio universitario también se tienen en cuenta ahora la especialidad elegida por el solicitante, la institución universitaria, el historial laboral e incluso las actividades en las redes sociales. En: ODINET, CHRISTOPHER

la fecha no se tenían en cuenta por parte de las organizaciones para obtener información útil en los procesos decisorios adquieren ahora una relevancia mayor ya que la conjunción de estos datos puede llegar a crear modelos más precisos que aquellos basados en datos tradicionales. Es por ello que cada vez más organizaciones acudan a la obtención de este tipo de datos ya que confían más en estos últimos que en los históricos.

	Finalidad	
Tipo de dato	**Seguros** **Cálculo de la póliza**[307]	**Bancario** **Solvencia financiera**[308]
Datos tradicionales	Historial médico, edad, sexo, domicilio, consumo de tabaco, de alcohol.	Nómina Edad Patrimonio
Datos alternativos	Compras online, actividad en las redes sociales, hábitos de consumo, situación financiera, actividad física, hábitos en la conducción de vehículos.	Actividad en las redes sociales Establecimientos donde se realizan compras (geolocalización)

Es por tanto esa nueva información valiosa presente en esos datos alternativos la que empuja a las organizaciones a no solamente acudir a los datos tradicionales para elaborar sus algoritmos. La recopilación de todo tipo de datos no convencionales trata de moldear mejor la realidad sobre la que se pretenden tomar decisiones, las cuales, pueden llegar a ser más precisas cuando se aportan estos nuevos datos. Y ello, pese a que dichos datos a priori no guarden relación con el objetivo que se pretende[309].

A) Ventajas

La primera ventaja que ofrece el uso de datos alternativos es que en un gran número de ocasiones, el uso de los mismos generan modelos al menos tan precisos

K: "The New Data of Student Debt". *Southern California Law Review*, December 8, 2019, pág.3. Disponible en: https://ssrn.com/abstract=3349478

[307] MUÑOZ PAREDES, M,L: "Big data y contrato de seguro: los datos generados por los asegurados y su utilización por los aseguradores". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*, op.cit., págs.140 y 141.

[308] La ICO ha señalado otros ejemplos de datos no convencionales como pueden ser los datos que se obtienen de los sensores, la batería o el dispositivo de geolocalización de un teléfono móvil o datos de texto recogidos de la actividad en las redes sociales, etc. En: ICO, *Explaining decisions made with AI*, 2020, Pág.65.

[309] Ello permite responder a preguntas del tipo ¿sería posible ofrecer seguros de salud en función del análisis de las preferencias alimentarias manifestadas por los usuarios y por sus redes de amigos en espacios sociales de internet? En: MARTÍNEZ MARTÍNEZ,R: "Big data, investigación en salud y protección de datos personales ¿Un falso debate?". *Revista valenciana d'estudis autonòmics*, Nº 62, 2017, pág.41.

como aquellos otros modelos que se basan en datos históricos o fuentes de información altamente fiables. Así, organizaciones como la OCDE o el FMI ya han indicado el potencial de estos datos alternativos a la hora de evaluar la solvencia financiera, sobre todo cuando estos se complementan con los datos históricos que ya ostentan las entidades bancarias[310]. Ello permite a las organizaciones obtener todo tipo de información valiosa a partir de canales menos costosos y más accesibles. Así, por ejemplo, en Francia se ha considerado constitucional el análisis masivo de datos personales de los ciudadanos disponibles públicamente en las plataformas virtuales como método para luchar contra el fraude fiscal[311]. Se obtiene así un nuevo canal para obtener información valiosa para la organización, en este caso, se reducen los gastos en inspección y además, se obtiene una hipotética mayor precisión en la detención de actividades ilícitas.

Por otro lado, también se ha señalado que los datos alternativos pueden servir como sustituto a la hora de obtener información respecto de datos históricos sobre los cuales, las personas muestran cierto recelo para facilitárselos a las organizaciones. Como ejemplo se señala que las personas son desconfiadas para conceder en determinadas ocasiones el número de cuenta o la tarjeta bancaria en entorno virtuales. Estos datos han resultado hasta la fecha elementales a la hora de valorar la posible financiación de un producto adquirido a través de la web[312]. Desde el momento que esos datos dejan de ser necesarios y la información que se pretendía recopilar se obtiene de

---

[310] La investigación ha comparado el rendimiento de los algoritmos para predecir la probabilidad de impago basándose en la puntuación FICO y la que se deriva del uso este indicador junto con el uso de datos alternativos. La puntuación FICO por sí sola tuvo una tasa de precisión del 68,3%, mientras que un algoritmo basado en datos alternativos tuvo una tasa de precisión del 69,6%. Utilizando ambos tipos de datos conjuntamente, la tasa de precisión se elevó al 73,6%.
Estos resultados sugieren que los datos alternativos *complementan*, más que sustituyen, la información de las agencias de crédito. Por lo tanto, un prestamista que utilice información tanto de fuentes tradicionales (FICO) como de datos alternativos puede tomar mejores decisiones de préstamo. En: En: BOOT A; HOFFMANN,P ; LAEVEN,L ; RATNOVSKI,L: "Financial Intermediation and Technology: What's Old, What's New?", *Fondo Monetario Internacional*, WP/20/161, 2020, págs.11 y 12.

[311] La Decisión del Consejo Constitucional Francés. Décision n° 2019-796 DC du 27 décembre 2019. Disponible en: https://www.conseil-constitutionnel.fr/decision/2019/2019796DC.htm . Un análisis de esta resolución puede encontrarse en COTINO HUESO,L: *Hacia la transparencia 4.0: el uso de la inteligencia artificial y big data para la lucha contra el fraude y la corrupción y las (muchas) exigencias constitucionales*. En RAMIÓ,C (coord): "Repensando la Administración Pública. Administración digital e innovación pública". Op.,cit,. Págs.169 a 196.

[312] BERG, T ; BURG,V; GOMBOVIĆ, A; MANJU,P: "On the Rise of FinTechs – Credit Scoring Using Digital Footprints". *Michael J. Brennan Irish Finance Working Paper Series*, Research Paper No. 18-12, July 15, 2019, págs. 12,29 y 30.Disponible en: https://ssrn.com/abstract=3163781
En este texto se hace referencia a la huella digital como dato fundamental a la hora de proceder a la financiación de los productos que compran los usuarios de una empresa de muebles. Entre los datos que se incorporan a esa huella digital desatacan: tipo de dispositivo con el que se accede a la web, hora a la que se realiza la compra, correo electrónico utilizado, uso de número o no en el correo electrónico, etc.

datos alternativos como el comportamiento que realiza el individuo en la red, la sensación que tiene este último es que a priori, ese dato alternativo es menos sensible y por tanto, será más proclive a concederlo.

Otra ventaja interesante está relacionada con la reducción de los sesgos. Uno de los principales inconvenientes achacable al *data mining* y al aprendizaje automático es que los modelos que se basan en estas técnicas utilizan para su entrenamiento datos históricos. Por tanto, si estos datos históricos presentan sesgos, los modelos replicarán los mismos una vez comiencen a adoptar decisiones. Pues bien, los datos alternativos, dado que en muchos casos se derivan o miden hechos y datos del presente, a priori, los mismos no presentan esos desequilibrios históricos presentes en la sociedad. De esta manera, a la hora de crear un algoritmo que determine el potencial riesgo de impago o de delincuencia de un individuo ya no habrá que acudir necesariamente a bases históricas crediticias o policiales. Los datos alternativos permiten utilizar información fresca que no deviene del pasado sesgado sino del presente y por tanto, tal modelo se puede ajustar de forma más adecuada a la realidad sobre la que se pretende que el mismo adopte decisiones.

En otros supuesto, el uso de datos alternativos puede permitir que sectores de la población que hasta la fecha veían limitado su acceso a determinados servicios por falta de datos fiables u objetivos puedan ahora recibir un mejor trato al poder valorar otras variables o características que también pueden reflejar el objetivo que se persigue con el modelo sin tener en cuenta o prestar menos relevancia a los datos históricos. Por ejemplo, en el sector público, se podría llegar a desarrollar un sistema en el que, a la hora de conceder una determinada beca o priorizar en la elección del acceso a una titulación universitaria no sólo se tuviera en cuenta la calificación de los alumnos sino que se pudieran obtener otras variables o características que pudieran representar también la variable objetivo de ese servicio a la hora de adoptar una u otra decisión. En el ámbito privado, concretamente en el sector crediticio de préstamos estudiantiles de EEUU está proliferando el uso de datos alternativos en la concesión de créditos. Ello puede permitir el acceso a un número importante de estudiantes que hasta la fecha encontraban dificultades para financiarse debido a que no ostentaban un historial crediticio solvente o cuanto menos sólido. Ahora, dado que se tienen en cuenta otros datos como la titulación universitaria o las calificaciones obtenidas hasta la fecha, estos

nuevos datos pueden mostrar signos que, para una entidad bancaria, resultan útiles a la hora de evaluar la solvencia financiera[313].

B) Riesgos

Los peligros que presenta el uso de estos datos hay que tenerlos en cuenta a la hora de llevar a cabo el análisis de riesgos[314]. Destacamos los siguientes:

En primer lugar se ha indicado que el uso de datos alternativos suele llevar consigo un mayor seguimiento y control del particular que solicita el servicio o la actividad sobre la que se pretende tomar decisiones automatizadas. Y es que, debido a que muchos de estos datos exigen una actividad y control total del particular para que dichos datos sean lo más precisos posibles, la intromisión en la esfera personal del individuo puede llegar a considerarse excesiva[315]. Además, al no tener en cuenta datos históricos y centrarse en datos relativos como por ejemplo el comportamiento del individuo en las redes, la posibilidad de que este último altere su conducta y su forma de interrelacionarse para moldearse a las exigencias requeridas por la organización aumenta[316]. Ello reduce la libertad de elección y comportamiento del individuo en pro de contentar al algoritmo que lo evalúa.

En segundo lugar, resulta habitual que el proceso de recopilación de esos datos alternativos se realice de forma oculta para el particular afectado[317]. Como consecuencia de ello, en muchas ocasiones, el particular desconoce las razones por las que obtiene un determinado servicio o se adopta una determinada decisión. Este elemento sí que se diferencia respecto de la recopilación de los datos tradicionales donde los particulares conocían en cierta medida las razones o los elementos que daban lugar a la adopción de

[313] ODINET, CHRISTOPHER K: "The New Data of Student Debt". *Southern California Law Review*, op.cit., pág. 4. Disponible en: https://ssrn.com/abstract=3349478

[314] Analizar los riesgos del tratamiento de datos alternativos será fundamental para valorar qué medidas se han de implementar para mitigar o reducir los riesgos detectados. Véase el Capítulo III, aparatados I y II de esta tesis.

[315] MUÑOZ PAREDES, M,L: "Big data y contrato de seguro: los datos generados por los asegurados y su utilización por los aseguradores". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed.Aranzadi, Navarra, 2020, pág. 143.

[316] Por ejemplo, el llamado crédito social que se utiliza en China castiga a las personas que están jugando a la video consola y favorece a las personas que compran pañales por entenderse que tienen hijos y son responsables. OCDE, "Artificial Intelligence in Society", 2019, pág.58.Disponible en: *https://doi.org/10.1787/eedfee77-en*.

[317] MUÑOZ PAREDES, M,L: "Big data y contrato de seguro: los datos generados por los asegurados y su utilización por los aseguradores". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed.Aranzadi, Navarra, 2020, pág. 143 a 148.

un tipo de decisión u otra. Pero es que además, no sólo no existe una ausencia total de información sobre las variables principales en las que se basa la decisión sino que a ello hay que sumarle un conocimiento mínimo de cómo funciona ese sistema. Por tanto, en el hipotético caso de que un particular pueda llegar a ser consciente de que su prima de riesgo depende de la actividad que este realice en sus redes sociales, todavía desconocería cómo la forma de interactuar en estas redes influye en el precio de la prima.

En tercer lugar también merece especial atención la fiabilidad de determinadas variables o datos alternativos que se recopilan para alimentar a los algoritmos[318]. Nótese que muchas de las variables alternativas que se suelen utilizar para la adopción de decisiones automatizadas son variables ficticias creadas por el que recopila el dato y que puedan cambiar de un momento a otro en función de cómo interactúe el particular en ese momento. Ello puede generar errores iniciales en esa recopilación del dato, o, mejor dicho, errores a la hora de valorar ese dato. Por ejemplo, en la tabla posterior se representan algunos datos que tienen en cuenta una empresa que se dedica a la venta de muebles[319]. Para aquellas compras que superen una cantidad de euros se habilita la opción de financiar el producto. Pues bien, la evaluación de la solvencia financiera que permite o no la financiación del producto se realiza enteramente por internet. Entre las variables que se tiene en cuenta destacan; el correo electrónico utilizado para solicitar la financiación, si el usuario se equivoca o no inicialmente al introducir el correo electrónico, si ese usuario introduce en el formulario el nombre de la ciudad donde reside en minúscula o en mayúscula, etc.

Dato Alternativo	**Valor del dato alternativo**	
	Mayor probabilidad de impago	**Menor probabilidad de impago**
Correo utilizado	Hotmail , Yahoo	Correo de clientes de otro servicios. (T-online)
Horario de pedido	A partir de las 6 de la tarde	Antes de las 6 de la tarde
Tipo de dispositivo en el que se realiza el pedido	Teléfono móvil	Ordenador o tableta

[318] A día de hoy, no está constatado que los datos en redes sociales sean confiables a la hora de evaluar la solvencia financiera. En: Experian. *The State of Alternative Credit Data. How the financial services industry is adopting and benefiting from these new data sources.* 2018, pág.7

[319] BERG, T ; BURG,V; GOMBOVIĆ, A; MANJU,P: "On the Rise of FinTechs – Credit Scoring Using Digital Footprints". *Michael J. Brennan Irish Finance Working Paper Series*, op.,2019, pág-12.Disponible en: https://ssrn.com/abstract=3163781

Sistema operativo	Android	iOS (Apple)
Uso de minúsculas en la primera letra del nombre y dirección de su domicilio	Uso de minúsculas a la hora de introducir el nombre o la ciudad	Uso de las mayúsculas en la primera letra del nombre y la ciudad.
Uso del nombre/apellidos en el correo electrónico	No uso del nombre/apellido en el correo electrónico	Uso del nombre/apellido en el correo electrónico
Error a la hora de introducir el correo electrónico	Se comete error al introducir el correo electrónico	No se comete el error al introducir el correo electrónico

Todos estos datos alternativos reflejan para esta empresa la posible solvencia futura financiera de una persona. Sin embargo, muchos de estos dependen en gran medida de cómo en ese momento el particular se interrelaciona con la web. Así, por ejemplo, cualquier persona puede equivocarse a la hora de introducir su correo electrónico, una misma persona puede tener distintos correos para realizar distintas transacciones, las personas pueden conectarse a una determinada hora porque concretamente en ese momento y en ese día les viene mejor, etc. La fiabilidad de estos datos puede en muchos casos ser limitada.

Por último, *en cuarto lugar*, y aunque decíamos que una de las ventajas de los datos alternativos es que pueden crear modelos precisos sin tener que acudir a los modelos históricos que reflejan sesgos históricos estructurales. En muchos casos, los datos alternativos o variables utilizadas pueden seguir presentando sesgos ocultos cuando dichas variables indirectamente reflejen también dichos sesgos estructurales. Por ejemplo, a la hora de valorar la solvencia financiera de un estudiante, se evita valorar su historial crediticio reduciendo la posibilidad de relacionar determinadas variables con la raza de la persona. Sin embargo, utilizar la variable de las calificaciones académicas obtenidas a lo largo de los años puede ser una variable que puede seguir presentando problemas ya que se ha demostrado que en EEUU la calificación obtenida y la raza pueden presentar fuerte interrelación. Así, existe un porcentaje mayor de personas de raza blanca con mejores notas que las de raza negra. Esto supone que, al utilizar la variable calificación académica, indirectamente se está utilizando la variable raza a la hora de conceder los créditos[320]. Es decir, el dato académico resultar ser *proxy* de la raza. En otras ocasiones, el uso de variables alternativas puede esconder una

[320] ODINET, CHRISTOPHER K: "The New Data of Student Debt". *Southern California Law Review* op.cit., págs 58 y 59.Disponible en: https://ssrn.com/abstract=3349478

realidad todavía más compleja. Nos referimos a aquellos supuestos donde, conociendo la interrelación entre una variable alternativa y una variable prohibida, los diseñadores utilizan dichas variables alternativas para evitar los usos limitados o prohibidos a los que se someten determinadas variables. Por ejemplo, el artículo 94.1 de la Ley 20/2015, de 14 de julio, de ordenación, supervisión y solvencia de las entidades aseguradoras y reaseguradoras establece determinadas exigencias a la hora de incorporar la variable sexo en los procesos de evaluación de las primas de seguro. Por otro lado, el artículo 9 del RGPD hace alusión a una serie de datos especialmente protegidos. Pues bien, un diseñador de un sistema automatizado podría utilizar datos alternativos con el único objetivo de evitar las limitaciones impuestas por la legislación por el uso de determinadas variables. Así, a priori, formalmente se evita utilizar la variable prohibida o limitada por la norma, pero, la variable alternativa realmente sigue estando reflejada en el modelo. Gracias a los avances en las técnicas de procesamiento de datos como el *big data*, las variables alternativas consiguen en muchos casos iguales o mejores resultados en los que implícitamente, y de forma encubierta, siguen valorando la variable prohibida o más protegida..

Dato alternativo utilizado	**Variable que se esconde (*proxy*)**	**Norma relevante**
Marca o gama del vehículo	Sexo	Artículo .94.1 de la Ley 20/2015, de 14 de julio
Notas académicas	Raza	Artículo 9 RGPD
Diarios o periódicos consultados	Opción política	Artículo 9 RGPD.

En definitiva, el uso de datos no comúnmente utilizados en procesos de toma de decisiones suponen una nueva fuente de conocimiento que pueden llevar a desarrollar modelos que reflejen de forma más fidedigna la realidad sobra la que dichos modelos adoptarán decisiones, sustituyendo o complementando a los datos que históricamente se han utilizado. Sin embargo, es importante destacar los riesgos que se pueden derivar de los usos de los mismos.

C) ¿Son los datos alternativos datos personales?

La respuesta a esta pregunta es afirmativa teniendo en cuenta la interpretación que ha realizado el GT29 al concepto de dato personal. Así, si la conjunción de datos

alternativos permite la singularización de por ejemplo un dispositivo, debemos entender que dichos datos son personales. Además, si bien, uno o varios datos alternativos no logran identificar a una persona, la conjugación con otros que sí vinculan a la misma los convertiría a todo el conjunto en datos personales. Por ejemplo, en el supuesto anterior, decíamos que la empresa evaluaba toda una serie de datos alternativos como el comportamiento en la página web, el dispositivo utilizado, el sistema operativo de dicho dispositivo o la hora a la que se accede al servicio. Aunque cada uno de esos datos por sí sólo no logra identificar a un individuo, si junto con todos estos datos también se le agrega por ejemplo el nombre y apellidos de una persona, entonces, todos esos datos alternativos también deben considerarse datos personales. Como luego se analizará, de acuerdo al principio de minimización de datos, teniendo en cuenta los riesgos que presentan el uso de datos alternativos, la pertinencia de su uso requerirá un mayor grado de justificación que el establecido para el resto de datos convencionales[321].

4. Dato inferido

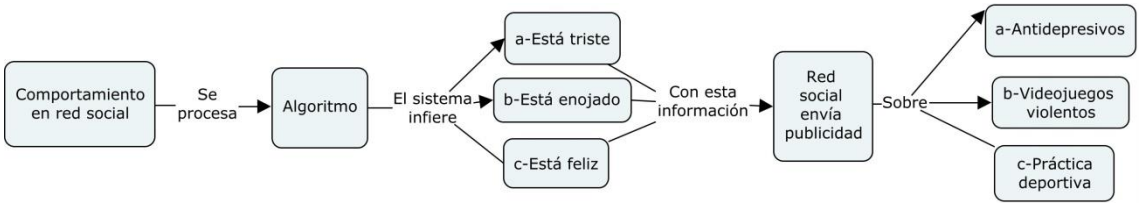
Cuando hablamos de inferencia o dato inferido nos referimos a aquella información relativa a una persona física identificada o identificable creada mediante deducción o razonamiento y no mediante la mera observación o recogida del interesado[322]. Estas inferencias atribuyen al particular uno o varios atributos que pueden o no estar presentes en el mismo. De esta manera, a través de la inferencia, aquel que la deduce realiza una estimación de que una persona presenta la característica y además, sobre dicha característica estimada y atribuida se establecen una serie de consecuencias. Por ejemplo, al introducir en una aplicación informática nuestro rostro, esta aplicación procesa los datos y puede tratar de estimar la edad que tenemos. A su vez, otro sistema, puede por ejemplo inferir nuestro nivel de felicidad midiendo nuestro comportamiento en una red social[323].

---

[321] Véase el Capítulo IV, apartado III, punto 1 de esta tesis.

[322] Sobre el concepto de dato inferido véase: WACHTER, S; MITTELSTADT, B: "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI". *Columbia Business Law Review*, 2019, pág. 2. Disponible en : https://ssrn.com/abstract=3248829

[323] Infiriendo el estado anímico de sus usuarios, una plataforma podría dirigir de forma más personalizada sus anuncios. Fuente de la noticia : REILLYARCHIVE,M: "Is Facebook Targeting Ads at Sad Teens?", *MIT Technology Review*, 01/05/2017. Disponible en: https://www.technologyreview.com/2017/05/01/105987/is-facebook-targeting-ads-at-sad-teens/
Como ha indicado el Grupo del artículo 29, los conocimientos obtenidos a raíz del análisis de la actividad de los usuarios por parte de una plataforma social son considerados datos inferidos. Grupo del Artículo 29. *Guidelines 8/2020 on the targeting of social media users*. Adoptadas el 2 septiembre de 2020, pág.8,



En los últimos tiempos, y gracias a la expansión de las técnicas del *big data*, la minería de datos o el *machine learning*, las inferencias a personas se han convertido en una práctica muy habitual realizada por todo tipo de organizaciones. Estas, tienen su base en la elaboración de perfiles, ya sean individuales o grupales. Así, ante la incorporación de toda una serie de datos en el algoritmo, este los procesa y posteriormente arroja un resultado, ese resultado se convierte en un dato inferido, el cual, tendrá uno u otro significado en función del valor que haya establecido la organización que procesa los datos. Lo dicho anteriormente queda ilustrado con la siguiente tabla. Imaginemos que una aseguradora pretende implantar un sistema que valore la probabilidad de que un potencial cliente pague o no la póliza en un futuro. De esta manera, primeramente se entrenaría el sistema con la base de datos que se tenga, asignando la etiqueta de buen o malo pagador a cada una de las muestras en función de la información que se ostente sobre esos ejemplos.

BASE DE DATOS DE LA ASEGURADORA. FASE DISEÑO				
Ejemplos	**Características**			**Etiqueta**
	Sexo	Nivel de retribución	Modelo de coche	
Cliente 1	Hombre	1000 €	Renault Clio	Mal pagador
Cliente 2	Mujer	1200 €	Renault Megane	Mala pagadora
Cliente 3	Mujer	1850 €	Seat Ibiza	Buena pagadora
Cliente 4	Hombre	1300 €	Volkswagen Golf	Buen pagador

Una vez creado el modelo, este se encuentra en perfecto estado para ser aplicado a nuevos clientes. Dicho sistema tendrá como objetivo principal inferir el valor pretendido, esto es, malos o buenos pagadores. De manera que, una vez procesados los datos por parte del sistema, este arrojará un resultado, el cual, asignará al particular una información, en este caso, que la persona es mala o buena pagadora.

apartado 22. Por otro lado, en el sector de seguros la consideración de datos inferidos la encontramos en el resultado derivado de la selección y tarificación de riesgos. En: Asociación empresarial del seguro. *Guía para el tratamiento de los datos personales por las entidades aseguradoras,* febrero 2019, pág.21.

NUEVOS CLIENTES. FASE TOMA DE DECISIONES				
Nuevos clientes	**Características/Datos aportados por el cliente**			**Dato inferido/Dato nuevo**
	Sexo	Nivel de retribución	Modelo de coche	
Cliente 1	Mujer	1200 €	Citroën C.5	Mala pagadora
Cliente 2	Mujer	1500 €	Renault Cactus	Buena pagadora
Cliente 3	Mujer	950 €	Seat Panda	Mala pagadora
Cliente 4	Hombre	1700 €	Audi A4	Mal pagador

La inferencia que realiza el modelo al establecer que una persona es mala o buena pagadora es un dato nuevo que se origina fruto del tratamiento al que se ven sometido los datos personales que previamente se han introducido sobre el sujeto sobre el que se crea ahora ese nuevo dato[324]. Estos datos inferidos pueden mostrar información que puede ser más o menos veraz referida un individuo y sobre la cual se pueden adoptar decisiones o actuaciones sobre el particular afectado por ese dato[325].

A) Riesgos

Son diversos los riesgos que presentan el uso de datos inferidos:

En primer lugar, se ha indicado que en la mayoría de las ocasiones los particulares desconocen qué datos infieren las organizaciones sobre ellos y cómo esos datos pueden influir en sus vidas. A diferencia de los datos recopilados u observados de los particulares, los datos inferidos no son aportados directamente por el particular sino que se infieren del procesamiento de los primeros. De esta manera, un particular puede ser consciente de que recopilan el dato personal de su color de ojos, pero no es consciente de que el análisis de ese dato por parte de un algoritmo junto con otros datos que ha aportado el mismo arrojan una inferencia específica sobre esa persona, inferencia de la que en muchas ocasiones el particular no tendrá conocimiento.

[324] Recomendación CM/Rec(2020)/13 del Comité de Ministros a los Estados Miembros sobre la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal en el contexto de la creación de perfiles. Disponible en: https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=09000016805cdd2a

[325] Por ejemplo, los resultados que se derivan de la evaluación de los datos PNR que realizan las Unidades de Información sobre Pasajeros (UIP) con la finalidad de evaluar a los pasajeros debemos considerarlos también datos inferidos ya que dichos resultados infieren una información sobre esa persona basada en los datos PNR. Véase el artículo 12 de la Ley Orgánica 1/2020, de 16 de septiembre, sobre la utilización de los datos del Registro de Nombres de Pasajeros para la prevención, detección, investigación y enjuiciamiento de delitos de terrorismo y delitos graves.
Visto en: https://www.boe.es/buscar/act.php?id=BOE-A-2020-10776

Datos que se recopilan del particular	**Datos inferidos por el algoritmo**	**Decisión que adopta la organización basada en el dato inferido**
Cliente 1: Edad 45, 1100€/mes, CP:12345, Librero	Cliente de dudoso cobro	Se fijan mayores intereses
Cliente 2: Edad 56, 1300€/mes, CP 12456, Arquitecto	Cliente solvente	Se concede el préstamo solicitado
Cliente 3: Edad 24, 800€/mes, CP 67895, Administrativo	Cliente nefasto	Se deniega el préstamo solicitado.

En esta tabla se evaluaba la solvencia de varios clientes que solicitaban un crédito. Los clientes conocen los datos de entrada sobre los que se realiza la evaluación, además, son conscientes de la decisión que toma la entidad bancaria. Sin embargo, en muchos supuestos desconocerán que inferencias se han generado tras el análisis de sus datos. Nótese que la inferencia no deja de ser una denominación ficticia que establece la organización sobre el resultado arrojado por el algoritmo y sobre la cual, se derivan toda una serie de consecuencias que afectan al particular sobre el que se deriva ese dato.

En segundo lugar, y fruto de ese desconocimiento, muchas de las inferencias ocultas que se obtienen tras la elaboración de los perfiles pueden resultar especialmente relevantes cuando las mismas generen efectos significativos negativos para los particulares[326]. Estas inferencias pueden tratarse por terceras organizaciones en un ocultismo total hacia los particulares sobre los que se realiza la inferencia. A causa de ello, en determinadas ocasiones se ha prohibido tratar datos con el objetivo de inferir determinados atributos. En este sentido, la AEPD, a través de una circular y para un contexto específico prohibió el tratamiento de datos personales cuya finalidad fuese la de inferir la ideología política de una persona[327].

[326] Center for Democracy and Technology. *Algorithmic Systems in Education: Incorporating Equity and Fairness When Using Student Data*, 2019, pág.18. Disponible en: https://cdt.org/insights/algorithmic-systems-in-education-incorporating-equity-and-fairness-when-using-student-data/

[327] Artículo 5.2 de la Circular 1/2019, de 7 de marzo, de la Agencia Española de Protección de Datos, sobre el tratamiento de datos personales relativos a opiniones políticas y envío de propaganda electoral por medios electrónicos o sistemas de mensajería por parte de partidos políticos, federaciones, coaliciones y agrupaciones de electores al amparo del artículo 58 bis de la Ley Orgánica 5/1985, de 19 de junio, del

Es por ello que se apuesta cada vez más por prever reglas especiales para el tratamiento de estos datos. Resumidamente cabe destacar la propuesta de Wachter, partiendo de las limitaciones del RGPD, se propone que los particulares se les reconozca un derecho a las inferencias razonables que generen alto riesgo. Este derecho englobaría una serie de garantías previas y posteriores al momento en que se realiza la inferencia. Por un lado, el responsable del tratamiento que quiera realizar inferencias debería justificar ex ante dicha inferencia. Esta justificación abarcaría los siguientes elementos; (1) por qué ciertos datos son base relevante para hacer inferencias; (2) por qué estas inferencias son relevantes para el propósito de procesamiento elegido o el tipo de decisión automatizada; y (3) si los datos y métodos utilizados para hacer las inferencias son precisos y estadísticamente fiables. La justificación *ex ante* se vería reforzada por un mecanismo adicional *ex post* que permite impugnar las inferencias irrazonables. Estas garantías que se proponen son muy acertadas ya que otorgarían a los titulares sobre las que se realizan esas inferencias un mayor protagonismo y además obligaría a los responsable de los tratamiento a prestar mayor atención sobre el tratamiento realizado. Dicho esto, consideramos que, una interpretación amplia de muchos de los artículos del RGPD podría conseguir prácticamente los mismos efectos que se proponen por esta autora. Para ello, en páginas posteriores realizaremos un análisis más profundo de los datos inferidos y las posibilidades que brinda la normativa de protección de datos en este contexto a los particulares.

B) ¿Son los datos inferidos datos personales?

Aunque a priori se pudiera pensar que un dato inferido es un dato personal, a día de hoy no está del todo claro. En este sentido, cabe destacar que el GT29 sí que lo considera de forma rotunda. Así, esta institución afirma que los datos inferidos son datos personales nuevos que no han sido facilitados por el interesado[328]. Sin embargo,

---

Régimen Electoral General. Esta circular fue emitida por la AEPD con la intención de evitar que el Tribunal Constitucional considerara inconstitucional el artículo 58 bis de la Ley Orgánica 5/1985, de 19 de junio, del régimen electoral general, incorporado por la Ley Orgánica 3/2018, de 5 de diciembre, de protección de datos personales y garantía de los derechos digitales. Esta última ley debería haber regulado suficientes medias de garantías. A pesar de que la AEPD trató de colmar dicho vacío por medio de la mentada circular. El TC finalmente consideró inconstitucional dicha reforma de la LOREG a través de la LOPD de 2018 debido a esa carencia de suficientes medias de garantía. Véase STC de 76/2019, de 22 de mayo de 2019. Resolución disponible en: https://www.boe.es/buscar/doc.php?id=BOE-A-2019-9548

[328] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*. Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.10. También así lo ha indicado el Consejo de Europa en: The protection of individuals with regard to automatic processing of personal data in the context of profiling.

cuando el TJUE ha tenido la oportunidad de valorar este problema, este no ha sido del todo claro a la hora de tildar los datos inferidos como datos personales[329]. De esta manera, son dos las ocasiones en las que el TJUE ha tenido la oportunidad de analizar esta problemática:

En la primera sentencia, asuntos acumulados *C-141/12 y C-372/12: YS. y M. y S* de 17 de julio de 2014 [330], el tribunal analizó si un solicitante tiene derecho a acceder al análisis jurídico de una acta relacionada con un permiso de residencia. Concretamente el tribunal debía valorar si dicho análisis jurídico puede ser considerado dato personal del particular sobre el que se realizaba dicho estudio. Para el tribunal, a pesar de que los datos relativos al solicitante del documento de residencia puedan constituir la base fáctica del análisis incluida en el acta[331], el análisis jurídico como tal no puede considerarse dato personal[332]. De manera que, a través de la normativa de protección de datos no se puede ejercer ni el derecho de acceso ni en su caso el derecho de rectificación del análisis jurídico establecido por esta legislación[333]. Como ha apuntado la doctrina[334], esta sentencia es interesante ya que por vez primera el TJUE valora si datos o informaciones referidas a una persona que no son hechos objetivos sino subjetivos han de considerarse o no datos personales del particular sobre el que se realiza ese análisis[335]. Si trasladamos este caso a la operatoria descrita hasta ahora sobre

---

Recommendation CM/Rec(2010)13 adopted by the Committee of Ministers of the Council of Europe on 23 November 2010 and explanatory memorandum, p-39. En este mismo sentido, WACHTER, S; MITTELSTADT, B: "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI". op.cit. págs 22 y ss. Así como LAZCOZ MORATINOS, G., & CASTILLO PARRILLA, J: "Valoración algorítmica ante los derechos humanos y el Reglamento General de Protección de Datos: el caso SyRI". *Revista Chilena de Derecho y Tecnología*, 2020, págs. 207-225. En el mismo sentido, GIL GONZÁLEZ,E : *Big data, privacidad y protección de datos*, op,cit.,, pág.46.
Disponible en: https://www.aepd.es/sites/default/files/2019-10/big-data.pdf .

[329] Wachter realiza una aproximación muy descriptiva y acertada sobre el análisis jurídico que ha realizado hasta la fecha el Tribunal de Justicia de la Unión Europea sobre la naturaleza de los datos inferidos. En: WACHTER, S; MITTELSTADT, B: "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI". op.cit., págs .29 y ss.

[330] Sentencia del TJUE (Sala Tercera) de 17 de julio de 2014, asuntos acumulados, C-141/12 y C-372/12, caso YS. y M. y S. Resolución disponible en: http://curia.europa.eu/juris/document/document.jsf;jsessionid=3038FAB16CCECF9E5D63441B61D0839C?text=&docid=155114&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=3694614

[331] Apartado 45 de la resolución comentada.

[332] Apartado 48 de la resolución comentada.

[333] Apartados 45 y 48 de la resolución comentada.

[334] WACHTER, S; MITTELSTADT, B: "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI". op.cit. pág.30.

[335] Los datos de carácter personal a los que se refiere la normativa de protección de datos son heterogéneos. Sentencia del TJUE de 7 de mayo de 2009, asunto C-553/07, caso asunto Rijkeboer, apartado 59. Resolución disponible en: https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX%3A62007CJ0553

los datos inferidos, la situación es perfectamente comparable. Esto es, existen una serie de datos personales que sirven como base para una evaluación y posteriormente se infiere un conjunto de información sobre la persona. La inferencia la realiza un algoritmo.

En la segunda resolución, el TJUE da un paso más aperturista. Así, en el asunto *C-434/16: Nowak* [336], el tribunal abordó la cuestión de si las anotaciones que realiza un evaluador sobre un examen realizado por una persona son datos personales relativos a esta última. Pues bien, para el órgano judicial, dado que las anotaciones realizadas por un examinador expresan la opinión o valoración de este sobre los resultados individuales del aspirante que ha realizado el examen y tales anotaciones pueden tener efectos para el aspirante, se ha de considerar que dicha información está relacionada con el aspirante y por tanto son datos personales del mismo, y ello, a pesar de que las mismas anotaciones también sean datos personales referidos a la examinador[337]. De esta manera, el aspirante de ese examen puede ejercer los derechos que reconoce la normativa de protección de datos personales, entre otros, el derecho de acceso, rectificación o el de supresión sobre esas anotaciones. Ahora bien, estas facultades encuentran su limitación, así, el tribunal indicó que el derecho de rectificación no habilitaba a que el aspirante pudiera ampararse en el mismo para rectificar *a posteriori* las respuestas incorrectas[338]. La rectificación sólo cabría en aquellos casos en los que *por error las hojas de los exámenes se hayan entremezclado de tal modo que las respuestas de otro aspirante se hayan atribuido al aspirante afectado, o cuando se haya perdido una parte de los folios que contienen las respuestas de ese aspirante, dando lugar a que esas respuestas queden incompletas, o incluso cuando las eventuales anotaciones del examinador no documenten correctamente la valoración que este ha dado a las respuestas del aspirante de que se trate*[339]. Teniendo en cuenta los ejemplos a los que alude el tribunal donde cabe el derecho de rectificación, no queda claro si tal rectificación puede abarcar el contenido de los comentarios del evaluador, contenido

[336]Sentencia del TJUE (Sala Segunda) de 20 de diciembre de 2017, asunto, C-434/16, caso Nowak. Resolución disponible en: http://curia.europa.eu/juris/document/document.jsf;jsessionid=6C361FE3B54FB5329197ADD52A787D08?text=&docid=198059&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=3496287

[337] Aparatados 43 y 44 de la resolución comentada.

[338] Apartado 52 de la resolución comentada.

[339]Apartado 54 de la resolución comentada.

que vendría a ser el ejemplo más parecido al dato inferido[340]. Sí que fue más clara la abogada general que se encargó de este asunto al señalar que el derecho de rectificación que se deriva de la normativa de protección de datos no puede abarcar la justificación -sea o no adecuada- sobre la que se realiza la valoración[341]. Así, se vino a decir que las posibles objeciones a los comentarios realizados por el evaluador deberían resolverse a través de la reclamación contra la valoración del trabajo pero no a través de la facultad de rectificación derivada de la normativa de protección de datos[342].

En definitiva, teniendo en cuenta las aportaciones realizadas por el GT29 y la jurisprudencia europea que han analizado los datos inferidos podemos llegar a la conclusión de que las inferencias a priori son datos personales. Ahora bien, las facultades que reconoce la normativa de protección de datos a los titulares de dichos datos quedan en parte limitadas cuando por ejemplo, los interesados pretendan ejercer los derechos de rectificación o de acceso[343].

En nuestra opinión, dado que un dato inferido refleja cierta información -que puede o no ser cierta- sobre una persona y dicha información genera ciertas consecuencias hacia la misma. Debemos entender que dichas inferencias han de considerarse datos personales desde que se produce esa vinculación de esa información con la persona a la que se le asigna. Cuando un sistema algorítmico infiere un determinado resultado y a ese resultado el responsable del tratamiento le asigna un valor, por ejemplo; mal pagador, orientación política de extrema derecha, presenta o no enfermedad, etc. Lo que el responsable está haciendo, es asignar a esa persona una serie de características que presumiblemente presenta el particular y además, por el hecho de ostentar esas características, establecer una serie de consecuencias. Es lógico pensar por tanto que dichas características asignadas, sean o no veraces o más o menos correctas sean considerados datos personales del particular.

---

[340] WACHTER, S; MITTELSTADT, B: "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI". op.cit., pág.42.

[341] Conclusiones de la Abogada General Sra. J. Kokott, presentadas el 20 de julio de 2017. Peter Nowak contra Data Protection Commissioner. Apartado 54. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:62016CC0187

[342] Apartado 55 de las conclusiones de la abogada general comentadas.

[343] En esta materia el elemento clave es conocer cómo se tipifican o perfilan a las personas. VAN DER HOLF,S & PRINS,C: "Personalisation and its Influence on Identities, Behaviour and Social Values" En: HILDEBRANDT, M & GUTWIRTH,S (eds): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008, op., p,140. Hildebrandt cita SCHAUER, F.: *Profiles Probabilities and Stereotypes*, Belknap Press of Harvard University Press, Ed, Cambridge, Mass. London, 2003.

Dato inicial	Creador de la inferencia	Inferencia
Respuestas a un examen	Examinador	Anotaciones en el examen
Edad, salario, nº de hijos, etc.	Algoritmo	Mal Pagador/Buen pagador

En contra de esta opinión, se podría argumentar que los datos inferidos difieren en parte de los datos estrictamente personales en varios elementos. Por un lado, en el hecho de que estos datos pueden ser en muchos casos no veraces o hipotéticos, siendo los datos personales objetivos y reales. Sin embargo, este enfoque cae por su propio peso cuando el propio RGPD permite la rectificación de los datos que pueden haber sido erróneamente recopilados por parte del responsable del tratamiento o facilitados de forma inadecuada por parte del particular, de manera que los mismos, pueden ser rebatidos y en su caso alterados para que dejen de ser inadecuados o incorrectos. En segundo lugar, se podría alegar que, dado que los datos inferidos son realmente creados por un algoritmo y además a dicho resultado el responsable le asigna un valor hipotético ficticio, dicho dato no pertenecería al particular afectado. Tampoco consideramos que este argumento sea adecuado, nuevamente indicamos que, dado que esa información inferida –sea más o menos atinada- es atribuida a esa persona de forma personalizada y además, sobre esa información se prevén una serie de efectos, el hecho de que dicha inferencia la haya realizado una máquina no puede ser obstáculo para evitar la consideración de dato personal, más si cabe, cuando la inferencia realizada se ha basado en todos o parte de los datos personales de esa persona.

Considerado así que los datos inferidos son datos personales. Lo cierto es que el ejercicio de las facultades reconocidas por el RGPD como el derecho de acceso, rectificación, supresión… podrán quedar en parte limitados. Estas limitaciones pueden obedecer a la propia naturaleza híbrida que presentan los datos inferidos, es decir, su origen deviene de los propios titulares de los datos y del procesamiento que haga de estos el algoritmo. Así, en muchos casos, el acceso a dichos datos puede quedar limitado por otros intereses legítimos como pueden ser los secretos comerciales ya que no podemos olvidar que las inferencias han sido creadas por el algoritmo. En estos supuestos, los datos inferidos presentan una doble naturaleza, datos personales y a su vez, secretos comerciales. No obstante, pese a que los datos inferidos pueden presentar ciertas diferencias respecto de los datos convencionales, estas diferencias como mucho habilitarán a determinadas limitaciones de las facultades, pero no a la negación de las mismas.

En los siguientes capítulos se analizaran las repercusiones que comporta considerar un dato inferido como personal, esta consideración por ejemplo supone reconocerle todas las facultades reconocidas en el RGPD con los efectos que ello supone.

5. Otras clasificaciones de datos

En este apartado se hace mención a otras tipologías de datos que también tiene especial relevancia para el objeto de esta tesis.

A) Datos sintéticos

Los datos sintéticos son aquellos que se generan artificialmente y no son recopilados del mundo real. Entre las ventajas que presentan estos datos es que los mismos pueden producirse en cualquier cantidad lo que permite simular realidades para entrenar a los modelos que posteriormente tomarán decisiones[344]. Se trata de una alternativa cada vez más solvente para entrenar sistemas de aprendizaje automático. Por ejemplo, pueden resultar muy adecuados para aportar datos de colectivos que están infrarrepresentados en las bases de datos. Se evita además el tratamiento de datos personales en estas fases iniciales. Su uso es frecuente en organizaciones que carecen de un número importante de datos. Gracias a este tipo de datos el sistema puede ser testado tanto en la fase del diseño como en fases posteriores donde dicho algoritmo esté interactuando con el entorno sobre el que adopta decisiones[345]. Entre sus desventajas, dado que estos datos no representan enteramente la realidad, es probable que el modelo no aprenda adecuadamente el entorno sobre el que se quiere que este último interactúe.

B) Datos estructurados y no estructurados

Como ya dijimos en el capítulo I de esta tesis, los datos que integran los sistemas de toma decisiones automatizadas están compuestos por datos estructurados, semi-

[344] Comité de ética alemán. Gutachten der Datenethikkommission, 2019, pág-59. https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.html

[345] OCDE, "Artificial Intelligence in Society", 2019, pág.100.Disponible en: *https://doi.org/10.1787/eedfee77-en*

estructurados y no estructurados[346]. Así, por datos estructurados entendemos aquellos datos que están organizados e incorporados en una base de datos con filas y columnas y que son fácilmente relacionables entre sí. Por su parte, los datos no estructurados se refieren a toda aquella información que no está organizada de una manera predefinida. Por ejemplo, los documentos que incorporan fechas, frases, números, imágenes, etc. El carácter o no estructurado de los datos puede presentar una fuerte incidencia jurídica. Así, dos elementos que resultan muy relevantes a la hora de aplicar el régimen de protección de datos son referidos a la estructuración de los mismos en ficheros o bases de datos, así como, especialmente, a la posibilidad de identificabilidad de los datos con relación a una persona[347]. Esto último resulta muy relevante ya que en muchas ocasiones estos datos son anonimizados.

Por lo que se refiere a los ficheros, conviene señalar que de acuerdo al artículo 2.1 del RGPD esta norma *se aplica al tratamiento total o parcialmente automatizado de datos personales, así como al tratamiento no automatizado de datos personales contenidos o destinados a ser incluidos en un fichero*. Según el considerando 15 de este texto, los ficheros o conjuntos de ficheros que no estén estructurados con arreglo a criterios específicos no deben entrar en el ámbito de aplicación de esa norma. A su vez, el artículo 4.6 de este mismo texto define al fichero como *todo conjunto estructurado de datos personales, accesibles con arreglo a criterios determinados, ya sea centralizado, descentralizado o repartido de forma funcional o geográfica.* Surge la duda de considerar si los datos no estructurados que se recopila a través de las herramientas de big data entran dentro del ámbito de aplicación del RGPD teniendo en cuenta que los mismos, cuando se integran en bases de datos, no están organizados de una manera predefinida y ordenada.

Pues bien, una primera interpretación nos podría llevar a considera que, tal y como establece el artículo 2.1 del RGPD, cualquier tratamiento de datos automatizado, independientemente de que exista un fichero o no, entra dentro del ámbito de aplicación de la norma de manera que, la existencia de un fichero más o menos ordenado es irrelevante a efectos de la normativa de protección de datos cuando se tratan datos personales de forma automatizada. Sólo sería relevante el carácter más o menos

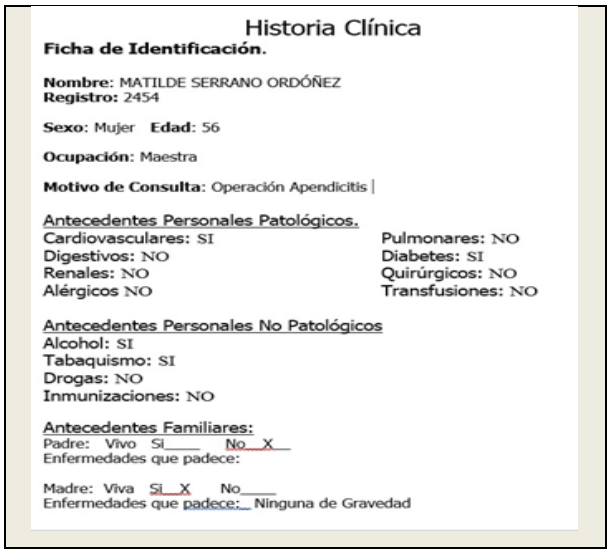
[346] ORTEGA JIMÉNEZ, A: *Las aplicaciones del Big Data en el ámbito asegurador y el tratamiento legal de sus datos. Una perspectiva desde el derecho internacional privado*, op.,cit , págs 31 y 32.Véase también la explicación realizada en el capítulo anterior de esta tesis sobre esta distinción. Capítulo I, II,2.C)

[347] COTINO HUESO,L: "El alcance e interactuación del régimen jurídico de los datos personales y *big data* relacionado con salud y la investigación biomédica", op.,cit , pág.75.

organizado del fichero cuando se traten datos de forma manual. Dado que el uso de sistemas de inteligencia artificial requiere de tratamientos automatizados, cualquier tratamiento de los mismos, con independencia de que estén más o menos estructurados entraría dentro del ámbito de aplicación del RGPD.

Por el contrario, una segunda interpretación puede dar lugar a que, cualquier tratamiento de datos personales, sea o no automatizado, requiere que los datos que se tratan estén incorporados a un fichero, es decir, a un conjunto estructurado de datos personales. Para esta segunda interpretación, los datos no estructurados podrían escapar a la aplicación de la normativa de protección de datos cuando los mismos no estén configurados de acuerdo a unos criterios determinados. En este sentido, a diferencia de los datos estructurados que se almacenan en tablas, los datos no estructurados suelen almacenarse en bases de datos no relacionales a través de documentos[348]. A modo de ejemplo, el historial clínico de un paciente puede estar o no estructurado en una base de datos. Así, estaría estructurado si la distinta información que contiene ese historial clínico se normaliza y se estructura en diferentes tablas de manera que cada atributo se incorpora a dicha tabla y se relacionan con cada paciente. A su vez, no estaría estructurado el historial clínico si el mismo se almacena en bruto y toda la información que contiene se mantiene inalterada sin clasificarse ni ordenarse.

	Sexo	**Edad**	**Fuma**
Sujeto 1	Hombre	20	Si
Sujeto 2	Mujer	34	No
Sujeto 3	Mujer	43	No
Sujeto 4	Hombre	60	Si



Historia Clínica

Ficha de Identificación.

Nombre: MATILDE SERRANO ORDÓÑEZ
Registro: 2454

Sexo: Mujer **Edad**: 56

Ocupación: Maestra

Motivo de Consulta: Operación Apendicitis |

Antecedentes Personales Patológicos.
Cardiovasculares: SI
Digestivos: NO
Renales: NO
Alérgicos NO
Pulmonares: NO
Diabetes: SI
Quirúrgicos: NO
Transfusiones: NO

Antecedentes Personales No Patológicos
Alcohol: SI
Tabaquismo: SI
Drogas: NO
Inmunizaciones: NO

Antecedentes Familiares:
Padre: Vivo Si____ No__X__
Enfermedades que padece:

Madre: Viva Si__X__ No____
Enfermedades que padece: Ninguna de Gravedad

Datos estructurados y no estructurados de una historia clínica

[348]Para más información sobre las bases de datos no relacionales consúltese: *Ayudaley*. "Base de datos no relacional. ¿Qué es? Características y ejemplos". Información disponible en: https://ayudaleyprotecciondatos.es/bases-de-datos/no-relacional/

Como se puede apreciar, mientras que en los datos estructurados la información está ordenada y responde a distintos criterios. En los datos en brutos, los distintos atributos que pueden ser útiles a la hora de analizar la información están incorporados al documento sin más, pero no responden a un tipo de organización específica. Cada historial clínico en bruto puede presentar distintos formatos y más o menos atributos. Pues bien, el TJUE, al analizar el concepto de fichero de la Directiva 95/46[349], el cual se define en términos similares al descrito en el RGPD, indicó que de ese texto no se deduce que para poder apreciar la existencia de un fichero los datos personales deban figurar en fichas, en catálogos específicos o en otro sistema de búsqueda. El requisito por el cual el conjunto de datos personales deba estructurarse conforme a criterios específicos solo tiene la finalidad de permitir que los datos relativos a una persona puedan recuperarse fácilmente[350]. Es decir, trasladado a nuestro contexto, y considerando que el concepto de fichero se aplica tanto tratamientos de datos automatizados como a manuales, los datos no estructurados no entrarían dentro del ámbito de aplicación del RGPD cuando no fuera sencillo para el responsable recuperar dichos datos de las bases de datos. Esta interpretación no obstante tiene sus limitaciones ya que el TJUE dictó está resolución analizando un tratamiento manual de datos

[349] Véase los considerandos 15 y 27 y el artículo 2.c) de la Directiva 95/46/CE del Parlamento Europeo y del Consejo, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respeta al tratamiento de datos personales y a la libre circulación de estos datos.

[350] En esta resolución judicial el TJUE valoró si la recogida de datos personales llevada a cabo por miembros de una comunidad religiosa en relación con una actividad de predicación puerta a puerta y el tratamiento posterior de esos datos se consideraba un tratamiento de datos a los efectos de la Directiva. Según las partes, dicho tratamiento quedaba fuera del ámbito de aplicación de dicha directiva ya que los datos tratados de forma manual no se incorporaban a un fichero en los términos descritos en la mencionada Directiva 95/46. Así, el TJUE consideró que el concepto de fichero comprende un conjunto de datos personales recogidos en relación con una actividad de predicación puerta a puerta, consistentes en nombres, direcciones y otra información relativa a las personas contactadas, siempre que los datos estén estructurados según criterios determinados que permitan, en la práctica, recuperarlos fácilmente para su utilización posterior. Para que dicho conjunto de datos esté comprendido en ese concepto no es preciso que incluya fichas, catálogos específicos u otros sistemas de búsqueda. SENTENCIA DEL TRIBUNAL DE JUSTICIA (Gran Sala) de 10 de julio de 2018, asunto C–25/17, caso Jehovan todistajat, FJº 57 y 58. Resolución disponible en:
https://curia.europa.eu/juris/document/document.jsf;jsessionid=AE4423E467C34CC980927AA34F1B8A35?text=&docid=203822&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=4096504
En sentido contrario, nuestro Tribunal Supremo consideró que un libro de bautismos no supone un conjunto organizado de datos y por tanto este no puede considerarse fichero a los efectos de la normativa de protección de datos. STS (Sala de lo Contencioso Administrativo) 4646/2008-ECLI: ES:TS:2008:4646, de 19/09/2008 ,FJº4. Resolución disponible en:
https://www.poderjudicial.es/search/doAction?action=contentpdf&databasematch=TS&reference=2468223&links=&optimize=20081009&publicinterface=true
Un análisis crítico de esta sentencia puede verse en: ARENAS RAMIRO, M: "Protección de datos personales y apostasía: la sentencia del Tribunal Supremo de 19 de septiembre de 2008". *Anuario de Derecho Eclesiástico del Estado*, vol. XXVI, 2010, pág. 690.

personales[351]. Por otro lado, entendemos, que tal y como señala el considerando 35 del RGPD, la técnica utilizada no debería ser un motivo para tratar de eludir la normativa de protección de datos. El hecho de que las bases de datos estén más o menos estructuradas u organizadas puede ser relevante a efectos de exigir mayores o menores exigencias en materia de protección de datos pero no ser un requisito sobre el cual se aluda la exclusión de la normativa. La no aplicación del RGPD debe centrarse en la posibilidad de que dicho dato no estructurado pueda vincularse o no a una determinada persona, estos es, se pueda identificar a esa persona con esa información. La identificabilidad de los datos se analizará más profusamente en el apartado sobre la anonimización y la seudonimización de datos personales del capítulo III de esta tesis[352].

C) Datos mixtos

Finalmente, por datos mixtos entendemos aquellas bases de datos que están conformadas por datos personales y no personales. Los datos mixtos son muy comunes en los contextos actuales donde se recopilan datos de diversas fuentes que pueden devenir tanto de personas como de todo tipo de máquinas y objetos que no analizan o tratan datos personales. Así, el artículo 2.2 del Reglamento Europeo de Datos No Personales indica que cuando exista un conjunto de datos mixtos, esta norma se aplicará únicamente al conjunto de datos no personales. Por otro lado, para el resto de datos, esto es, los datos personales, se habrán de aplicar las reglas establecidas en el RGPD. En el caso de que ambos tipos de datos estén totalmente unidos y sea imposible su separación, será aplicable el RGPD a todo el conjunto de datos[353]. Como es lógico, si ese conjunto de datos se puede dividir, lo normal es que cada conjunto de datos se almacene en distintos lugares para así poder aplicar distintas legislaciones a dichas bases de datos.

[351] Recordemos que el tratamiento de datos consistía en ir recopilando información puerta a puerta.
[352] Capítulo III, apartado IX de esta tesis.
[353] Esta norma tiene como objetivo principal la libre circulación de datos no personales, para ello, a lo largo de sus disposiciones se establecen una serie de mecanismos que reducen las trabas que hasta la fecha han limitado este principio. 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

III. LOS SUJETOS INTERVINIENTES DURANTE EN EL CICLO DE VIDA DE LOS SISTEMAS DE TOMA DE DECISIONES AUTOMATIZADAS

A lo largo de todo el proceso que abarca el desarrollo y despliegue de los sistemas automatizados es frecuente que participen multitud de agentes y distintas organizaciones. Cada una de estas organizaciones pueden asumir distintos roles y ostentar una mayor o menor presencia durante todo ese proceso y por tanto, sus implicaciones pueden resultar más o menos relevantes.

Desde el punto de vista de la normativa de protección de datos, es esencial que queden claros los distintos agentes que participan en los diversos tratamientos de datos personales que están presentes durante todas las fases que comprenden el desarrollo e implementación de los sistemas de toma de decisiones automatizadas. Para ello, la normativa de protección de datos establece distintas figuras a las cuales les asigna toda una serie de consecuencias legales relacionadas con el cumplimento de la normativa de protección de datos. Nos estamos refiriendo, al responsable, al corresponsable y al encargado del tratamiento.

1. El responsable del tratamiento

De acuerdo al artículo 4.7 del RGPD el responsable del tratamiento es:

> *la persona física o jurídica, autoridad pública, servicio u otro organismo que, solo o junto con otros, determine los fines y medios del tratamiento;*

El concepto de responsable resulta elemental en la normativa de protección de datos ya que este se convierte en el centro de imputación sobre el que recaen la mayoría de las responsabilidades y obligaciones en materia de protección de datos[354].

De esta definición se desprende que el elemento esencial que atribuye a un ente la condición de responsable es el hecho de valorar quién es el agente encargado de determinar los fines y medios del tratamiento de los datos personales. En este sentido, por lo que se refiere a *los fines del tratamiento*, esencialmente la norma se refiere al propósito u objetivo que persigue el responsable con las actividades que dan lugar a los distintos tratamientos de datos. Por tanto, una Administración Pública que

[354] LORENZO CABRERA,S: "Posición jurídica de los intervinientes en el tratamiento de datos personales. Medidas de cumplimiento". En: MURGA FERNÁNDEZ, J,P; FERNÁNDEZ SCAGLIUSI, M,A ; ESPEJO LERDO DE TEJADA,M: (dirs): *Protección de datos, responsabilidad activa y técnicas de garantía*. *Curso de delegado de protección de datos.* Ed. Reus, Madrid, 2018, pág. 118.

pretenda implantar un sistema de reconocimiento de voz basado en un *chat bot* para ofrecer distintos servicios públicos será la responsable del tratamiento de esos datos personales ya que establece la finalidad de dichos tratamientos. Por lo que se refiere a *los medios del tratamiento*, el Grupo del Artículo 29 hace una distinción de los mismos[355]. Así, los elementos esenciales ligados a los medios técnicos y organizativos han de ser establecidos por el responsable, estos pueden ser: los datos o categoría de datos que se utilizarán, el plazo de duración de los tratamiento, personas sobre las que se recopilarán los datos, los distintos tratamientos que se llevarán a cabo, etc. Por otro lado, el resto de decisiones ligadas a los medios técnicos u organizativos no esenciales, como pueden ser el modo de almacenamiento de los datos o la seguridad sobre los mismos podrán ser asumidos por el encargado del tratamiento y por tanto, la determinación de los mismos, no llevará a la consideración de ese agente como responsable.

A) El corresponsable del tratamiento

El RGPD ha decidido regular expresamente también la figura del corresponsable en el artículo 26[356]. En este sentido, este texto es consciente de que en muchos supuestos no sólo es una organización la que fija los fines para un mismo tratamiento sino que en determinadas ocasiones pueden actuar de forma conjunta varias entidades. Esta participación conjunta puede presentar un mayor o menor grado de interacción de manera que las combinaciones a la hora de fijar los fines y los medios pueden resultar diversos, y ello pese a que en algunos casos uno de los responsables ni si quiera tenga acceso a los datos personales en cuestión[357]. Por ejemplo, uno de los responsables puede fijar parte de los fines del tratamiento, mientras que el segundo establece los medios esenciales del tratamiento. En el ámbito del uso de sistemas automatizados esta figura puede adquirir gran relevancia ya que presumiblemente será frecuente que la organización que diseña el sistema marque las pautas principales del tratamiento de

[355] Grupo del Artículo 29. *Opinion 1/2010 on the concepts of controller and processor*. Adoptada el 16 de febrero de 2010, págs. 14 y 15. Recientemente, esta guía ha sido actualizada por el CEPD a través de las "Guidelines 07/2020 on the concepts of controller and processor in the GDPR" .Adoptadas el 2 de septiembre de 2020, apartados 37 a 39, págs. 13 a 15.

[356] Esta figura ya se establecía en la Directiva 95/46 en la definición de responsable. Sin embargo, ahora, el legislador europeo considera que es necesario un tratamiento regulatorio específico. Véase el artículo 2.d) de la Directiva 95/46.

[357] Sobre la figura del Corresponsable véase la Sentencia del TJUE (Sala Segunda) de 29 de julio de 2019. asunto C-40/17, CASO Fashion ID. Apartado 69. Resolución disponible en: https://curia.europa.eu/juris/document/document.jsf?text=&docid=216555&doclang=ES

datos como por ejemplo los datos personales a tratar pero, existirá una tercera organización que es la que usa el sistema, la cual, indica las finalidades del mismo.

2. El encargado del tratamiento

La figura de encargado se encuentra regulada en el artículo 28 del RGPD y definida en el artículo 4.8 del mismo texto jurídico. Así, se entiende por encargado a:

> *la persona física o jurídica, autoridad pública, servicio u otro organismo que trate datos personales por cuenta del responsable del tratamiento;*

Esta figura se diseñó con el objetivo de otorgar una respuesta adecuada y eficaz a los supuestos en los que el responsable descentraliza sus actividades de tratamiento de datos. En la definición de encargado se vislumbra el carácter de subordinación que tiene el mismo respecto del responsable del tratamiento, el cual, determina y establece las instrucciones bajo las cuales el encargado debe tratar los datos personales. Es decir, la única razón por la que el encargado trata esos datos personales es porque el responsable decide externalizar un servicio o actividad propia que conlleva en su caso uno o varios tratamientos de datos. Como se ha indicado anteriormente, el encargado podrá prever y establecer las medidas técnicas y organizativas no esenciales para cumplir con las exigencias establecidas por el responsable. La figura del encargado del tratamiento es muy común en el despliegue de sistemas automatizados. Así, por ejemplo, una Administración Pública puede contratar con terceras empresas un determinado sistema para desarrollar un concreto servicio público durante un determinado periodo de tiempo. En estos supuestos, esa tercera empresa ofrece el algoritmo mientras que la Administración Pública lo utiliza.

3. La interrelación de estas figuras y su dudosa incardinación en el ecosistema del ciclo de vida de los sistemas de toma de decisiones automatizadas

La implantación de estas figuras en la práctica jurídica no siempre ha sido tarea fácil y en muchos casos pueden existir dudas de si un encargado realmente realiza labores de un responsable de tratamiento o viceversa. Estas complicaciones pueden resultar aún más complejas cuando tratamos de encajarlas en los tratamientos de datos que están presentes durante las fases de desarrollo y despliegue de los sistemas de toma de decisiones automatizadas. Así, la participación de distintas organizaciones es muy

frecuente durante todo el ciclo de vida de un sistema de inteligencia artificial y además, la presencia de estos agentes puede ser más o menos relevante, lo que puede llevar a que, tal y como ha señalado la ICO, en distintas fases los roles de responsable y encargado se intercambien[358]. En la tabla siguiente se muestran algunos de los posibles tratamientos de datos personales que pueden aparecer implicados en las dos grandes fases que comprenden la materialización de sistemas inteligentes, esto es, las fases de desarrollo e implementación de los mismos.

	Posibles tratamientos de datos personales implicados
Fase de diseño	Planificación del proyecto y recopilación de datos Pre procesamiento de datos y limpieza de los mismos Desarrollo del modelo (fase de entrenamiento) Evaluación y elección del modelo
Fase de despliegue	Generación del resultado/Inferencia Adopción de la decisión Evolución del modelo

Teniendo en cuenta por tanto este conjunto de tratamientos y los posibles roles que pueden reflejarse en cada una de estas fases, resulta esencial valorar cuándo estamos ante la presencia de un responsable, encargado o corresponsables. Como antes hemos comprobado, la clave para considerar que un agente es responsable se deriva de la valoración de los criterios ligados a quién establece la finalidad y los medios esenciales del tratamiento.

Por lo que se refiere a la finalidad, a menudo resultará sencillo detectar al responsable a través de este criterio, y es que, aquella organización que marque o establezca el propósito para el cual se tratan los datos personales durante las distintas fases que comprenden la elaboración de estos sistemas será considerado responsable. Así, será considerado responsable aquél que establece como objetivo tratar los datos para realizar perfiles, tratar los datos para entrenar el modelo, tratar los datos para adoptar las decisiones, etc.

No resultará tan sencillo valorar o distinguir la persona que establece los medios esenciales para que se lleve a cabo el tratamiento. En este sentido, y trasladado al objeto de nuestro estudio, en muchos casos, distintas actuaciones relevantes relacionadas con

[358] Así lo ha indicado la Information Commissioner's Office, disponible en:
. https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/what-are-the-accountability-and-governance-implications-of-ai/#howshouldweunderstand

el desarrollo y despliegue del algoritmo pueden ser tomadas por diversas organizaciones y la posibilidad de que ambas sean consideradas corresponsables es relativamente alta. Así, en el capítulo anterior pudimos comprobar como a lo largo de todo el proceso que aglutina el desarrollo implementación de los sistemas de decisiones automatizados las organizaciones adoptan multitud de decisiones que son sumamente relevantes durante los diversos tratamientos de datos que se llevan a cabo. Muchas de estas decisiones consideramos que pueden ser relevantes e indiciarias para considerar que aquel agente que las tome, será considerado responsable del tratamiento en cuestión, entendiendo que las mismas son medidas técnicas y organizativas relevantes en relación con el artículo 4.7 del RGPD. En este orden de cosas, la ICO considera que estas decisiones pueden abarcar por ejemplo las siguientes: establecer la fuente y la naturaleza de los datos[359], elegir el algoritmo o algoritmos que desarrollarán el modelo[360], seleccionar las características o variables que en su caso se utilizarán para el modelo, adoptar las decisiones más relevantes relacionadas con los algoritmos[361] o decidir las técnicas y métodos de evaluación[362]. Junto a estas, nosotros también consideramos que dentro de este abanico de decisiones que pueden ser consideradas medidas técnico/organizativas encontramos la forma en que se recopilarán los datos[363], la alteración de variables incompletas o atributos atípicos[364], justificación de las variables elegidas y las correlaciones existentes entre las mismas[365], realización de pruebas piloto previas a la venta o despliegue del sistema[366],qué umbral de decisión se establece, qué interpretación se le otorga a las inferencias, se opta o no por la plena automatización o no del resultado, etc.

[359] También así lo ha indicado la Agencia Española de Protección de Datos . *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, págs.17 a 19.
[360] Árbol de decisión, naives, bosque aleatorio, regresión logística, máquina de soporte vectorial, etc.
[361] Cuántas ramas tendrá el árbol de decisión, cuántos árboles compondrán el bosque aleatorio, cuántas capas conformarán la red neuronal, número de épocas para la optimización de la red, en cuántos grupos se dividirá el análisis del algoritmo k-means, etc.
[362] Se utiliza la validación cruzada, métricas de la matriz de confusión, etc.
[363] Por ejemplo, si los datos provienen directamente de los individuos, de sensores, de aparatos basado en el internet de las cosas, etc.,
[364] Decidir qué hacer con los atributos atípicos, eliminar muestras repetidas o incompletas, sustitución de los valores faltantes de una muestra por aquellos que presentan mayor frecuencia absoluta, etc.
[365] Realizar operaciones que demuestren la pertinencia de esas correlaciones, al menos, descartando aquellas que pueden resultar por azar o espurias.
[366] Realizar pruebas piloto por ejemplo con los llamados equipos rojos o simulación real del entorno al que se enfrentará el sistema.

Fases	Decisiones que pueden considerarse medidas técnicas y organizativas esenciales a efectos de considerar a una organización como responsable del tratamiento
Planificación del proyecto y recopilación de datos	-Qué datos se van a utilizar para crear el modelo. -Cómo van a ser recopilados esos datos.
Pre procesamiento de datos y limpieza de los mismos	-Alteración de variables inadecuadas/incompletas. -Reducir el número de variables. -Análisis de los datos.
Desarrollo del modelo (fase de entrenamiento)	-Algoritmo/s que se van a utilizar para crear el modelo. -Decisiones relevantes sobre el/los algoritmos elegidos.
Evaluación y elección del modelo	-Métricas utilizadas para testar el sistema. -Evaluación del sistema y separación de los datos. -Elección del modelo/pruebas piloto.
Generación del resultado/Inferencia	-Quién/qué interpreta el resultado/inferencia. -Conversión automática o no del resultado. -Umbral de decisión.
Adopción de la decisión	-Quién adopta la decisión. -Papel del humano/máquina en el resultado.
Evolución del modelo	-Quién controla las actualizaciones y adaptaciones del modelo.

Elaboración propia

Por tanto, en presencia de varios agentes en distintas fases, resultará esencial valorar todos elementos necesarios que permitan encuadrar a cada uno de los agentes en las figuras jurídicas del RGPD a las cuales se les imponen distintas obligaciones en materia de cumplimiento de esta normativa.

Cabe mencionar que la PRAI define todo un conjunto de sujetos que están presentes durante toda la cadena de valor o ciclo de vida de los sistemas de inteligencia artificial catalogados de alto riesgo. Estos son: los proveedores, importadores, distribuidores, representantes autorizados y usuarios. Especialmente conviene prestar atención a dos, estos son, los proveedores y los usuarios. Así, resumidamente, los proveedores son aquellos que desarrollan los modelos algorítmicos y los usuarios son aquellos que los utilizan[367]. Sorprendentemente, esta norma no hace referencia en

[367] Así, por proveedor se entiende: *una persona física o jurídica, autoridad pública, agencia u otro organismo que desarrolle un sistema de IA o que haga desarrollar un sistema de IA con el fin de comercializarlo o ponerlo en servicio bajo su propio nombre o marca, ya sea a cambio de una remuneración o de forma gratuita*; A su vez, usuario es: *cualquier persona física o jurídica, autoridad pública, organismo o cualquier otra entidad que utilice un sistema de IA bajo su autoridad, excepto cuando el sistema de IA se utilice en el curso de una actividad personal no profesional*; Véase el artículo 3 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

ningún momento a los afectados por estos sistemas, es decir, las personas sobre las que se toman las decisiones algorítmicas. El encaje de estos sujetos y las obligaciones que impone a los mismos esta propuesta normativa puede generar ciertas distorsiones sobre la normativa de protección de datos y las figuras de responsable, encargado o corresponsable[368].

	Sujetos presentes durante las fases del ciclo de vida de los sistemas	
	Diseño	**Despliegue**
PRAI	Proveedor	Usuario o Proveedor[369]
RGPD	Responsable, encargado o corresponsable	Responsable, encargado o corresponsable

En las siguientes líneas mantendremos el foco puesto en el análisis de las interrelaciones que se generan entre los agentes descritos por el RGPD, si bien, será necesario ir realizando referencias a las previsiones normativas dispuestas por la PRAI que puedan ser relevantes para el objeto de estudio de esta tesis.

A) Implicaciones y roles en la fase de diseño

Para el diseño de los sistemas resulta habitual que muchas organizaciones acudan a terceras empresas para crear sus modelos. Así, como dijimos en páginas anteriores, a día de hoy existe un número importante de empresas que ofrecen todo tipo de herramientas alojadas en la nube para que terceras organizaciones, ya sean públicas o privadas, puedan hacer uso de los mismas para crear sus propios modelos[370]. Teniendo en cuenta que el tratamiento de datos que se lleva a cabo en estos entornos da lugar al entrenamiento y desarrollo de modelos. Cabría preguntarse qué papel pueden jugar esas empresas que ofrecen servicios *cloud* y aquellos que en su caso los explotan. En este

[368] COTINO HUESO,L; CASTILLO PARRILLA, J,A; SLAZAR,I; BENJAMINS,R; CUMBRERAS,M; ESTEBAN,A,M: "Un análisis crítico constructivo de la Propuesta de Reglamento de la Unión Europea por el que se establecen normas armonizadas sobre la Inteligencia Artificial". *Diario La Ley*, 2 de Julio de 2021, pág.8.

[369] En la fase de despliegue, el usuario se considerará proveedor cuando: i) comercialice o ponga en servicio un sistema de IA de alto riesgo bajo su nombre o marca; ii) modifique la finalidad prevista de un sistema de IA de alto riesgo ya comercializado o puesto en servicio; iii) realice una modificación sustancial del sistema de IA de alto riesgo. Artículo 28 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

[370] Herramientas como *TensorFlow* de Google, *Microsoft Azure* de Microsoft, *Amazon web series* de Amazon o *Deep learning as a service* de IBM. Como dijimos, estos servicios ofrecen a un coste relativamente bajo algoritmos ya desarrollados dispuestos a procesar los datos que aporten los clientes. En: PANESAR, A: *Machine Learning and AI for Healthcare. Big Data for Improved Health Outcomes*. Ed. Apress, Berkeley, 2021, pág.152. Disponible en: https://link.springer.com/book/10.1007/978-1-4842-6537-6#toc

sentido, resulta necesario acudir nuevamente al concepto de responsable y encargado. Así, a priori, el responsable del tratamiento ha de considerarse aquella organización que hace uso de estos servicios ya que normalmente son estas las que deciden qué modelos crear y qué datos aportar a estas herramientas, fijando en su caso los objetivos que pretende con ese entrenamiento. En el otro lado, la empresa de *cloud*, dado que en principio sólo facilita la herramienta para que se puedan desarrollar y entrenar los modelos, sus funciones serían meramente de encargado. Y es que, si bien se tratan datos personales en dichas herramientas, el tratamiento de los mismos responde a las finalidades que hayan previsto los usuarios que hacen uso del servicio *cloud*. Dicho lo anterior, a medida que la empresa que ofrece el servicio *cloud* tenga mayor impronta durante el tratamiento de datos, las posibilidades de corresponsabilidad aumentan ya que se puede llegar a entender que también participa en la imposición de los fines y los medios[371]. En este sentido conviene indicar que la AEPD ya ha señalado que a priori, el hecho de que esa empresa *cloud* ofrezca sus propios datos para los tratamientos que lleven a cabo terceras organizaciones, no será motivo para considerar a dicha empresa *cloud* como responsable, siempre que dicho tratamiento de los mismos tengo como finalidad los fines del responsable, esto es, de la organización usuaria del servicio *cloud*[372].

En otros supuestos, las implicaciones entre el que ofrece la tecnología y el que la explota suelen ser más intensas y las decisiones que se van tomando sobre los distintos tratamientos son generalmente adoptadas de forma conjunta[373]. A modo de ejemplo ya comentado, en 2016, la empresa privada *DeepMind* comenzó a trabajar con la *Royal Free London NHS Foundation Trust* para desarrollar una aplicación que ayudara con el diagnóstico de la lesión renal aguda[374]. Este proyecto acabó implantándose en los hospitales de la Royal NHS. Esta última fundación facilitó en torno a 1,6 millones de

[371] El artículo 33.2 párrafo segundo de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales establece que *tendrá asimismo la consideración de responsable del tratamiento quien figurando como encargado utilizase los datos para sus propias finalidades*.

[372] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.17

[373] El Tribunal de Justicia de la Unión Europea ha considerado que *una persona física o jurídica que, atendiendo a sus propios objetivos, influye en el tratamiento de datos personales y participa, por tanto, en la determinación de los fines y los medios del tratamiento puede ser considerada responsable del tratamiento en el sentido del artículo 2, letra d), de la Directiva 95/46.* Sentencia del TJUE de 10 de julio de 2018, asunto Jehovan todistajat, C-25/17, EU:C:2018:551, apartado 68. Disponible en: http://curia.europa.eu/juris/document/document.jsf;jsessionid=13BFBBCF55443B41C12D157261F5611C?text=&docid=203822&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=697640

[374] Más información sobre el proyecto y la aplicación creada en: https://www.royalfree.nhs.uk/patients-visitors/how-we-use-patient-information/our-work-with-deepmind/

historiales de pacientes a *DeepMind* para crear tal sistema. Durante este proyecto, ambas organizaciones, tanto pública como privada, trataron todo tipo de datos personales y en muchas ocasiones las decisiones sobre cómo debían tratarse esos datos y que procesos había que aplicarles a los mismos se adoptaron de forma conjunta. Posiblemente en no todas las etapas la presencia fue conjunta, pero sí en algunas de ellas. Esta situación puede resultar sumamente compleja a hora de asignar los roles de responsable y encargado.

B) Implicaciones y roles en la fase de aplicación y toma de decisiones

Los problemas a la hora de asignar la figura jurídica de responsable y encargado de tratamiento también están presentes en la fase temporal en la cual, una organización distinta a la que diseño el sistema, lo acaba adquiriendo y utilizando el algoritmo para la toma de decisiones. Ello es así debido a la fuerte interrelación que existe entre la fase del diseño del sistema y la aplicación del mismo para tomar decisiones. Así, resultará habitual que la misma organización que diseñó el sistema no sólo ofrezca la explotación del mismo, sino además, también se facilite el mantenimiento una vez que comienza a desplegar sus efectos el algoritmo en el entorno real. Pues bien, hay que distinguir distintas situaciones.

En *primer lugar*, los roles quedarán meridianamente claros en aquellos supuestos en los que la organización que desarrolló el sistema únicamente ofrece el soporte técnico del sistema pero no interviene en las decisiones fundamentales que se derivan de los tratamientos de datos personales. Tareas como facilitar la herramienta, controlar su correcto funcionamiento, mantenerlo actualizado e incluso prever determinadas medidas de seguridad para evitar ataques al sistema que adopta decisiones automatizadas entrarían dentro de las posibles funciones básicas que llevaría a cabo una organización considerada como encargado del tratamiento. Esta práctica puede ser habitual entre Administraciones Públicas y terceros cuando los mismos conciertan el suministro de este tipo de herramientas[375].

[375] A modo de ejemplo, el Consorcio de la Corporación Sanitaria del Parque Tulí de Sabadell licitó un contrato con el objetivo de que empresas privadas ofrecieran una herramienta que ayude en el soporte de decisiones en la unidad de asistencia de resolución avanzada mejorando el sistema de triaje de urgencias hospitalarias. Se pretende así reducir el tiempo de espera de los pacientes de urgencias, así como mejorar su experiencia durante la estancia en el servicio. Véase: https://contractaciopublica.gencat.cat/ecofin_pscp/AppJava/es_ES/notice.pscp?idDoc=72768054&reqCode=viewCn

Dicho lo anterior, no obstante, cabe destacar que la PRAI descarga sobre los proveedores/fabricantes de sistemas de alto riesgo, los cuales serán frecuentemente aquellos que desarrollaron el modelo, toda una serie de obligaciones ligadas al mantenimiento del sistema una vez que este se comercialice. Concretamente, el artículo 61 del proyecto legislativo mencionado obliga a estos proveedores a realizar un seguimiento posterior de estos sistemas tras su comercialización, permitiendo así una evaluación continua de los mismos. Esto supone que, en muchas situaciones, aquella organización que desarrolló el sistema tendrá la obligación de mantener un seguimiento del algoritmo que está utilizando una tercera organización[376]. Esas tareas asignadas a dicho proveedor generalmente conllevarán el tratamiento de datos personales[377]. Pues bien, en la medida que tales obligaciones son directamente asignadas a los proveedores por parte de la norma[378], hay que entender que, para ese concreto tratamiento de datos, el responsable del tratamiento será considerado el proveedor/fabricante del sistema, el cual, en algunos supuestos será la organización que diseño el sistema y no la que lo está utilizando.

En *segundo lugar*, el asunto se vuelve más complejo en los supuestos en los que entre el que ofrece el sistema y el que lo explota se mantienen determinadas relaciones que van más allá del mero servicio técnico o mantenimiento del sistema. Por ejemplo, una organización puede ofrecer un determinado sistema a una Administración Pública

[376] La mencionada propuesta establece diferentes agentes presentes en este ciclo. Por un lado, define al proveedor como la *persona física o jurídica, autoridad pública, agencia u otro organismo que desarrolle un sistema de IA o que haga desarrollar un sistema de IA con el fin de comercializarlo o ponerlo en servicio bajo su propio nombre o marca, ya sea a cambio de una remuneración o de forma gratuita*. Por otro lado, el usuario es *cualquier persona física o jurídica, autoridad pública, organismo o cualquier otra entidad que utilice un sistema de IA bajo su autoridad, excepto cuando el sistema de IA se utilice en el curso de una actividad personal no profesional*. Véase los artículos 3 y 61 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

[377] El artículo 61.2 de la propuesta de Reglamento mentada establece que: *El sistema de seguimiento postcomercialización recogerá, documentará y* ***analizará*** *de forma activa y* ***sistemática los datos pertinentes proporcionados*** *por los usuarios o recogidos a través de otras fuentes sobre el rendimiento de los sistemas de IA de alto riesgo a lo largo de su vida útil, y permitirá al proveedor evaluar la conformidad continua de los sistemas de IA con los requisitos establecidos en el capítulo 2 del título III.*(la negrita y la cursiva son nuestras).

[378] El CEPD ha considerado que la consideración de responsable puede devenir de la propia norma cuando esta impone a algún sujeto la obligación de tratar determinados datos. En estos casos, la ley suele determinar la finalidad del tratamiento. Por tanto, el responsable del tratamiento será normalmente el designado por la ley para la realización de esta finalidad. En: Comité Europeo de Protección de Datos. "Guidelines 07/2020 on the concepts of controller and processor in the GDPR" .Adoptadas el 2 de septiembre de 2020, pág, 11, apartado 22. En el mismo sentido véase: DURÁN CARDO, B: *La figura del responsable en el derecho a la protección de datos*. Ed. Wolters Kluwer, Madrid, 2016, págs. 140 a 142.

para que esta preste un determinado servicio público. En principio, ese sistema seguiría en propiedad de la empresa, la cual, permite su explotación a la Administración Pública. En estos supuestos, a priori, la Administración Pública es el responsable del tratamiento de datos que se lleva a cabo, en este caso el servicio que se esté prestando. A su vez, el encargado sería la empresa tercera si únicamente facilita la herramienta y en su caso trata datos personales para cumplir con las exigencias previamente establecidas por la Administración Pública. Ahora bien, en aquellos supuestos en los que esa empresa se beneficie de los tratamientos llevados a cabo por la Administración como por ejemplo a través de la mejoría del algoritmo, en la medida que esta también interviene en ese tratamiento para obtener un beneficio con un fin distinto para el cual la Administración Pública ofrece el servicio, podrá considerarse a esa empresa como corresponsable del tratamiento de datos. En este sentido, tanto la AEPD como la autoridad de protección de datos de Portugal ya han dejado claro que un encargado puede convertirse en responsable cuando, ofreciendo un servicio de IA para su explotación a un tercero, utilice los datos de ese tratamiento para fines propios distintos a los que se derivan del tratamiento que lleva a cabo el usuario[379]. Por ejemplo, ello puede ser frecuente en aquellos sistemas dinámicos o adaptativos que están continuamente mejorando ya que la organización que ofrece el sistema puede beneficiarse de la mejora en la precisión del mismo. Desde el momento que esa organización que facilita la herramienta algorítmica se aproveche de los nuevos desarrollos del sistema y por tanto se estén tratando los datos para otras finalidades, esta organización deberá considerarse responsable del tratamiento[380].

En definitiva, queda claro por tanto que, los roles establecidos por la normativa de protección de datos pueden quedan difuminados en muchas ocasiones por las distintas funciones que ejercitan diversos agentes durante todo el ciclo de vida de un

[379] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, págs. 17 y 18. Véase el cuadro específico que aparece en esas páginas. En el mismo sentido se ha pronunciado la autoridad de protección de datos portuguesa, esta última consideró responsable del tratamiento a la empresa que había puesto a disposición de una Administración Pública una aplicación de reconocimiento facial. Esta empresa no sólo ofrecía el producto a la Administración sino que además se beneficiaba de los datos recopilados para mejorar el sistema algorítmico. Comissão Nacional de Proteção de Dados. Deliberação 622/2021. Apartados 56 y 57. Resolución disponible en: https://www.cnpd.pt/decisoes/deliberacoes/

[380] De acuerdo al artículo 33.2 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. *Tendrá asimismo la consideración de responsable del tratamiento quien figurando como encargado utilizase los datos para sus propias finalidades*. (la cursiva es nuestra).

sistema de inteligencia artificial. En este sentido, será necesario que las distintas responsabilidades que se asignan a cada uno de los agentes participantes durante estas fases queden en todo momento determinadas y clarificadas. Así, será muy relevante que en los contratos de encargo entre responsable y encargado de tratamiento estos elementos aparezcan fijados tal y como prevé el artículo 28.3 del RGPD. Junto al contrato de encargo, en los casos en los que se considere que están presentes varios responsables, también será necesario que las responsabilidades y decisiones que cada uno de estos adopta en los distintos tratamientos se fijen adecuadamente en el acuerdo que formalicen (Artículo 26 del RGPD). Esto último podrá clarificar las posibles responsabilidades en caso de incumplimiento de la normativa de protección de datos[381]. Dicho lo anterior, e independientemente de las figuras que se hayan asignado a cada uno de los agentes en estos acuerdos o contratos, dado que los conceptos de responsable y encargado son funcionales[382], esto es, responden al papel real que estos desempeñan, habrá que estar a la funciones concretas que de facto realizan durante el ciclo de vida de los sistemas automatizados para considerar que nos encontramos realmente ante las figuras indicadas.

[381] Como ha indicado el Tribunal de Justicia TJUE (Sala Segunda) de 29 de julio de 2019 . ECLI:EU:C:2019:629 sobre la figura del Corresponsable, la responsabilidad conjunta no supone necesariamente que, con respecto a un mismo tratamiento de datos personales, los diversos agentes tengan una responsabilidad equivalente. Bien al contrario, los agentes pueden estar implicados en distintas etapas del tratamiento y en distintos grados, de modo que el nivel de responsabilidad de cada uno de ellos debe evaluarse teniendo en cuenta todas las circunstancias pertinentes del caso concreto. (apartado 76). Véase en este mismo sentido la Sentencia del TJUE de 10 de julio de 2018, asunto C-25/17, caso Jehovan todistajat. Apartado 66. Resolución disponible en: https://curia.europa.eu/juris/document/document.jsf?text=&docid=203822&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=6143211

[382] Comité Europeo de Protección de datos. *Guidelines 07/2020 on the concepts of controller and processor in the GDPR* .Adoptadas el 2 de septiembre de 2020, aparatados 12 y 49, págs. 9 a 17.

CAPÍTULO III. LAS MEDIDAS DE RESPONSABILIDAD ACTIVA BASADAS EN EL ENFOQUE DEL RIESGO DURANTE EL USO DE SISTEMAS DE TOMA DE DECISIONES AUTOMATIZADAS

El capítulo IV del RGPD configura un nuevo estatuto de cumplimiento de la normativa de protección de datos para el responsable y el encargado del tratamiento basado en el principio de responsabilidad proactiva o *accountability*[383]. Este principio aparece reconocido en el artículo 5.2 del RGPD señalando que el responsable del tratamiento no sólo será responsable del cumplimiento de los principios relativos al tratamiento de datos sino que además deberá ser capaz de probar el cumplimiento de los mismos. Este principio supone un cambio de paradigma en la concepción del enfoque que han de adoptar las organizaciones a la hora de cumplir con la normativa de protección de datos, pasando de una visión reactiva, cuyo principal objetivo es actuar ante el incumplimiento, a una perspectiva proactiva. Para afrontar este cambio, el RGPD configura toda una serie de obligaciones y medidas que han de implantar los responsables con el doble objetivo de, por un lado, prever y por consiguiente reducir los posibles riesgos en los derechos de los individuos que pueden generar los tratamientos de datos personales que llevan a cabo y por otro lado, demostrar que efectivamente se está cumpliendo con la normativa de protección de datos.

Materialización principio de responsabilidad activa/accountability. Artículo 5.2	
Medidas técnico/organizativas de responsabilidad activa	-Obligación general de establecer medidas. Artículo 24. -Análisis del riesgo. Artículo 24 y Considerando 76. -Privacidad desde el diseño y por defecto. Artículo 25. -Nombrar representante. Artículo 27. -Elección adecuada de encargados de tratamiento. Artículo 28. -Llevanza registro de actividades de tratamiento. Artículo 30. -Seguridad del tratamiento. Artículo 32. -Notificación violación de seguridad. Artículo 33. -Comunicación violación de seguridad. Artículo 34. -Evaluación de impacto y consulta previa. Artículo 35 y 36. -Designar Delegado de protección de datos. Artículo 37. -Códigos de conducta y certificaciones. Artículos 40 y 42.

[383] El principio de responsabilidad activa en materia de protección de datos fue reconocido por vez primera por la OCDE en sus "Directrices sobre protección de la privacidad y flujos transfronterizos de datos personales" de 1980. Texto disponible en: https://www.oecd.org/sti/ieconomy/15590267.pdf

Así, la articulación jurídica del principio de responsabilidad activa comienza con el artículo 24 del RGPD al señalar que:

> *Teniendo en cuenta la naturaleza, el ámbito, el contexto y los fines del tratamiento así como los riesgos de diversa probabilidad y gravedad para los derechos y libertades de las personas físicas, el responsable del tratamiento aplicará medidas técnicas y organizativas apropiadas a fin de garantizar y poder demostrar que el tratamiento es conforme con el presente Reglamento. Dichas medidas se revisarán y actualizarán cuando sea necesario.*

De esta manera, de este precepto se denota claramente el doble objetivo indicado previamente, es decir, el responsable ha de prever toda una serie de medidas que, por un lado, garanticen que el tratamiento cumple con la normativa de protección de datos y por otro lado, que demuestren que se cumple dicha normativa. Corresponde por tanto al responsable decidir qué medidas implantar, si bien, las mismas quedan condicionadas al riesgo que en su caso puede conllevar el tratamiento de datos proyectado para los derechos y libertades de los interesados. Para valorar ese riesgo se ha de tener en cuenta la naturaleza, el ámbito, el contexto y los fines del tratamiento (considerando 78 RGPD). Se evita así un enfoque de "talla única" que deriva normalmente en el completado de una *checklist*[384]. A modo de ejemplo, la antigua normativa española de protección de datos preveía diferentes medidas de seguridad en función de tres niveles, estos eran; básico, medio y alto. Dependiendo del tipo de dato y la finalidad del tratamiento, cada fichero era asignado a un nivel. Cada nivel tenía estipuladas una serie de medidas similares aplicables a los datos que formaban parte de dicho nivel. De esta manera, dichas medidas no eran personalizadas en función del riesgo que pudiera presentar ese concreto tratamiento sino del nivel asignado[385]. Con el enfoque basado en el riesgo esto cambia. Así, no presentará los mismos riesgos para los derechos de los particulares un sistema de recomendación de películas a los que se pueden encontrar en una aplicación que se encarga de predecir enfermedades. Es por ello que las medidas previstas también

[384] Grupo del Artículo 29. *Dictamen 3/2010 sobre el principio de responsabilidad.* Adoptado el 13 de julio de 2010, apartado 45, pág.14.

[385] Véase los artículos 79, 80 y 81 del Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal.

diferirán. El llamado "enfoque basado en el riesgo[386]" no es un desconocido en la normativa de protección de datos, este ya se reconocía en la Directiva 95/46/CE a la hora de regular las medidas de seguridad del tratamiento y las obligaciones de control previo para legitimar determinados tratamiento[387]. Lo que ocurre ahora es que el RGPD lo vehicula como uno de los elementos basilares sobre el cual se han de diseñar unas medidas u otras de responsabilidad activa que le permitan al responsable aplicar y demostrar el cumplimiento de las normas jurídicas abstractas que se derivan del RGPD de forma adecuada y eficaz[388]. De esta manera, resulta sumamente relevante identificar los posibles riesgos que presenta un tratamiento. Así, el RGPD hace referencia a un conjunto de tratamientos que entrañan riesgos para los derechos y libertades de los individuos como pueden ser aquellos que entre otros; den lugar a discriminación, pérdidas financieras, suplantación de identidad, revelación e datos inferidos sensibles, elaboración de perfiles, tratamiento de grandes cantidades de datos, etc. (Considerando 75)

Junto a estos riesgos expresamente mencionados por la norma, corresponde al responsable realizar una evaluación objetiva que trate de identificar y en su caso valorar los riesgos que puedan presentar los tratamientos proyectados (Considerando 76 RGPD). Dicha evaluación objetiva vendrá de la mano de la herramienta del análisis de riesgos, la cual, será posteriormente analizada.

Además de servir como elemento básico para establecer medidas técnico/organizativas genéricas, el enfoque del riesgo también está presente para activar otra serie de mecanismos específicos que derivan del principio de responsabilidad activa. De esta manera, el factor riesgo también es relevante a la hora de implantar medidas como la privacidad desde el diseño y por defecto, el nombramiento de representantes, la llevanza del registro de actividades de tratamiento, la implantación de mayores o menores medidas de seguridad o las notificaciones de seguridad. A su vez, en los casos en los que se considere que un tratamiento entraña un grave riesgo, se requerirán de la implantación de medidas

[386] El enfoque del riesgo aparece reflejado en un gran número de preceptos y considerandos del RGPD. Así, véase entre otros, los considerandos 51,71, 74,75,76,77,80,83,84,85,86,89,90,91,96,98. Artículos 24, 25, 27, 30, 32, 33,34, 35, 36,39.

[387] Artículos 17.1 y 20.1 de la Directiva 95/46 del Parlamento Europeo y del Consejo, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos.

[388] QUELLE, C: "Enhancing Compliance under the General Data Protection Regulation: The Risky Upshot of the Accountability and Risk-based Approach". *European Journal of Risk Regulation*, *9*(3), 2018, pág.503.

adicionales como la evaluación de impacto o la comunicación a los interesados de una violación de seguridad. Este enfoque resulta esencial en todos los tratamientos de datos personales implicados durante el ciclo de vida de desarrollo y despliegue de los sistemas de decisiones automatizadas ya que permite anticiparse a las posibles amenazas que pueden generarse por el uso de los mismos tanto en el diseño como en la fase previa al despliegue de los mismos[389]. Además, dado que los sistemas de toma de decisiones automatizadas pueden incorporarse a entornos muy diferentes y los propósitos de los mismos y las consecuencias de estos pueden presentar diversos peligros, el enfoque el riesgo permite la adaptación de estas medidas de cumplimiento a las exigencias del mismo. En definitiva, el principio de responsabilidad activa en relación con el enfoque del riesgo obliga a los responsables y encargados de tratamiento que desarrollen y desplieguen sistemas de toma de decisiones automatizadas a diseñar todo un *traje a medida* de aquellas medidas –valga la redundancia- técnico organizativas que logren reducir los riesgos potenciales que estos sistemas pueden generar en los derechos de los particulares.

Cabe destacar no obstante que independientemente del riesgo que un tratamiento pueda generar en los derechos y libertades de los individuos, los derechos de acceso, rectificación, supresión, oposición, información, olvido, o portabilidad de datos entre otros no pueden verse restringidos o limitados por el menor o mayor riesgo que dichos tratamientos puedan comportar. Las obligaciones materiales derivadas de la normativa de protección de datos se mantienen intactas con independencia del nivel de riesgo que se haya podido identificar[390]. Así, no sería lícito reemplazar cualquiera de los principios del RGPD por medidas técnicas y organizativas encaminadas a sustituir dichos principios o a mitigar las posibles consecuencias que dicha falta de cumplimiento pudiera tener sobre los interesados afectados[391].

Finalmente, junto con el enfoque del riesgo, existen otra serie de elementos que también son considerados relevantes a la hora de establecer mayores obligaciones que pretendan demostrar el cumplimiento de la normativa de protección de datos. Así,

[389] Anticiparse a los riesgos resulta esencial en el desarrollo y despliegue de los sistemas de inteligencia artificial. Anticiparse a los riesgos resulta esencial en el desarrollo y despliegue de los sistemas de inteligencia artificial. PANESAR, A: *Machine Learning and AI for Healthcare. Big Data for Improved Health Outcomes*, op cit., pág. 232 in fine**.**

[390] Grupo del Artículo 29. *Statement on the role of a risk-based approach in data protection legal frameworks.* Resolución adoptada el 30 de mayo de 2014. Págs 3 y 4.

[391] Agencia Española de Protección de Datos. *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales.* Junio 2021, pág.18.

factores como el estado de la técnica con relación a los avances tecnológicos[392], el coste económico de aplicación de las medidas organizativas, el ámbito donde se ejecute el tratamiento, esto es, el sector público o privado, el tratamiento de datos de categoría espacial o en su caso el tratamiento de los mismos a gran escala también resultan indicadores que orientarán la incorporación de unas u otras medidas de cumplimiento.

	Elementos a tener en cuenta para implantar las medidas			
Medidas a implementar	**Riesgo para los derechos y libertades de las personas físicas con relación a la naturaleza ámbito, contexto y fines del tratamiento**	**Riesgo para los derechos y libertades de las personas físicas con relación al estado de la técnica, coste de aplicación la naturaleza ámbito, contexto y fines del tratamiento**	**Alto riesgo para los derechos y libertades de las personas físicas**	**Datos de categoría especial**
Medidas y obligaciones generales. (art.24)	✖			
Privacidad desde el diseño. (art.25)	✖	✖		
Nombrar representante. (art.27)	✖			✖
Registro de actividades de tratamiento. (art.30)	✖			✖
Seguridad del tratamiento. (art.32)	✖	✖		
Notificación violación de seguridad. (art.33)	✖			
Comunicación violación de seguridad. (art.34)	✖		✖	
Evaluación de impacto. (art.35)	✖		✖	
Delegado de protección de datos.(art.37)				✖

De esta manera, tanto de las exigencias generales que se derivan del artículo 24, como del conjunto de medidas que aparecen expresamente incorporadas en el

[392] El responsable debe estar al día de los avances tecnológicos que puedan generar riesgos para los tratamientos de datos. Comité Europeo de Protección de Datos. *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*. Directrices adoptadas el 13 de Noviembre de 2019. Apartado 19, pág.8.

capítulo IV del RGPD se desprende claramente como el legislador europeo ha trasladado al responsable toda una serie de obligaciones que exigen del mismo un cumplimiento de la normativa de protección de datos que hasta la fecha no estaba presente en legislaciones previas. Es decir, el responsable ya venía obligado a cumplir con las exigencias materiales derivadas de la normativa de protección de datos como pueden ser el cumplimiento de los principios en materia de protección de datos o el atendimiento adecuado al ejercicio de los derechos de los interesados. Ahora, el principio de responsabilidad proactiva exige además un plus de medidas que responden al cómo se protegen precisamente los derechos y libertades de los titulares cuyos tratamientos de datos se llevan a cabo. Estas medidas de cumplimiento escalables en función de los factores indicados previamente constituyen una base mínima de cumplimiento de la normativa de protección de datos en materia de responsabilidad activa. Ahora bien, nada impide que un responsable pueda implantarlas cuando no esté obligado[393]. Por ejemplo, se puede designar un delegado de protección de datos aunque un responsable no esté obligado a ello.

Como ahora se analizará, la mayoría de los tratamientos que están presentes durante el ciclo de vida de los sistemas de decisiones automatizados requerirán la implementación de prácticamente la totalidad de medidas de responsabilidad activa que prevé el RGPD en su capítulo IV. Ya sea porque dichos tratamientos entrañan riegos o en su caso alto riesgo, o porque tratan una gran cantidad de datos, lo cierto es que en la mayoría de la ocasiones la medidas que deberán implantarse requerirán de la mayor protección que el RGPD brinda para los tratamientos de datos. Es turno por tanto de analizar las distintas medidas y obligaciones que se desprenden del principio de responsabilidad activa y el enfoque del riesgo centrándonos en las principales implicaciones que se pueden derivar cuando se llevan a cabo tratamientos de datos personales durante todo el ciclo de vida de un sistema de toma de decisiones automatizadas.

[393] Grupo del Artículo 29. *Dictamen 3/2010 sobre el principio de responsabilidad.* Adoptado el 13 de julio de 2010, apartado 14, pág.6.

I. EL ANÁLISIS DE RIESGOS PARA LOS DERECHOS Y LIBERTADES DE LAS PERSONAS FÍSICAS

La gestión del riesgo para los derechos y libertades requiere de toda una serie de pautas y pasos a seguir por parte del responsable del tratamiento. La técnica que se muestra esencial en estos contextos es el llamado análisis de riesgos focalizado en el riesgo que para los derechos y libertades puede ocasionar un tratamiento de datos a los interesados (Artículo 24 y Considerando 76 del RGPD).

Así, esta herramienta es un proceso que permite establecer una serie de salvaguardas que minimicen a niveles aceptables los riesgos que pueden sufrir los derechos y libertades de las personas con relación al tratamiento de sus datos personales[394]. Como señala la AEPD[395], el análisis de riesgos presenta esencialmente tres fases fundamentales; en la primera se identifican las amenazas existentes, en la segunda se evalúa el nivel de riesgo existente y en la tercera se gestiona ese riesgo mediante la implementación de medidas técnicas y organizativas para eliminar o mitigar dichos riesgos, reduciendo así el impacto o la probabilidad de que se materialicen las amenazas detectadas. Una vez que han sido implantadas las medidas adecuadas se ha de evaluar el riesgo residual remanente para mantenerlo controlado. Conviene señalar que el RGPD en ningún momento establece que tipo de metodología o análisis ha de implantarse sino que más bien identifica los posibles riesgos que pueden considerarse que están patentes en los tratamientos de datos y deja a los responsables del tratamiento la gestión de los mismos, gestión que puede seguir distintas metodologías como son las Normas ISO[396], MAGERIT v.3[397] , las guía sobre gestión de riesgo creadas por la

[394] DE LA PRADA, ESPINA, D: "Análisis y gestión de riesgos de los tratamientos de datos personales". En: MURGA FERNÁNDEZ, J,P; FERNÁNDEZ SCAGLIUSI, M,A ; ESPEJO LERDO DE TEJADA,M: (dirs): *Protección de datos, responsabilidad activa y técnicas de garantía*. *Curso de delegado de protección de datos.* Ed. Reus, Madrid, 2018, pág.349.

[395] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág 30.

[396] Las normas ISO son un conjunto de documentos que especifican requerimientos que pueden ser empleados en organizaciones para garantizar que los productos y/o servicios ofrecidos por dichas organizaciones cumplen con su objetivo. Estas normas son establecidas por el Organismo Internacional de Estandarización (ISO) y, concretamente, la serie de normas ISO 3100 y 27001 se encargan de fijar y establecer toda una serie de principios que ayudan a la gestión de los riesgos. Visto en: https://www.isotools.org/normas/

[397] MAGERIT, acrónimo de "Metodología de Análisis y Gestión de Riesgos de los Sistemas de Información" , es la metodología de análisis y gestión de riesgos elaborada por el antiguo Consejo Superior de Administración Electrónica (actualmente Comisión de Estrategia TIC). La razón de ser de MAGERIT está directamente relacionada con la generalización del uso de las tecnologías de la información en las Administraciones Públicas, las cuales, si bien supone unos beneficios evidentes para los ciudadanos, también generan ciertos riesgos que deben minimizarse con medidas de seguridad.

AEPD[398], etc. Precisamente estas últimas resultan muy interesantes ya que están específicamente pensadas para la gestión de los riesgos que se derivan del tratamiento de datos personales[399]. Serán estas guías la que se seguirán para el desarrollo de los siguientes epígrafes.

1. Descripción sistemática de las operaciones

El primer paso que se requiere a la hora de llevar a cabo el análisis de riesgos es conocer y definir claramente el tratamiento de datos personales al que va a ser sometido tal evaluación. Ello nos llevará a obtener una radiografía completa de tal tratamiento a través de la cual el responsable podrá identificar de forma más precisa las amenazas y los riesgos a los que se exponen los derechos y libertades de las personas e implantar las medidas para minimizar tales riesgos.

A) El ciclo de vida de los datos

De esta manera, es necesario primeramente tener en cuenta el ciclo de vida de los datos que abracará las fases principales de los tratamientos objeto de análisis. El ciclo de vida de los datos ha sido ampliamente analizado por la doctrina y además viene desarrollado en diversas guías y directrices de las autoridades de control. Nosotros pondremos el acento en los aspectos que pueden ser más relevantes durante las fases que comprenden el desarrollo y despliegue de los sistemas de toma de decisiones automatizadas. Estas fases son: recogida y clasificación de los datos, almacenamiento, uso de los datos y supresión de los mismos.

En primer lugar, por lo que se refiere a la *recogida*, el responsable debe indicar cómo ha obtenido los datos y los instrumentos utilizados para ello. Así, en la fase de diseño será esencial que queden claras las bases de datos que se pretende utilizar para la creación del modelo algorítmico. Por otro lado, por lo que se refiere a la fase de despliegue, se ha de indicar si por ejemplo los inputs que alimentarán el algoritmo se

[398] Agencia Española de Protección de Datos. *Guía práctica de análisis de riesgos en los tratamientos de datos personales sujetos al RGPD*. 2018. A ello hay que sumarle una nueva guía actualizada de esta misma Agencia titulada: *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales*. Junio de 2021.

[399] El considerando 77 del RGPD apuesta porque las autoridades de control desarrollen guías o directrices para facilitar la identificación de riesgos específicos presentes en tratamientos de datos personales.

obtienen a través de los sensores que pueden llevan integrado el sistema de toma de decisiones automatizadas. En ambas fases, esto es, la de diseño y despliegue, se deberá indicar además si dichos datos se han obtenido directamente de los interesados o de terceros. En esta última categoría se enmarcarían aquellos datos obtenidos de portales de datos abiertos, de terceras organizaciones, a través de sensores, etc.

En segundo lugar, en lo que respecta a la *clasificación de los datos*, también conviene realizar una distinción entre las diversas categorías de datos que pueden ser relevantes tanto en el diseño como en el despliegue de los sistemas de toma de decisiones automatizadas. Así, podría distinguirse entre datos facilitados por los interesados, datos facilitados por terceros, datos obtenidos a través de sensores, datos obtenidos de fuentes públicas[400], datos de categoría especial, datos personales inferidos, datos alternativos[401], etc. Esto permitirá al responsable valorar para cada una de estas categorías las medidas que se han de implantar para un correcto cumplimiento de la normativa de protección de datos. De esta manera, en función de la tipología de datos, una base de legitimación u otra podrá ser necesaria tal y como ocurre con los datos de categoría especial. A su vez, cuando por ejemplo nos encontremos ante datos alternativos, puede ser recomendable justificar su uso cuando se dude de su pertinencia. Además, centrándonos en la fase de despliegue de estos sistemas, en este momento también puede resultar coherente indicar aquellos datos o variables que más inciden en la decisión que adopta el algoritmo. En esta fase, el responsable ya dispone de esa información ya que durante la fase de diseño se habrán realizado las correspondientes pruebas para justificar las posibles correlaciones existentes entre los datos de entrada y de salida.

En tercer lugar, también debe quedar claro dónde se *almacenan los datos* que se utilizan durante la fase de diseño y despliegue de los sistemas. Respecto a la fase de diseño de los sistemas, será relevante indicar dónde se almacenan las bases de datos, si estas son o no estructuradas y quién tienen acceso a ellas. Por otro lado, también será necesario indicar si los datos se almacenan en data centers propios o por el contrario se contratan servicios en la nube para su almacenamiento.

[400] La obtención de datos a través de las distintas plataformas de open data es cada vez más frecuente por parte de todo tipo de organizaciones. En España destaca la plataforma del Gobierno de España. https://datos.gob.es/

[401] Autoritat Catalana de Protecció de dades. *Guía práctica. Evaluación de impacto relativa a la protección de datos*. Enero de 2018, versión 2.0, pág.46.

En cuarto lugar y por lo que se refiere al uso de los datos, en esta etapa se ha de indicar el conjunto de operaciones al que se van a someter los datos previamente recogidos y clasificados que corresponderán en muchos casos con las actividades más relevantes del tratamiento que se pretende llevar a cabo.

En quinto lugar, es posible que durante el desarrollo y despliegue de los sistemas se produzcan *cesiones de datos a terceros***.** Así, la Propuesta de Reglamento Europeo sobre la gobernanza europea de datos prevé la reutilización de información personal en manos de las Administraciones Públicas por parte de terceros a través de entornos seguros bajo la supervisión de un organismo[402]. En estos supuestos se deberá prever que dichos datos pueden cederse a ese tercero para que los analice bajo la supervisión de dicho organismo.

Finalmente, *la supresión de los datos supondrá* la obligación de prever los plazos máximos de conservación de los datos personales. Por ejemplo, durante la fase del diseño se pueden establecer compromisos de supresión por parte del responsable para eliminar aquellos datos y variables que se vayan descartando conforme avanza el proceso de diseño del sistema. Es decir, aunque para estos supuestos no se puede establecer un plazo específico temporal, sí que se ha de prever esa obligación de supresión cuando dichos datos ya no sean pertinentes.

B) Descripción detallada de las operaciones del tratamiento

Para continuar con la radiografía del tratamiento, es turno de analizar específicamente las actividades o procesos más relevantes que componen los tratamientos de datos presentes durante el ciclo de vida de un sistema de toma de decisiones automatizadas. Estos son, la fase de diseño y la fase de despliegue.

Tratamientos de datos	Posibles operaciones
Fase de diseño	Planificación del proyecto y recopilación de datos Pre procesamiento de datos y limpieza de los mismos Desarrollo del modelo (fase de entrenamiento) Evaluación y elección del modelo

[402] Artículos 6 y 7 de la Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos).

Fase de despliegue	Incorporación de los datos/inputs Generación del resultado/Inferencia/outputs Adopción de la decisión Evolución del modelo

Cada una de estas fases está comprendida a su vez por distintas operaciones, las cuales, se han de describir con la suficiente precisión para posteriormente evaluar los riesgos aparejados a las mismas y las medidas para mitigar dichos riesgos. Así, por ejemplo, durante la fase de despliegue ha de quedar claro si el proceso decisorio es completo o parcialmente automatizado ya que los riesgos no serán los mismos si un humano visualiza la salida del algoritmo antes de que se tome una decisión a que dicho resultado automáticamente se convierta en una decisión que afecta a un particular. Si se opta por automatizar completamente la decisión, se ha de valorar si este tratamiento queda encajado en el ámbito de aplicación del artículo 22 del RGPD. En este sentido, si el responsable, a pesar de incorporar la plena automatización en su tratamiento no considera que el mismo se incluye en el artículo 22 del RGPD, este último deberá justificar las razones por las que no ha considerado que las decisiones totalmente automatizadas que adopta su sistema no generan efectos relevantes en los particulares afectados por las mismas. Además, en los casos en los que el responsable opte por introducir la presencia humana previa a la toma de decisión, dicha intervención humana deberá justificarse que es real. Para ello, el responsable del tratamiento debería describir el nivel de control e implicación humana de la persona, así como los momentos en los que existirá dicha intervención[403]. Es decir, la incorporación del humano no puede enmascarar una posible decisión automatizada fáctica que tenga como fin la elusión del artículo 22 del RGPD.

C) Intervinientes

Es necesario indicar cada uno de las personas que estarán presentes a lo largo de todas las operaciones que engloban el tratamiento de los datos personales objeto de

[403] Véase el listado de recomendaciones que propone el Grupo de expertos sobre inteligencia de la Unión Europea sobre el papel del human en la toma de decisiones total o parcialmente automatizadas. En: Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Una definición de la inteligencia artificial: Principales capacidades y disciplinas científicas*, 2019, pags.33 y 34.

análisis. Las funciones de cada uno de los agentes deberán quedar claras, así como las responsabilidades en las distintas operaciones presentes en tal tratamiento.

Así, por lo que respecta a la fase de despliegue de estos sistemas, debe indicarse la persona que se encargue de recopilar e ingresar en el algoritmo el *input*. También será relevante indicar qué persona es la competente para interpretar la decisión cuando la misma se utilice como soporte a la decisión final. A su vez, se habrá de fijar el grado de discrecionalidad que esta persona tiene para alterar o no la decisión que ha tomado el algoritmo y si es necesario o no consultar a terceras personas para adoptar una decisión final. Además, en los tratamientos basados en decisiones plenamente automatizadas también se debe precisar e indicar las personas de la organización que interpretarán y valorarán la decisión cuando un particular la impugne. Como se ha señalado previamente, cada una de estas actuaciones debe quedar claramente delimitada. Se recomienda por tanto que estos intervinientes tengan la suficiente cualificación para realmente desarrollar un cumplimiento efectivo de las responsabilidades que se les encomiendan.

D) Tecnología

Los elementos tecnológicos que se van a utilizar durante el tratamiento de datos también deben quedar identificados. Ahora bien, tal identificación no debe de ser exageradamente pormenorizada sino que bastará con una identificación general[404]. Teniendo en cuenta que durante el ciclo de vida de los sistemas algorítmicos el factor tecnológico está muy presente, será necesario prestar cierta atención. Por lo que respecta a la fase de diseño, será relevante el tipo de algoritmo o algoritmos que se utilizarán para crear el modelo, así como la técnica de aprendizaje automático, si es que la hay. También puede incluirse los métodos de validación que se realizan del modelo, así como las técnicas que se utilizarán durante su desarrollo. Toda esta información será realmente útil para la evaluación de impacto que posteriormente se realizará antes del despliegue del sistema. Recordemos además que la PRAI obliga a los desarrolladores/proveedores de los sistemas a facilitar a las organizaciones que los

[404] Agencia Española de Protección de Datos. *Guía práctica para las Evaluaciones de Impacto en la Protección de los Datos sujetas al RGPD*, 2018, pág.16.

utilizarán todo un conjunto de información referido al modelo algorítmico. Esta información será esencial para practicar la evaluación de impacto[405].

2. Identificación los riesgos

Una vez que el responsable ha indicado de forma sistemática el tratamiento de datos objeto de análisis. Este ya cuenta con un grado de información suficiente sobre el tratamiento en cuestión para llevar a cabo la detección y evaluación de los riesgos que puede presentar, así como en su caso las medidas que debe implantar para combatirlos o minimizarlos. En este sentido, aunque es posible realizar la evaluación de riesgos del conjunto del tratamiento, es recomendable identificarlos en cada una de las operaciones, evitando así que pasen desapercibidos importantes riesgos en operaciones que se hayan considerado de poca relevancia en el tratamiento. La identificación de los riesgos como es lógico variará del tipo de tratamiento que se pretenda implantar. A modo de orientación, las distintas autoridades de control han clasificado distintos tipos de riesgos en función de distintas variables. En este sentido y ayudándonos de la clasificación de escenarios que ha establecido la Autoridad Catalana de Protección de Datos, vamos a distinguir varios[406]:

Riegos que afectan al derecho fundamental de la protección de datos:	**Riesgos que pueden afectar a otros derechos y libertades:**
Escenarios de riesgo que pueden afectar a: • los principios de tratamiento de los datos personales. • los derechos que se derivan del RGPD. • las obligaciones, incluyendo aquí tanto las obligaciones generales derivadas del RGPD, como las obligaciones específicas en materia de seguridad de los tratamientos de datos.	Escenarios de riesgo que pueden afectar a otros derechos y libertades.

[405] El artículo 29 de la PRAI establece que *los usuarios de los sistemas de IA de alto riesgo utilizarán la información facilitada en virtud del artículo 13 para cumplir con su obligación de realizar una evaluación de impacto sobre la protección de datos en virtud del artículo 35 del Reglamento (UE) 2016/679 o del artículo 27 de la Directiva (UE) 2016/680, cuando proceda.* El artículo 13 de la PRAI hace referencia a toda una serie de información relacionada con los sistemas algorítmicos desarrollados. Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

[406] Autoritat Catalana de Protecció de dades. *Guía práctica. Evaluación de impacto relativa a la protección de datos*. Enero de 2018, versión 2.0, pág.89.

Tal y como indica la AEPD[407], un riesgo se puede definir como la combinación de la posibilidad de que se materialice una amenaza y sus consecuencias negativas. A su vez, una amenaza es cualquier factor de riesgo con potencial para provocar un daño o perjuicio a los interesados sobre cuyos datos de carácter personal se están tratando. Visto así, las amenazas y los riesgos asociados a estas están directamente relacionados. En consecuencia, identificar y evaluar los riesgos siempre implica considerar la amenaza que los puede originar. Aunque las amenazas que pueden estar presentes en las distintas operaciones que forman parte del ciclo de vida de un sistema de toma de decisiones automatizadas pueden ser amplísimas, en el siguiente cuadro hacemos mención a algunas de las que, a modo de ejemplo podrían presentarse.

Identificación de amenazas en la fase del diseño de los sistemas	
Amenaza	**Derecho/Principio/Obligación afectada**
Reversión de los procesos de seudonimización/anonimización	Principio de minimización de datos Principio de seguridad del tratamiento
Sobre ajuste del modelo durante la fase de entrenamiento	Principio de minimización de datos Principio de lealtad Principio de no discriminación
Revelación de datos inferidos, sobre todo sensibles	Principio de limitación de la finalidad Principio de licitud
Presencia de variables o datos no necesarios para el objetivo que se pretende resolver con el modelo	Principio de minimización de datos
Existencia de correlaciones espurias	Principio de minimización de datos Principio de no discriminación
Presencia de datos desbalanceados	Principio de lealtad Principio de minimización de datos Principio de no discriminación
Acceso no autorizado a las bases de datos	Principio de seguridad del tratamiento
Ataques contradictorios al modelo	Principio de seguridad en el tratamiento

Identificación de amenazas en la fase despliegue de los sistemas	
Amenaza	**Derecho/Principio/Obligación afectada**
Falta de interpretabilidad del modelo	Principio de transparencia Derecho de información. Derecho de impugnación de la decisión
Re identificación de los datos personales utilizados durante la fase de entrenamiento	Principio de seguridad Principio de licitud

[407] Agencia Española de Protección de Datos. *Guía práctica para las Evaluaciones de Impacto en la Protección de los Datos sujetas al RGPD*, 2018, pág.24.

Errores a la hora de adoptar las decisiones Falsos positivos y negativos.	Principio de exactitud
El sistema puede adoptar decisiones discriminatorias	Principio de lealtad Principio de exactitud Principio de no discriminación
El modelo queda afectado debido a ataques contradictorios	Principio de seguridad del tratamiento
Dificultad para interpretar los resultados expulsados por el algoritmo	Derecho a la intervención humana Derecho a informar sobre la lógica del tratamiento Principio de transparencia
El sistema puede comenzar a adoptar decisiones que generen efectos relevantes	Derecho a no ser sometido a la toma de decisiones automatizadas relevantes Principio de licitud
La persona encargada de interpretar los resultados del sistema no es crítica con los mismos	Derechos de información Derecho a la intervención humana Derecho a expresar el punto de vista

Nótese que estas amenazas únicamente son referidas a algunas de las que se pueden presentar durante la fase del ciclo de vida de estos sistemas automatizados. Sin embargo, existen otra serie de amenazas propias de cualquier tratamiento de datos que también están presentes cuando se llevan a cabo este tipo de operaciones.

3. Análisis y valoración de los riesgos

Una vez se han detectado las amenazas que son susceptibles de generar riegos resulta fundamental tener en cuenta los elementos y características que puedan entrar en juego a la hora de determinar el nivel de riesgo. Son muchos los factores relevantes, algunos de los que podemos señalar son: establecer la plena o parcial automatización de todo el proceso de toma de decisiones, finalidad que se pretenda obtener con el despliegue del sistema, efectos de las decisiones en los derechos de los particulares sometidos a los tratamientos, tipo de algoritmo elegido, consecuencias de los falsos positivos y negativos, entorno en el que se despliega el sistema, esto es, dinámico, estático, presencia de muchos individuos interactuando con el sistema, etc.

En cuanto a la finalidad, no presentará el mismo riesgo un sistema automatizado cuya objetivo sea la recomendación de películas a un sistema que tenga como finalidad la prescripción de un tratamiento médico. A su vez, incluso en un mismo contexto, las finalidades y el grado de afectación de los derechos pueden presentar diversos riesgos.

Así, por ejemplo, los cuerpos y fuerzas de seguridad pueden utilizar todo tipo de sistemas que quedan englobadas en la llamada policía predictiva. Cada uno de estos usos presenta distintos grados de incidencia en los derechos de los particulares en función de la finalidad para la que se implementen esos modelos algorítmicos[408]. Por un lado, las aplicaciones que menos incidencia generan son aquellas destinadas a determinar en qué zonas de una ciudad es probable que se cometa un delito. En un segundo nivel de afectación encontramos aquellos sistemas que evalúan el nivel de riesgo para predecir qué personas en un futuro pueden cometer un delito o sufrirlo[409]. Finalmente, los más relevantes serían todos aquellos sistemas y herramientas automatizadas que se utilizan con el objetivo de fijar medidas cautelares, denegar o conceder la libertad condicional o incluso configurar la condena/pena de los sujetos[410].

El tipo de algoritmo o sistema elegido también puede resultar relevante a la hora de valorar el riesgo. Así, los riesgos para los derechos de los particulares que presente un sistema más o menos complejo o más o menos interpretable serán diferentes.[411]. A su vez, a través de algunos modelos algorítmicos se puede acceder a los datos personales que se utilizaron durante la fase de entrenamiento, de manera que, ese acceso podría no estar autorizado, los riesgos en este tipo de sistemas son mayores[412].

También pude ser relevante el entorno donde se implanta el sistema. Así, en aquellos entornos dinámicos o adaptativos donde continuamente el sistema está mejorando, los riesgos pueden ser mayores que cuando ese mismo sistema se ingresa en

---

[408] M O,DONELL,R:"Challenging racist predictive policing algorithms under the equal protection clause". ***New*** *York University Law Review*. Vol. 94:544, 2019, pág.542.

[409] Este tipo de sistemas llevan a la policía a realizar sobre los individuos señalados por el sistema determinados cacheos o controles rutinarios. Véase todo lo relativo al "National Data Analytics Solution (NDAS) Project" del Reino Unido. Visto en:
https://www.westmidlands-pcc.gov.uk/archive/ethics-committee-meeting-november-2019/

[410] El caso paradigmático es el caso Loomis. Sentencia State v.Loomis, 881.N.W.2d 749 (Wisc.2016). Disponible en: https://www.courts.ca.gov/documents/BTB24-2L-3.pdf
El Consejo de Europa también ha indicado el especial impacto que pueden causar estas técnicas automatizadas en la esfera de los particulares sometidos a las mismas. En: Consejo de Europa. *European ethical charter on the use of artificial intelligence in judicial systems and their environment*, 2018. apartado 133, pág.55.

[411]Para el Parlamento Europeo, la complejidad del algoritmo podría servir como parámetro a la hora establecer mayores o menores medidas. European Parliamentary Research Service.*A governance framework for algorithmic accountability and transparency*, op.cit., pág.49.

[412] Los modelos como el de máquinas de soporte vectorial o el k-vecinos más cercano pueden presentar estos riesgos. Así lo ha indicado la ICO en:
https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/how-should-we-assess-security-and-data-minimisation-in-ai/

un entorno poco dinámico no adaptivo. Con relación al contexto, también hay que destacar como factor esencial el número de afectados y el grado de interrelación que pueden tener esos afectados con el sistema. No supondrá el mismo riesgo un sistema que utiliza una entidad bancaria para conceder crédito a un sistema que utiliza la plataforma Facebook para controlar la difusión de *fakes news* o recomendar contenido informativo[413]. Mientras que en el primer algoritmo únicamente tienen acceso al mismo los trabajadores de la organización bancaria, en el segundo resulta incalculable el número de persona que pueden interactuar con el sistema y en su caso pretenden modificarlo, y ello pesar de que exista un importante control por parte de la plataforma que lo gestiona.

Criterios que influyen en la valoración del riesgo durante el ciclo de vida de los sistemas de toma de decisiones automatizadas	Plena o parcial automatización de todo el proceso Finalidad que se pretenda obtener con el despliegue del sistema Base de datos elegida Efectos de las decisiones en los derechos de los particulares Tipo de algoritmo/modelo elegido Consecuencias de los falsos positivos y negativos Entorno en el que se despliega el sistema, esto es: Dinámico/estático/número de personas que interactúan con el sistema Grupos especialmente vulnerables afectados

Analizados algunos de los factores relevantes susceptibles de generar más o menos riesgos, se ha de proceder a la evaluación de los mismos. La evaluación de los riesgos consiste en valorar y estimar la probabilidad y el impacto de que el riesgo se materialice. A diferencia de un análisis de riesgos general donde se evalúan los riesgos que puede comportar una operación para la entidad. A través del análisis de riesgos que lo que se valora son los riesgos que puede generar ese tratamiento de datos en la esfera

[413] Precisamente, desde la Unión Europea, uno de los elementos básicos sobre los que se apoya la nueva directiva de servicios digitales para imponer mayores medidas de responsabilidad y cumplimiento en materia de uso de algoritmos es precisamente el número de usuarios que ostenta una determinada plataforma. Véase los artículos 25 y 26 de la 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. Resolución de 15 de diciembre de 2020. Entre los riesgos que se pueden originar en este tipo de plataformas se señala por ejemplo la explotación automatizada del servicio, con un efecto negativo real o previsible sobre la protección de la salud pública, los menores, el discurso cívico o efectos reales o previsibles relacionados con procesos electorales y con la seguridad pública.

de los particulares sometidos en este caso al conjunto de operaciones presentes durante el ciclo de vida de un sistema de toma de decisiones automatizadas[414].

Pues bien, dado que toda operación o actividad presenta un riesgo inherente, el siguiente paso es valorar ese riesgo y en función del mismo establecer las medidas adecuadas. Así, el riesgo inherente se calcula teniendo en cuenta el posible impacto que pueda generar una determinada amenaza detectada en el tratamiento y la probabilidad de que se materialice la misma. De esta manera, la AEPD establece una metodología para calcular el impacto y la probabilidad en cuatro niveles, esto es, despreciable, limitada, significativa y máxima.

Probabilidad de que suceda la amenaza	
1. Probabilidad despreciable → casos fortuitos.	**3. Probabilidad significativa** → es muy posible que ocurra la amenaza.
2. Probabilidad limitada → la posibilidad de que ocurra la amenaza es baja.	**4. Probabilidad máxima** → es muy posible que ocurra la amenaza.

Grado de impacto de relevancia sobre los interesados	
1. Impacto despreciable → impacto muy bajo.	**3. Impacto significativo** → impacto alto.
2. Impacto limitado →impacto bajo.	**4. Impacto máximo** → impacto muy alto.

A modo de ejemplo, vamos a analizar la probabilidad y el impacto que puede causar una amenaza para calcular el riesgo.

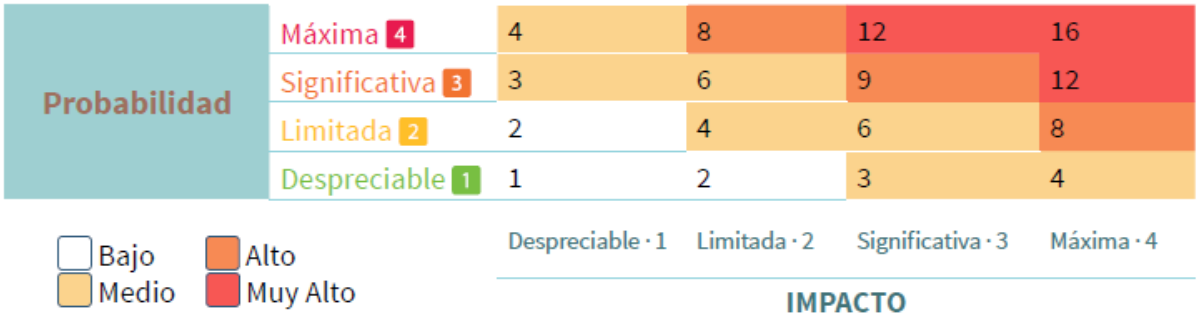
Amenaza: Sistema comienza a adoptar decisiones inadecuadas a la hora de interactuar con el entorno		
Impacto	**Despreciable**	Si la finalidad del sistema es recomendar películas/anuncios.
	Limitado	Si la finalidad del sistema es ofrecer un anuncio para un colectivo específico vulnerable.
	Significativo	Si la finalidad del sistema es controlar los videos, noticias que se suben a una plataforma.
	Máximo	Si la finalidad es recomendar un tratamiento médico.
Probabilida	**Despreciable**	El sistema interactúa en un entorno estático.
	Limitada	El sistema interactúa en un entorno estático pero pueden intervenir terceros en su configuración que no están los suficientemente formados.
	Significativa	El sistema interactúa en un entorno dinámico/adaptativo.

[414] Autoritat Catalana de Protecció de dades. *Guia Pràctica Avaluació d'impacte relativa a la protecció de dades*, 2019, pág.23.

	Máxima	El sistema interactúa en un entorno dinámico/adaptativo donde intervienen multitud de participantes: personas, *bots*, empresas, etc[415].

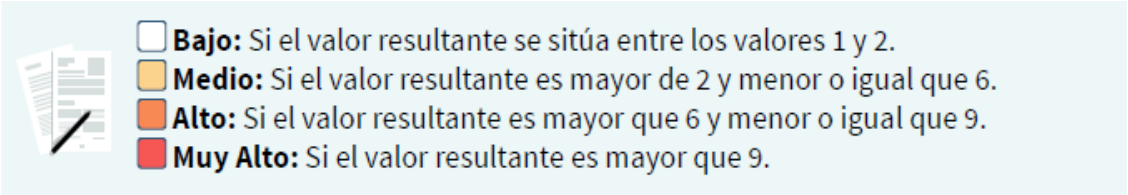
Fuente AEPD 2018.

Como se puede apreciar, existen diferentes criterios que se han de tener en cuenta para valorar el nivel de impacto y la probabilidad de las amenazas indicadas. El siguiente paso será asignar a cada uno de los niveles de la escala de probabilidad e impacto distintos valores. Esto es, desde el valor 1, en el caso de que la magnitud sea despreciable, hasta el valor 4 en el caso donde la magnitud es máxima. Seguidamente, al enfrentar los valores asignados al impacto y a la probabilidad, se crea una matriz de riesgo.



Matriz de riesgo. Fuente AEPD 2018.

Finalmente, con la matriz de riesgo ya se puede valorar la estimación de riesgo que puede derivarse de esa amenaza. Para ello, la AEPD propone que los valores resultantes de la probabilidad e impacto se multiplique y así se obtenga el riesgo inherente teniendo en cuenta que:



Fuente AEPD 2018.

En función de esa estimación, bajo, medio, alto o muy alto, se deciden implantar las medidas que en su caso puedan mitigar el riesgo inherente a ese tratamiento. Es decir, en este momento el responsable ya ha podido identificar los posibles riesgos y

[415] Este supuesto sería el caso de las grandes plataformas sociales donde multitud de personas y empresas interactúan entre sí y donde los algoritmos cumplen todo tipo de funciones.

además los ha evaluado. Se encuentra así ya en disposición de adoptar las medidas adecuadas que eviten que tales amenazas se generen o, produciéndose, el riesgo de las mismas sea aceptable. Para materializar lo indicado hasta ahora vamos a hacer mención a dos ejemplos donde se valoran varias amenazas y el riesgo aparejado a las mismas dentro del ciclo de vida de un sistema de toma de decisiones automatizadas.

Tipo de sistema	**Amenaza**	**Riesgo Inherente**	**Medidas a implantar**
Control del contenido ilícito en una plataforma con una cuota de usuarios muy alta	El sistema comienza a adoptar decisiones inadecuadas a la hora de interactuar en un entorno altamente adaptativo.	Probabilidad Máxima (4) Impacto significativo (3) Riesgo inherente: 4*3=12 Riesgo Muy alto	-Realización de testeos continuos. -Auditorías cada cierto tiempo. -Reentrenamiento adecuado de datos.

De esta manera, las medidas que se implanten para reducir o evitar los riesgos deben ser realmente efectivas. Al igual que se establece el cifrado de datos o la autenticación reforzada para evitar posibles accesos no autorizados a las bases de datos, las medidas que traten de reducir los riesgos propios aparejados al usos de sistemas de toma de decisiones automatizadas también deberán conseguir ese fin. En consecuencia, si una organización no acredita el despliegue de medidas adecuadas para reducir los riesgos inherentes a este tipo de sistemas, se entenderá que, por un lado, se está incumpliendo la obligación de realizar una gestión adecuada de los riesgos (artículo 24 RGPD) y a su vez, se estará incumpliendo el principio de responsabilidad activa reconocido en el artículo 5.2 RGPD.

II. LA EVALUACIÓN DE IMPACTO EN PROTECCIÓN DE DATOS

1. Concepto y finalidad

La evaluación de impacto en la protección de datos personales, en adelante, EIPD, supone uno de los mecanismos esenciales de cumplimiento y por tanto de representación del principio de responsabilidad activa que más relevancia encuentra en el RGPD.

Para el grupo del artículo 29, la EIPD es un proceso concebido para describir el tratamiento de datos, evaluar su necesidad y proporcionalidad y ayudar a gestionar los riesgos a los que se expone el mismo con el objetivo de garantizar los derechos y libertades de las personas físicas[416]. La gestión de los riesgos para los derechos de los interesados se convierte así en elemento esencial de esta herramienta[417]. Tal y como ha señalado la AEPD, la gestión del riesgo y la EIPD son procesos íntimamente vinculados. Esta herramienta forma parte de la gestión del riesgo para aquellos tratamientos de datos que impliquen un alto riesgo para los particulares. En estos casos, la normativa obliga al responsable del tratamiento a desarrollar una serie de requisitos adicionales que ayuden a mitigar los riesgos presentes en tales tratamientos. De esta manera, mientras que la gestión del riesgo a través del análisis de riesgos es preceptiva para todo tratamiento de datos, las obligaciones concretas que se establecen para la EIPD son únicamente obligatorias para los tratamientos de datos que supongan un alto riesgo[418]. Se podría decir así que la EIPD es un análisis de riesgos agravado[419].

Fuente AEPD 2021

Las evaluaciones de impacto no son una herramienta novedosa de la normativa de protección de datos sino que su origen se remonta a las normas medio ambientales ligadas a los estudios de impacto que pueden generar determinados proyectos en el

[416]Grupo del Artículo 29. *Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679***.** Adoptadas el 4 de abril de 2017. Revisadas por última vez y adoptadas el 4 de octubre de 2017, pág.4.

[417] Information Commisioner´s Office. *Data Protection Impact Assessments (DPIAs)*. Documento adoptado el 22 de marzo de 2018, pág.5.

[418] Agencia Española de Protección de Datos. *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales***.** Junio de 2021. Pág.159.

[419] MARTÍNEZ MARTÍNEZ, R: "Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo". *Revista Catalana de Dret Públic*, (58). 64-81, 2019, pág.76. Disponible en: https://doi.org/10.2436/rcdp.i58.2019.3317

medio ambiente[420]. Así, esta técnica comenzó a extenderse en el entorno de la privacidad en países como Canadá, Australia o EEUU[421]. En Europa, esta herramienta ha sido acogida en el ámbito de la normativa de protección de datos a través del RGPD.

Cabe señalar que la obligación de realizar una EIPD corresponde al responsable del tratamiento con el apoyo y la colaboración del encargado del tratamiento y el delegado de protección de datos en el caso de que hubiese[422].

Aunque es lógico que la EIPD se realice respecto de un tratamiento en concreto, existen distintas circunstancias que pueden habilitar al responsable para realizar una única EIPD para evaluar múltiples operaciones de tratamiento que sean similares en términos de naturaleza, alcance, contexto, fines y riesgos[423].

En primer lugar, si un mismo tratamiento es llevado a cabo por diversas organizaciones que forman parte de un mismo grupo de empresas o por órganos de una misma Administración, la EIPD realizada globalmente puede ser suficiente para el conjunto de tratamientos. Trasladado a nuestra esfera, se pueden prever que una misma EIPD sea aplicable a una serie de empresas que instalen el mismo sistema de toma de decisiones automatizadas o Administraciones Públicas que implanten estos sistemas a escala nacional[424], autonómica o municipal. Ahora bien, en aquellos supuestos en los que el tratamiento inicialmente previsto sufra alteraciones relevantes, se deberá realizar una nueva EIPD para ese nuevo tratamiento, y ello, pese a que se esté utilizando el mismo sistema de toma de decisiones automatizadas sobre el que se realizó inicialmente la EIPD. Así, si una Administración Pública decide implantar un sistema de toma decisiones pero deja margen de maniobra a sus distintos organismos para que decidan si

[420] Véase por ejemplo la evaluación de impacto establecida por la Ley 21/2013, de 9 de diciembre, de evaluación ambiental.

[421] A través del *E-Governence Act of 2002* se estableció la obligación de realizar una EIPD a todas las agencias del gobierno federal que desarrollen o adquieran nueva tecnología de información que involucre la recolección, mantenimiento de información en forma identificable. A su vez, existen guías para la implementación de estas EIPD. OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002**.** Más información en: https://digital.gov/resources/guidance-for-implementing-the-privacy-provisions-of-the-e-government-act-of-2002-m-03-22/

[422] Agencia Española de Protección de Datos. *Guía práctica para las Evaluaciones de Impacto en la Protección de los Datos sujetas al RGPD*, 2018, pág.9.

[423] Artículo 35.1 del RGPD.

[424] Un ejemplo de algoritmo automatizado que se aplica en toda España de forma similar es el sistema VioGen. Este sistema se aplica para todos los casos de supuestos delitos de violencia de género. Entendemos que para este caso una evaluación de impacto general de este tratamiento en concreto podría ser suficiente. Para más información véase: http://www.interior.gob.es/ca/web/servicios-al-ciudadano/violencia-contra-la-mujer/sistema-viogen

el sistema adoptará dichas decisiones de forma total o parcialmente automatizada o permite que estos organismos alteren las finalidades para las que se previó inicialmente el despliegue del sistema, tales alteraciones se han de considerar relevantes y por tanto se requerirá de una nueva EIPD para ese tratamiento[425].

En segundo lugar, tampoco será necesario realizar distintas EIPD en aquellos supuestos donde los distintos tratamientos de datos que se llevan a cabo están estrechamente relacionados entre sí y forman parte de un tratamiento más global. Traducido a nuestro estudio, en muchos supuestos será recomendable que únicamente se realicen dos EIPD que aglutinen los dos principales tratamientos de datos que están presentes durante el ciclo de vida de los sistemas de toma de decisiones automatizadas, esto es, una que englobe la fase de diseño y la otra la de despliegue.

2. La valoración de realizar o no una evaluación de impacto. El alto riesgo

A) Valoración del alto riesgo

La EIPD se basa en un enfoque basado en el riesgo tal y como se deriva del RGPD. De esta manera, esta herramienta sólo resulta obligatoria cuando el tratamiento que se pretende ejecutar presente un altor riesgo para los derechos y libertades de los particulares. Le corresponde por tanto al responsable valorar si efectivamente el tratamiento que pretende llevar a cabo genera o no dicho riesgo. En este sentido, a día de hoy los responsables del tratamiento tienen un abanico de instrumentos legales y técnicos que le facilitan la labor analítica a través de la cual pueden valorar si es o no necesario implantar una EIPD. Estos instrumentos son:

En primer lugar, el propio RGPD en su artículo 35.1.3 RGPD alude a determinados tratamientos que son considerados de alto riesgo y que por tanto requieren de la EIPD. Entre ellos se hace referencia a la elaboración de perfiles sobre cuya base se tomen decisiones que generen efectos relevantes en la esfera de los particulares. Como se puede apreciar, el RGPD no distingue entre si la decisión sobre la que se basa la elaboración de perfiles es total o parcialmente automatizada. Consecuencia de ello,

[425] Tal y como señala el considerando 92 del RGPD. *Hay circunstancias en las que puede ser razonable y económico que una evaluación de impacto relativa a la protección de datos abarque más de un único proyecto, por ejemplo, en el caso de que las autoridades u organismos públicos prevean crear una aplicación o plataforma común de tratamiento, o si varios responsables proyecten introducir una aplicación o un entorno de tratamiento común en un sector o segmento empresarial o para una actividad horizontal de uso generalizado*.

siempre que se pretenda realizar una elaboración de perfiles sobre la cual se adopten decisiones relevantes se deberá realizar una EIPD. Esta inclusión resulta sumamente positiva ya que la EIPD se amplía a tratamientos que exceden del ámbito de aplicación específico del artículo 22 del RGPD. Supone así una protección jurídica reforzada para las decisiones relevantes parcialmente automatizadas que se derivan la elaboración de perfiles. A su vez, también resulta necesaria la realización de la EIPD cuando se lleven a cabo la *evaluación sistemática a gran escala de datos especialmente sensibles o aquellos referidos a condenas e infracciones penales*, así como *la observación sistemática a gran escala de una zona de acceso público*. En ambos supuestos también pueden entrar en juego el desarrollo y despliegue de sistemas de toma de decisiones automatizadas[426]. Tal y como ocurre con el desarrollo de proyectos de investigación científica basadas en técnicas de aprendizaje automático cuando se analizan datos de salud o el uso de sistemas de videovigilancia en los que se incorporan tecnologías de reconocimiento facial.

En segundo lugar, el Grupo del artículo 29 ha mencionado una serie de criterios que pueden evidenciar un elevado riesgo presente en las actividades de un concreto tratamiento. De esta manera, si al menos dos criterios de los señalados por el GT29 forman parte de las actividades u operaciones de ese tratamiento, el responsable debería considerar la necesidad de realizar una EIPD. Cuantos más criterios converjan en el tratamiento que se pretende implantar, más probable es que el mismo represente un alto riesgo para los derechos y libertades de los interesados[427]. Dichos criterios son:

> *Evaluación o scoring, toma de decisiones automatizadas con efectos jurídico o significativo, monitorización sistemática, tratamiento de datos sensibles o datos muy personales, combinación o coincidencia de conjunto de datos, tratamiento de datos a gran escala, tratamientos de datos de personas vulnerables[428], uso innovador o aplicación de nuevas soluciones tecnológicas,*

---

[426] BASTIDAS CID, Y,V: "El cumplimiento de los principios del tratamiento de datos personales establecidos en el reglamento general de protección de datos de la unión europea en proyectos de Big Data". *Informática y Derecho: Revista Iberoamericana de Derecho Informático* (segunda época), Nº. 6, 2019, pág.37. Texto disponible en: http://www.fiadi.org/wp-content/revista_fiadi_segunda_epoca/historial_de_revistas/00006/REVISTA-FIADI-0006.pdf

[427] Grupo del Artículo 29. *Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679***.** Adoptadas el 4 de abril de 2017. Revisadas por última vez y adoptadas el 4 de octubre de 2017, pág.12.

[428] Numerosas autoridades de control también establecen como criterio a tener en cuenta el alto riesgo que puede suponer para ese tratamiento la presencia de colectivos más vulnerables. Véase Information

tratamientos que impidan a los interesados ejercer un derecho, acceder a un servicio o ejecutar un contrato[429].

Como puede observarse, la mayoría de los criterios a los que hace referencia el GT29 pueden estar presentes en el despliegue e implementación de sistemas de toma de decisiones automatizadas. En ninguno de los criterios señalados la plena automatización de la decisión resulta relevante. A modo de ejemplo, en la tabla siguiente se realiza un análisis del sistema SyRI ya previamente comentado[430]. Este sistema fue implantado por el gobierno neerlandés con el objetivo de detectar el fraude en la obtención de prestaciones sociales por parte de determinados sectores de la población. Al aplicar los criterios establecidos por el GT29 para considerar si es necesaria la evaluación de impacto, se puede llegar a la conclusión de que en este supuesto, es necesario abordar una EIPD. Dicha EIPD fue en este supuesto realizada, si bien, el propio tribunal indicó que se tendría que haber realizado dicha evaluación para cado uno de los proyectos que integraba este programa. Este tribunal además indicó que la EIPD es necesaria cuando se pretende implantar este tipo de sistemas con el objetivo de garantizar un adecuado respeto del derecho fundamental a la privacidad reconocido por el Convenio Europeo de Derechos Humanos[431].

Análisis de la necesidad de realizar la EIPD		
Finalidad	**Criterios aplicables**	
SyRI Sistema utilizado como	Evaluación	Elaboración de perfiles con el objetivo de predecir el posible fraude al sistema
	Toma de decisiones automatizadas con efectos	Si el informe del riesgo se utiliza para iniciar un procedimiento sancionatorio,

Commisioner´s Office. *Data Protection Impact Assessments (DPIAs)*. Documento adoptado el 22 de marzo de 2018, pág.44. Autoridad de Protección de Datos Sueca (datainspektionen). *List regarding Data Protection Impact Assessments according to article 35.4 of the Data Protection Regulation*. Documento adoptado el 16 de enero de 2019, pág.3. Autoridad de Protección de Datos Eslovena (Information Commisioner). The list of the kind of processing operations1 which are subject to the requirement for a Data Protection Impact Assessment under the Article 35(4) of the General Data Protection Regulation (EU) 2016/679 (GDPR). Resolución adoptada el 21 de diciembre de 2018,pág.4. Incluso, la Autoridad de Protección de datos Irlandesa establece como obligatorio la realización de la EIPD cuando se traten datos de personas vulnerables. (Data protection comission). *List of Types of Data Processing Operations which require a Data Protection Impact Assessment*, pág.4.

[429] Gran parte de los criterios a los que hace alusión el Grupo del artículo 29 se basan en distintos considerandos del RGPD donde se hace referencia a las actividades que pueden ser consideradas de alto riesgo. Entre otros, véase el considerando 75 del RGPD.

[430] Véase el Capítulo II, apartado I, punto 3 de esta tesis.

[431] Para llegar a esta conclusión el tribunal de la Haya se amparó en el artículo 8.2 del Convenio Europeo de Derechos Humanos. Véase la Sentencia del Tribunal de la Haya de 5 de febrero de 2020. Apartados 6.103 a 6.106. Disponible en: https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878

indicador de fraude potencial a ayudas de la seguridad social	relevantes	también se da este criterio.
	Tratamiento de datos sensibles o de naturaleza altamente personal	En principio, del conjunto de datos que se utilizaban, ninguno de ellos era sensible o de naturaleza altamente personal
	Combinación o coincidencia de conjuntos de datos	Se combinaban distintas bases de datos procedentes de varias administraciones.
	Tratamientos de datos relativos a las personas vulnerables	El sistema se focalizó en sectores de la población con escasos recursos económicos.
	Tratamientos de datos a gran escala	Se tratan datos de un importante número de personas y además con una variedad de categorías muy amplia.
	Uso innovador o aplicación de nuevas soluciones tecnológicas	Se utilizan tecnologías de aprendizaje automático para llevar a cabo el tratamiento de datos personales.
	Tratamientos que impidan a los interesados ejercer un derecho o acceder a un servicio	Una vez detectado el fraude por el sistema, se dejaban de percibir las ayudas que se hasta ese momento se estaban obteniendo.

Por tanto, como regla general, todos aquellos sistemas que se utilizan con el objetivo de detectar posibles patrones o conductas de forma global o por segmentos de la población que tengan en cuenta una gran variedad y tipologías de datos requerirán normalmente de la realización de la EIPD ya que generalmente cumplirán bastantes de los criterios indicados por el Grupo del artículo 29. Como es lógico, en no todos los supuestos será tan evidente el encaje de estos criterios en el tratamiento que se pretende llevar a cabo por lo que el responsable debe valorar si es o no necesaria la realización de la EIPD[432].

En tercer lugar, las distintas autoridades de control de los países europeos han desarrollado en parte los criterios designados por el GT29 elaborando un listado de criterios y tratamientos que también han de tenerse en cuenta a la hora de valorar el alto riesgo de un tratamiento[433]. Así, por ejemplo, la AEPD ha optado por establecer distintos criterios que han de estar presentes en un tratamiento para considerarlo de alto

[432] El responsable puede llegar a considerar ese alto riesgo y por tanto llevar a cabo una EIPD aunque el tratamiento sólo comporte uno de los criterios mencionados anteriormente. Grupo del Artículo 29. *Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679***.** Adoptadas el 4 de abril de 2017. Revisadas por última vez y adoptadas el 4 de octubre de 2017. Pág.12

[433] El artículo 35.4 del RGPD prevé que la Autoridad de Control establecerá y publicará una lista de los tipos de operaciones de tratamiento que requerirán una evaluación de impacto relativa a la protección de datos.

riesgo. Concretamente, y por lo que a nuestro estudio nos interesa, la AEPD señala que un criterio relevante está presente en aquellos tratamientos *que impliquen la toma de decisiones automatizadas o que contribuyan en gran medida a la toma de tales decisiones, incluyendo cualquier tipo de decisión que impida a un interesado el ejercicio de un derecho o el acceso a un bien o un servicio o formar parte de un contrato*[434]. Una interpretación amplia de este criterio permite ampliar la obligación de realizar la EIPD a aquellos sistemas de toma de decisiones total o parcialmente automatizados que no generen efectos relevantes en la esfera de los particulares. No obstante, no podemos olvidar que este es únicamente un criterio que revela el posible alto riesgo, de manera que será necesario que al menos estén presentes otros criterios para considerar la necesidad de realizar la EIPD. A su vez, también será indicio de alto riesgo aquellos tratamientos de datos *que permitan determinar la situación financiera o de solvencia patrimonial o deducir información sobre las personas relacionada con categorías especiales de datos*. En nuestra opinión, la AEPD se está refiriendo al tratamiento de datos personales que tengan como objetivo la elaboración de perfiles y sobre la misma se obtengan datos inferidos que puedan expresar categorías de datos especiales.

En cuarto lugar, también será relevante valorar si por vía normativa se ha previsto la obligación de realizar la EIPD para un concreto tratamiento de datos personales[435]. O si directamente una determinada autoridad de control específica que ese tratamiento es considerado de alto riesgo[436]. En este sentido, la PRAI establece un listado de sectores y finalidades donde se ha de considerar que un sistema de inteligencia artificial genera un alto riesgo para los derechos y libertades de las personas[437]. Para el CEPD, esta lista supone una presunción de que dichos tratamiento

[434] Agencia Española de Protección de datos. *Listas de tipos de tratamientos de datos que requieren evaluación de impacto relativa a protección de datos (art 35.4).* Resolución adoptada el 4 de septiembre de 2019, pág. 2.

[435]La ley eslovena de protección de datos realiza una mención explícita a la necesidad de realizar una EIPD cuando se pretenda llevar a cabo un tratamiento de toma de decisiones automatizadas. En: MALGIERI,G: "Automated decision-making in the EU Member States: The right to explanation and other "suitable safeguards" in the national legislations". *Computer Law & Security Review*, Volume 35, Issue 5, October 2019, pág.18.

[436] En España, por ejemplo, la AEPD ha indicado que los tratamientos de datos personales relativos a opiniones políticas y envío de propaganda electoral por medios electrónicos o a través de sistemas de mensajería por parte de partidos políticos, federaciones, coaliciones requerirán la realización de una EIPD previa al tratamiento. Véase el artículo 7.1. apartado 4 de la Circular 1/2019, de 7 de marzo, de la Agencia Española de Protección de Datos.

[437] Véase el artículo 6.3 y anexo III de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

para esas concretas finalidades generan un alto riesgo a los efectos de la normativa de protección de datos[438].

Finalmente, e incluso en aquellos supuestos en los que el tratamiento no se incluya en ninguno de los instrumentos previamente indicados, el responsable debe valorar el posible riesgo que puede implicar dicho tratamiento teniendo en cuenta la naturaleza, alcance, contexto o fines del mismos[439], así como especialmente si se implementan innovaciones tecnológicas en dicho tratamiento[440].

Por tanto, de una interpretación global de todos estos instrumentos orientativos y legales podemos llegar a concluir que en la mayoría de los supuestos donde se desarrollen o desplieguen sistemas de toma de decisiones que traten datos personales se requerirá necesariamente de la realización de una EIPD.

Tipo de tratamiento	**Necesidad de realizar la EIPD**
Elaboración de perfiles +Decisión plenamente automatizada que genera efectos relevantes	Si, artículo 35.3.a) RGPD
Elaboración de perfiles + decisión parcialmente automatizada que genera efectos relevantes	Si, artículo 35.3.a) RGPD
Decisión plenamente automatizada que genera efectos significativos	Probablemente sí, suma de criterios del grupo del artículo 29 y AEPD.
Decisión parcialmente automatizada que genera efectos relevantes	Es probable que sí, suma de criterios del grupo del artículo 29 y AEPD.
Decisión total o parcialmente automatizada que no genera efectos relevantes	No es tan probable, hay que valorar la suma de criterios del grupo del artículo 29 y AEPD.
Elaboración de perfiles	Es muy probable, suma de criterios del grupo del artículo 29 y AEPD.

[438] Comité Europeo de Protección de Datos. *EDPB-EDPS. 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).* Resolución de 18 de junio de 2021. Apartado 21, pág.9.

[439] Agencia Española de Protección de Datos. *Guía práctica para las Evaluaciones de Impacto en la Protección de los Datos sujetas al RGPD,* 2018, págs.13 y 14. En el mismo sentido se ha pronunciado la Autoritat Catalana de Protecció de dades. *Guía práctica. Evaluación de impacto relativa a la protección de datos*. Enero de 2018, versión 2.0, págs. 37 y ss.

[440] La tecnología es un elemento radical y sobre el que hace hincapié el RGPD para considerar que es necesaria la EIPD. Los sistemas automatizados que se basen en nuevas técnicas como el *machine learning* o *deep learning* aunque no generen efectos relevantes en el particular pueden presentar riesgos ya que los mismos integran dichas tecnologías. Véase también, la resolución de la AEPD Nº E/02666/2020, pág.6. Disponible en: https://www.aepd.es/es/documento/e-02666-2020.pdf

B) Valoración negativa de que no existe alto riesgo

Existen determinados tratamientos que por sus características no requieren de la realización de una EIPD por no entrañar un alto riesgo. El propio RGPD establece en su artículo 35.10 que no será necesario realizar una EIPD para aquellos tratamientos que sean necesarios para el cumplimiento de una obligación legal, el cumplimiento de una misión realizada en interés público o en el ejercicio de poderes públicos conferidos al responsable, siempre que se haya realizado una EIPD en el contexto de la adopción de dicha base jurídica[441]. De esta manera, si a través de una norma se implanta un sistema de decisiones automatizadas y se realiza una EIPD previa sobre el concreto tratamiento que se pretende implantar con esa legislación habría que entender que ya no es necesario realizar una nueva EIPD. Esta eximente para no realizar la EIPD se ha de utilizar con cautela[442]. Y es que, dado que la EIPD que se realiza durante la fase legislativa puede no haber tenido en cuenta todos los elementos técnicos específicos que son relevantes previos a la puesta en marcha del tratamiento que se legitima en este caso, el uso de un algoritmo para la toma de decisiones podrá requerir la realización de una nueva EIPD previa a la puesta en funcionamiento del concreto tratamiento. Además, hay que tener en cuenta que, dada la variabilidad que caracteriza el funcionamiento de muchos modelos algorítmicos a la hora de adoptar decisiones automatizadas, la necesidad de realizar EIPD posteriores a la puesta en funcionamiento de estos sistemas será obligatoria independientemente de que se haya realizado una EIPD en la fase legislativa que habilitó al mismo[443].

Junto al Artículo 35.10 del RGPD, tanto el GT29 como las distintas autoridades de control han hecho referencia a otra serie de tratamientos que no son considerados de alto riesgo[444]. El responsable del tratamiento deberá en su caso valorar que

[441] Agencia Española de Protección de Datos. *Lista orientativa de tipos de tratamientos que no requieren una evaluación de impacto relativa a la protección de datos según el artículo 35.5 RGPD*. Resolución del 4 de septiembre de 2019. Véase también: Grupo del Artículo 29. *Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679***.** Adoptadas el 4 de abril de 2017. Revisadas por última vez y adoptadas el 4 de octubre de 2017, pág.14

[442]Information Commisioner´s Office. *Data Protection Impact Assessments (DPIAs)*. Documento adoptado el 22 de marzo de 2018, pág..27.

[443] Ya mencionamos en páginas anteriores la diferencia entre sistemas estáticos y dinámicos. Estos últimos tienen una mayor capacidad para alterarse una vez que interactúa con el entorno sobre el cual se adoptan las decisiones automatizadas.

[444] La AEPD ha desarrollado una aplicación sencilla denominada "Facilita" que ayuda a las empresas cumplir con la normativa de protección de datos cuando llevan a cabo tratamientos de datos personales

efectivamente dicho tratamiento no genera un alto riego, debiendo justificar tal decisión[445]. Será especialmente interesante esta justificación en aquellos supuestos dudosos donde el responsable haya implementado un sistema de toma de decisiones automatizadas y haya considerado que el mismo no genera alto riesgo para los derechos y libertades de los interesados.

Finalmente, cabe indicar que pese a que el responsable llegue a la conclusión inicial de que su tratamiento no genera un altor riesgo, ello no significa que en un momento futuro no pueda alcanzar ese nivel. Esto obliga al responsable a estar atento, sobre todo, cuando se hayan alterado determinadas circunstancias que pueden hacer que efectivamente ahora y no antes existe tal nivel de riesgo[446]. Otros cauces que pueden avisar al responsable de esa alteración del riesgo pueden devenir de los propios particulares afectados por las decisiones o de las auditorías que se realicen del algoritmo y que muestren el riesgo potencial no apreciado inicialmente.

3. La realización de la evaluación de impacto

Una vez el responsable ha considerado que la evaluación de impacto es necesaria realizarla porque el tratamiento de datos conlleva un alto riesgo, el siguiente paso es iniciarla. Es en este momento cuando el responsable debe decidir qué metodología ha de utilizar para llevar a cabo tal EIPD. Así, el RGPD concede cierta flexibilidad a los responsables para, en función de distintos factores, establecer una metodología u otra a la hora de llevar a cabo dicha EIPD.

Pese a lo dicho, el RGPD establece un contenido mínimo que ha de integrar toda EIPD, el cual, esencialmente ha de integrar un descripción sistemáticas de las operaciones previstas, una evaluación de la necesidad y de la proporcionalidad de

---

de escaso riesgo. Más información en: https://www.aepd.es/es/guias-y-herramientas/herramientas/facilita-rgpd

[445] Grupo del Artículo 29. *Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679***.** Adoptadas el 4 de abril de 2017. Revisadas por última vez y adoptadas el 4 de octubre de 2017, pág.13

[446] Artículo 35.11 del RGPD. Así, los cambios pueden suponer tanto que disminuya el nivel de riesgo como que aumente, llegando incluso a ser obligatoria la realización de una EIPD cuando inicialmente no lo era. Agencia Española de Protección de Datos. *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales*. Junio 2021,pág.128.

dichas operaciones, un análisis de los riesgos que estas generan en los derechos y libertades y las medidas previstas para afrontar esos riesgos[447].

El objetivo que se pretende con este contenido mínimo es que, por una lado se logre una cierta estandarización a la hora de llevar a cabo la EIPD por parte de distintas organizaciones y, por otro lado, se consiga una aplicación efectiva de esta medida ya que se fijan los parámetros mínimos que el legislador considera necesarios para conseguir una protección adecuada de los derechos y libertades de los interesados.

Pues bien, antes de entrar a analizar los aspectos más importantes de la evaluación de impacto que pueden tener especial incidencia en la toma de decisiones automatizadas conviene indicar dos elementos relevantes relacionados con el análisis de esta herramienta:

En primer lugar, y en coherencia con el estudio que hasta ahora hemos desarrollado sobre los sistemas de toma de decisiones automatizadas, en las siguientes páginas analizaremos la incorporación de la EIPD aplicada a las dos grandes fases que comprende el ciclo de vida de un sistema algorítmico. Esto es, la fase de diseño y la de aplicación. Somos conscientes que dentro de cada una de esas fases existen a su vez otra serie de etapas, estas últimas son consideradas en este trabajo como operaciones específicas incorporadas a cada uno de las dos grandes fases a las que hemos aludido.

En segundo lugar, la metodología que se utilizará para explicar los distintos procesos que componen tal medida de responsabilidad activa será una mezcla de las diversas propuestas metodológicas que han establecido algunas de la autoridades de protección de datos europeas y regionales. Nos centraremos sobre todo en las diseñadas por la AEPD[448].

---

[447] El artículo 35.3 del RGPD establece que la evaluación de impacto relativa a la protección de los datos requerirá de: *a) una descripción sistemática de las operaciones de tratamiento previstas y de los fines del tratamiento, inclusive, cuando proceda, el interés legítimo perseguido por el responsable del tratamiento; b) una evaluación de la necesidad y la proporcionalidad de las operaciones de tratamiento con respecto a su finalidad; c) una evaluación de los riesgos para los derechos y libertades de los interesados a que se refiere el apartado 1, y d) las medidas previstas para afrontar los riesgos, incluidas garantías, medidas de seguridad y mecanismos que garanticen la protección de datos personales, y a demostrar la conformidad con el presente Reglamento, teniendo en cuenta los derechos e intereses legítimos de los interesados y de otras personas afectadas.*

[448]Concretamente, se han tenido en cuenta las Guías de la ICO (Británica), CNIL (Francesa), AEPD (2018 y 2021), APDCAT(Cataluña) y Grupo del Artículo 29.

A) Momento para realizar la EIPD

La EIPD se ha de realizar antes de iniciar el tratamiento de datos personales. Trasladado al ciclo de vida de los sistemas de decisiones automatizadas resulta conveniente que al menos se realice una antes de comenzar el desarrollo del sistema y otra previa a la puesta en marcha del mismo.

En la fase del diseño, pese a que algunas operaciones del tratamiento no estén todavía fijadas o se desconozcan, la evaluación de impacto deberá realizarse. En estos supuestos, la actualización de la EIPD será continua. Ello requerirá en muchas ocasiones la repetición de los pasos concretos que están comprendidos dentro de la EIPD ya que estas primeras fases están sometidas a continuos cambios[449]. Así, por ejemplo, puede que inicialmente se hayan previsto el tratamiento de unos determinados datos personales pero, conforme va avanzando el desarrollo del proyecto algunos de estos datos dejen de ser necesarios. En otros casos puede ser inverso el proceso, se proyecta el tratamiento de unos datos, pero, a raíz de las inferencias que se pueden obtener se comienzan a tratar nuevos datos que pueden llegar a considerarse especialmente sensibles[450]. A su vez, la elección de determinados algoritmos para elaborar un modelo puede ser posteriormente alterada al comprobar que los resultados esperados por la precisión de los mismos no resultan óptimos, por lo que se acaba optando por el uso de otras técnicas. Es por ello que una EIPD debe ser revisada y reevaluada con regularidad. Sobre todo cuando exista una variación relevante de alguna de las operaciones que forman parte del tratamiento de datos, tal y como ocurre en aquellos ámbitos donde los sistemas automatizados continuamente se están alterando[451].

[449] Grupo del Artículo 29. *Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679***.** Adoptadas el 4 de abril de 2017. Revisadas por última vez y adoptadas el 4 de octubre de 2017, pág.16.

[450] Aplicar esto en la práctica significa que las evaluaciones de impacto deben realizarse durante cada fase del ciclo de vida algorítmico. Durante la etapa de diseño y desarrollo, las evaluaciones de impacto deben evaluar cómo es probable que funcione un algoritmo, así como garantizar que funcione según lo previsto e identificar cualquier proceso o supuesto problemático. Esto brinda la oportunidad de modificar el diseño de un algoritmo en una etapa temprana, incorporar el cumplimiento de los derechos humanos, incluidos los mecanismos de monitoreo, desde el principio o detener el desarrollo si no se pueden abordar los problemas de derechos humanos. Las evaluaciones de impacto también deben llevarse a cabo en la etapa de implementación a fin de monitorear los efectos durante la operación. En: MCGREGOR, L; MURRAY, D.; NG, V: "International human rights law as a framework for algorithmic accountability". *International and Comparative Law Quarterly*, 68(2), 2019, pág.330.

[451] Grupo del Artículo 29. *Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679***.** Adoptadas el 4 de abril de 2017. Revisadas por última vez y adoptadas el 4 de octubre de 201, .pág.16.

4. Evaluación de los riesgos

Tal y como señala el Artículo 35.7.c) del RGPD, la EIPD debe incluir una evaluación de los riesgos para los derechos y libertades de los interesados. Nos remitimos a lo ya señalado en el apartado referido al análisis de riesgos[452].

5. Evaluación de la necesidad y proporcionalidad

Con el objetivo de seguir recopilando mayor información relacionada con el tratamiento que se pretende llevar a cabo, el responsable deberá también valorar la proporcionalidad de las operaciones que conforman el tratamiento con respecto la finalidad del mismo[453]. Para ello se ha de tener clara la finalidad u objetivo que se pretende con el tratamiento de datos[454]. Junto a la finalidad del tratamiento, también será relevante establecer la base legitimadora del tratamiento, la cual, estará estrechamente relacionada a la finalidad prevista para el mismo[455].

El principio de proporcionalidad requiere de la superación de un triple test. Esto es, el juicio de idoneidad o adecuación, el juicio de necesidad y el juicio de proporcionalidad en sentido estricto[456]. Es turno de analizarlos.

En primer lugar, en el *juicio de idoneidad* se debe valorar si las distintas operaciones que conforman el tratamiento logran cumplir con la finalidad del mismo. Por lo que se refiere a este primer análisis, será habitual que el responsable pueda justificar tal juicio ya que como regla general la adecuación de las operaciones responderá al tratamiento que se pretenda llevar a cabo. En esta fase por ejemplo el responsable deberá valorar si el sistema algorítmico que pretende implantar es adecuado

[452] Véase el Capítulo III, aparatado I de esta tesis.
[453] Artículo 35.7.b) del RGPD.
[454] Autoritat Catalana de Protecció de dades. *Guía práctica. Evaluación de impacto relativa a la protección de datos*. Enero de 2018, versión 2.0, pág.54.
[455] Tanto la finalidad del tratamiento como la licitud del mismo se analizarán en los respectivos apartados del capítulo IV. Véase: principio de licitud (Capítulo IV, apartado I) y principio de limitación de la finalidad (Capítulo IV, apartado II).
[456] Sobre el principio de proporcionalidad véase entre otros: DOMÉNECH PASCUAL, G: *Derechos fundamentales y riesgos tecnológicos*. Ed. Centro de Estudios Políticos y Constitucionales. 2006. Págs.159 y ss. Específicamente, la aplicación del principio de proporcionalidad en la normativa de protección de datos puede verse en: Agencia Española de Protección de Datos. *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales.* Junio 2021,págs 138 a 146.

para el contexto donde dicho sistema desplegará sus efectos. Como indicábamos en páginas anteriores, es posible que el objetivo para el que se diseñó un sistema automatizado acabe desplegando sus efectos en entornos distintos a los que se modelaron durante esa primera fase de desarrollo. En estos supuestos, aunque estos sistemas pueden seguir ofreciendo resultados óptimos, conviene prestar especial atención en este momento. También en este momento será necesario valorar si el sistema ha tenido en cuenta a todos los sectores de la población sobre la que puede recaer la toma de decisiones automatizadas[457]. Así, dado que estos grupos de la sociedad podrían no estar representados en la realidad que ha perfilado el algoritmo, se podría llegar a considerar que el sistema introducido no es adecuado para la finalidad que se persigue con la implementación del mismo. Junto a estos elementos, también será recomendable justificar la adecuación de estos sistemas teniendo como base los resultados de precisión que se deriven de la evaluación de la bondad de los modelos algorítmicos[458]. De esta manera, si la precisión es mayor, más idóneo será el sistema para el objetivo que se pretende.

En segundo lugar, por lo que respecta al *juicio de necesidad*, se ha de analizar si las operaciones previstas en el tratamiento son necesarias. Esto es, que no existan otras operaciones más moderadas o menos restrictivas que consigan el objetivo marcado con la misma eficacia[459]. Por ejemplo, la AEPD analizó la proporcionalidad de utilizar técnicas de reconocimiento facial inteligentes a la hora de realizar exámenes online durante los meses más graves de la pandemia[460]. Así, consideró prevalente esta técnica frente a otras medidas como la evaluación presencial, la cual, era prácticamente

[457] En este sentido, la propuesta de Reglamento europeo sobre la regulación de los sistemas de inteligencia artificial obliga a los proveedores de sistemas de IA de alto riesgo a incorporar en la documentación técnica de dicho modelo algorítmico, las personas o grupos de personas con los que se pretende utilizar el sistema. Véase el artículo 11 y aparatado 2.b) del Anexo IV de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

[458] SORIANO, ARNANZ,A: "Decisiones automatizadas y discriminación: aproximación y propuestas generales". *Revista General de Derecho Administrativo,* 56, 2021, pág.24.

[459] Especial atención siempre ha merecido el test de necesidad cuando un responsable pretende implantar una nueva tecnología en el tratamiento de datos. Así, el Tribunal Supremo ha considerado que un proyecto que pretendía implantar una empresa consistente en obligar al trabajador a aportar su teléfono móvil con la finalidad de instalar una App y poder geolocalizarlos pudiendo conocer en todo momento la situación de los pedidos que los mismos reparten vulnera el derecho a la protección de datos ya que existían otros medios menos invasivos para alcanzar la finalidad del tratamiento de datos que pretendía el responsable del tratamiento. En: STS 518/2021 sala de lo social, - ECLI:ES:TS:2021:518. FJº 2º.

[460] Agencia Española de Protección de Datos. Informe de 8 de mayo de 2021. Informe N/REF: 0036/2020.

imposible en esos momentos. Ahora bien, este tipo de tratamientos algorítmicos en ningún momento podían considerarse por defecto sino que deberían quedar limitados a aquellas enseñanzas y asignaturas concretas que, por su importancia, complejidad u otras circunstancias de especial incidencia no aconsejaran acudir a otras opciones como la evaluación continua o hicieran excesivamente gravoso la adopción de otros medios como el control por videocámara o la realización de exámenes orales. De esta manera, en nuestra opinión, la automatización del proceso decisorio debería responder a una auténtica necesidad de esa organización y no sólo a un deseo de utilizar esta tecnología porque la misma es vanguardista o está de moda.

Existen distintos elementos que se pueden valorar a la hora de analizar el test de necesidad. Centrándonos nuevamente en la fase de despliegue de un sistema algorítmico podemos destacar varios:

- En primer lugar se debe justificar las razones por las que se ha optado por automatizar el proceso decisorio, esto es, la sustitución del humano por la máquina. En este sentido, el responsable del tratamiento deberá evidenciar que su inclusión no obedece a un simple capricho[461]. Para ello puede alegar que el sistema al menos es tan preciso como el humano cuando realiza dichas labores[462]. También se puede justificar esa sustitución parcial de la persona por la máquina cuando se demuestre que la conjugación entre el factor humano junto con la salida emitida por el sistema ofrezca resultados mejores que si únicamente fuera el humano el que participa en el proceso[463]. Será necesario por tanto realizar un análisis comparativo entre el rendimiento obtenido por un operador humano cualificado frente a los resultados arrojados por el sistema algorítmico[464]. En otros supuestos, esta justificación puede devenir de la propia imposibilidad del humano para realizar dichas operaciones o porque las misma resultan altamente costosas e ineficaces. En este sentido, cada vez se establecen mayores obligaciones de control

[461] Como indica la ICO, si los sistemas de inteligencia artificial complementan o reemplazan la toma de decisiones humana, en la EIPD se ha de documentar cómo el proyecto podría comparar la precisión humana y algorítmica en paralelo para justificar mejor su uso. En: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/what-are-the-accountability-and-governance-implications-of-ai/

[462] BERMAN, E:"A government of laws and not of machines". *B.U. L. RE*, 98, 2018, pág.22.

[463] Por ejemplo, el sistema VeriPol que utiliza la policía nacional es más preciso cuando los resultado que emite el algoritmo son analizados por la persona que finalmente adopta la decisión. En: QUIJANO-SÁNCHEZ,L; LIBERATORE,F; CAMACHO-COLLADOS,J, CAMACHO-COLLADOS,M: "Applying automatic text-based detection of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the police", op.cit., págs 160 y 161.

[464] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.45.

por parte de las grandes plataformas para controlar el contenido ilícito que se vierte en estas. Aunque los humanos conservan un papel relevante en el control del contenido, resulta lógico que estas organizaciones acaben acudiendo a la automatización para cumplir con esa finalidad de control[465]. Por ejemplo, el artículo 17.4 de la Directiva de la UE sobre los derechos de autor establece que plataformas como YouTube deben ser lo más diligentes posibles a la hora de evitar la publicación de contenido ilícito que atente contra los derechos de autor. En el caso de que no se realice esa diligencia adecuada, las responsabilidades de los contenidos que se suban recaerán sobre estas plataformas[466]. Teniendo en cuenta que en YouTube se publican cerca de 35 horas de vídeo por minuto, lo que representa varios centenares de miles de vídeos cada día[467], el control de estas publicaciones requerirá de la presencia de uno o varios algoritmos[468].

- A su vez, la necesidad también puede evaluarse tomando como referencia las distintas funcionalidades que se han implantado en el sistema algorítmico para que genere el menor impacto en la esfera de los particulares una vez que adopta la decisión. A modo de ejemplo, un sistema utilizado por una plataforma social que tiene como finalidad detectar el control de contenido que es potencialmente ilícito puede estar configurado de diversas formas una vez que infiere automáticamente

[465] La normativa de la Unión Europea cada vez está imponiendo mayores deberes de control a las plataformas sobre el contenido que se difunde en sus entornos. Véase la Propuesta DEL PARLAMENTO EUROPEO Y DEL CONSEJO sobre un Mercado Único de Servicios Digitales (Ley de Servicios Digitales) y por la que se modifica la Directiva 2000/31/CE.

[466] 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. Artículo 17.4.

En este mismo sentido, el TJUE se ha pronunciado recientemente en su sentencia de 22 de junio de 2021, asuntos acumulados C-682/18 (YouTube) C-683/18 (Cyando). Concretamente se analizaba si YouTube realiza o no una comunicación al público en el sentido de la Directiva 2001/29/CE del Parlamento Europeo y del Consejo, de 22 de mayo de 2001, sobre derechos de autor indicando que, tal comunicación existirá y por tanto se pueden exigir responsabilidades a esta plataforma del contenido ilegal que se suba cuando esta se abstenga de aplicar las medidas técnicas apropiadas que cabe esperar de un operador normalmente diligente en su situación con el fin de combatir de forma creíble y eficaz violaciones de los derechos de autor en esa plataforma (FJº 102). Entre esas medidas destaca el uso de algoritmos de control del contenido.(fº 94). Texto de la sentencia disponible en:
https://curia.europa.eu/juris/document/document.jsf;jsessionid=E544BB1B1477BE990C7B56B32A026656?text=&docid=243241&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=21309305

[467]Conclusiones del abogado general Sr. Henrik Saugmandsgaard presentadas el 16 de julio de 2020, Asuntos acumulados Frank Peterson contra Google LLC, C-682/18 y YouTube LLC, YouTube Inc., Google Germany GmbH C-683/18,y Elsevier Inc. Contra Cyando AG (C-683/18). Apartado 16. Resolución disponible en:
https://curia.europa.eu/juris/document/document.jsf;jsessionid=8667950A1A86F73CCDE040C63B998688?text=&docid=228712&pageIndex=0&doclang=ES&mode=req&dir=&occ=first&part=1&cid=1884929

[468] El control del contenido que se ejecuta por esta plataforma lo lleva a cabo a día de hoy el algoritmo denominado Content ID. Visto en: https://support.google.com/youtube/answer/2797370?hl=ca

que dicho contenido es ilícito. Las actuaciones posteriores que en su caso se hayan planificado pueden afectar en mayor o menor medida a los usuarios que subieron ese contenido y que están ejerciendo su derecho de expresión. Así, no será el mismo grado de afectación para los derechos de los usuarios la retirada automática del contenido a que en su caso se opten por medidas menos restrictivas. Medidas como: i) marcar el contenido como potencialmente ilícito o falso, ii) proporcionar enlaces de otras noticias que contradicen ese contenido tildado de falso a los usuarios que visualizan esa noticia potencialmente ilícita, iii) bajar la puntuación de la noticia falsa para que no se comparta con tanta facilidad por parte del resto de usuarios, iv) generar fricción en las publicaciones para evitar que se interactúen con ellas o se compartan[469]. De esta manera, a través de estas alternativas se consigue el objetivo marcado con el despliegue del sistema algorítmico restringiendo en menor medida los derechos de los particulares afectados por las decisiones.

- Por otro lado, el juicio de necesidad también debe abarcar las razones por las que el responsable ha optado por un modelo más o menos interpretable. A priori, y teniendo en cuenta este juicio de necesidad, el responsable debería acudir a la medida menos restrictiva de derechos, en este caso, apostar por un modelo más interpretable[470]. Sin embargo, entendemos que este test podría superarse a pesar de que se opte por un algoritmo menos trasparente cuando se den una serie de circunstancias acumulativas; a) la precisión entre el modelo que es más interpretable y el que menos es muy considerable, b) se han previsto otras medidas para compensar esa falta de interpretabilidad y además, c) no existe un requisito indispensable que obligue al responsable del tratamiento a prever una alta explicabilidad de la decisión.

En tercer lugar, el *juicio de proporcionalidad en sentido estricto* exige valorar si las operaciones implantadas resultan suficientemente equilibradas en la medida que

[469] El Consejo asesor de contenido de Facebook ha recomendado a esta plataforma social el despliegue de medidas menos restrictivas que no consistan en la eliminación del contenido cuando el mismo resulte controvertido. Decisión del caso 2021-008-FB-FBR. 19 de agosto de 2021. Apartado 10.1 Resolución disponible en: https://oversightboard.com/decision/FB-B6NGYREK/

[470] Piénsese por ejemplo en el principio de transparencia o los derechos de información sobre la lógica del tratamiento, los cuáles, se ven más limitados cuando un algoritmo es poco interpretable respecto de aquellos que son más transparentes. Una organización que adquiere un sistema basado en la técnica del *deep leanring* deberá en su caso justificar por qué ha optado por utilizar un sistema algorítmico tan opaco respecto de otros más interpretables.

tales operaciones conceden más beneficios o ventajas para el interés general, que perjuicios sobre otros bienes o valores en conflicto, en este caso, el derecho a la protección de datos personales y otros derechos que puedan resultar afectados por ese tratamiento. En este último punto se valora estrictamente la proporcionalidad del tratamiento con relación al resto de bienes o valores que pueden verse perjudicados. Concretamente será necesario por un lado valorar qué derechos e intereses están amparados en la implantación de ese sistema y, por otro, los bienes y derechos afectados por el despliegue del mismo. Una vez se hayan valorado esos intereses en juego se debe analizar si la injerencia realizada en los derechos afectados por este sistema generan más beneficios para los primeros que perjuicios para los segundos. En ningún caso, se puede asumir la negación absoluta del derecho a la protección de datos y vaciarle de su contenido esencial[471]. Retornando al supuesto del sistema VeriPol que utiliza la policía nacional. Lo que deberíamos de ponderar es los intereses generales en los que se ampara esta medida, entre los que se encuentra la lucha contra la delincuencia y la racionalización de los recursos públicos y, los derechos y libertades de los particulares que se someten a tales decisiones. Del resultado meditado que se realice de esa ponderación deberemos considerar que se supera o no tal test de proporcionalidad. Por ejemplo, para justificar tal nivel de proporcionalidad, dado que se ha demostrado que existe un alto índice de denuncias falsas de robos de móviles y ello comporta un número importante de horas de dedicación de la policía a esclarecer los hechos. Un sistema algorítmico como VeriPol podría llegar a considerarse proporcionado en la media que, los beneficios para el interés general, en este caso, la gestión adecuada de recursos policiales, son mayores que los perjuicios que en su caso puede genera en los particulares cuando dichas denuncias son analizadas por este sistema. Por supuesto, será necesario valorar las garantías y medias que se hayan implantado para proteger los derechos de los particulares que se ven perjudicados por este sistema teniendo en cuenta la restricción prevista, ya que si no, difícilmente se superará dicho test de proporcionalidad.

En otros supuestos, también puede ser recomendable acudir a la legislación que habilita o sobre la cual se justifica el despliegue de un sistema de decisiones automatizadas para poder discernir cuales son los posibles derechos e intereses que

[471] Agencia Española de Protección de Datos. *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales.* Junio 2021,pág.139.

entran en conflicto y cómo esa norma potencia o dar cierta preponderancia a unos u otros objetivos. Así, por ejemplo, ya hemos indicado que la Directiva de la UE sobre derechos de autor de forma indirecta legitima el uso de sistemas automatizados de control del contenido debido a las obligaciones que dicha norma impone a las plataformas[472]. Tal y como establece esta norma, los derechos que entran en conflicto cuando se incorpora un sistema de control de contenido en estas plataformas son por un lado, el derecho a la propiedad y específicamente la intelectual y por otro, los derechos de información y libertad de expresión[473]. De una lectura de los distintos considerandos y preceptos de esta norma se puede interpretar que el legislador europeo, a la hora de ponderar los derechos que entran en conflicto, acaba inclinándose por el derecho de propiedad intelectual en defecto de los derechos libertad de expresión e información, imponiendo a estas plataformas mayores deberes de control sobre la información que circula en dichos entornos. Así lo ha entendido también el abogado general SR Henrik en los asuntos acumulados C-682/18 y C-683/18 Frank Peterson contra Google LLC al considerar que la irrupción de esta nueva Directiva ha alterado el equilibrio de los intereses que hasta la fecha había establecido la normativa previa[474]. Esta priorización por controlar el contenido que se vierte en estas plataformas y por tanto priorizar el

[472] No existe una habilitación expresa al uso de sistemas de toma de decisiones automatizadas a las plataformas en todo el texto de la Directiva. Sin embargo, teniendo en cuenta las nuevas obligaciones de control del contenido que se derivan de esta norma y la gran cantidad de información que puede circular por las mismas, el control de dicho contenido requerirá habitualmente del uso de sistemas automatizados.

[473] Considerando 70 de 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.

[474] Así, el abogado general señaló que mediante la Directiva 2000/31 (norma que ha sido sustituida por la Directiva de 2019 sobre derechos de autor), el legislador de la Unión pretendía favorecer el desarrollo de los prestadores de servicios intermediarios, con el fin de estimular de forma más general el crecimiento del comercio electrónico y de los «servicios de la sociedad de la información» en el mercado interior. Por tanto, se pretendía evitar imponerles una obligación general de diligencia que pudiera poner en peligro la rentabilidad de sus servicios. Se debían salvaguardar los intereses de los titulares de los derechos y ponderarlos con la libertad de expresión de los usuarios de dichos servicios en el marco de mecanismos de notificación y retirada. (Apartado 245). Es indudable que las circunstancias han cambiado desde la adopción de estas Directivas. Los prestadores de servicios intermediarios ya no tienen las mismas características y es posible que este equilibrio ya no esté justificado. (Apartado 246). Pues bien, debo recordar que el legislador de la Unión precisamente acaba de reevaluar de cara al futuro el equilibrio de los derechos e intereses en materia de derechos de autor. En efecto, en el transcurso de los presentes procedimientos prejudiciales ha entrado en vigor la Directiva 2019/790. (Apartado 247). CONCLUSIONES DEL ABOGADO GENERAL SR. HENRIK SAUGMANDSGAARD ØE presentadas el 16 de julio de 2020. Asuntos acumulados C-682/18 y C-683/18, caso Frank Peterson contra Google LLC, YouTube LLC, YouTube Inc., Google Germany GmbH (C-682/18) y Elsevier Inc. Contra Cyando AG (C-683/18). Apartados 245 a 248.Texto disponible en: https://curia.europa.eu/juris/document/document.jsf;jsessionid=8667950A1A86F73CCDE040C63B998688?text=&docid=228712&pageIndex=0&doclang=ES&mode=req&dir=&occ=first&part=1&cid=1884929
En este mismo sentido, véase: BARRIOS,ANDRÉS,M: "La sentencia del TJUE YouTube y la responsabilidad de los operadores de plataformas digitales". *Diario La Ley*, Nº 53, Sección Ciberderecho, 19 de Julio de 2021, pág.3.

derecho a la propiedad intelectual respecto de otros como la libertad de expresión se ha de trasladar al algoritmo que se pretenda implantar en ese ecosistema. Es decir, aquel algoritmo que refleje de forma más adecuada la realidad que se pretende diseñar a través de la norma será el que se acerque o tenga más posibilidades de superar el test de proporcionalidad en sentido estricto. Una vez más, en estos supuestos será esencial valorar los resultados que se hayan obtenido del sistema en la fase de desarrollo, siendo relevante evaluar las métricas como la tasa de verdaderos positivos, la tasa de verdaderos negativos, el valor predictivo positivo o el valor predictivo negativo. Además, la priorización de unos derechos o intereses respecto de otros requerirá la compensación con otras garantías.

Finalmente procede indicar que la proporcionalidad en sentido estricto a la hora de incorporar un sistema automatizado ya ha sido analizada por diversos tribunales. Así, el Tribunal Constitucional Alemán valoró la constitucionalidad de una norma que legitimaba el uso de un sistema de reconocimiento automático de matrículas[475]. Este órgano judicial, tras considerar que gran parte de las disposiciones de esta norma eran conforme a la Constitución[476], señaló que un sistema que realizaba un control indiscriminado y encubierto de todas las matrículas de vehículos sin establecer criterios específicos sería contrario a la Constitución ya que la injerencia realizada sobre el derecho a la autodeterminación informativa no resultaría proporcional debido a que los intereses jurídicos sobre los que se basaba la implantación de esta medida no eran de suficiente peso comparable a los perjuicios que causaba dicha injerencia en los particulares[477]. En España, la Audiencia Provincial de Barcelona ha denegado a una importante cadena de supermercados el uso de sistemas de reconocimiento facial para detectar a condenados por delito de robo y cuya entrada a los establecimientos de esta empresa está prohibida. Para este tribunal, la medida no es proporcional ya que esta no persigue intereses generales sino meros intereses privados[478]. De estas resoluciones judiciales se desprende que un sistema automatizado que realice controles masivos de la población sin establecer medidas específicas que justifiquen ese control discriminado

[475] BVerfG, Order of the First Senate of 18 December 2018 - 1 BvR 142/15 -, Resolución de 18 de diciembre de 2018paras. 1-176. Resolución disponible en: http://www.bverfg.de/e/rs20181218_1bvr014215en.html

[476] Véase los párrafos 113 y ss de la resolución judicial comentada previamente.

[477] Párrafos 104 y ss de la resolución judicial comentada previamente.

[478] Audiencia Provincial de Barcelona, Sección 9ª, Auto 72/2021 de 15 Feb. 2021, Rec. 840/2021, FJº3. En el mismo sentido también se ha pronunciado la AEPD sobre la proporcionalidad de dicho sistema. En: Agencia Española de Protección de Datos. Resolución Nº: PS/00120/2021,págs.80, 83 y ss.

podría no superar ese test de proporcionalidad[479]. Ello es relevante teniendo en cuenta que cada vez se está apostando por la implementación de sistemas de control de grandes grupos de la población con fines de todo tipo como el control del fraude fiscal o de la seguridad social[480].

6. La participación de los interesados

El RGPD establece que cuando proceda, el responsable recabará la opinión de los interesados o de sus representantes con relación al tratamiento o tratamiento de datos sobre el que se está realizando la EIPD (Artículo 35.9). A priori no se trata de una medida obligatoria, sin embargo, si se procede a la misma, es posible que la protección de los derechos de los individuos se vea más reforzada y con ello el respeto del principio de responsabilidad activa y por consiguiente de la normativa de protección de datos. En este sentido, el Grupo de expertos de la Comisión Europea apuesta por la participación de las partes interesadas que se pueden ver afectadas de manera directa o indirecta por los sistemas algorítmicos a lo largo de todo su ciclo de vida[481]. La EIDP puede ser por tanto un momento esencial para dicha participación de la sociedad[482]. En nuestra opinión, puede resultar altamente recomendable que en este momento los responsables consulten o recaben la opinión de asociaciones u organizaciones especializadas. Estas últimas podrán aportar sus enfoques orientados al cumplimiento ético y normativo del tratamiento de datos que se está evaluando y sobre el cual se adoptarán decisiones automatizadas y se elaborarán perfiles.

---

[479] En este mismo sentido se ha pronunciado el Tribunal de Apelación del Reino Unido con relación al uso de un sistema de reconocimiento facial utilizado por la policía de Gales del Sur para prevenir conductas delictivas. En palabras del tribunal, no se establece ningún requisito normativo sobre dónde puede tener lugar el despliegue de este tipo de aplicaciones algorítmicas. Es decir, el alcance es muy amplio y sin límites aparentes. Tampoco se indica por ejemplo el lugar dónde es posible que las personas puedan razonablemente resultar identificadas. Se deja en manos de la policía los lugares de control. En: Court of appeal (civil division) on appeal form the high court of justice queen´s bench division (administrative court). Case No: C1/2019/2670, (2020) EWCA Civ 1058. Apartado 130, pág.28. Texto disponible en: https://www.judiciary.uk/wp-content/uploads/2020/08/R-Bridges-v-CC-South-Wales-ors-Judgment-1.pdf

[480] Tal y como ocurre en Francia a través del artículo 154 de la Loi de finances pour 2020, sous le n° 2019-796 DC, le 20 décembre 2019 que legitima el análisis masivo de datos personales de los ciudadanos disponibles públicamente en las plataformas virtuales como método para luchar contra el fraude fiscal. Texto disponible en: https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000039683923

[481] Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Directrices éticas para una IA fiable*.2019, pág.23.

[482] MIRÓ-LLINARES,F: "Predictive Policing: Utopia or Dystopia? On attitudes towards the use of big data algorithms for law enforcement", *IDP. Revista de Internet, Derecho y Política.* No. 30, 2020, págs 12 y 13. Texto disponible en http://dx.doi.org/10.7238/idp.v0i30.3223

7. La consulta a la autoridad de control

En los supuestos en los que las medidas que se hayan decidido implantar no consigan mitigar el riesgo a un grado tolerable, el responsable del tratamiento deberá consultar a la autoridad de control, está última en su caso propondrá las medidas a implantar o la posibilidad de que incluso no puede llevarse a cabo dicho tratamiento. Esta consulta queda por tanto condicionada a que el responsable no prevea las medidas adecuadas.

8. Deficiencias que presenta la EIPD en el contexto de la toma de decisiones automatizadas y propuestas para mejorar su funcionalidad

Se han destacado por la doctrina algunas de las limitaciones que presenta la EIPD en el contexto de la toma de decisiones automatizadas.

En primer lugar, el responsable del tratamiento no tiene la obligación de divulgar o hacer pública la EIPD. Ello limita las posibilidades para que la sociedad en general y los propios particulares afectados por las decisiones puedan analizar si efectivamente se ha desplegado una EIPD adecuada y por tanto se han evaluado correctamente los riesgos presentes en un sistema algorítmico.

En segundo lugar, la mayor ventaja que presenta el principio de responsabilidad activa, esto es, diseñar más o menos medidas basadas en el enfoque del riesgo, presenta una gran hándicap cuando el responsable del tratamiento no se toma en serio este enfoque o tomándoselo no idee las medias adecuadas para reducir los riesgos que presentan la inclusión de sistemas automatizados[483]. En estos supuestos, dado que no existe una obligación directa para establecer unas medidas específicas para mitigar los riesgos presentes en el tratamiento analizado, la posibilidad de incumplir la normativa de protección de datos y por tanto vulnerar los derechos de los particulares sometidos a estos sistemas aumenta considerablemente.

En tercer lugar, y dado que el control del cumplimiento queda en manos de las autoridades de control, una dotación inadecuada de recursos y personal especializado

[483] El RGPD confía demasiado en el comportamiento de las empresas. En: MARGOT E KAMINSKI,M; GIANCLAUDIO,M: "Algorithmic impact assessments under the GDPR: producing multi-layered explanations", op.cit., pág.19.

asignado a estas para que realicen controles del funcionamiento de estos sistemas puede ahondar aún más en los inconvenientes señalados anteriormente[484].

Para combatir estas deficiencias y situar a la EIPD como el instrumento basilar sobre el que se estructuran el desarrollo y cumplimento de la normativa de protección de datos en este contexto sería necesario adoptar una serie de medidas.

Así, en primer lugar resultaría adecuado dotar de mayores recursos tanto humanos como económicos a las autoridades de control de protección de datos. En este sentido, sería necesario establecer una unidad específica que se focalice en el análisis de los sistemas algorítmicos y sus implicaciones cuando dichos sistemas traten datos personales. De esta manera, estos órganos especializados dedicarían parte de sus labores a evaluar y valorar las EIPD que se hayan realizado por las organizaciones que hayan desplegado estos sistemas.

Por otro lado, en segundo lugar, dado que las EIPD no son públicas, y la publicación de las mismas puede llevar a mostrar información sensible o que afecte a derechos e intereses protegidos de las organizaciones que las desarrollan, sería recomendable que terceros externos pudieran en determinados momentos visualizar y realizar estudios de cómo estos sistemas están operando. Aquí sería ideal que desde los sectores específicos de cada uno de los ámbitos donde estos sistemas irradian sus efectos se estructuren este tipo de organismos. Como dijimos, la participación de estas organizaciones puede realizarse durante la fase de la elaboración de la EIPD.

En tercer lugar, en nuestra opinión, los responsables del tratamiento deberían consultar a las autoridades de control siempre con carácter previo al despliegue de un sistema de decisiones automatizadas cuando se haya detectado un riesgo inherente alto o muy alto en la operación u operaciones que engloban dicho tratamiento de datos personales. De esta manera, la autoridad de control podría en su caso analizar dichas medidas y valorar si realmente son efectivas a la hora de reducir a o mitigar los altos riesgos que se derivan del uso de este tipo de sistemas automatizados[485]. Esta situación

[484] SORIANO, ARNANZ, A: "Decisiones automatizadas: problemas y soluciones jurídicas. más allá de la protección de datos". *Revista de Derecho Público: Teoría y Método.* Vol. 3 | 2021, *págs.115 y 116.*

[485] Así, la AEPD obliga a los partidos políticos que pretendan elaborar perfiles sobre la orientación política de las personas a consultar previamente a las autoridades sobre el tratamiento que pretende llevar a cabo o en su caso a facilitar la documentación sobre el análisis de riesgos y la EIDP realizada sobre dicho tratamiento. Artículo 7.1.5º de la Circular 1/2019, de 7 de marzo, de la Agencia Española de Protección de Datos, sobre el tratamiento de datos personales relativos a opiniones políticas y envío de propaganda electoral por medios electrónicos o sistemas de mensajería por parte de partidos políticos,

ya se prevé en diversas normas tanto en el plano europeo como nacional. Así, la normativa española que regula los tratamientos de datos personales con fines de prevención e investigación de infracciones penales establece la consulta previa obligatoria cuando en el tratamiento pretendido se usen tecnologías, mecanismos o procedimientos nuevos[486]. En la Unión Europea, el Reglamento contra los abusos sexuales de menores en línea obliga a las plataformas a detectar y retirar contenido pedófilo por medio de sistemas automatizados. Pues bien, antes de que se ponga en marcha el uso de estos algoritmos, se ha de consultar a la autoridad de control en materia de protección de datos sobre el despliegue de esas tecnologías[487].

Finalmente, también sería recomendable que las autoridades de protección de datos facilitarán guías o documentos que se centrarán específicamente en las medidas que en su caso los responsables de tratamiento de sistemas automatizados deberían implantar en función de los principales riesgos que pueden presentar este tipo de sistemas. Conceder un valor preponderante a la EIPD como medida de cumplimiento normativo se hace sumamente necesario dadas las virtualidades que esta puede ofrecer. Además, ello puede reforzar los derechos de los particulares que se ven sometidos a estos sistemas ante un ordenamiento jurídico que, a excepción del RGPD, prácticamente no ha regulado este fenómeno. .

9. La relevancia de la evaluación de impacto durante el ciclo de vida de los sistemas de toma de decisiones automatizadas

Del proceso que engloba la EIPD se puede desprender la importancia de esta herramienta a la hora de clarificar y establecer las medidas más relevantes que el responsable del tratamiento debe desplegar para mitigar los riesgos que acompañan todo tratamiento presente durante el ciclo de vida de un sistema de decisiones automatizadas. Recordemos que el análisis de riesgos se ha de realizar para cualquier tratamiento, sin embargo, la EIPD queda reservada a aquellos tratamientos que generen un alto riesgo.

federaciones, coaliciones y agrupaciones de electores al amparo del artículo 58 bis de la Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General.

[486] Artículos 35 y 36.1.b) de 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.

[487] Artículo 3.1.c) del REGLAMENTO (UE) 2021/1232 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 14 de julio de 2021 por el que se establece una excepción temporal a determinadas disposiciones de la Directiva 2002/58/CE en lo que respecta al uso de tecnologías por proveedores de servicios de comunicaciones interpersonales independientes de la numeración para el tratamiento de datos personales y de otro tipo con fines de lucha contra los abusos sexuales de menores en línea.

Para estos últimos casos, ambas herramientas quedan integradas la una con la otra. Como dijimos, la EIPD no deja de ser un análisis de riesgos reforzado.

Es turno de analizar a modo de resumen las principales virtudes que ofrecen estos dos mecanismos cuando ambos han de ejecutarse de forma conjunta por parte de los responsables que pretenden desarrollar o utilizar estos sistemas algorítmicos.

En primer lugar, la EIPD puede resultar sumamente oportuna como herramienta de evaluación o validación de un sistema algorítmico en relación con el cumplimiento no sólo de la normativa de protección de datos sino de otras legislaciones en la medida que, a través del tratamiento de datos que se analiza, puedan protegerse otros derechos y principios contemplados por el ordenamiento jurídico. De esta manera, antes de llevar a cabo el tratamiento, el responsable puede valorar los riesgos que puede generar ese sistema y decidir no implantar el mismo. Hemos de partir del carácter instrumental del derecho fundamental a la protección de datos de manera que estos dos mecanismos se pueden convertir en las herramientas que detecten dichos riesgos generales para los individuos. La EIPD puede también formar parte de otras evaluaciones de impacto o análisis de riesgos más globales centrados en los riesgos específicos sobre los que se pretende focalizar la atención[488]. Es decir, en estos casos, las implicaciones sobre protección de datos forman parte del análisis global de riesgos, sin embargo, la parte centrada en protección de datos continúa adquiriendo un papel central dada la relevancia de los datos tanto en el diseño como en el despliegue de los mismos.

En segundo lugar, la EIPD podría establecerse como requisito previo a la adquisición del modelo algorítmico. En esa fase previa a la adquisición se valoraría la viabilidad de dicho sistema en el entorno donde se pretende implantar y, en el caso de que tal sistema no sea lo suficientemente adecuado en ese entorno, no proceder a su adquisición. Es decir, la EIPD funcionaría como sistema de alerta sobre la viabilidad del

[488] El artículo 26 de la Propuesta de Reglamento de la Directiva de servicios digitales establece que las plataformas en línea de gran tamaño identificarán, analizarán y evaluarán al menos una vez al año cualquier riesgo sistémico significativo derivado del funcionamiento y uso de sus servicios. Esta evaluación de riesgos será específica de sus servicios e incluirá los siguientes riesgos sistémicos: a)la difusión de contenidos ilícitos a través de sus servicios; b) los efectos negativos para el ejercicio de los derechos fundamentales ;c) Manipulación intencionada de su servicio, incluso mediante el uso no auténtico o la explotación automatizada del servicio. 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.

mismo en la actividad donde se pretende que despliegue sus efectos. Esta visión de la EIPD podría implantarse por ejemplo en los procesos de contratación pública que lleven a cabo las Administraciones Públicas cuando pretende adquirir sistemas algorítmicos del mercado. De esta manera, la Administración Pública adquirente podría testar el sistema y en su caso optar por aquel que más se adecúe a la finalidad que se pretende con la inclusión del mismo. En este sentido, y dada la especial relación que existe entre los principios relativos al tratamiento de datos (Artículo 5 RGPD) y el grado de robustez de un sistema algorítmico, la propia EIPD no sólo estaría valorando si el tratamiento que se pretende llevar a cabo con ese algoritmo es acorde al cumplimiento normativo sino que además, también se estaría valorando la robustez del sistema automatizado. En estos supuestos también sería adecuado que la organización que ha desarrollado el sistema facilite la EIPD que realizó al potencial adquirente[489]. Ello resulta esencial dada la especial relación que existe entre la fase del diseño y despliegue de los sistemas de toma de decisiones automatizadas. Así, el adquirente del algoritmo puede hacerse una idea global de los principales riesgos que se detectaron en la fase de desarrollo y cómo estos se lograron mitigar, permitiendo a su vez a este último poder valorarlos durante la implantación del sistema en su contexto específico.

En tercer lugar, el enfoque del riesgo presente en la EIPD obliga a los responsables a prever todas aquellas medidas que mitiguen el riesgo hasta una fase tolerable. Este enfoque permite que tratamiento de datos que pueden quedar fuera del ámbito de aplicación del artículo 22 sigan contando con un nivel óptimo de protección y de garantías de cumplimiento de la normativa de protección de datos. Así, gracias a la EIPD, será habitual que tratamientos de datos donde no esté presente la plena automatización, gocen de las mismas exigencias de cumplimiento que en su caso estén previstas para un tratamiento de datos que encaje con la definición del artículo 22 del RGPD.

En cuarto lugar, la información obtenida a lo largo de todo el proceso que engloba la EIPD permite al responsable obtener un gran conocimiento del tratamiento en cuestión. Ello facilita la labor a la hora de diseñar los canales sobre los que

[489] Grupo del Artículo 29. *Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679***.** Adoptadas el 4 de abril de 2017. Revisadas por última vez y adoptadas el 4 de octubre de 2017, pág.9.

posteriormente los interesados podrán ejercer las facultades que se derivan del RGPD[490]. Así, por ejemplo, dado que durante la EIPD se ha de analizar el impacto que puede generar un sistema en la esfera de los particulares, la información que se derive de dicho análisis podrá ser útil para valorar por ejemplo si dicho tratamiento se encuadra o no dentro del artículo 22 del RGPD. El análisis de ese impacto también puede resultar práctico a la hora de facilitar la información que exigen los artículos 13,14 y 15 sobre las consecuciones jurídicas que se pueden derivar de ese tratamiento cuando se adopte la decisión automatizada. Así, y dado que previamente se valoraron los principales impactos que en su caso dicho sistema podía generar en determinados individuos, ahora, y con esa información recopilada, el derecho de información puede materializarse de forma efectiva. Visto así, toda la información obtenida y recopilada no sólo permite un mejor despliegue de los derechos, sino que además, facilita la demostración del cumplimiento de la normativa. Esta información durante la fase del diseño del sistema resulta sumamente importante ya que puede servir también como una especie de manual básico de los principales riesgos detectados y sobre los que hay que tomar especial atención una vez que el sistema comience a desplegar sus efectos[491]. A su vez, la EIPD no sólo puede ayudar al diseño de los canales que facilitan el ejercicio de los derechos de los interesados sino también para registrar y en su caso justificar las limitaciones que pueden establecerse a los mismos.

Fases de la EIPD	**Información recogida de la EIPD**	**Derechos y garantías derivadas de la normativa de protección de datos**
Descripción detallada del tratamiento	Papel del humano tras la adopción de la decisión automatizada. Formación, capacidad para alterar la decisión, etc.	Derecho a la intervención humana. Derecho a expresar el punto de vista. Artículo 22.
Evaluación de los riesgos. Impactos	Se valoran los daños materiales, morales o físicos que se pueden generar si se materializa la amenaza detectada	Derecho a informar sobre las consecuencias prevista del tratamiento de datos. Artículos 13,14 y 15.
Tecnología	Modelo algorítmico elegido.	Derecho a informar sobre la

[490] MARGOT E KAMINSKI,M; GIANCLAUDIO,M: "Algorithmic impact assessments under the GDPR: producing multi-layered explanations". *International Data Privacy Law*, ipaa 020, 06 December 2020 , pág.17. Disponible en: https://academic.oup.com/idpl/advance-article/doi/10.1093/idpl/ipaa020/6024963

[491] Es muy relevante que en la fase del diseño del sistema se registren todos los procesos llevados a cabo y en su caso sea elaborado un manual de funcionamiento del prototipo. European Parliamentary Research Service. *A governance framework for algorithmic accountability and transparency*, op.cit., p-63. También en: Red iberoamericana de protección de datos personales. *Orientaciones Específicas para el Cumplimiento de los Principios y Derechos que Rigen la Protección de los Datos Personales en los Proyectos de Inteligencia Artificial.*2019, pág.25.

utilizada Descripción detallada del tratamiento	Principales funcionalidades.	lógica del tratamiento. Artículos 13,14 y 15. Límites para reducir la información a facilitar
Ciclo de vida de los datos	Se hace mención a las variables que se utilizan para predecir la salida y además se indican aquellas que presentan un mayor peso.	Información sobre la lógica del tratamiento. Arts. 13,14 y15 Derecho a la explicación. Artículo 22 Derecho a la intervención humana. Artículo 22
Descripción detallada del tratamiento Valoración de los riesgos	Se especifica si las decisiones que adopta el sistema son o no relevantes conforme a la descripción de tratamiento establecida en el artículo 22 RGPD.	Derecho a no ser sometido a toma de decisiones plenamente automatizadas. Artículo 22.
Evaluación de la necesidad y proporcionalidad del tratamiento	Se justifica la incorporación del sistema automatizado. Se apuesta por modelos menos invasivos con la privacidad. Por ejemplo, más transparentes.	Información sobre la lógica del tratamiento. Artículos . 13, 14 y 15. Derecho a la explicación. Artículo 22

En quinto lugar, la configuración de la EIPD basada en el enfoque del riesgo que se ha descrito en las páginas anteriores permite que la misma pueda utilizarse como estándar sobre la cual se pueda monitorear el funcionamiento del sistema algorítmico[492]. Tal y como señala el grupo del artículo 29, la EIPD es un proceso continuo, sobre todo, cuando una operación de tratamiento es dinámica y está sujeta a cambios permanentes[493]. Esa actualización resulta obligatoria cuando se produzca un cambio del riesgo lo suficientemente relevante en la operación u operaciones a las que se somete el tratamiento de datos[494]. Trasladado a nuestra realidad, resulta esencial establecer y fijar los cauces a través de los cuales se pueden detectar esas alteraciones. Así, en muchas ocasiones, será el propio responsable el que, debido a los cambios intencionados que en

[492] L,JANSSEN,H:, "An approach for a fundamental rights impact assessment to automated decision-making", *International Data Privacy Law*, Volume 10, Issue 1, February 2020, pág.79. Disponible en: https://doi.org/10.1093/idpl/ipz028

[493] Grupo del Artículo 29. *Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679***.** Adoptadas el 4 de abril de 2017. Revisadas por última vez y adoptadas el 4 de octubre de 2017, pág.16.

[494] Artículo 35.11 del RGPD. Así lo ha aclarado el Consejo de Estado Francés en su resolución Nº 434376 de 6 de noviembre de 2019. Considerando 13. ECLI: FR: CECHR: 2019: 434376.20191106.
Disponible en: https://www.legifrance.gouv.fr/ceta/id/CETATEXT000039335911/
En este mismo sentido, la CNIL ha indicado que la no actualización de los nuevos cambios en el tratamiento llevarán consigo un incumplimiento del artículo 35 del RGPD. Véase la resolución de la Autoridad de protección de datos francesa (CNIL), Decisión MED-2020-015 de 15 julio de 2020. Disponible en: https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000042125452/

su caso aplique sobre el propio modelo algorítmico, acabe alterando el resigo inicial establecido. Algunos ejemplos pueden ser; cambios en las variables sobre las que se toman las decisiones, alteración de los pesos de esas variables, reentrenamiento y reajuste, actualización del modelo algorítmico, etc[495]. En estos supuestos la realización de una auditoría externa o la realización de testeos que valoren esos nuevos cambios introducidos pueden resultar medidas adecuadas. En otras situaciones, las alteraciones en el riesgo de los tratamientos no son achacables directamente a los cambios establecidos por el responsable sino que se derivan de factores externos. Por ejemplo: interacción del sistema algorítmico con el entorno, ataques intencionados para alterar el sistema, etc. En estos supuestos, la forma de detectar esos riesgos imprevistos exigirá de continuos testeos del modelo algorítmico.

En definitiva, la integración conjunta de la EIPD y el análisis de riesgos se convierten en un mecanismo excelente y básico para que los responsables valoren la adecuación del sistema algorítmico en el entorno en el que pretenden que el mismo despliegue sus efectos. La información recopilada durante todo el proceso no sólo facilitará la elaboración del mismo sino que resultará de suma importancia para el correcto despliegue de los derechos de los interesados que se deriva de la normativa de protección de datos. Además, y fruto de la continua adaptación a la que se someten muchos de estos sistema algorítmicos, esta herramienta puede servir como base sobre la que se deberán realizar las correspondientes actualizaciones de la misma para que mantengan en un nivel tolerable los riesgos que se vayan alterando.

III. LA PRIVACIDAD DESDE EL DISEÑO Y POR DEFECTO

La llamada privacidad desde el diseño y por defecto, también conocida como la *privacy by design*, fue desarrollada por Ann Cavoukian en los años noventa[496]. Para esta

[495] Existen algorítmicos que están continuamente actualizándose, sin embargo, los cambios más relevantes suelen realizarlos los propios técnicos que lo han desplegado cuando actualizan estos sistemas. Por ejemplo, el algoritmo de Facebook ha sufrido varios cambios muy significativos desde que se desplegó por vez primera. En: https://blog.hootsuite.com/facebook-algorithm/

[496] Los siete principios fundamentales del Privacy by Design son los siguientes: 1.- Protección Preventiva y Proactiva. 2.- Privacidad "por Defecto" 3.- Privacidad integrada en el Diseño 4.- Funcionalidad Plena "Win-Win" en lugar de "Suma cero" 5.- Protección durante todo el Ciclo Vital: "End to End" 6.- Visibilidad y Transparencia: "Trust but Verify". 7.- Respeto y Empoderamiento del Usuario. El Usuario en el Centro. "User-centric". En: CAVOUKIAN,A: "Privacy by Design The 7 Foundational Principles Implementation and Mapping of Fair Information Practices".Disponible en:

autora, el cumplimiento de la normativa de protección de datos nunca sería completo si la privacidad no se incorpora en el diseño de todo sistema, tecnología o práctica organizativas de las distintas organizaciones que pretenden llevar a cabo tratamientos de datos personales. Para ello, se hace necesario que dichas organizaciones incorporen el elemento de la privacidad desde las primeras fases de desarrollo de los sistemas y procesos tecnológicos que engloben tratamientos de datos personales. Desde el punto de vista normativo, la Directiva 95/46 ya preveía de forma tímida este principio al establecer que el responsable debía *adoptar medidas técnicas y de organización apropiadas, tanto en el momento de la concepción del sistema de tratamiento como en el de la aplicación de los tratamientos mismos*[497]. Fue sin embargo el GT29 el que en 2009, y a través de una consulta realizada por la Comisión[498], planteó la necesidad de apostar por nuevos principios que sin dejar de lado los ya implantados y reconocidos históricamente[499], reforzaran aún más la protección de los titulares de los datos. Así, principios como el de rendición de cuentas o la privacidad desde el diseño fueron ampliamente defendidos por parte del GT29 como nuevos principios que debían reconocerse en el futuro marco normativo que pretendiera regular la normativa sobre protección de datos fruto de los avances tecnológicos que habían surgido desde que se implantara en su momento la Directiva 95/46. Ha sido finalmente el RGPD en su artículo 25 el que ha reconocido el carácter vinculante del principio de privacidad desde

---

https://www.iab.org/wp-content/IAB-uploads/2011/03/fred_carter.pdf

[497] La referencia de la Directiva "al momento de la concepción" ha de entenderse como que, desde la fase del diseño se debe adoptar un enfoque que integre medidas respetuosas con la protección de datos. Véase el considerando 46 y en menor medida el artículo 17 de la directiva 95/46. Desde las primeras propuestas normativas de la Directiva ya se vislumbraba ese enfoque de la privacidad desde el diseño. Así, la primera propuesta de Directiva indicaba en su considerando 17 que, *en lo referente a los datos personales, la protección de la intimidad exige la adopción de medidas de seguridad apropiadas, tanto en la fase de concepción del tratamiento como con relación a la tecnología de este, a fin de impedir todo tratamiento no autorizado.* (la cursiva es nuestra). Comunicación de la Comisión sobre la protección de las personas en lo referente al tratamiento de datos personales en la Comunidad y a la seguridad de los sistemas de Información. 24 de septiembre de 1990. Disponible en:
https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:51990DC0314&from=EN.
Más explícita fue posteriormente la Comisión Europea al establecer en el considerando 21 *que la protección de los derechos y libertades de los interesados en lo que respecta a los tratamientos de datos personales exige la adopción de medidas técnicas apropiadas tanto en el momento de la concepción de las tecnologías como en el de la aplicación de los tratamientos, sobre todo con objeto de garantizar la seguridad e impedir, por tanto, todo tratamiento no autorizado*. (la cursiva es nuestra). En: Propuesta modificada de DIRECTIVA DEL CONSEJO relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. 15 de octubre de 1992. Disponible en:
https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:51992PC0422&from=EN

[498] Grupo del Artículo 29 y Grupo de Policía y Justicia. *The Future of Privacy Joint contribution to the Consultation of the European Commission on the legal framework for the fundamental right to protection of personal data*. Resolución adoptada el 1 de diciembre de 2009. Apartados 41 y ss , pags. 12 a 15.

[499] Es decir, principios como el principio de transparencia, minimización de datos, exactitud, limitación de la finalidad, etc.

el diseño y por defecto. Este principio obliga a los responsables del tratamiento a incorporar la protección de la privacidad a lo largo de todo el ciclo de vida del sistema, producto servicio o proceso[500], esto es, desde el momento inicial de su concepción hasta que finalmente se despliega o se implanta[501]. Traducido a los sistemas de toma de decisiones automatizadas basados en inteligencia artificial, el artículo 25 es una disposición que obliga a los responsables del tratamiento a implantar y prever la privacidad desde las fases iniciales a las que hacíamos referencia en el capítulo I de esta tesis, estas son; la planificación del proyecto, recopilación y obtención de datos, pre procesamiento, elección de variables, construcción y evaluación del modelo, etc. El principio de privacidad desde el diseño además ha servido de inspiración para que otras normas acaben imponiendo la obligación de incorporar la privacidad o el cumplimiento normativo desde las fases iniciales del desarrollo de productos que incorporen nuevas tecnologías. A modo de ejemplo, el Reglamento europeo contra los abusos sexuales de menores en línea autoriza a las plataformas a desplegar sistemas automatizados para controlar el contenido o actuaciones pedófilas que se vierten en dichas plataformas, esta norma obliga a esas organizaciones a implementar la privacidad desde las fases iniciales del desarrollo de esos modelos algorítmicos[502]. A su vez, la Carta de Derechos Digitales Española ha reconocido el principio de cumplimiento normativo desde el diseño, el cual, ha de incorporarse desde el inicio del proceso de cualquier desarrollo tecnológico[503].

De esta manera, el artículo 25 del RGPD resulta esencial dada la fuerte interrelación que existe entre las fases de desarrollo y despliegue de los sistemas de decisiones automatizadas. En este sentido, un importante número de exigencias que en materia de protección de datos se exigen durante la toma de decisiones automatizadas

[500] Agencia Española de Protección de Datos. *Guía de Privacidad desde el Diseño*. Octubre de 2019. Pág.6

[501] Ello exige por ejemplo que, desde el inicio, el equipo que pretende desarrollar el producto o sistema esté capacitado y formado en esta materia y que sea conocedor de las implicaciones que puede irradiar este sistema en el derecho fundamental a la protección de datos. Autoridad de Protección de Datos de Noruega. Datatilsynet. *Software development with Data Protection by Design and by Default.* Pág.4

[502] Artículo 3.1.b) del REGLAMENTO (UE) 2021/1232 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 14 de julio de 2021 por el que se establece una excepción temporal a determinadas disposiciones de la Directiva 2002/58/CE en lo que respecta al uso de tecnologías por proveedores de servicios de comunicaciones interpersonales independientes de la numeración para el tratamiento de datos personales y de otro tipo con fines de lucha contra los abusos sexuales de menores en línea.

[503] La propia Carta indica en sus consideraciones previas que la misma se ha inspirado en el principio de privacidad desde el diseño reconocido en la normativa de protección de datos. Carta de Derechos Digitales.

deben haber sido programadas o desarrolladas durante la fase de desarrollo de estos sistemas. Por tanto, si en esas fases iniciales no se han perfilado esos elementos normativos, muy probablemente este sistema presentará defectos de cumplimiento normativo en materia de protección de datos cuando se incorpore al entorno donde adoptará decisiones. A modo de ejemplo, los artículos 13 y 14 del RGPD obligan a los responsables del tratamiento a informar sobre la lógica implícita del tratamiento que hay detrás de la adopción de decisiones y perfiles automatizados.. Pues bien, si a la hora de desarrollar el modelo que adopta decisiones no se opta por modelos interpretables, este derecho no podrá llegar a ser ejercido eficazmente lo que llevará a dicho responsable que adopta las decisiones al muy probable incumplimiento de la normativa de protección de datos. Otro ejemplo sería la falta de justificación de las correlaciones extraídas por el modelo algorítmico[504]. Así, si en la fase del diseño no se indican las razones o mecanismos que justifican la elección de unas variables en defecto de otras[505], muy probablemente y una vez que el sistema comience a desplegar sus efectos, este podrá ser puesto en entredicho alegando que el mismo no ha cumplido con el principio de minimización de datos al no haberse justificado la pertinencia de las variables, esto es, los datos personales elegidos para la adopción de las decisiones que se están llevando a cabo.

FASE DE DISEÑO	**FASE DE DESPLEIGUE**
Se tienen que elegir modelos más o menos interpretables que informen sobre las principales características del sistema[506].	Artículos 13 y 14. Se ha de informar sobre la lógica del tratamiento.

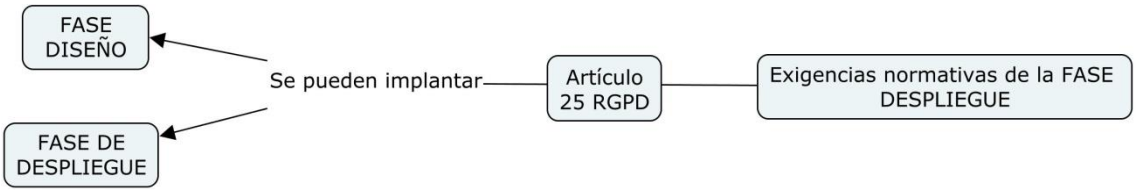
[504]Por ejemplo, en la fase del diseño debe quedar meridanamente diseñado el ámbito de irradiación donde actuará el sistema. Así, si un sistema pretende controlar el posible contenido terrorista que se emita en una determinada plataforma, desde esa fase inicial ya deben quedar claras las variables que posteriormente se explicarán a los potenciales usuarios de esas plataformas. Artículo 2.7 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.

[505] Dentro de las estrategias de privacidad desde el diseño propuestas por la AEPD se encuentra la de minimizar, la cual tiene como objetivo evitar el procesamiento de datos que no sean necesarios para las finalidades perseguidas en el tratamiento De esta manera, se ha de seleccionar *la muestra de individuos relevante y los atributos necesarios siguiendo una actitud conservadora al establecer el criterio de selección y realizar el tratamiento únicamente sobre los datos que respondan a dicho criterio (lista blanca)*. En: Agencia Española de Protección de Datos. *Guía de Privacidad desde el Diseño*. Octubre de 2019, pág.18.

[506]Desde el diseño se pueden desarrollar modelos algorítmicos que favorezcan la interpretabilidad de los sistemas. SELBST,A & BAROCAS,S: "The Intuitive Appeal of Explainable Machines", 87 Fordham L. Rev. 1085, 2018, pág. 1110. Disponible en: https://ir.lawnet.fordham.edu/flr/vol87/iss3/11/

Se han de justificar las correlaciones existentes entre las variables elegidas para desarrollar el modelo.	Artículo 5. Principio de minimización de datos en relación con la pertinencia de las variables escogidas.
Se tienen que elegir modelos altamente interpretables para poder explicar las decisiones que adoptan los algoritmos.	Considerando 71 y artículo 22. Explicación de las decisiones.

Visto así, el artículo 25 se convierte en un precepto *bisagra* que une las fases de diseño y despliegue de los sistemas automatizados ya que este último obliga al responsable a tener en cuenta desde la fase del diseño las implicaciones jurídicas que en materia de protección de datos posteriormente se exigirán en la fase de aplicación del mismo. Es decir, la privacidad desde el diseño no sólo logra que se cumpla la protección de datos desde las fases iniciales sino sobre todo, se instrumentan y perfilan las posibles implicaciones de privacidad que en un futuro el sistema algorítmico puede generar una vez que el mismo se implante y comienza a adoptar decisiones. Y es que, no podemos olvidar que en muchas situaciones el impacto legal directo de estos sistemas sobre todo estará presenten en la fase de despliegue y no en el desarrollo de los mismos[507].



El artículo 25 del RGPD establece que:

> *Teniendo en cuenta el estado de la técnica, el coste de la aplicación y la naturaleza, ámbito, contexto y fines del tratamiento, así como los riesgos de diversa probabilidad y gravedad que entraña el tratamiento para los derechos y libertades de las personas físicas, el responsable del tratamiento aplicará, tanto en el momento de determinar los medios de tratamiento como en el momento del propio tratamiento, medidas técnicas y organizativas apropiadas, como la seudonimización, concebidas para aplicar de forma efectiva los principios de protección de datos, como la minimización de datos, e integrar las garantías necesarias en el tratamiento, a fin de cumplir los requisitos del presente Reglamento y proteger los derechos de los interesados*.

[507] Comisión Europea. *Ethics and data protection* , noviembre de 2018, pág.17.

Lo primero que conviene indicar es que el RGPD obliga primeramente al responsable a, con carácter previo a la adopción de medidas que potencia la privacidad del diseño, realizar un análisis de riesgos sobre las implicaciones que las operaciones de los distintos tratamientos, en términos de probabilidad y gravedad, suponen para los derechos y libertades de las personas. Debiendo tener en cuenta el estado de la técnica, el coste de aplicación (económico), la naturaleza, ámbito, contexto y fines del tratamiento. Así, desde el inicio será recomendable que aquellos que pretendan diseñar sistema de decisiones automatizadas realicen dicho análisis de riesgos para valorar ya desde ese momento los derechos y libertades principalmente afectadas. A partir de ahí, ya podrán comenzar a planificar aquellas medidas que puedan ser relevantes para implementar la privacidad desde el origen. Nótese que el RGPD no fija ningún tipo de medida más allá de la seudonimización o la minimización de datos[508]. Esto permite al responsable del tratamiento elegir aquellas herramientas o medidas que considere más oportunas y que mejor se adapten al desarrollo del sistema que en su caso se pretenda implantar[509].

Por otro lado, estas medidas se pueden implantar tanto a la hora de determinar los medios del tratamiento como en el momento del propio tratamiento. Si traducimos esta redacción del mencionado precepto a las fases que comprende la creación y despliegue de los sistemas automatizados podemos llegar a interpretar que las medidas ligadas a la privacidad desde el diseño pueden integrarse no sólo durante la fase de desarrollo de los sistemas sino también una vez que dicho sistema se haya implantado. En nuestra opinión, y teniendo en cuenta la relevancia de la fase del diseño en el ciclo de vida de los sistemas de toma de decisiones automatizadas, las diferentes medidas de privacidad desde el diseño deberían focalizarse en esta primera fase. Todo ello sin perjuicio de que durante la fase de aplicación también puedan integrarse otras medidas que complementen las ya previstas inicialmente. Como ejemplo de esto último, podrían desplegarse medidas de privacidad desde el diseño en una fase posterior a la adquisición

[508] Al responsable no se le obliga a establecer medidas específicas, sino que él debe decidir cuál es la adecuada. Comité Europeo de Protección de Datos. *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*. Directrices adoptadas el 13 de Noviembre de 2019. apartado 15, pág.7.

[509] La AEPD hace referencia a toda una serie de patrones de diseño que pueden ser utilizados por los responsables en el cumplimiento de esta obligación. Agencia Española de Protección de Datos. *Guía de Privacidad desde el Diseño*. Octubre de 2019, págs. 34 y ss. A su vez, el CEPD también establece todo un elenco de medidas que potencian la privacidad desde el diseño en relación con los principio de tratamiento de datos personales. Comité Europeo de Protección de Datos. *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*. Directrices adoptadas el 13 de Noviembre de 2019.

del sistema, pero anterior a la toma de decisiones automatizadas. No obstante, como hemos dicho anteriormente, dado que la fase de despliegue depende en gran medida de cómo se diseñe el algoritmo previamente, las medidas de privacidad conviene implantarlas en esas fases iniciales.

Por último y no menos importante, de lo indicado por este precepto se vislumbra que estas medias resultan de obligado cumplimiento únicamente para los responsables del tratamiento. La vinculación de la privacidad desde el diseño no afecta ni a los diseñadores ni a los proveedores de los sistemas cuando los mismos no traten datos personales durante el desarrollo de los mismos. En este sentido, resulta habitual en este ámbito que la organización que desarrolló el sistema automatizado sea distinta a la que finalmente lo utiliza para la toma de decisiones. Ello puede llevar a que el artículo 25 quede en parte *descafeinado* al dejar fuera del ámbito de la norma a aquellas organizaciones sobre las que especialmente se les debería dirigir la implantación de medidas de privacidad desde el diseño[510]. Para cubrir este vacío, proponemos varias alternativas dependiendo de la situación a la que se enfrente aquel responsable del tratamiento que implante un sistema de decisiones automatizadas y pretenda cumplir con las exigencias en materia de privacidad desde el diseño. Se analizan distintos supuestos:

A) La organización usuaria del sistema de decisiones automatizadas es la misma organización que desarrolló el sistema. En estos supuestos, resulta lógico pensar que independientemente de que en la fase de diseño se traten o no datos personales y dada la dependencia que existe entre la fase de desarrollo y despliegue, el algoritmo debería diseñarse teniendo en cuenta las posibles incidencias que en materia de protección de datos pueden derivarse una vez que el sistema comience a adoptar decisiones. Obviar la normativa potencial de protección de datos en la fase del diseño por aquel que posteriormente usará el sistema supone en nuestra opinión un incumplimiento claro del artículo 25 del RGPD.

[510] El Grupo del Artículo 29 propuso que el principio de privacidad desde el diseño debería ser vinculante para los diseñadores y productores de las tecnologías. Sin embargo, esta propuesta no fructificó en la redacción del actual RGPD. Grupo del Artículo 29 y Grupo de Policía y Justicia. *The Future of Privacy Joint contribution to the Consultation of the European Commission on the legal framework for the fundamental right to protection of personal data*. Resolución adoptada el 1 de diciembre de 2009. Apartado 46, pág.13.

B) La organización que utiliza el sistema es distinta a la que lo desarrolló, caben dos situaciones:

Si durante la fase de diseño del sistema/algoritmo se tratan datos personales. Hay que entender que el productor o desarrollador del mismo, al ser responsable del tratamiento durante la fase del diseño debe cumplir con las exigencias en materia de protección de datos y por tanto, derivado de lo establecido en el artículo 25, también debe prever e implantar medidas de privacidad desde el diseño que potencialmente puedan posteriormente estar presentes una vez que el sistema adopte decisiones.

Si en cambio, durante la fase del diseño del sistema/algoritmo no se tratan datos personales. El diseñador del sistema no es considerado responsable ni encargado de tratamiento, por tanto, no tiene la obligación de aplicar medidas de privacidad desde el diseño. De esta manera, la organización que adquiere el sistema únicamente podría incitar o alentar a los diseñadores y productores de los modelos algorítmicos a que en su caso implanten dichos elementos que favorezcan la privacidad en los algoritmos que están diseñando. Para ello, la organización que compra el algoritmo puede exigir como condición previa a la adquisición que estos sistemas cumplan con las exigencias previstas por el RGPD. En este sentido, la AEPD considera que es deber del responsable *ceñirse a la selección de productos y de encargados capaces de garantizar el cumplimiento de los requisitos del RGPD, y en particular, la obligación de garantizar la protección de datos desde el diseño y por defecto*[511]. Para reforzar esta idea, el artículo 28.1 del RGPD obliga al responsable a elegir a aquellos encargados que ofrezcan garantías suficientes que aseguren que los tratamientos que se llevarán a cabo sean conforme a las exigencias normativas previstas por el RGPD[512]. Es por ello que el responsable deberá ser especialmente diligente a la hora de optar por uno u otro encargado ya que las posibles responsabilidades posteriores que se deriven del incumplimiento de la normativa de

[511]Agencia Española de Protección de Datos. *Guía de Privacidad desde el Diseño*. Octubre de 2019,págs 11, 12 y 33.

[512] Tal y como señala el CEPD, el responsable, a la hora de evaluar si el encargado proporciona las suficientes garantías ha de tener en cuenta entre otros elementos: i) los conocimientos especializados del encargado (por ejemplo, los conocimientos técnicos en relación con las medidas de seguridad y las violaciones de datos). ii) La fiabilidad del encargado del tratamiento. iii) Los recursos del encargado del tratamiento. iv) La reputación del encargado del tratamiento en el mercado. En: Comité Europeo de Protección de Datos. *Guidelines 07/2020 on the concepts of controller and processor in the GDPR* .Adoptadas el 2 de septiembre de 2020, apartado 95, pág.30. Véase también el Considerando 81 del RGPD.

protección de datos no sólo recaerán sobre el encargado sino también respecto del responsable en virtud del principio legal de *culpa in eligendo*[513]. Traducido a nuestro contexto, el responsable, en aplicación de dicho deber de diligencia estará obligado a elegir a aquellas organizaciones que ofrezcan sistemas algorítmicos respetuosos con la normativa de protección de datos. Para ello, el responsable podría entre otras alternativas: i) realizar auditorías o evaluaciones previas del sistema algorítmico que se ofrece o pretende implementar, ii) comprobar que la organización que ha diseñado el sistema está sometida a códigos de conducta, iii) chequear que los modelos algorítmicos que se ofrecen han superado algún proceso de certificación en materia de protección de datos[514].

Junto a este deber de diligencia, existen otra serie de normas que directa o indirectamente promueven el despliegue de la privacidad desde las fases iniciales del diseño de los sistemas algorítmicos. Así, en el sector privado, la Directiva de contratos de suministros digitales considera como falta de conformidad que el producto que se facilita al usuario no contemple por ejemplo la privacidad desde el diseño[515]. A su vez, en el ámbito público, la Ley de contratos del sector público establece que dentro del contenido mínimo de los pliegos de condiciones del contrato se incluya la obligación del futuro contratista del cumplimiento de la normativa en materia de protección de datos[516]. Entendemos por tanto que aquellas Administraciones Públicas que pretendan la adquisición de algoritmos deban exigir a los diseñadores y productores de los mismos que tales sistemas se adecúen a la normativa de protección de datos. Es más, el propio RGPD así lo contempla también

---

[513] Entre otros: RODRÍGUEZ AYUSO, J,F: *Garantía administrativa de los derechos del interesado en materia de protección de datos personales*. Ed. Bosch, Barcelona, 2021, pág.128. También véase: ÁLVAREZ RIGAUDIAS, C: "El nuevo reglamento de desarrollo de la LOPD". *Actualidad Jurídica Uría y Menéndez*. Nº 21. 2008, pág.27. Texto disponible en: https://www.uria.com/es/revista/23

[514] Se trata de medidas de responsabilidad activa que se analizan en esta tesis. Véase el Capítulo III, apartado VII (códigos de conducta), Capítulo III apartado VIII (Mecanismos de certificación), Capítulo III, apartado X (Monitoreo y evaluación de los sistemas automatizados).

[515]Considerando 48 y artículos 6.7 y 8 de la DIRECTIVA (UE) 2019/770 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 20 de mayo de 2019 relativa a determinados aspectos de los contratos de suministro de contenidos y servicios digitales.

[516] Artículos 35.d), 122.2 y D.A 25 de la Ley 9/2017, de 8 de noviembre, de Contratos del Sector Público, por la que se transponen al ordenamiento jurídico español las Directivas del Parlamento Europeo y del Consejo 2014/23/UE y 2014/24/UE, de 26 de febrero de 2014. Sobre los cambios introducidos en esta norma sobre protección de datos se recomienda la lectura de los siguientes trabajados. SOBRINO GARCÍA,I: "Desafíos y limitaciones en la contratación pública: el impacto de la protección de datos tras los últimos cambios legislativos". *Revista General de Derecho Administrativo*, número 56, enero 2021. También véase: MARTÍNEZ MARTÍNEZ, R; "El laberinto de la contratación pública en protección de datos". *Diario La Ley*, sección Ciberderecho, núm. 35, 2019. A su vez, PIÑAR MAÑAS,J,L: "Contratación pública y protección de datos". En: ORTEGA BURGOS,E; PASTOR RUIZ,F (dir): *Derecho administrativo 2021*. Ed. Tirant lo Blanch. Valencia, 2021, págs 413 a 419.

en su considerando 78 *in fine* respecto de la contratación pública[517]. En este mismo sentido se ha pronunciado el Gobierno del Reino Unido recomendando a las autoridades públicas incorporar en los pliegos contractuales las exigencias a los contratistas para que sus modelos algorítmicos sean más interpretables y explicables[518]. Por otro lado, no podemos pasar por alto la PRAI, la cual, obliga a los desarrolladores/proveedores de algoritmos catalogados de alto riesgo a implementar en estos sistemas toda una serie de garantías y técnicas que en gran parte coinciden con muchas de las medidas que se derivan del principio de privacidad desde el diseño[519]. De esta manera, el vacío legal que muestra el RGPD en la fase de diseño de los sistemas cuando no se tratan datos personales y que hemos tratado de solventar a través de una interpretación amplia del principio de privacidad desde el diseño resultará en gran parte cubierto por las obligaciones que se imponen a los proveedores durante el desarrollo de los sistemas de inteligencia artificial de alto riesgo. A modo de ejemplo, el artículo 13 de la PRAI obliga a los proveedores de sistemas de alto riesgo a desarrollar sistemas que garanticen que durante todo el funcionamiento de los mismos resulten lo suficientemente transparentes para que los usuarios puedan interpretar los resultados. Esto implica que desde la fase de diseño y aunque no se

[517] Este considerando establece que: *La protección de los derechos y libertades de las personas físicas con respecto al tratamiento de datos personales exige la adopción de medidas técnicas y organizativas apropiadas con el fin de garantizar el cumplimiento de los requisitos del presente Reglamento. A fin de poder demostrar la conformidad con el presente Reglamento, el responsable del tratamiento debe adoptar políticas internas y aplicar medidas que cumplan en particular los principios de protección de datos desde el diseño y por defecto. Dichas medidas podrían consistir, entre otras, en reducir al máximo el tratamiento de datos personales, seudonimizar lo antes posible los datos personales, dar transparencia a las funciones y el tratamiento de datos personales, permitiendo a los interesados supervisar el tratamiento de datos y al responsable del tratamiento crear y mejorar elementos de seguridad. Al desarrollar, diseñar, seleccionar y usar aplicaciones, servicios y productos que están basados en el tratamiento de datos personales o que tratan datos personales para cumplir su función, ha de alentarse a los productores de los productos, servicios y aplicaciones a que tengan en cuenta el derecho a la protección de datos cuando desarrollan y diseñen estos productos, servicios y aplicaciones, y que se aseguren, con la debida atención al estado de la técnica, de que los responsables y los encargados del tratamiento están en condiciones de cumplir sus obligaciones en materia de protección de datos.* ***Los principios de la protección de datos desde el diseño y por defecto también deben tenerse en cuenta en el contexto de los contratos públicos.*** (la cursiva y la negrita son nuestras)

[518]Gobierno del Reino Unido. *Guidelines for AI procurement*. junio de 2020. Véase el punto octavo de las principales consideraciones a las que se hace referencia en esta guía. Disponible en: https://www.gov.uk/government/publications/guidelines-for-ai-procurement/guidelines-for-ai-procurement

[519] A la hora de desarrollar los sistemas de inteligencia artificial de alto riesgo, los diseñadores tienen que implantar una política adecuada de gobernanza de datos, elaborar la documentación técnica del sistema, llevar a cabo un registro de los eventos (logs), desarrollar sistemas transparentes y precisos, etc. Véase artículos 8 y ss. Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

traten datos personales, los proveedores deberán implantar técnicas que favorezcan la transparencia del sistema. Estas herramientas que facilitan la transparencia posteriormente permitirán el despliegue de algunas de las facultades básicas exigibles por la normativa de protección de datos como la información sobre la lógica del algoritmo (Artículos 13, 14 y 15 del RGPD) o el derecho de explicación de las decisiones adoptadas por el mismo (Considerando 71 y artículo 22 del RGPD). Es por ello que, pese a que la PRAI no obliga a los proveedores a implantar medidas que favorezca la privacidad desde el diseño, dado que muchas de estas exigencias en parte coinciden con las propias que se derivan de la normativa de protección de datos, la protección que se brindará en estos supuestos resultará muy relevante a pesar de que el RGPD no se aplique. Finalmente, el Reglamento Europeo que fija el *Programa Europa Digital* para los próximos años establece expresamente que las soluciones basadas en la IA y los datos que se faciliten amparados en este programa deberán respetar el principio de privacidad desde el diseño en particular y la normativa sobre protección de datos en general[520].

PRAI **Obligaciones a los proveedores de sistemas de alto riesgo**	**RGPD** **Privacidad desde el diseño que afecta a:**
Prácticas adecuadas de gobernanza y gestión de datos.	Principio de minimización de datos. Principio de exactitud.
Elaboración de documentación técnica sobre el sistema.	Principio de transparencia. Derechos de información, lógica del tratamiento, explicación de la decisión.
Registro automático de eventos ("logs").	Principio de transparencia. Explicación de la decisión. Responsabilidad activa.
Transparencia y suministro de información a los usuarios.	Principio de transparencia Derechos de información, lógica del tratamiento, explicación de la decisión.
Supervisión humana.	*Derechos* de información, derecho a obtener intervención humana por parte del responsable.
Precisión, solidez y ciberseguridad de los sistemas.	Principio de exactitud, seguridad en el tratamiento de datos, responsabilidad activa.

[520] Artículo 5.1 del REGLAMENTO (UE) 2021/694 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 29 de abril de 2021 por el que se establece el Programa Europa Digital y por el que se deroga la Decisión (UE) 2015/2240. Con este programa la Unión Europea pretende alcanzar toda una serie de objetivos, entre esos objetivos está el fomento de la inteligencia artificial.

En definitiva, independientemente de que la organización que desarrolla un algoritmo implante o no medidas que potencien la privacidad desde las primeras fases, lo cierto es que, la dinámica general llevará irremediablemente a que estas acaben implantándola o al menos a tener en cuenta determinadas normas relacionadas con la protección de datos. Así, en primer lugar, existe un deber del responsable por el cual ha de ser debidamente diligente a la hora de elegir a los encargados que hayan implantado el factor privacidad en sus modelos algoritmos. A su vez, en segundo lugar, existen otra serie de normas tanto en el sector público como en el privado que sancionan a aquellos responsables que no incorporan el factor privacidad en los productos algorítmicos que pretenden utilizar. En tercer lugar, la PRAI, obliga a los desarrolladores de los sistemas automatizados a desplegar toda una serie de medias, las cuales, se corresponden en gran parte con las exigencias de la normativa de protección de datos. De esta manera, incorporar el factor de privacidad en las fases de desarrollo supondrá por un lado una ventaja competitiva a estas organizaciones y por otro, reducirá los riesgos de incumplimiento de la normativa de protección de datos una vez que el sistema comienza a adoptar decisiones. A su vez, como iremos viendo en páginas posteriores, muchas de las exigencias técnicas de robustez que aplican los desarrolladores de sistemas algorítmicos coinciden en gran parte con las propias exigencias que se derivan del cumplimiento de los principios en materia de protección de datos. Así, por ejemplo, resulta habitual que los diseñadores de los sistemas de toma de decisiones automatizadas lleven a cabo toda una serie de técnicas para justificar las correlaciones existentes entre las variables elegidas. Pues bien, esa misma exigencia técnica de robustez que se deriva del propio interés del desarrollador de presentar un producto solvente también coincide con el principio de minimización de datos que exige que exista una correlación justificada o adecuada entre las variables elegidas y el tratamiento que se pretende llevar a cabo. Es decir, el cumplimiento de la normativa de protección de datos en muchos casos convergerá con las exigencias internas mínimas que un sistema debe contener para considerase que la tecnología analizada presenta una grado de madurez suficiente para desplegar efectos en el mercado. La necesidad de acompasar estas medidas resulta fundamental.

Supuesto	Fase de diseño	Fase de despliegue	Aplicación artículo 25 RGPD	
			Diseño	Despliegue
La organización que diseña y despliega el algoritmo es la misma.	Organización A trata o no datos personales.	Organización A implanta el algoritmo	SI	SI
La organización que desarrolla y despliega el algoritmo **no** es la misma.	Organización A trata datos personales.	Organización B implanta el algoritmo.	SI	SI
La organización que desarrolla y despliega el algoritmo **no** es la misma.	Organización A **no** trata datos personales.	Organización B implanta el algoritmo.	**NO**. El vacío se cubre con: -Deber de diligencia del responsable a la hora de elegir al encargado. -Pliego contractual. -PRAI. -Programa Europa Digital -Otras normas.	SI

En los siguientes capítulos, conforme se vayan analizando los principios y garantías que corresponden a los particulares, se aludirá a todo tipo de medidas que se han de incluir durante la fase de diseño producto de las exigencias que se derivan de la normativa de protección de datos.

IV. EL REGISTRO DE ACTIVIDADES DE TRATAMIENTO

Una de las obligaciones de cumplimiento derivadas de la normativa de protección de datos más criticadas en el pasado por parte tanto de la doctrina como de los operadores jurídicos era referida a la obligación que tenían los responsables de tratamientos de notificar a la AEPD la creación de ficheros de carácter personal[521]. Con la entrada en vigor del RGPD la notificación previa de las actividades de tratamiento de datos ha sido sustituida por el registro de las mismas. Se reduce así el elemento burocrático que caracterizaban estas notificaciones previas al tratamiento por un sistema de registro a través del cual, los responsables y encargados han de documentar los principales elementos que componen los tratamientos que llevan a cabo.

[521] Estas obligaciones se reconocían en el Artículo 26 de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal y artículo 55 del Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal.

Tal y como establece el artículo 30 del RGPD, no todas las organizaciones están obligadas a llevar un registro de actividades. Concretamente, la norma prevé dos criterios para exigir dicha medida responsabilidad activa. Por un lado, un criterio cuantitativo por el cual, todas las empresas superiores a 250 trabajadores tienen la obligación de llevar este registro. A su vez, y basado en un criterio cualitativo, también estarán obligadas a llevar ese registro de tratamiento aquellas organizaciones que sin contar en su plantilla con 250 lleven a cabo; i) tratamientos no ocasionales, ii) cuando los tratamientos generen un riesgo para los derechos o, iii) cuando se traten datos especialmente sensibles o de categoría penal[522]. En este sentido, prácticamente todas las organizaciones que pretendan desarrollar o desplegar sistemas de toma de decisiones automatizadas deberán implementar un registro de las actividades de tratamiento que se realicen. Estas obligaciones recaen tanto en el responsable como en el encargado del tratamiento.

Por lo que se refiere al contenido de los registros, estos deberán incluir las principales características del tratamiento y los datos de contacto del responsable o en su caso del encargado del tratamiento[523]. Es por ello que en este registro deba aparecer al menos la indicación de que se llevan a cabo la elaboración de perfiles[524], la adopción de decisiones automatizadas o en su caso los principales datos inferidos que proyectan obtener tras haber realizado el algoritmo el análisis de los mismos.

Como regla general, el registro de actividades de tratamiento no es público, este ha de ser conservado por el responsable o encargado de tratamiento. Queda por tanto el mismo a disposición de las autoridades de control cuando así lo soliciten. Ahora bien, tratándose de Administraciones Públicas, el artículo 31.2 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales, en adelante LOPD de 2018, sí que impone a estas últimas la obligación de publicar a través de medios electrónicos un inventario de las actividades de tratamiento que llevan a cabo donde se incluya la información a la que hace referencia el artículo 30 del

[522] Tal y como indica la CNIL francesa, en la práctica, la obligación de llevar un registro de actividades será prácticamente exigible a la mayoría de organizaciones que traten datos personales teniendo en cuenta los criterios que prevé el artículo 30 del RGPD. Visto en: https://www.cnil.fr/en/record-processing-activities

[523] Artículo 30 apartados primero y segundo del RGPD.

[524] El artículo 10 de la Ley francesa n° 78-17 del 6 de enero de 1978 relativa a la informática, archivos y libertades establece la obligación de registrar expresamente el tratamiento de datos con fines de elaboración de perfiles por parte de las organizaciones. Visto en: https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000886460/

RGPD[525]. Dicha obligación de transparencia también se prevé para las autoridades públicas en el ámbito de los tratamientos de datos personales con fines de prevención e investigación penal, en estos casos además, se ha de indicar expresamente si dichas autoridades llevan a cabo la elaboración de perfiles[526]. Estas obligaciones en el ámbito público se convierte en una herramienta mínima de transparencia que debe facilitar la visualización de la actividad algorítmica opaca a la que hasta ahora nos han acostumbrado los poderes públicos cuando hacen usos de sistemas para la toma de decisiones parcial o totalmente automatizadas. Se complementa así en parte las deficiencias relacionadas con la falta de publicidad que caracterizan a las evaluaciones de impacto ya que a través del registro se puede al menos conocer qué tratamientos se están llevando a cabo. De esta manera, una vez que se conoce que una Administración está llevando a cabo este tipo de tratamientos, las posibilidades de que los particulares o la sociedad civil puedan acudir a otros cauces para obtener más información sobre el uso de estos sistemas se amplían. Estas obligaciones se convierten en el primer peldaño que puede llevar a pasos ulteriores a través de otras herramientas como el ejercicio del derecho de acceso a la información derivado de la normativa de transparencia o la realización de denuncias o comunicaciones a la AEPD[527].

Información a incorporar en el registro de actividades de tratamiento (Art.30 RGPD)	**Información relevante a efectos del usos de sistemas de toma de decisiones automatizadas**
Datos de contacto del responsable Fines del tratamiento Categorías de interesados, destinatarios y de datos personales Existencia de transferencias internacionales Cuando sea posible, las medidas de seguridad establecidas y plazos de supresión de los	Se llevan a cabo elaboración de perfiles Se adoptan decisiones parcial o totalmente automatizadas Fines de la elaboración de perfiles o la toma de decisiones Datos inferidos proyectados

[525] Véase también el artículo 6 bis de la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno.

[526] Artículo 32.1.e) de 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.

[527] Por ejemplo, se podría ejercer el derecho de acceso solicitando información del sistema algorítmico en cuestión a través del artículo 12 de la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno. Por su parte, el artículo 57.1.f) RGPD fija como función de las autoridades de control de protección de datos "tratar las reclamaciones presentadas por un interesado o por un organismo, organización o asociación de conformidad e investigar, en la medida oportuna, el motivo de la reclamación".

datos. Base de legitimación del tratamiento[528]	Medidas específicas de seguridad aplicables al sistema algorítmico

V. LAS MEDIDAS DE SEGURIDAD DEL TRATAMIENTO. NOTIFICACIÓN Y COMUNICACIÓN DE LAS VIOLACIONES DE SEGURIDAD

La información constituye un activo valioso de todas las organizaciones. Es por ello que resulte esencial desarrollar e implementar una política adecuada de seguridad frente a posibles brechas. Esa información puede contener o no datos personales. Es por ello que, el RGPD en su artículo 5.1.f) reconozca como principio del tratamiento de datos personales la seguridad. Estos es, los datos han de ser tratados de tal manera que se garantice una seguridad adecuada. Consagrado este principio, este es posteriormente materializado en los artículos 32 y ss. del RGPD. Así, estos preceptos establecen una serie de obligaciones. Por un lado, tanto al responsable como al encargado del tratamiento han de implementar todo tipo de medidas técnicas y organizativas que resulten apropiadas para garantizar un nivel de seguridad adecuado al riesgo. Por otro lado, se establece un mecanismo para notificar y en su caso comunicar las brechas de seguridad a las autoridades de control y titulares de los datos respectivamente cuando las mismas generen un cierto riesgo para los derechos y libertades de los interesados.

De esta manera y siguiendo con el artículo 32.1 del RGPD , al responsable una vez más le corresponde acudir a la herramienta de análisis de riesgos para valorar las principales amenazas que pueden afectar a la seguridad de los datos personales. Este análisis ya se habrá realizado a la hora de gestionar el riesgo de los tratamientos de datos que se pretenden llevar a cabo[529], sin embargo, será necesario realizar un análisis específico de los riesgos derivados de la seguridad[530]. En este sentido, incumbe al responsable decidir la metodología adecuada para llevar a cabo este análisis de los riesgos, pudiendo optar entre otras opciones por: i) las guías publicadas por las autoridades de control públicas[531], ii) la adhesión a un código de conducta aprobado por

[528] La publicidad del registro referida a la base de legitimación del tratamiento únicamente es exigible a las Administraciones Públicas. Artículo 31.2 Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.
[529] Sobre el análisis de riesgos véase el Capítulo III, apartado I de esta tesis.
[530] Agencia Española de Protección de Datos. *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales.* Junio 2021, págs 40 y ss. y 116 y ss.
[531] Agencia Española de Protección de Datos. *Guía Práctica de análisis de riesgos en los tratamientos de datos personales sujetos al RGPD*. 2018. Aunque ya más antigua, véase también la *Guía de seguridad de datos* de la AEPD de 2010.

las autoridades de control o, iii) la obtención de algún tipo de certificación del sector al que pertenezca la organización. Para el caso de las Administraciones Públicas o los contratistas de estas últimas, las medidas a establecer por parte de los responsables o encargados del tratamiento han de adecuarse a las desarrolladas por el Esquema Nacional de Seguridad[532].

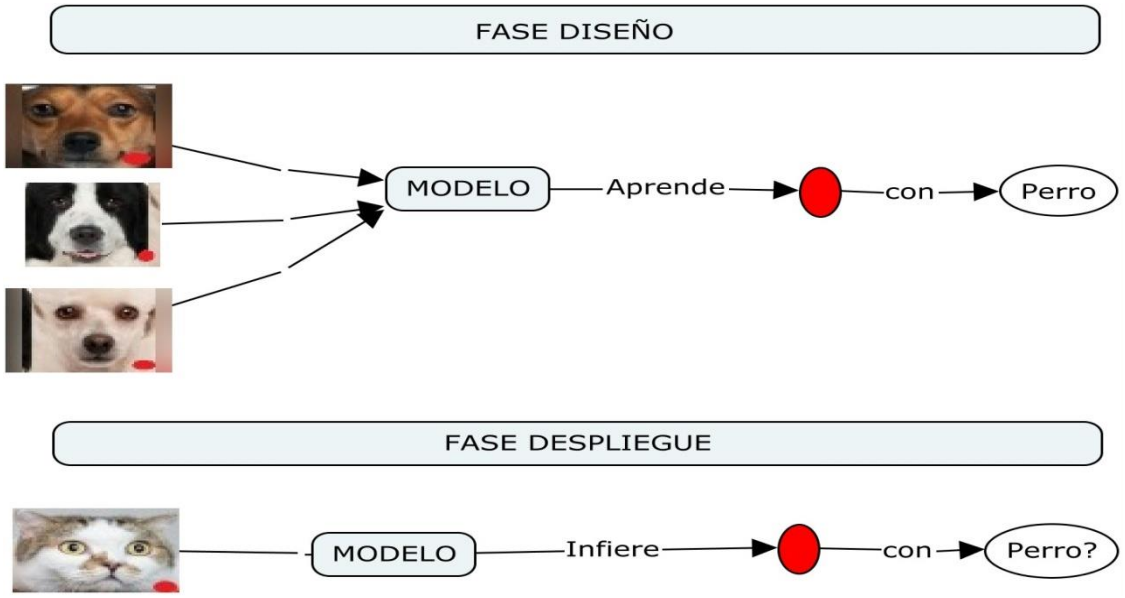
Elaboración propia

1. Amenazas que afectan a la seguridad de los sistemas de toma de decisiones

Es turno de analizar algunas de las amenazas principales que pueden afectar a la seguridad de los datos personales en el desarrollo y despliegue de los sistemas de toma de decisiones automatizadas. Como ahora se verá, un grado deficiente de medidas de seguridad afectará a la robustez del sistema y por tanto, a las decisiones que este pueda llegar a tomar con las consecuencias negativas que ello comporta para los particulares sometidas a las mismas. Así, los ataques en estos entornos pueden ir dirigidos esencialmente contra los datos o contra los modelos algorítmicos[533]. Para mantener una coherencia con este trabajo, a la hora de presentar los posibles ataques a sistemas de toma de decisiones diferenciaremos entre ataques durante la fase de desarrollo y despliegue.

[532] Tal y como establece la Disposición Adicional Primera de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.
[533] Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. Directrices éticas para una IA fiable.2019, pág.20.

Por lo que se refiere a la *fase del diseño*, las amenazas pueden venir directamente de accesos no autorizados o filtraciones a los *data lakes* o bases de datos sobre las cuales se pretenden desarrollar posteriormente los modelos. Esas bases de datos pueden contener información muy diversa y de forma más o menos estructurada sobre multitud de individuos. También se ha destacado la posibilidad de que los datos de entrenamiento sean manipulados antes de su procesamiento a través de las conocidas como *técnicas de envenenamiento de los datos*[534]. Así, a través de este proceso se corrompen el conjunto de datos de entrenamiento con el objetivo de que el modelo genere posteriormente decisiones erróneas, es decir, durante la fase del diseño los datos que se introducen son intencionadamente mal etiquetados[535].



Podemos ver un ejemplo de *envenenamiento* de datos en la imagen de arriba. Al entrenar un modelo que pretenda distinguir entre gatos y perros este tratará de conectar los valores de píxeles de cada una de las imágenes con las etiquetas asignadas. De esta manera, el modelo tratará de ajustar sus parámetros a los datos sin que dicho ajuste responda necesariamente a una lógica. Así, de forma intencionada se pueden etiquetar todos los datos de perros con una marca roja debajo de la foto. El modelo puede detectar esa correlación y considerar que todas las imágenes que contienen un punto rojo son consideradas perros. Al ingresar una nueva imagen, en este caso de gato, con ese punto rojo, el sistema lo considera un perro, ya que ha aprendido que dicha marca coincide

[534] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.42.

[535] YINGZHE,H ; GUOZH, M; KAI,CH;, XINGBO,H; JINWEN,H:" Towards Security Threats of Deep Learning Systems: A Survey". *IEEE Transactions on software engineering*, Octubre 2020, pág 10 y ss. Disponible en: https://arxiv.org/abs/1911.12562

con la etiqueta de perro[536]. Aunque este ejemplo pueda resultar trivial, en otras ocasiones estos problemas pueden ser relativamente graves. En este sentido, se demostró que un algoritmo identificaba riesgo de melanoma en todas las imágenes que contenían marcas de las reglas que se utilizaban para medir las lesiones cutáneas que aparecen en la piel[537]. Es decir, el sistema aprendió que siempre que apareciera la regla, la instancia había que clasificarla como melanoma. Pese a que en ambos ejemplos los ataques no fueron intencionados, la posibilidad de que un tercer atacante interfiera en la fase de diseño se muestra como un riesgo que se ha de contemplar y evaluar.

Por otro lado, en relación con la *fase de despliegue*, encontramos esencialmente cuatro tipos de ataques.

En primer lugar, todos aquellos ataques que buscan *acceder a los perfiles o inferencias* que estén o se hayan generado durante la fase de toma de decisiones. Estas violaciones de seguridad pueden derivarse del acceso por parte del personal de la organización que no está autorizado al mismo o puede venir de la mano de ataques externos, es decir, de terceros interesados en esos perfiles e inferencias generadas. El riesgo puede aumentar por ejemplo si dicha toma de decisiones automatizadas se realiza utilizando un modelo bajo la prestación de servicios en la nube.

En segundo lugar, una segunda familia está formada por los llamados *ataques contradictorios*[538]. Es decir, un ejemplo contradictorio es una instancia o dato que se introduce en el modelo que presenta pequeñas perturbaciones intencionales de características que hacen que el sistema realice una predicción falsa al no conseguir identificar esa pequeña diferencia existente en el dato introducido[539]. Los ejemplos contradictorios hacen que los modelos de aprendizaje automático sean vulnerables a los ataques. A modo de ejemplo, en 2017 se consiguió probar la facilidad que existe a la

[536] Se trata de un mero ejemplo pero el mismo refleja las posibles ataques o fallas que puede sufrir un sistema de toma de decisiones automatizados.
Más información en: https://thenextweb.com/news/what-is-machine-learning-data-poisoning-syndication
[537] AKHILA, N; BRETT, K; KAVITA,S; ROBERTO,N; JUSTIN, K: "Automated Classification of Skin Lesions: From Pixels to Practice". *Journal of Investigative Dermatology*, 138, 2018. Disponible en: https://www.jidonline.org/article/S0022-202X(18)32293-0/fulltext
[538] También conocidos como ataques por adversial machine learning. En: OCDE, "Artificial Intelligence in Society", 2019, pág.99 .Disponible en: *https://doi.org/10.1787/eedfee77-en*
[539] MOLNAR,CH*: Interpretable Machine Learning. A Guide for Making Black Box Models Explainable*. Libro disponible on line. ISBN 9780244768522. Véase el apartado 6.2. Libor disponible en: https://christophm.github.io/interpretable-ml-book/

hora de manipular la visión artificial de un coche autónomo[540]. Concretamente, este coche dejó de identificar la señal viaria de STOP al no reconocerla por el mero hecho de añadir a dicha señal pegatinas de distintos colores[541].

En tercer lugar, y muy relacionado con los ataques contradictorios, también es posible que los *atacantes traten de alterar la funcionalidad del sistema* en entornos sumamente complejos y adaptativos. En este sentido, se ha demostrado cómo a través de grupos más o menos organizados se puede alterar o dirigir las decisiones de los algoritmos al resultado esperado por el atacante. Ejemplo de ello lo podemos encontrar en las plataformas sociales como Facebook donde ejércitos enteros de personas y *bots* se dedican a difundir noticias falas y a compartirlas. Ello lleva a que en muchas ocasiones los algoritmos de esta plataforma acaben potenciando la difusión de este contenido ya que su objetivo primordial es mantener a los usuarios el mayor tiempo posible en estas redes. Los riesgos en estas situaciones resultan sumamente graves[542].

[540] Fuente de la noticia: MERINO,M: "Conceptos de inteligencia artificial: qué es la inteligencia artificial antagónica y cómo puede manipular a otras IAs", *Xataka*. 18/08/2019. Disponible en: https://www.xataka.com/inteligencia-artificial/conceptos-inteligencia-artificial-que-inteligencia-artificial-antagonica-como-puede-manipular-a-otras-ias

[541] Una prueba similar fue realizada con otro algoritmo que no conseguía distinguir entre un rifle y una tortuga. La tortuga presentaba unos colores muy parecidos a los del rifle, de manera que, cada vez que se mostraba la tortuga el algoritmo la consideraba un rifle. Piénsese en las implicaciones que podría aparejar este tipo de problemas en sistemas de reconocimiento facial en controles de fronteras. En: ANISH ATHALYE,A; ENGSTROM,L; ILYAS,A; KWOK,K: "Synthesizing Robust Adversarial Examples". *Computer Vision and Pattern Recognition*, 2018.

[542] Se ha demostrado que a través de la red social Facebook se potenciaron las revueltas iniciadas por el gobierno birmano para masacrar al pueblo musulmán de los *rohinyás* a partir de 2016. De esta manera, el ejército comenzó a crear perfiles falsos de personas supuestamente famosas. Poco a poco se fueron distribuyendo en dichos perfiles noticias falas y publicaciones incendiarias en contra de los musulmanes birmanos. El objetivo era que estas noticias e imágenes se fueran compartiendo y difundiendo con el objetivo de generar un ambiente hostil hacia esa comunidad por parte del pueblo birmano. Fuente de la

Finalmente, en cuarto lugar, encontramos otra serie de ataques focalizados en la re identificación de los datos personales que se utilizaron para desarrollar el modelo. En este sentido, un primer ataque consiste en la llamada *inversión de modelos.* Es decir, un atacante tiene acceso a determinados datos personales presentes en los datos de entrenamiento y puede inferir información personal adicional de esas personas al observar los datos de entrada y salida de los modelos[543]. Por otro lado encontramos los llamados ataques centrados en la *inferencia de pertinencia*, en estos casos, el atacante puede llegar a saber si los datos personales de un individuo han formado parte del conjunto de datos de entrenamiento. A diferencia del supuesto anterior, el atacante no recupera los datos de entrenamiento sino que obtiene información sobre si un individuo en particular estaba o no en el conjunto de entrenamiento[544]. Nótese que estos ataques resultan más dañinos cuando el modelo es adquirido por una tercera organización, la cual, puede tratar de realizar dichos ataques una vez que ostente el modelo.

Fase	**Tipo de ataque**	**Violación de seguridad de los datos personales**[545]
Diseño	Acceso o filtración de la información presente en los *data lakes* o bases de datos	Violación de la confidencialidad
	Envenenamiento de los datos	Violación de la integridad
Despliegue	Acceso o filtración a los perfiles e inferencias generadas durante la fase de despliegue del sistema	Violación de la confidencialidad
	Ataques contradictorios a los modelos	Violación de la integridad
	Ataques a la funcionalidad del sistema	Violación de la integridad
	Inversión de modelos	Violación de la confidencialidad
	Inferencia de pertinencia	Violación de la confidencialidad

Elaboración propia

noticia: MOZUR,P: "Los militares que usaron Facebook para incentivar un genocidio". *The New York Time*. 18/10/2018. Disponible en: https://www.nytimes.com/es/2018/10/18/espanol/facebook-violencia-rohinya-birmania.html

[543] Fuente ICO: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/how-should-we-assess-security-and-data-minimisation-in-ai/

[544] FINLAY,S: "Predictive Analytics, Data Mining and Big Data. Myths, Methods and Misconceptions". *Palgrave Macmillian*. 2014, pág.4. Disponible en: http://www.odbms.org/2015/02/ethical-risk-assessment-automated-decision-making-systems/

[545] El Grupo del Artículo 29 hizo referencia a los tres criterios de seguridad clásicos, estos son: «violación de la disponibilidad», es decir, la destrucción accidental o ilegal o a la pérdida de datos personales, «violación de la integridad», la alteración de datos personales, y la «violación de la confidencialidad», esto es, revelación no autorizada de datos personales o al acceso no autorizado a los mismos. Dictamen 03/2014 sobre la notificación de violación de datos personales. Adoptado el 25 de marzo de 2014, pág.4.

2. Medidas de seguridad a implantar

Conocidas las principales amenazas, resulta conveniente valorar algunas de las medidas tanto técnicas como organizativas de seguridad que pueden ayudar a reducir o mitigar los riesgos presentes. En este sentido, se hacen referencia a algunas de estas medidas, las cuales, habrán de ser adoptadas por el responsable y el encargado teniendo en cuenta entre otros elementos el estado de la técnica, los costes de aplicación, y la naturaleza, el alcance, el contexto y los fines del tratamiento. Pues bien, en primer lugar es necesario determinar claramente quiénes son las personas autorizadas para poder acceder a las bases de datos y en su caso a los perfiles e inferencias que se vayan generando durante el proceso de toma de decisiones[546]. También se han de reflejar las circunstancias bajo las cuales se pueden acceder a los datos de manera que pueda quedar un registro de dichos accesos[547]. En segundo lugar, es conveniente mantener una monitorización de los resultados para en su caso valorar si el sistema puede estar generando un número de errores excesivo, producto por ejemplo de determinados tipos de ataques. Así, la propuesta de Reglamento de Servicios Digitales que ya hemos comentado en otras páginas obliga a las grandes plataformas a realizar evaluaciones de los principales riesgos sistemáticos que puedan aparecer en estas redes sociales, el objetivos es anticiparse a los mismos y mitigarlos en la medida de lo posible[548]. En tercer lugar, resulta relevante hacer testeos del sistema atacando deliberadamente el modelo a través de ejemplos contradictorios que son probables que se utilicen por terceros atacantes en un futuro[549]. En este sentido, y relacionado con las amenazas a las que se exponen estos sistemas indicadas previamente, la PRAI obliga a los desarrolladores de sistemas de alto riesgo a desplegar soluciones técnicas que ayuden a prevenir y controlar los ataques que intenten manipular el conjunto de datos de entrenamiento o las entradas diseñadas para hacer que el modelo cometa un error[550].

[546] Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. Directrices éticas para una IA fiable.2019, pág.20.
[547] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*. Febrero de 2020, pág.43.
[548] Artículo 26 de la 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.
[549] MOLNAR,CH*: Interpretable Machine Learning. A Guide for Making Black Box Models Explainable*. Libro disponible on line. ISBN 9780244768522. Véase el apartado 6.2. Libor disponible en:https://christophm.github.io/interpretable-ml-book/
[550] Artículo 15.4 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

Finalmente, en cuarto lugar, también convendrá realizar auditorías periódicas de seguridad informática, evaluaciones de vulnerabilidad y pruebas de penetración para detectar por adelantado estos posibles ataques[551].

El objetivo esencial de todas estas medidas por tanto es mitigar o reducir los potenciales incidentes de seguridad de datos personales que sufran los sistemas automatizados o, cuando estos suceden, que puedan ser detectados a tiempo y la respuesta a los mismos sea rápida y eficaz. En este sentido, y teniendo en cuenta que nos movemos en un ámbito donde cada vez surgen nuevas herramientas y técnicas que tratan de sabotear los sistemas de toma de decisiones automatizadas, el factor del estado de la técnica y de coste de aplicación de implantación de las medidas resulta sumamente relevante. Así, en nuestra opinión, y pese a que las organizaciones han de prever un nivel alto de protección frente a posibles incidentes de seguridad que afectan a los datos personales, resulta coherente que, en el caso de determinados ataques plenamente desconocidos y novedosos, las posibles sanciones que se derivan de estas brechas de seguridad queden mitigadas por la novedad de las técnicas o herramientas utilizadas para generar dichos ataques. Resultará más importante en estos supuestos valorar por tanto cómo la organización ha actuado tras ese incidente de seguridad para reducir los posibles impactos en los derechos y libertades de las personas.

3. La materialización de los ataques. Las brechas de seguridad

A pesar de que se hayan podido evaluar los riesgos y se hayan previsto medidas adecuadas, sigue siendo probable que estos, previstos o no, acaben materializándose. En este último caso, se habrá producido una brecha de seguridad. El RGPD en sus artículos 33 y 34 distingue dos mecanismos a la hora de comunicar los incidentes de seguridad que sufre una organización. Estos mecanismo son las notificaciones a las autoridades de control y las comunicaciones a los particulares afectados por dichas brechas. La activación de estos mecanismos dependerá esencialmente del riesgo que entrañe esa específica violación de seguridad para los derechos y libertades de los individuos[552]. Es importante destacar que estos mecanismos únicamente se activarán

[551] Comité Europeo de Protección de Datos. *Guidelines 01/2021 on Examples regarding Data Breach Notification*. Resolución adoptada el 14 de junio de 2021, apartado 51. pág.15.

[552] Así, la notificación a las autoridades de control se requerirá únicamente cuando dicho incidente suponga un riesgo para los derechos y libertades. Por otro lado, cuando dicho riesgo sea alto, la brecha de seguridad se deberá comunicar también a los titulares de los datos. Artículos 33 y 34. RGPD.

cuando la violación de seguridad afecte a los datos personales[553]. De esta manera, cualquier incidente de seguridad que sufra un determinado sistema automatizado no implicará necesariamente la notificación y/o comunicación de la brecha de seguridad. Así, por ejemplo, si se detecta que un vehículo autónomo no es capaz de diferenciar una señal de STOP porque la misma tiene varias pegatinas y por tanto queda limitada su capacidad de reconocimiento artificial, tal incidente de seguridad, al menos, desde la vertiente de la protección de datos es irrelevante[554]. A la inversa, si por ejemplo, a través de una violación de seguridad se logran conocer determinados datos inferidos o los perfiles correspondientes a una persona, el responsable muy posiblemente deberá comunicar tal situación a los interesados afectados. Piénsese por ejemplo si salen a la luz las inferencias realizadas por una entidad bancaria de todo un conjunto de interesados a los que se les ha denegado el crédito debido a que el algoritmo los consideró personas poco solventes o malos pagadores.

En definitiva, la seguridad se convierte en el RGPD en un elemento esencial que han de tener en cuenta los responsables del tratamiento que diseñen y desplieguen sistemas de toma de decisiones automatizadas. Una vez más, el enfoque propuesto de interpretación de los principios en materia de protección de datos con las exigencias de robustez de estos sistemas algorítmicos queda patente. Así, una correcta adecuación del principio de seguridad reconocido en el artículo 5 del RGPD derivará en una protección adecuada de ese sistema frente ataques. Esto se materializa tanto en una ventaja corporativa para la organización como en un adecuado cumplimiento de la normativa de protección de datos.

VI. EL DELEGADO DE PROTECCIÓN DE DATOS

Una de las figuras novedosas previstas por el RGPD es el relativo al delegado de protección de datos, en adelante DPD. Este agente forma parte del conjunto de

[553] Artículo 4.12 RGPD y Grupo Artículo 29. *Directrices sobre la notificación de las violaciones de la seguridad de los datos personales de acuerdo con el Reglamento 2016/679*. Adoptadas el 3 de octubre de 2017. Revisadas por última vez y adoptadas el 6 de febrero de 2018, págs. 7 y 8.

[554] Lo que no quiere decir que dicho incidente de seguridad no haya de notificarse en virtud de otras normas. Ello puede ocurrir con los incidentes de seguridad que pudiera sufrir un sistema de toma de decisiones automatizadas en la Administración Pública. Artículo 36 y Deposición Adicional Cuarta del 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. Dicha notificación del incidente de seguridad ha de realizarse de acuerdo al procedimiento contenido en la Resolución de 13 de abril de 2018, de la Secretaría de Estado de Función Pública, por la que se aprueba la Instrucción Técnica de Seguridad de Notificación de Incidentes de Seguridad.

medidas de garantía que han de implantar los responsables del tratamiento durante el ciclo de vida de los sistemas de toma de decisiones automatizadas. Como ahora veremos, además de facilitar el cumplimiento mediante la aplicación de instrumentos de rendición de cuentas, el DPD actúa como intermediarios entre las partes interesadas correspondientes[555].

Por lo que a nosotros nos interesa, la AEPD ha señalado que el uso de sistemas automatizados no implica, *per se*, una obligación para designar un DPD[556]. En este sentido, el artículo 37.1 del RGPD establece los casos en los que resultará obligatoria dicha designación, estos son: i) cuando el tratamiento lo lleve a cabo una autoridad u organismo público, ii) cuando las actividades principales requieran una observación habitual y sistemática de interesados a gran escala, o iii) cuando las actividades principales consistan en el tratamiento a gran escala de categorías especiales o relativos a condenas e infracciones penales. De esta manera, cuando el desarrollo o el despliegue de los sistema automatizados se lleve a cabo por Administraciones Públicas o en los tratamientos implicados en esas fases se traten datos personales sensibles o de incidencia penal, será necesario designar un DPD. Por otro lado, resultará habitual la presencia de este agente en aquellas organizaciones que se dediquen a la analítica masiva de datos personales para desarrollar modelos algorítmicos, tal y como ocurre en la investigación científica. El tratamiento de datos a gran escala y de forma sistemática en estos contextos resulta evidente. En este sentido, la LOPD de 2018 establece que los comités de ética de la investigación, en el ámbito de la salud, biomédico o del medicamento deberán estar integrados por un DPD[557]. A su vez, y dado que una de las funciones principales de los sistemas de toma de decisiones automatizadas es que se apliquen a un número importante de personas, el uso de los mismos en dichos tratamientos comportará también la habilitación del DPD. Piénsese por ejemplo en los sistemas dedicados a la recomendación de contenidos en las plataformas digitales o la elaboración de perfiles a efectos de mercadotecnia. Así, el artículo 34 de la LOPD de 2018 obliga a los prestadores de servicios de la sociedad de la información a que implemente esta figura cuando elaboren a gran escala perfiles de los usuarios del servicio.

[555] Grupo del artículo 29. *Directrices sobre los delegados de protección de datos (DPD). Adoptadas el 13 de diciembre de 2016.* Revisadas por última vez y adoptadas el 5 de abril de 2017, pág.4.
[556] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.35.
[557] Disposición Adicional. 17.2.h) de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

Es importante destacar que la implementación de un DPD corresponde tanto al responsable del tratamiento como al encargado. Todo dependerá de quién cumpla los criterios de designación obligatoria descritos anteriormente. En este sentido, normalmente, durante todo el ciclo de vida de los sistemas automatizados suele ser frecuente que varios agentes intervengan en calidad de responsables y encargados. En estos casos, los DPDs de cada una de estas organizaciones deberán cooperar entre sí focalizando su atención en sus respectivas tareas. A modo de ejemplo, en la fase de diseño, una organización (responsable) puede estar utilizando los servicios de *cloud* que le ofrece una empresa (encargado) para analizar masivamente datos personales y desarrollar modelos algorítmicos. Cuando se requiera, estos DPDs deberían estar en contacto.

1. El rol del delegado de protección de datos en el contexto de los sistemas de toma de decisiones automatizadas

El RGPD asigna toda una serie de funciones y competencias en favor del DPD con el objetivo de que su papel en el seno de la organización sea relevante. Por lo que se refiere a los sistemas de toma de decisiones automatizadas, estas tareas pueden concretarse en las siguientes:

Con relación a la *fase de diseño*, el DPD debería asesorar a los desarrolladores de los modelos algorítmicos de todos aquellos aspectos que puedan tener incidencia en materia de protección de datos durante todas las etapas que engloba la creación de estos sistemas[558]. En concreto, puede tener un papel muy notable durante la configuración de la EIPD o a la hora de decidir sobre la compatibilidad de los usos posteriores de los datos proyectados, tal y como puede ocurrir con el uso secundario de los datos para fines de analítica[559]. A su vez, en este momento el DPD puede ayudar a implementar técnicas que favorezcan la privacidad desde el diseño con el objetivo de que después, y una vez los sistemas comiencen a adoptar decisiones, estos cumplan con las exigencias normativas en materia de protección de datos.

[558] Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág.20. Tal y como ha señalado el Grupo de Expertos de la Comisión Europea, el DPD debería estar presente desde las fases iniciales de los proyectos algorítmicos. Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Directrices éticas para una IA fiable*.2019, pág.36.

[559] Agencia Española de Protección de Datos. *El delegado de protección de datos en las administraciones públicas*, 2018 pág.3.

En lo que respecta a la *fase de despliegue*, el DPD puede tener una gran relevancia a la hora de relacionarse con las partes interesadas durante el proceso de toma de decisiones automatizadas. Así, por un lado, cuando las autoridades de control se dirijan al responsable o encargado del tratamiento, el DPD actuará como intermediario facilitando los documentos o información que requieran[560], por ejemplo; los registros de funcionamiento del sistema, la monitorización que hasta la fecha se está llevando a cabo, información técnica del modelo algorítmico, etc. Por otro lado, el DPD también podrá comunicarse con los afectados por la toma de decisiones o la elaboración de perfiles. Este agente puede ser la primera persona de la organización con la que el particular puede tener contacto una vez que se ha tomado una decisión plenamente automatizada relevante y exija una intervención humana conforme a lo previsto en el artículo 22 del RGPD. Finalmente, para facilitar las comunicaciones tanto con las autoridades de control como con los interesados, el propio DPD ha de mantener también una relación estrecha con el resto de personal de la organización que forma parte del proceso de formación o elaboración de las decisiones automatizadas.

En definitiva, el DPD se configura en el RGPD como una herramienta básica que tanto el responsable como encargado han de designar. Pese a que su despliegue no es obligatorio en todos los supuestos. Aquellas organizaciones que hagan uso de sistema automatizados o los desarrollen deberán normalmente contar con esta figura fruto del análisis masivo de datos personales que acompaña a estos tratamientos. Su conocimiento sobre la normativa de protección de datos le permitirá asesorar y supervisar el cumplimiento de este derecho fundamental durante el ciclo de vida de estos sistemas. En este sentido, el Tribunal Constitucional Alemán ha señalado la importancia que presenta el DPD en la supervisión de los sistemas de toma de decisiones automatizadas, sobre todo, cuando dichas decisiones ni siquiera se conozcan por parte de las personas afectadas por las mismas[561].

[560] Véase el artículo 39. epígrafes d) y e) del RGPD.

[561] Este órgano judicial se refiere al uso por parte de la policía alemana de sistemas automatizados de reconocimiento de matrículas. BVerfG, Order of the First Senate of 18 December 2018 - 1 BvR 142/15 -, Resolución de 18 de diciembre de 2018, FJº 157. Resolución disponible en: http://www.bverfg.de/e/rs20181218_1bvr014215en.html

VII. LOS CÓDIGOS DE CONDUCTA

Los códigos de conducta son una herramienta de rendición de cuentas voluntaria en las que se establecen toda una serie de disposiciones específicas en materia de protección de datos destinadas a responsables y encargados de tratamiento que, mediante su adhesión a los mismos, se comprometen a cumplir lo indicado en tales disposiciones[562]. Una vez elaborados, estos han de presentarse a la autoridad de control correspondiente o al Comité Europeo de Protección de datos para que los aprueben.

Son dos los las características esenciales de esta herramienta, por un lado facilitan la correcta aplicación del RGPD en un sector específico ya que permiten una adaptación más adecuada de la normativa europea a los problemas detectados en esos ámbitos de una forma práctica y, por otro lado sirven como mecanismo que ayuda a las organizaciones adheridas voluntariamente al mismo a demostrar que cumplen con las obligaciones establecidas por el RGPD de acuerdo al principio de responsabilidad activa, permitiendo que, dichas adhesiones pueden ser tenidas en cuenta a la hora de imponer mayores o menores sanciones por incumplimiento de esa normativa[563].

Impronta de los códigos de conducta como mecanismo que demuestra el cumplimiento de las obligaciones previstas en el RGPD	Principio de responsabilidad activa. Artículo 24.3 Encargado del tratamiento. Artículo 28.5 Seguridad de los datos. Artículo 32.3 Evaluaciones de impacto. Artículo 35.8

Por tanto, los códigos de conducta se manifiestan como una herramienta que resulta muy interesante para sectores específicos que dedican habitualmente parte de sus esfuerzos económicos a tratamientos de datos personales cuyas actividades están ligadas al diseño y despliegue de sistemas automatizados. Escenarios como la investigación sanitaria, el sector bancario o seguros son ejemplos donde podrían elaborarse una serie de disposiciones que aporten soluciones prácticas y específicas a los principales

[562] Comité Europeo de Protección de Datos. *Directrices 1/2019 sobre códigos de conducta y organismos de supervisión con arreglo al Reglamento (UE) 2016/679***.** Versión 2.0. Resolución aprobada el 4 de junio de 2019. Apartado 7, pág.7.

[563] Artículo 83.2.j) del RGPD. También en: Grupo del Artículo 29. *Directrices del grupo del artículo 29 sobre la aplicación y la fijación de multas administrativas a efectos del Reglamento 2016/679*. Adoptadas el 3 de octubre de 2017. Pág.16. Un estudio muy profundo sobre las sanciones en materia de protección de datos personales puede verse en: MARZAL RAGA, R: *El apercibimiento. Una nueva sanción en materia de protección de datos de carácter personal.* Ed. Tirant lo Blanch, Valencia, 2015.

problemas detectados durante las distintas operaciones que engloban los tratamiento de datos personales implicados en el ciclo de vida de estos sistemas automatizados.

Son varios los criterios que ha indicado el Comité Europeo de Protección de Datos que deberán tener en cuenta las autoridades de control a la hora de considerar la aprobación de un determinado código de conducta, estos son[564]:

En primer lugar, el código ha de satisfacer una necesidad específica de ese sector o actividad sobre el cual se pretende proyectar las mencionadas disposiciones. Es decir, los titulares de esta herramienta deben demostrar que es necesario establecer un código en ese contexto. Para ello, se han de identificar los problemas detectados que se pretenden abordar justificando que las soluciones ofrecidas no sólo son beneficiosas para los miembros del mencionado código sino para los titulares de los datos[565].

En segundo lugar, el código de conducta ha de facilitar la aplicación adecuada del RPGD estableciendo por ejemplo definiciones claras de los elementos más relevantes de ese ámbito[566]. Así, en dicho código podrían diferenciarse claramente las distintas fases que engloban el ciclo de vida de los sistemas automatizados o aquellas fases específicas que están implicadas en el sector concreto sobre el que irradiaría sus efectos dicho código. De esta manera, en el sector ligado a la investigación sanitaria podrían establecerse claramente cuáles son las etapas habituales presentes durante la analítica masiva de datos[567]. Estas aclaraciones también puede ayudar a vislumbrar y discernir si determinadas herramientas u operaciones que se llevan a cabo en concretos sectores pueden o no considerarse inteligencia artificial a efectos de las distintas normativas aplicables.

En tercer lugar y ligado a la anterior, se ha de especificar la aplicación práctica del RGPD. Es decir, se han de aportar mejoras realistas y asequibles al sector con relación al cumplimiento de la normativa de protección de datos. En palabras del CEPD, el código no debe limitarse a reiterar el contenido del RGPD sino que debe codificar

---

[564] Comité Europeo de Protección de Datos. *Directrices 1/2019 sobre códigos de conducta y organismos de supervisión con arreglo al Reglamento (UE) 2016/679***,** op.cit., apartados 32 y ss y pág.s14 y ss.

[565] Comité Europeo de Protección de Datos. *Directrices 1/2019 sobre códigos de conducta y organismos de supervisión con arreglo al Reglamento (UE) 2016/679***,** op.cit., apartados 33 y 34, pág..15.

[566] Por ejemplo, el Código Ético de Comercio Electrónico y Publicidad Interactiva que está registrado en la AEPD establece determinadas definiciones útiles aplicables a ese concreto marco. Véase la página 10 del mentado código. Disponible en: https://www.aepd.es/sites/default/files/2019-11/ct-confianza-online.pdf

[567] Aunque no es un código de conducta, un buen ejemplo de buenas prácticas en el ámbito de la analítica de datos en el ámbito sanitario lo encontramos en el playbook realizado por Fundación 29 ya comentado en otras páginas. Playbook Health Data: Disponible en: https://www.healthdata29.org/playbook

cómo ha de aplicarse esta norma de forma específica, práctica y precisa en ese sector[568]. El código de conducta ha de aportar un valor añadido al cumplimiento normativo de ese sector[569].Así, puede ser recomendable que el código se focalice en las principales fricciones que presentan el desarrollo y despliegue de sistemas automatizados en relación con las exigencias del RGPD. Los precursores de estos códigos se encuentran en la mejor posición para detectar esos problemas, aportar soluciones y a la vez, mantener un nivel alto de respeto de la normativa. Así, los puntos más críticos que ya hemos ido detectando a lo largo de este estudio podrían trasladarse a los contenidos de estas disposiciones para aclararlos. En este sentido, entre otros elementos, se podrían establecer; i) las principales decisiones automatizadas que generan efectos jurídicos en un concreto sector[570], ii) la forma en que se realiza la elaboración de perfiles y cómo dichos perfiles deben adecuarse a las exigencias del RGPD, iii) fijar reglas claras de los procesos de anonimización en relación con los específicos análisis de datos que se llevan a cabo, iv) establecer los datos que se recopilan, v) las formas lícitas de recopilación y los tipos de datos que normalmente se infieren, vi) fijar criterios a la hora de limitar algunas facultades como el derecho de acceso, rectificación … por motivos relacionadas con los secretos comerciales, propiedad intelectual, etc.

En cuarto lugar, estos códigos han de aportar suficientes garantías adecuadas que sean conforme al RGPD. Así, centrándonos específicamente en los derechos de los particulares, podría ser útil prever los canales específicos que pueden habilitarse para los titulares de los datos en las fases implicadas durante el desarrollo y despliegue de los sistemas algorítmicos. Por ejemplo, el derecho a la intervención humana que aparece reconocido en términos generales en el artículo 22 del RGPD puede ser instrumentalizado para ese concreto sector, teniendo en cuentas el tipo de sistemas o entornos habituales en los que se puede ejercitar tal facultad por parte de los interesados. También se puede hacer referencia a la implicación de los principios de protección de datos en las principales fases afectadas en ese concreto sector. Por

[568] Comité Europeo de Protección de Datos. *Directrices 1/2019 sobre códigos de conducta y organismos de supervisión con arreglo al Reglamento (UE) 2016/679***,** op.cit., apartado 37, pág.16.

[569] Agencia Española de Protección de datos. Informe N/REF: 0089/2020. Págs. 1, 10 y 155. Texto disponible en: https://www.aepd.es/es/documento/2020-0089.pdf

[570] Por ejemplo, también se puede especificar y justificar que en ese concreto sector los tratamientos que se llevan a cabo no toman decisiones plenamente automatizadas que generan efectos relevantes para los interesados. Así se indica en el código de conducta para el tratamiento de datos de carácter personal por organizaciones de investigación de mercado, social, de la opinión y el análisis de datos elaborado por Aneimo y Aedemo. (pág. 23). Este código ha sido aprobado por la AEPD. Disponible en: https://www.aepd.es/es/documento/ct-organizaciones-investigacion-mercados.pdf

ejemplo, cómo el principio de minimización de datos tiene su influencia en la fase de desarrollo de los sistemas a la hora de decidir entre unas u otras correlaciones presentes o deducidas durante la fase de entrenamiento.

Finalmente, los códigos de conducta han de prever mecanismos eficaces que permitan una adecuada supervisión del cumplimiento de los mismos. En este sentido, el artículo 41 del RGPD establece que la supervisión de los códigos se realizará no sólo por las autoridades de control sino también por un organismo que, ostentando el nivel adecuado de pericia en relación con el objeto del código, haya sido acreditado para tal fin por parte de la autoridad de control competente. El organismo designado por el código debe contar con herramientas eficaces para que en su caso pueda ser acreditado[571]. Se destaca la independencia que el mismo ha de ostentar respecto del resto de organizaciones que se adhieren al código[572]. Estos órganos pueden ser externos o internos. A su vez, dichos órganos pueden crearse ad hoc o en su caso y respecto de organismo ya creados para otras funciones, implantar unidades que se encarguen específicamente de la supervisión de los códigos de conducta. En los contextos que estamos analizando, los comités de ética basados en la privacidad podrían resultar muy adecuados para supervisar los códigos de conducta aprobados[573]. Este organismo podría establecer distintas unidades encargados de supervisar las distintas fases que pueden estar comprendidas durante el ciclo de vida de los sistemas algorítmicos y que hayan sido recogidas en esos códigos. Por ejemplo, en la fase de desarrollo, como ya hemos señalado, los comités de ética pueden ser muy relevantes ya que pueden aportar un enfoque de respeto de la normativa en estas primeras fases de diseño de los sistemas

[571] Estos comités debería ser multidisciplinares, con especialistas técnicos, usuarios de servicios especializados y que se incluya al menos un miembro con formación filosófica, ética o jurídica. En: COTINO HUESO, L: "Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho". *Revista Catalana de Dret Públic*, núm. 58, 2019, pág.42.

[572] Tal y como indica el EDBP, el proyecto de código debe designar un órgano adecuado que disponga de mecanismos que le permitan velar por la supervisión eficaz del cumplimiento del código. Además de dicho órgano, también puede ser relevante que se fijen otros mecanismos entre los que se pueden incluir la realización de auditorías periódicas, presentación de informes, la gestión clara y transparente de las reclamaciones y los procedimientos de solución de litigios, sanciones específicas y medidas correctivas en caso de infracción del código, así como mecanismos para denunciar las infracciones de sus disposiciones. En: Comité Europeo de Protección de Datos. *Directrices 1/2019 sobre códigos de conducta y organismos de supervisión con arreglo al Reglamento (UE) 2016/679***,** op.cit., apartado 40, pág.17.

[573] La figura de los Comités de ética en el ámbito de la investigación sanitaria es obligatoria. Artículo 12 de la Ley 14/2007, de 3 de julio, de Investigación biomédica. La LOPD de 2018 ha atribuido un papel relevante a dichos comités en materia de protección de datos. Concretamente, esta norma obliga a los responsables del tratamiento a recabar informe favorable de dicho comité cuando la reutilización de datos personales tenga como finalidad la investigación en materia de salud y biomédica. D.A. 17.2.c) de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

automatizados. En un segundo momento, esto es, la fase de despliegue del sistema, podría asignarse la supervisión a comités especializados en interpretar y resolver problemas que se deriven de la toma de decisiones que adoptan los sistemas algorítmicos. Así, a modo de ejemplo, Facebook ha creado un organismo independiente que tiene como objetivo analizar las posibles denuncias que se derivan de la retirada de contenido ilícito que comparte los usuarios de esta plataforma, esta retirada es habitualmente realizada por sistemas algorítmicos[574]. Este tipo de organismos podrían asumir también las funciones de supervisión de los códigos de conducta. Por otro lado, y teniendo en cuenta que el diseño e implantación de sistemas automatizados se encuentra actualmente en una fase continua de desarrollo tecnológico, estos organismos podrían facilitar listas actualizadas de nuevas técnicas, investigaciones científicas o tecnologías que puedan resultar de utilidad o ayuden a una mayor protección de la privacidad durante las fases que comprenden el ciclo de vida de estos sistemas algorítmicos. Por ejemplo, durante la fase del desarrollo, desde hace ya algún tiempo se vienen proponiendo todo tipo de técnicas que favorecen la privacidad en la analítica de datos[575]. A su vez, dentro del mundo de los científicos de datos y programadores especializados en IA existen varios enfoques centrados en el estudio de diferentes ramas como la explicabilidad de las decisiones[576], la transparencia de los algoritmos o la prevención frente ataques intencionados contra los sistemas automatizados[577]. En todas estas disciplinas estrechamente relacionadas con la toma de decisiones automatizadas, es incesante el desarrollo de nuevas técnicas orientadas a adecuar los usos de los sistemas automatizados a los ordenamientos jurídicos en los que estos se implementan. La posibilidad de que exista un órgano que en parte pueda facilitar a las organizaciones

[574] Consejo asesor de contenidos de Facebook. https://oversightboard.com/

[575] Una recopilación de las técnicas de privacidad durante el proceso de *data mining* puede encontrarse en: YOUSRA ABDUL ALSAHIB, A; MAZLEENA, S; MOHAMMAD ABDUR, R: "A comprehensive review on privacy preserving data mining". *SpringerPlus* **4,** 694 (2015). https://doi.org/10.1186/s40064-015-1481-x

[576] La Inteligencia artificial explicable tiene como objetivo crear un conjunto de técnicas de aprendizaje automático modificadas que producen modelos que, cuando se combinan con una explicación eficaz, permiten a los usuarios finales comprender, confiar y gestionar de forma eficaz los resultados emitidos por los de sistemas de IA. El objetivo de la IA explicable es que el usuario final de quién dependen las decisiones o recomendaciones producidas por un sistema de IA pueda comprenderlas. En: GUNNING, D., & AHA, D: "DARPA's Explainable Artificial Intelligence (XAI) Program". *AI Magazine*, Vol. 40 No. 2: Summer 2019, pág,45. Texto disponible en: https://doi.org/10.1609/aimag.v40i2.2850

[577] Adversarial Machine Learning, es una rama del *machine learning* que trata de averiguar qué ataques puede sufrir un modelo en la presencia de un adversario malicioso y cómo protegerse de los ataques del mismo. Fuente de la noticia: HERNÁNDEZ, M; IGNACIO ESCRIBANO,J:: "Adversarial Machine Learning: una introducción. ¿Motivo de preocupación?", *BBVA Next Technologies,* 15/09/2020. Disponible en**:** En: https://www.bbvanexttechnologies.com/adversarial-machine-learning/

adheridas a dichos códigos de conducta las actualizaciones de las técnicas más novedosas y respetuosas con la normativa se muestra esencial en este contexto.

VIII. LOS MECANISMOS DE CERTIFICACIÓN

Al igual que ocurre con los códigos de conducta, la certificación también es voluntaria y puede servir como elemento para demostrar el cumplimiento de las obligaciones por parte del responsable. Así, la certificación consiste en la obtención de un distintivo que acredita que los tratamientos de datos que se llevan a cabo por parte del responsable y el encargado del tratamiento cumplen los criterios establecidos por el proceso de certificación. Esos criterios en materia de protección de datos estarán estrechamente relacionados con las disposiciones normativas previstas en el RGPD.

Los mecanismos de certificación en el desarrollo y despliegue de sistemas de tomas de decisiones automatizadas presentan una serie de ventajas, entre otras:

En primer lugar, los certificados pueden actuar como auténticos sellos de calidad que certifiquen que ese modelo algorítmico cumple con las exigencias derivadas de la normativa de protección de datos. Así, cuando una organización pretenda adquirir o utilizar un determinado sistema, la certificación del mismo permitirá que dicha organización lo obtenga con ciertas garantías de cumplimiento normativo. A modo de ejemplo, las empresas que ofrecen servicios de *cloud* para que otras organizaciones puedan procesar sus datos y elaborar sus propios modelos algorítmicos les convendrá certificar que los servicios que facilitan cumplen con la normativa de protección de datos[578]. De esta manera, estas certificaciones otorgan a esos sistemas algorítmicos o servicios una ventaja competitiva respecto de otros sistemas que no ostentan dichos sellos, resultando un factor esencial cuando estos modelos o servicios se pretendan ofrecer a las Administraciones Públicas. En este sentido, el artículo 128 de la Ley de contratos del sector público establece que los órganos de contratación podrán exigir a los futuros contratistas certificados que pruebe que el producto que se ofrece cumple con las prescripciones técnicas exigidas en el contrato público. Aunque la norma

[578] Como ya dijimos al inicio de esta tesis, las grandes tecnológicas ofrecen algoritmos ya desarrollados dispuestos a procesar los datos que aporten los clientes. La certificación de todos esos productos o paquetes que se ofrecen resulta sumamente importante para estas organizaciones. Ejemplos como *TensorFlow* de Google, *Microsoft Azure* de Microsoft, *Amazon web series* de Amazon o *Deep learning as a service* de IBM son herramientas de este tipo. Véase el Capítulo I, apartado I, punto 4 de esta tesis.

permite la posibilidad de aportar otras pruebas que suplan la certificación, lo cierto es que la misma facilita tanto la labor de la Administración contratante como la del potencial contratista a la hora de demostrar esas especificaciones técnicas[579].

En segundo lugar, la certificación puede reducir los problemas ligados a los derechos de propiedad intelectual y secretos comerciales que se derivan de las exigencias de trasparencia que se propugnan sobre los sistemas algorítmicos[580]. Así, los mecanismos de certificación permiten que la sociedad en general pueda confiar en los tratamientos que llevan a cabo estos algoritmos sin necesidad de publicitar los mismos ya que una tercera organización ha certificado que estos cumplen con las exigencias previstas por la normativa[581]. Aunque los mecanismos de certificación pueden ayudar a conjugar la dicotomía existente entre transparencia y secretos comerciales de los algoritmos, nunca deberían suponer un salvoconducto para establecer un total oscurantismo u opacidad de tales sistemas, sobre todo, en el sector público.

Tal y como establece el RGPD, existen distintos órganos que pueden expedir los certificados en materia de protección de datos. En primer lugar, la certificación puede ser expedida por las propias autoridades de control[582]. En segundo lugar, también está legitimado el Comité Europeo de Protección de Datos, en este caso, dicha certificación recibirá la denominación de Sello Europeo de Protección de Datos. En tercer lugar, también se habilita para la emisión de certificados a aquellos organismos que tenga un nivel adecuado de pericia en materia de protección de datos. Para que estos organismos puedan certificar previamente deberán de ser acreditados por la autoridad de control, por el CEPD o por la Entidad Nacional de Acreditación (ENAC)[583].

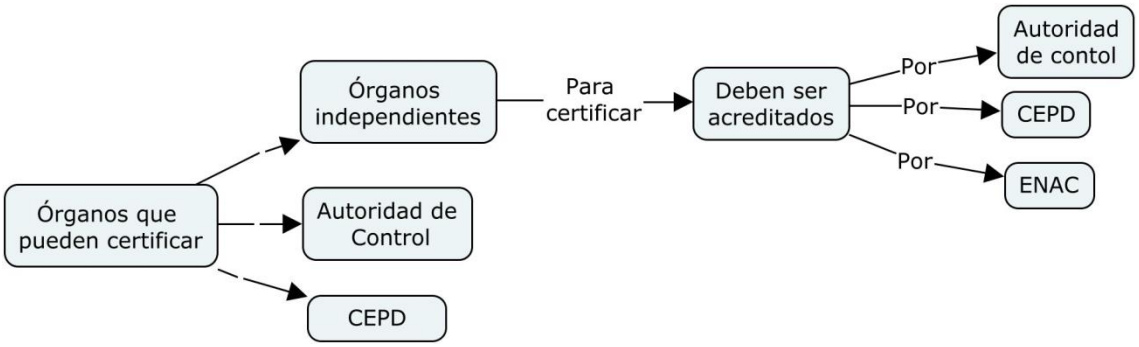
[579] Artículo 128 de la Ley 9/2017, de 8 de noviembre, de Contratos del Sector Público, por la que se transponen al ordenamiento jurídico español las Directivas del Parlamento Europeo y del Consejo 2014/23/UE y 2014/24/UE, de 26 de febrero de 2014.

[580] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, 2020, pág.34. En este mismo sentido, AZUAJE PIRELLA,M; FINOL GONZALEZ,D: "Transparencia algorítmica y la propiedad intelectual e industrial: tensiones y soluciones". *Revista la propiedad inmaterial*, n.º 30 - julio -diciembre de 2020, pág.131.

[581] Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. Directrices éticas para una IA fiable.2019.Pág.28

[582] Artículo 42.5. del RGPD.

[583] Tal y como se establece en el artículo 43.1 del RGPD y en el artículo 39 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.



Proceso de certificación establecido por el RGPD. Elaboración propia.

La certificación requiere que los responsables o encargados que pretendan la misma sometan sus actividades de tratamiento al organismo de certificación o autoridad de control para que esta pueda llevar a cabo adecuadamente el proceso de certificación. Una vez se haya concedido la certificación, el responsable y encargado no quedan eximidos de las posibles responsabilidades que se deriven del incumplimiento de la normativa de protección de datos[584]. En nuestra opinión, aunque las certificaciones focalizadas en la normativa de protección de datos resultan muy garantistas para este derecho. Este mecanismo puede tener poca aplicabilidad práctica en el desarrollo y despliegue de sistemas automatizados. En ese sentido, en el ciclo de vida de los sistemas automatizados no sólo queda afectada la normativa de protección de datos. Por ello, un mecanismo de certificación que contemple no sólo esta perspectiva sino otros elementos de cumplimiento normativo permitirán que los mecanismos de certificación pueda tener una mejor acogida a la hora de implementarlos.

IX. LA ANONIMIZACIÓN Y SEUDONIMIZACIÓN

1. La anonimización de datos personales

A) Concepto

A día de hoy no existe ningún texto legal que regule expresamente los procesos de anonimización. Más bien, existen un conjunto de normas que aluden a este proceso

[584] Comité Europeo de Protección de Datos. *Directrices 1/2018 sobre la certificación y la determinación de los criterios de certificación de conformidad con los artículos 42 y 43 del Reglamento*. Versión 3.0. Adoptadas el 4 de junio de 2019. Apartado 13, pág. 8.

pero sin establecer un desarrollo del mismo[585]. En este sentido, cabe destacar el Considerando 26 del RGPD el cual viene a indicar que:

> *los principios de protección de datos no deben aplicarse a la información anónima, es decir información que no guarda relación con una persona física identificada o identificable, ni a los datos convertidos en anónimos de forma que el interesado no sea identificable, o deje de serlo. En consecuencia, el presente Reglamento no afecta al tratamiento de dicha información anónima, inclusive con fines estadísticos o de investigación*

Debido a ese déficit de regulación, las distintas autoridades de protección de datos se han encargado de elaborar toda una serie de orientaciones y guías sobre el modo de proceder ante un proceso de anonimización de dato y las principales técnicas y garantías que en su caso se han de establecer por aquellas organizaciones que pretendan llevar a cabo dicho tratamiento[586]. Así, tal y como indica la AEPD, la anonimización es el proceso que permite eliminar o reducir al mínimo los riesgos de reidentificación de un individuo a partir de sus datos personales**,** eliminando toda referencia directa o indirecta a su identidad, pero manteniendo la veracidad de los resultados del tratamiento de los mismos[587]. De esta manera, una vez que una organización anonimiza los datos, esta puede disponer de ellos sin requerir de las exigencias establecidas por la normativa de protección de datos, elemento que resulta sumamente relevante en aquellas organizaciones que se dedican a desarrollar modelos algorítmicos que posteriormente se utilizarán para adoptar decisiones automatizadas.

Señalado lo anterior, resulta por tanto conveniente valorar cuándo un dato personal deja de serlo y se transforma en anónimo. En este sentido, no existe actualmente un criterio unánime sobre cuál ha de ser el umbral exigido a la hora de considerar que un dato ha dejado de ser personal[588]. Así, para el GT29, la anonimización será efectiva únicamente si se consigue que la reidentificación de los datos resulte irreversible, es decir, una situación similar al hecho de que el dato personal

---

[585] Considerando 26 del RGPD. Puede verse una definición de este concepto en el artículo 2.7 de 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.

[586] Para desarrollar parte del contenido de este apartado nos valdremos de los documentos elaborados por: i) Grupo del Artículo 29. Dictamen 05/2014 sobre técnicas de anonimización. Adoptado el 10 de abril de 2014. ii) Agencia Española de Protección de Datos. *Orientaciones y garantías en los procedimientos de anonimización de datos personales*. Adoptadas en 2016. iii) Information Commissioner´s Office. *Anonymisation: managing data protection risk code of practice*.2012.

[587] Agencia Española de Protección de Datos. *Orientaciones y garantías en los procedimientos de anonimización de datos personales*. Adoptadas en 2016,pág.2.

[588] GIL GONZÁLEZ,E : *Big data, privacidad y protección de datos*, op,cit.,, pág..86.

se hubiese borrado porque resulta imposible volver a vincularlo con la persona[589]. Esta interpretación ha sido criticada por la doctrina al considerarla extremadamente estricta y poco realista ya que hoy día resulta muy complicado reducir a cero el riesgo de reidentificación de un determinado dato que en su caso ha sufrido un proceso adecuado de anonimización[590]. Esta interpretación estricta puede cortocircuitar el desarrollo de la analítica masiva de datos anónimos[591]. En este sentido se afirma que el elemento a tener en cuenta a efectos de reidentificar a una persona se ha de centrar en el juicio de razonabilidad que prevé el considerando 26 del RGPD. Concretamente, a través de este juicio se han de valorar los medios que razonablemente podría utilizar el responsable o cualquier otro individuo para tratar de reidentificar a una persona[592]. Para ello, este precepto exige que se tengan en cuenta todos los factores objetivos que puedan ser relevantes para esa posible reidentificación, estos son; los costes y el tiempo necesarios para la identificación, la tecnología disponible y los avances tecnológicos del momento. De esta manera, el proceso de anonimización no se configura como una obligación de resultados sino de medios ya que el responsable ha de adoptar los medios razonables que reduzcan al máximo la posibilidad de reidentificación[593]. En nuestra opinión, este último enfoque resulta adecuado ya que permite reducir las posibilidades de identificación de un dato con una persona y a su vez mantiene la posible utilidad de los datos para su análisis en la búsqueda de nuevo conocimiento escondido en estos.

B) Los riesgos de reidentificación. El análisis de riesgos

Derivado de la conclusión anterior, será conveniente realizar un análisis de los principales riesgos a los que puede verse sometida la base de datos que se utilizará de entrenamiento para elaborar modelos algorítmicos y que se pretende anonimizar para en su caso establecer los medios razonables que traten de evitar la reidentificación de los

[589] Grupo del Artículo 29. Dictamen 05/2014 sobre técnicas de anonimización. Adoptado el 10 de abril de 2014. Pág.6

[590] EL EMAM,K; ÁLVAREZ RIGAUDIAS,C: "A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques", *International Data Privacy Law*, Volume 5, Issue 1, February 2015, pág.74. Disponible en: https://doi.org/10.1093/idpl/ipu033. Véase también, MARTÍNEZ MARTÍNEZ,R: "Big data, investigación en salud y protección de datos personales ¿Un falso debate?". *Revista valenciana d'estudis autonòmics*, op,cit., págs. 257 a 259.

[591] GIL GONZÁLEZ,E : *Big data, privacidad y protección de datos*, op,cit., pág.86.

[592] EL EMAM,K; ÁLVAREZ RIGAUDIAS ,C: "A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques", *International Data Privacy Law*, op.cit., pág.75.

[593] LOZA, CORERA,M,L: *De los microdatos a los datos masivos. Cuestiones legales.* Tesis doctoral. Universitat de València. 2017, pág.343.

datos contenidos en esos repertorios[594]. En este sentido, el grupo del artículo 29 ha identificado los tres riesgos esenciales a los que se ven expuestas las bases de datos anonimizadas y sobre las cuales puede seguir existiendo la posibilidad de reidentificación[595].

Estas son: la singularización, la vinculabilidad y la inferencia.

Por lo que se refiere a la *singularización*, esta se entiende como la posibilidad de aislar algunos o todos los registros que identifican a un individuo en el conjunto de datos.

Por otro lado, la *vinculabilidad* o asociación, es la capacidad de vincular como mínimo dos registros de un único interesado o de un grupo de interesados, ya sea en la misma base de datos o en dos bases de datos distintas. Por ejemplo, la reidentificación de una base de datos supuestamente anonimizada a través del riesgo de vinculabilidad es reflejada por el grupo del artículo 29 en un supuesto real sucedido en la década de los 90. Así, una empresa estadounidense del sector sanitario publicó un conjunto de datos supuestamente anonimizados con el fin de que los investigadores pudieran analizarlos y sacar conclusiones. La anonimización consistía en eliminar los nombres de los interesados, sin embargo, la información aún contenían datos como el código postal, el sexo y la fecha de nacimiento completa. Estos tres atributos también estaban contenidos en otros registros de acceso público, concretamente en el censo electoral donde, junto a esos tres atributos, figuraban otros datos como el nombre y los apellidos de esas personas. Por ello, una investigadora pudo vincular la identidad de determinados interesados con los atributos del conjunto de datos publicado al combinar los datos coincidentes en la base de datos supuestamente anonimizada con la base de datos correspondiente al censo electoral.

Finalmente, la *inferencia* es la posibilidad de deducir, con probabilidad

[594] La Disposición Adicional 17.2, f)),1º de la LOPD de 2018 a evaluar específicamente los riesgos de reidentificación vinculados a la anonimización o seudonimización de los datos cuando se pretenda lleva a cabo un tratamiento con fines de investigación en salud pública y, en particular, biomédica. Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.
[595] Grupo del Artículo 29. Dictamen 05/2014 sobre técnicas de anonimización. Adoptado el 10 de abril de 2014, pág 12.

significativa, el valor de un atributo a partir de los valores de un conjunto de otros atributos[596]. Por ejemplo, si una persona sabe que otra persona está contenida en una base de datos y conoce alguno o varios atributos que la definen, es probable que pueda extraer atributos desconocidos de esa persona al acceder a esa base de datos sin que se conozca la identidad de la persona. En la tabla situada a la izquierda, si una persona sabe que otra persona está contenida en esa base de datos y que tiene un Renault Clio, también sabrá que esa persona obtiene un salario de 1000€.

Atributos		
Sexo	Nivel de retribución	Modelo de coche
Hombre	1000 €	Renault Clio
Mujer	1200 €	Renault Megane
Mujer	2850 €	Seat Ibiza
Hombre	1700 €	Volkswagen Golf

Pues bien, estos riesgos pueden verse acrecentados o reducidos en función de toda una serie de factores. En función de los mismos, se habrán de establecer unas u otras medidas que aseguren de forma más sólida el proceso de anonimización. Entre ellos podemos destacar los siguientes:

En primer lugar, resulta fundamental que quede clara la finalidad o el objeto del proceso de anonimización ya que influirá radicalmente en el tipo de medidas que posteriormente se deberán implantar. Así, no es lo mismo que los datos se anonimicen con el objetivo de que una tercera organización pueda consultarlos a que esos mismos datos anonimizados se publiquen en abierto. En el primer caso, el riesgo de reidentificación radica en que esa tercera organización pueda ir obteniendo cada vez más información a raíz de las distintas consultas que realiza a la base de datos supuestamente anonimizada[597]. En el segundo supuesto, la probabilidad de reidentificación aumenta en la medida que cualquier organización o tercero podría combinar esos datos anonimizados con otras bases de datos públicas o no para identificar dichos datos[598].

[596] Hay que distinguir la inferencia como posible riesgo de re-identificación de una base de datos con las inferencias o datos inferidos que se pueden derivar del análisis masivos de los datos cuando se utilizan técnicas de *machine learning* y minería de datos. En: GIL GONZÁLEZ,E : *Big data, privacidad y protección de datos*, op,cit.,, pág.100.

[597] Grupo del Artículo 29. Dictamen 05/2014 sobre técnicas de anonimización. Adoptado el 10 de abril de 2014, pág.27

[598] Information Commissioner´s Office. *Anonymisation: managing data protection risk code of practice*.2012, pág.19

En segundo lugar, también resulta relevante valorar el tipo de atacante o adversario que hipotéticamente puede proceder a la reidentificación. Por atacante o adversario hay que entender cualquier persona u organización que de forma intencionada pretende reidentificar una base de datos anonimizada[599]. Ahora bien, es necesario realizar una contextualización adecuada del potencial adversario en relación con el proceso de anonimización en cuestión. En este sentido, el atractivo de los datos que se anonimizan, el nivel de conocimientos técnicos o informáticos del atacante, el coste de los procesos de reidentificación, la motivación por reidentificar o el entorno en el que se facilitan los datos anonimizados pueden ayudar a perfilar a ese potencial adversario y diseñar técnicas de anonimización más o menos robustas[600]. Así, el TJUE en el *asunto C-582/14 Breyer* ha señalado que, cuando la reidentificación implique un esfuerzo desmesurado en cuanto a tiempo, costes y recursos humanos para ese potencial atacante, hay que entender que el riesgo de reidentificación es insignificante[601]. No resultaría adecuado fijar el umbral de potencial adversario en una persona que por ejemplo es un ingeniero informático especializado en sabotear procesos de anonimización cuando una base de datos anonimizada de un hospital es facilitada a un grupo de investigadores para tratar de desarrollar un modelo que tiene como objetivo prevenir una determinada enfermedad. Los posibles atacantes a esa base de datos, esto es, los investigadores, no ostentan a priori el conocimiento técnico suficiente ni la motivación para tratar de identificar a las personas que forman parte de esa base de datos. Tal y como señala la doctrina, desarrollar técnicas de anonimización pensadas contra todos y cada uno de los posibles adversarios en cualquier momento que no sean necesariamente los destinatarios previstos de los datos resulta problemático y poco realista ya que obligaría al titular de esa base de datos a hacer siempre las peores suposiciones posibles, incluso si no son relevantes para el contexto específico[602].

[599] Tal y como indica el considerando 26 del RGPD *para determinar si una persona física es identificable, deben tenerse en cuenta todos los medios, como la singularización, que razonablemente pueda utilizar el responsable del tratamiento o cualquier otra persona para identificar directa o indirectamente a la persona física*. Es por ello, que un adversario puede ser tanto el responsable como cualquier otra persona.

[600] El valor comercial de la información anonimizada para un adversario puede aumentar las posibilidades de re identificación. En: Agencia Española de Protección de Datos. Orientaciones sobre protección de datos en la reutilización de la información del sector público.2016, pág.8.

[601] SENTENCIA DEL TRIBUNAL DE JUSTICIA (Sala Segunda) de 19 de octubre de 2016, asunto C-582/14, caso Breyer. Véase los FJ 42 a 46, en especial el 46. Resolución disponible: https://curia.europa.eu/juris/document/document.jsf?text=&docid=184668&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=17198716

[602] EL EMAM,K; ÁLVAREZ RIGAUDIAS,C: "A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques", *International Data Privacy Law*, op.cit., pág.83. Así lo ha

Distinto sería el supuesto si esos mismos datos sanitarios anonimizados del hospital se publicaran en abierto, en este caso, los potenciales adversarios pueden estar realmente interesados en conocer la identidad de esas personas a efectos de por ejemplo ofrecer determinados seguros médicos personalizados. Para este caso, las técnicas de anonimización deberían ser más robustas ya que el potencial adversario para su reidentificación es probable que, ostente un fuerte interés por esos datos identificados y además cuente con los medios técnicos e informáticos para reidentificarlos. Esta interpretación del potencial adversario resulta en nuestra opinión adecuada ya que permite afrontar procesos de anonimización coherentes a los riesgos probables de reidentificación y se asimila a la lógica hasta ahora descrita para afrontar estos problemas.

En tercer lugar, los riesgos pueden acrecentarse en muchos casos por una inadecuada implantación de los propios procedimientos de anonimización ya sea porque aquellos que los implementan carezcan de la suficiente formación o porque las técnicas establecidas no consiguen reducir adecuadamente ese riesgo de reidentificación[603].

C) Técnicas de anonimización

Una vez que hayan sido evaluados y valorados los riesgos, el siguiente paso será establecer aquellas técnicas y medias de anonimización adecuadas que reduzcan al máximo la probabilidad de reidentificación de la base de datos que se pretende anonimizar. De acuerdo al GT29, en términos generales existen dos enfoques diferentes de la anonimización, esto es, por un lado la *aleatorización*, formada por toda una serie de técnicas que modifican la veracidad de los datos con el fin de eliminar el estrecho vínculo existente entre los mismos y la persona y, por otro lado, la *generalización*, cuyo principal objetivo es generalizar o diluir los atributos de los interesados modificando las

señalado también Ricard Martínez. De manera que, si se toma como referencia a la hora de valorar a un potencial adversario al premio noble de matemáticas o un ingeniero informático para todo tipo de procesos de anonimización , jamás o muy difícilmente se podrá llegar a alcanzar la plena anonimización de la mayoría de las bases de datos que se pretendan anonimizar. En: VII Congreso del avance del Gobierno Abierto. Mesa 3. Límites al acceso y convergencia normativa. 21 de Mayo de 2020. Ponencia disponible en: (véase a partir de del minuto 15:25).
https://www.youtube.com/watch?v=IIalxHImBBU&list=PLMqWVbmUv8TiAWnWEmQ-5UFg0dCH48gCK&index=6

[603] Hasta no hace mucho se pensaba que la seudonimización de datos era un método de anonimización lo que suponía que en muchos casos las bases de datos que supuestamente estaban anonimizadas realmente no lo eran, con los perjuicios que ello podría comportar.

respectivas escalas u órdenes de magnitud[604]. En las siguientes líneas analizaremos algunas de estas técnicas y las principales implicaciones en el desarrollo de modelos algorítmicos que posteriormente se utilizarán para la toma de decisiones automatizadas.

c.1) Aleatorización

Como se ha indicado anteriormente, a través de la aleatorización se modifican los valores reales con el objetivo de impedir la vinculación de los datos anonimizados con los valores originales. Las técnicas de aleatorización que vamos a subrayar son la adición de ruido y la permutación.

En primer lugar, la *adición de ruido* consiste en modificar los datos de forma que sean menos precisos pero, manteniendo la distribución general de los datos. Por ejemplo, si queremos añadir ruido a los valores del atributo peso, podríamos alterar esos valores varios kilos, por ejemplo, entre ± 1 y 3 kilogramos.

Atributos		
Sexo	**Salario**	**Peso**
V	1000 €	74,5
M	1200 €	65,6
M	1850 €	87,4
V	1700 €	93,6

Atributos		
Sexo	**Salario**	**Peso con adición de ruido**
V	1000 €	73,5
M	1200 €	64,6
M	1850 €	86,4
V	1700 €	92,6

La adición de ruido requiere a su vez de otra serie de técnicas para que completen el proceso de anonimización. La utilidad de esta técnica a la hora de diseñar modelos algorítmicos variará en función del tipo de variables a los que se le aplique el ruido y en su caso la cantidad que se adhiera.

Una segunda técnica que encontramos es la *permutación*, esta técnica consiste en mezclar los valores de los atributos en una tabla para que algunos de ellos puedan vincularse artificialmente a distintos interesados. Por ejemplo, en la tabla siguiente se permutan los valores correspondientes al atributo salario, es decir, de forma aleatoria se cambian los valores de cada una de las muestras.

[604] Grupo del Artículo 29. Dictamen 05/2014 sobre técnicas de anonimización. Adoptado el 10 de abril de 2014,pág.27.

Atributos		
Sexo	**Salario**	**Marca**
V	1000 €	Renault
M	1200 €	Renault
M	1850 €	Mercedes
V	1700 €	Volkswagen

Atributos		
Sexo	**Salario Permutado**	**Marca**
V	1700 €	Renault
M	1850 €	Renault
M	1100 €	Mercedes
V	1200 €	Volkswagen

Al igual que ocurre con la adición de ruido, la permutación tampoco puede aplicarse aisladamente para conseguir una anonimización robusta. Esta técnica puede afectar gravemente al desarrollo de modelos o sistemas algorítmicos ya que resultará frecuente que las correlaciones detectadas entre los distintos atributos pueden resultar erróneas al alterarse en cierta medida los valores en cada uno de las muestras que se utilizan. Por ejemplo, tomando como referencia la tabla anterior, imaginemos que en los datos originales existe una correlación muy alta entre los atributos sexo y salario. Si se altera aleatoriamente el atributo salario para anonimizar la base de datos, es posible que al alterarlo, las posibles correlaciones que pudieran existir entre los atributos sexo y salario desaparezcan y por tanto, el modelo aprenda unas correlaciones erróneas y no representativas de la realidad que reflejaban los datos iniciales no anonimizados.

c.2) Generalización

La generalización es la segunda familia de técnicas de anonimización. Este enfoque generaliza o diluye los atributos de los interesados modificando las respectivas escalas u órdenes de magnitud, por ejemplo, sustituyendo una ciudad por una región[605].

Concretamente, las *técnicas de agregación y anonimato k* tienen el objetivo de impedir que un interesado sea singularizado cuando se le agrupa junto con, al menos, un número k de personas. Para lograrlo, los valores de los atributos se generalizan hasta el punto de que todas las personas acaban compartiendo el mismo valor[606]. Es decir, al menos deberían existir dos miembros que presenten los mismos valores en la base de datos, lo que supone que ante un potencial atacante o adversario, exista un 50 % de

[605] Grupo del Artículo 29. Dictamen 05/2014 sobre técnicas de anonimización. Adoptado el 10 de abril de 2014, pág.18

[606] Agencia española de protección de datos. La k-anonimidad como medida de la privacidad Anonimización. 2019, pág.3.

probabilidad de que se pueda volver a identificar a esa persona, es decir, ya no se puede reidentificar a esa persona al 100% en un primer ataque. Si pese a la generalización de los valores, aún quedan algunos que pueden ser individualizados, lo ideal es que se eliminen o se aumente aún más la generalización.

Cliente	**Atributos**	
	Sexo	**Nacimiento**
1	V	Abril 1957
2	M	Junio 1959
3	M	Marzo 1960
4	V	Enero 1958
5	V	Agosto 1985

Cliente	**Atributos**	
	Sexo	**Nacimiento generalizado**
1	V	55-65
2	M	55-65
3	M	55-65
4	V	55-65
~~5~~	~~V~~	~~35-45~~

Como vemos, en este supuesto se han generalizado todos los atributos correspondientes a los distintos clientes que forman parte de la base de datos. Así, a excepción del cliente 5, el resto de clientes han sido generalizados adecuadamente y en principio sus datos han quedado anonimizados ya que al menos, existen dos clientes que presentan los mismos valores en los distintos atributos. Sin embargo, para el cliente 5, dado que todavía es posible su reidentificación, muy probablemente, la única vía que quedará será la eliminación de ese registro ya que a través de la generalización no se puede conseguir la plena anonimización de ese sujeto. La otra opción sería aumentar las horquillas para generalizar aún más, por ejemplo, rangos de hasta 20 años, pero claro, la precisión y utilidad de las bases de datos podría reducirse. Este ejemplo también pone de manifiesto que en la práctica, la generalización afecta a la utilidad de los datos[607]. Es precisamente uno de los principales inconvenientes de la generalización de datos, de esta manera, como decíamos en páginas anteriores, una base de datos que pretende utilizarse para su análisis y creación de modelos debería ser lo más representativa posible de la realidad que pretende modelar, debiendo reflejar también los casos más aislados o anómalos ya que puede reflejar un grupo de personas infra representado que ha de estar presente en esa modelización. Tanto la eliminación como

[607] Information Commissioner´s Office. *Anonymisation: managing data protection risk code of practice*.2012, págs. 86 y 87.

la generalización en la que se basan estas técnicas irían en contra precisamente de esa necesidad de tratar de modelar de la manera más adecuada esa realidad.

D) Los procesos de anonimización en el desarrollo de modelos algorítmicos. La no aplicación de la normativa de protección de datos

Una vez que los datos son anonimizados, la normativa de protección de datos deja de ser aplicable. Los usos de estos datos anonimizados para diseñar modelos algorítmicos que posteriormente se utilizarán para adoptar decisiones sobre individuos presentan diversos escenarios que conviene representar[608].

En primer lugar, resulta habitual que la misma organización que ostenta los datos personales los anonimicen y a partir de ahí aplique las técnicas de *machine learning* y minería de datos para poder crear modelos y obtener inferencias. Es sin duda el modo de proceder de multitud de organizaciones ya que les permite adquirir conocimiento oculto en sus propias bases de datos y aplicar ese conocimiento posteriormente al desarrollo de sus propios modelos sobre los cuales la misma organización adoptará decisiones automatizadas. En estos supuestos por tanto, la analítica de los datos anonimizados se realiza en la esfera propia de las mismas organizaciones que recopilaron los datos[609], las cuales, suelen añadir datos agregados de otras fuentes internas o externas.

Una segunda opción es la que las organizaciones anonimicen los datos que ostentan y posteriormente los faciliten a terceras organizaciones. La forma en que se suministren esos datos varía enormemente en función del fin que pretende la primera organización sobre el uso de dichas bases de datos. Así, dentro de esta modalidad podemos encontrarnos con aquellos aquellas organizaciones que comercializan sus bases de datos anonimizados para que terceras organizaciones puedan en su caso adquirirlas. De manera que estas últimas luego las utilizan para desarrollar sus propios modelos algorítmicos. Como indicábamos al principio de este trabajo, muchas organizaciones desarrollan todo tipo de aplicaciones y productos donde el principal valor de las mismas, no es el servicio que ofrecen, sino más bien los datos que se

[608] Las operatorias descritas en las siguientes líneas son el reflejo de algunos métodos a través de los cuales las organizaciones hacen uso de los datos anonimizados para desarrollar modelos algorítmicos.

[609] Según un informe publicado por el Parlamento Europeo, la gran mayoría de los datos son generados y analizados internamente por la empresa o por un subcontratista de esta. Por lo tanto, en general, los datos tienden a permanecer dentro de una organización y no se comercializan con terceros. En: BALCKMAN,C; FORGE,S: "Data Flows- Future Scenarios". Directorate general for internal policies policy department a: economic and scientific policy. Parlamento Europeo,201 ,págs 12 y 13.

recopilan. En muchas ocasiones el producto sí que es relevante y los datos que se obtienen son necesarios para que el mismo pueda funcionar pero esos mismos datos pueden resultar sumamente útiles para terceras organizaciones y por tanto, resulta rentable su venta a terceros[610].

Una tercera vía es que los datos anonimizados se faciliten a través de un intermediario de confianza. En este caso, la organización que pretende hacer uso de esos datos accede a ese entorno seguro que le facilita la organización que posee los datos anonimizados. Generalmente, ese entorno es habilitado por un intermediario, el cual, puede anonimizar los datos o en su caso facilitar su acceso a esa tercera organización para que realice el análisis de los mismos[611]. Esta operatoria es parecida a la que pretende implantar la Unión Europea para fomentar la reutilización de datos que ostentan las Administraciones Públicas en favor de terceras organizaciones[612]. Ahora bien, ese entorno seguro también podría utilizarse a la inversa, es decir, una empresa privada permite que las Administraciones puedan analizar los datos anonimizados que ostentan y desarrollar sus propios modelos que posteriormente servirán para adoptar decisiones en el seno de estos organismos públicos[613].

Por último, aunque todavía con un margen importante de recorrido, cada vez es más frecuente que las organizaciones publiquen datos anonimizados en abierto. El objetivo esencial es que la sociedad pueda aprovecharse del potencial que puede estar escondido tras las grandes masas de datos que ostenta el sector público. De esta manera, una vez que se publican esos datos las organizaciones puede desarrollar sus modelos

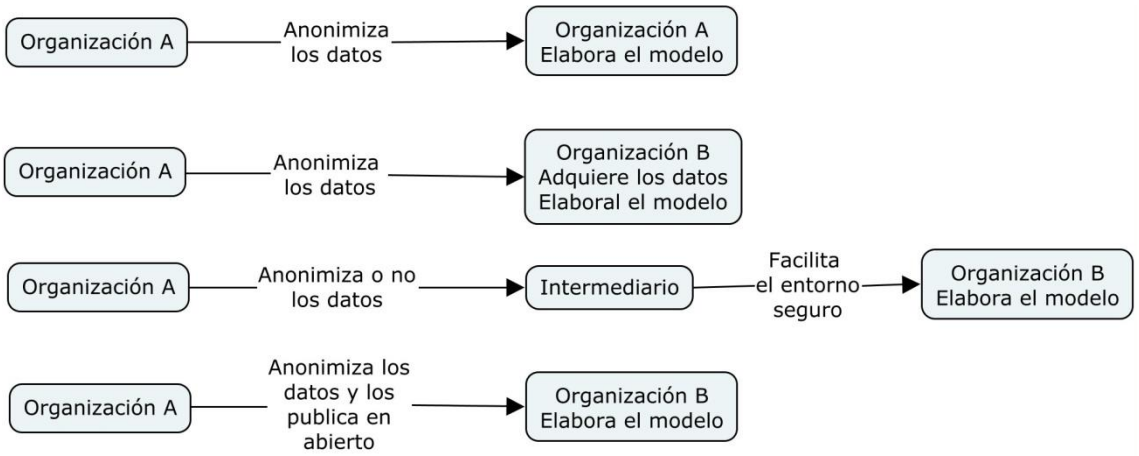
[610] Es el caso de los robots aspiradoras. Este tipo de sistemas recopilan datos del interior de las viviendas que limpian. Dichos datos son posteriormente vendidos a terceros con fines de mercadotecnia. Fuente de la noticia: JONES,R: "Roomba's Next Big Step Is Selling Maps of Your Home to the Highest Bidder", *Gizmodo*:24/7/2017. Disponible en: **https://gizmodo.com/roombas-next-big-step-is-selling-maps-of-your-home-to-t-1797187829**

[611] En el ámbito de la investigación sanitaria a través de la analítica de datos se han desarrollado ya algunas guías y directrices que pueden facilitar el tratamiento de estos datos. En: Health Data: el Playbook, realizado por Fundación 29. Disponible en: https://www.healthdata29.org/playbook

[612] Artículo 5 de la Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos). Resolución adoptada el 25 de noviembre de 2020.

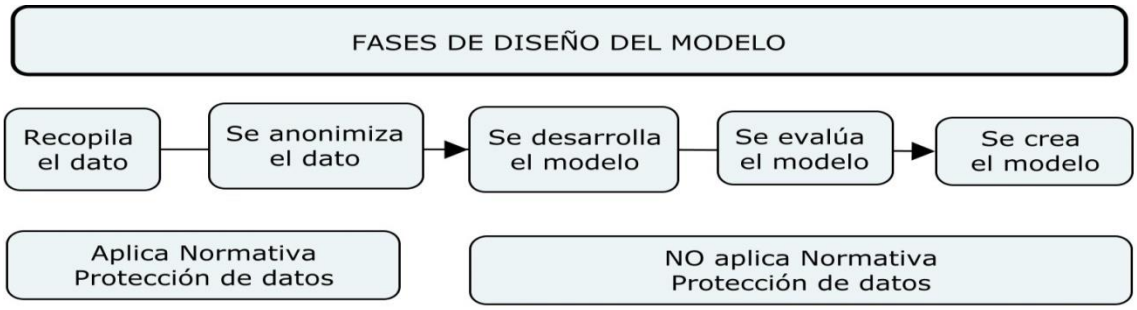
[613] Por ejemplo los datos de movilidad que recopilan los operadores de telefonía pueden resultar de gran utilidad para las Administraciones Públicas a la hora de adoptar todo tipo de políticas relacionadas con sus competencias. Congreso "Regulación y explotación de big data para los servicios públicos". Jornada "La explotación de datos para la inspección y control de la ejecución de políticas públicas. Inteligencia Artificial, políticas urbanas y sostenibilidad". Martes 9 de marzo de 2021. Título de la ponencia: Datos dinámicos y agregados para mejorar la visión de las Administraciones Públicas en materia de turismo y gestión urbana. Autor: Juan Murillo Arias.
Ponencia disponible en YouTube. https://www.youtube.com/watch?v=sGExSLP6ZmU

algorítmicos junto con otros datos que ya ostenten o estén disponibles públicamente[614]. Como ya indicamos al principio de este estudio, en el ámbito de la Unión Europea existe un interés aun poco explotado de fomentar el uso de datos abiertos que ostenten las Administraciones Públicas en favor de todo tipo de organizaciones[615].



E) Garantías jurídicas que ofrece el derecho fundamental a la protección de datos previas al proceso de anonimización

Gran parte del proceso de diseño de los sistemas y modelos algorítmicos que posteriormente se utilizarán para adoptar decisiones queda al margen del derecho a la protección de datos personales, ya que, una vez que los datos son anonimizados, el ámbito de aplicación de esta normativa no es aplicable. Dicho lo anterior, sigue aun quedando una parte de este proceso donde las normas de protección de datos continúan irradiando sus efectos, es concretamente el periplo que va desde el momento que se recopilan los datos personales hasta que estos se anonimizan de forma completa.



Proceso habitual de anonimización de datos durante el desarrollo de modelos algorítmicos.

[614]Agencia Española de Protección de Datos. *Orientaciones sobre protección de datos en la reutilización de la información del sector público*, 2016, pág.7

[615] Véase 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.

En primer lugar, por lo que se refiere a la recopilación de los datos, esta ha de realizarse cumpliendo el principio de licitud[616]. Es decir, el responsable ha de contar con alguna de las bases legítimas que le habilitan para recopilar esos datos personales que en un primer momento y por regla general fueron recopilados para una finalidad primaria distinta a la anonimización. En estos supuestos, resulta recomendable que, si las organizaciones ya son conscientes de que dichos datos pueden sufrir un proceso de anonimización desde la recogida, ello debe quedar expresamente reflejado. Por otro lado, en cumplimiento del principio de minimización de datos[617], los datos que se recopilen inicialmente han de ser estrictamente necesarios para la finalidad inicial para la que se recopilaron. De manera que, si se recopilan datos con visos únicamente de ser obtenidos para posteriormente ser anonimizados, ello se ha de especificar, ya que si no habrá que entender que la recopilación de los mismos incumple este principio.

En segundo lugar, una vez que se han recopilado esos datos, normalmente para una finalidad distinta de la anonimización de los mismos, las organizaciones deberían valorar la compatibilidad entre la finalidad inicial para la que se recopilaron inicialmente los datos y la secundaria, en este caso la anonimización, todo ello, en cumplimiento del principio de limitación de la finalidad[618]. Tenemos que partir de que la anonimización de los datos es una tratamiento de datos personales específico en sí mismo[619], por lo tanto, se hace necesario valorar si la compatibilidad inicial y el proceso posterior de anonimización es compatible. En este sentido, existe un criterio bastante unificado en considerar dicha compatibilidad. Así lo ha considerado el grupo del artículo 29 que establece tal compatibilidad siempre que el proceso de anonimización genere fiablemente información anonimizada[620], esto es, que se adopten todas las medias adecuadas para evitar la reidentificación[621]. En estos supuestos, el responsable del tratamiento que procederá a la anonimización de datos no requerirá de la necesidad

[616] Artículo 5.1.a) y 6 del RGPD. Véase también el Capítulo IV, apartado I de esta tesis.
[617] Conforme al artículo 5.1.c), los datos personales serán adecuados, pertinentes y limitados a lo necesario en relación con los fines para los que son tratados.
[618] Este principio se analizará más detenidamente en los siguientes capítulos. Véase los Artículos 5.b) y 6.4 del RGPD.
[619] Grupo del Artículo 29. Dictamen 05/2014 sobre técnicas de anonimización. Adoptado el 10 de abril de 2014, pág.7
[620] Grupo del Artículo 29. Dictamen 05/2014 sobre técnicas de anonimización. Adoptado el 10 de abril de 2014, pág.8.
[621] Tal y como señala la AEPD, *el uso de métodos apropiados para la anonimización de datos personales en combinación con un riguroso análisis de riesgos son una potente herramienta para proteger la privacidad de la información y garantizar los principios de protección de los datos que van a ser utilizados para una finalidad distinta a la que inicialmente fueron recabados*. Agencia Española de Protección de Datos. *Orientaciones y garantías en los procedimientos de anonimización de datos personales*. Adoptadas en 2016, pág.8.

del consentimiento[622]. Entre los elementos que establece el artículo 6.4 del RGPD para valorar la compatibilidad entre el fin inicial y el posterior es el referido *a las posibles consecuencias para los interesados del tratamiento ulterior previsto*[623]. Esto último es muy importante porque hasta ahora, esencialmente, a la hora de valorar la compatibilidad o no de un proceso de anonimización se ha puesto el acento en los riesgos que se pueden derivar de la propia reidentificación, esto es, un tercero descubre la identidad de una persona de la base de datos supuestamente anonimizada. Sin embargo, con la entrada en juego del *big data* sobre datos anonimizados, los riesgos ya no sólo se derivan de la posible reidentificación sino también de los usos posteriores que se pueden hacer de los modelos una vez que estos se crean basados en esos datos. Por tanto, una interpretación extensiva del principio de limitación de la finalidad en relación con los usos posteriores podría establecer que el responsable del tratamiento cuando tenga previsto anonimizar los datos valore las posibles consecuencias negativas que se pueden derivar ya no sólo de la anonimización en sí a efectos de reidentificación, sino también de las consecuencia negativas que se pueden derivar del uso de esos datos para genera modelos que posteriormente afecten de forma negativa a las personas. Es decir, lo que proponemos es extender el juicio de compatibilidad entre la finalidad inicial y la ulterior a la valoración de los riegos que se pueden derivar una vez que esos datos anonimizados sean facilitados a terceros o en su caso se utilicen por ese mismo responsable que ha alterado la finalidad inicial de los datos. Como es lógico, ese juicio de compatibilidad basado en los riesgos en algunas ocasiones resultará más complicado, piénsese por ejemplo cuando los datos son publicados en abierto, es difícil que la organización que los publica pueda prever los posibles usos mal intencionados que pueden tener terceras organizaciones a la hora de usar esos datos para diseñar modelos algorítmicos. Sin embargo, existen multitud de supuestos en los que sí podría realizarse ese juicio de compatibilidad más adecuado como puede ser el caso en el que una organización con o sin intermediario facilita temporalmente el acceso a una base de datos anonimizada a terceras organizaciones. Aquí, se podría exigir por ejemplo la imposición de condiciones contractuales relacionadas con el uso de los datos cuando

[622] Tal y como señala la ICO, obtener el consentimiento para la anonimización de datos personales puede ser logísticamente muy oneroso cuando hay un número muy importante de personas en los registros que se pretenden anonimizar. Incluso podría llegar a ser imposible la solicitud de ese consentimiento cuando por ejemplo, los datos personales son antiguos y no existe un medio fiable de contacto con los titulares de esos datos. Information Commissioner´s Office. *Anonymisation: managing data protection risk code of practice*.2012, págs. 28 y 29.

[623] Artículo 6.4.c) del RGPD.

estos se cedan a un tercero[624]. Algo parecido ocurriría en aquellos supuestos en los que la misma organización que recopiló los datos, ahora pretende anonimizarlos para llevar a cabo tratamientos de datos masivos con el objetivo de diseñar modelos algorítmicos. A la hora de valorar ese juicio de compatibilidad, dichos propósitos no deberían tener como objetivo crear modelos que posteriormente arrojen decisiones discriminatorias, injustas, etc[625]. Dicho lo anterior, somos conscientes no obstante que esta interpretación que proponemos que trata de extender las obligaciones en materia de protección de datos sobre los tratamientos posteriores que pueden sufrir los datos ya anonimizados y por tanto no personales puede entenderse que excede del ámbito de aplicación de la normativa de protección de datos. El principio de limitación de la finalidad en relación con el análisis masivo de datos se estudiará más detenidamente en el capítulo siguiente[626].

En tercer lugar, otra garantía que se deriva del derecho a la protección de datos estrechamente relacionado con el anterior es que todo el proceso que abarca la anonimización de datos desde que se inicia hasta que se entiende completada ha de realizarse de la forma más adecuada posible y con el objetivo de evitar la reidentificación. Será necesario por tanto realizar el mentado análisis de riesgos. Corresponde al responsable del tratamiento que pretende la anonimización prever todas esas medidas, resultando recomendable la reevaluación periódica del riesgo residual existente con el fin de introducir parámetros de mejora de la calidad del proceso de anonimización[627]. Recordemos además que la LOPD de 2018 considera como infracción muy grave la reversión deliberada de un procedimiento de anonimización a fin de permitir la reidentificación de los afectados[628]. Por tanto, las autoridades de protección de datos tienen competencia para actuar sobre este ámbito y chequear los procesos de anonimización.

En definitiva, más allá de estas garantías, el derecho fundamental a la protección de datos pierde su irradiación durante las siguientes fases que comprenden el diseño de

[624] Agencia Española de Protección de Datos. *Orientaciones y garantías en los procedimientos de anonimización de datos personales*. Adoptadas en 2016, pág.11.
[625] EL EMAM,K; ÁLVAREZ RIGAUDIAS,C: "A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques", *International Data Privacy Law*, op.cit., pág.81.
[626] Capítulo IV, aparatado II de esta tesis.
[627] Agencia Española de Protección de Datos. *Orientaciones y garantías en los procedimientos de anonimización de datos personales*. Adoptadas en 2016, pág.8.
[628] Artículo 72.1.p) Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

sistemas algorítmicos basados en datos anonimizados. En estas fases son desarrollados los modelos y se elaboran por ejemplos los perfiles grupales. Perfiles que en una fase posterior ya sí se aplicarán a personas y colectivos más o menos definidos que presentan los mismos o parecidos atributos y donde la normativa de protección de datos volverá a irradiar sus efectos.

F) Otras garantías jurídicas ligadas a evitar la reidentificación de los datos anonimizados

Junto a las exigencias derivadas del derecho a la protección de datos, se han planteado otra serie de garantías centradas en evitar precisamente la reidentificación de los datos personales. Todas ellas se centran esencialmente en imponer todo tipo de obligaciones y compromisos a las organizaciones que utilizan los datos anonimizados. Así, todo aquel que trate datos anonimizados ha de ser consciente del riesgo residual de reidentificación de dichos datos y las medidas para evitarlo[629]. También resulta necesario realizar acuerdos vinculantes donde los destinatarios de los datos anonimizados se comprometan a la no reidentificación de los datos y a la confidencialidad de los mismos donde por ejemplo, la organización destinataria se comprometa públicamente a no volver a identificar los datos[630]. En este sentido, hay que tener en cuenta que a través de las técnicas de IA se pueden rastrear el origen de los datos y reidentificar las bases de datos anonimizadas[631]. Por tanto, esos acuerdos de no reidentificación resultan esenciales. Por otro lado, se puede imponer también la exigencia de realizar auditorías para chequear los posibles riesgos de reidentificación[632]. Así, en el ámbito del sector público, una herramienta ideal para prever condiciones específicas a la hora de reutilizar información procedente de las Administraciones

[629] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.14. Véase también: Consejo de Europa. *Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data*, 2017, apartado 6.4, pág.5.

[630] TENE, O; POLONETSKY,J: "Big Data for All: Privacy and User Control in the Age of Analytics". *Northwestern Journal of Technology and Intellectual Property*. Volume 11 | Issue 5, 2013, pág.259. Así también lo contempla el artículo 8.f) de la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público.

[631] Comisión Europea. *Libro Blanco sobre la inteligencia artificial. Un enfoque europeo orientado a la excelencia y la confianza*, Bruselas, 19.2.2020,pág.14.

[632] Agencia Española de Protección de Datos. Orientaciones sobre protección de datos en la reutilización de la información del sector público.2016, pág.7.

Públicas es la concesión de licencias específicas donde se especifiquen claramente los compromisos jurídicos del uso de dichos datos[633].

G) Garantías jurídicas ligadas al diseño de modelos algorítmicos con datos anonimizados

.

Es turno de analizar algunas de las garantías que se han propuesto por parte de la doctrina y las autoridades cuando se elaboran modelos algorítmicos basados en datos anonimizados. Algunas salvaguardas se han centrado en focalizar la atención en las principales implicaciones éticas que se pueden generar durante la fase de desarrollo de estos sistemas, implicaciones que posteriormente tendrán su relevancia cuando estos sistemas adopten decisiones. En este sentido, se ha indicado la necesidad de realizar evaluaciones de impacto centradas en los derechos humanos[634], es decir, valorar las implicaciones legales y ética de los proyectos que se pretende llevar a cabo y sus implicaciones en la esfera de los derechos de los particulares, así como en su caso de los grupos que pueden resultar más afectados[635]. Estas implicaciones éticas también podrán ser valoradas por grupos de expertos internos o externos de las organizaciones que utilizan datos anonimizados cuando desarrollan estos sistemas[636]. En este sentido, los Comités de ética en el ámbito de la investigación sanitaria son un ejemplo adecuado de órgano colegiados que precisamente se encarga de velar por los riesgos asociados a las metodologías y estudios que se llevan a cabo en una determinada investigación científicas[637]. Estos mismo comités, de acuerdo a las características propias de cada organización, podrían desarrollarse en el resto de organizaciones que pretenda desarrollar sistemas algorítmicos donde únicamente se utilicen datos anonimizados. Cabe señalar que la LOPD de 2018 ha previsto la necesidad de que estos Comités de

[633] Así, el artículo 9.2 de Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. establece que; "en los casos en los que se otorgue una licencia, esta deberá reflejar, al menos, la información relativa a la finalidad concreta para la que se concede la reutilización, indicando igualmente si la misma podrá ser comercial o no comercial, para la que se concede la reutilización, la duración de la licencia, las obligaciones del beneficiario y del organismo concedente, las responsabilidades de uso y modalidades financieras, indicándose el carácter gratuito o, en su caso, la tarifa aplicable".

[634] MANTELERO,A: "AI and big data: a blueprint for a human rights, social and ethical impact assessment", Computer Law & Security Review, Volume 34, Issue 4, August 2018, pág.764. Texto disponible en: https://doi.org/10.1016/j.clsr.2018.05.017

[635] Comisión Europea. *Ethics and data protection*. Texto aprobado el 14 de noviembre de 2018, pág.7.

[636] Por ejemplo a través de los Comités de ética. En: EL EMAM,K; ÁLVAREZ RIGAUDIAS,C: "A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques", *International Data Privacy Law*, op.cit., pág.78.

[637] Artículo 12 de la Ley 14/2007, de 3 de julio, de Investigación biomédica.

ética emitan un informe previo antes de que se lleve a cabo un proyecto de investigación científica donde puede entrar en juego el desarrollo de sistemas algorítmicos en los que se utilicen datos personales o seudonimizados[638].

2 La seudonimización

El RGPD define la seudonimización en su artículo 4.5 como *el tratamiento de datos personales de manera tal que ya no puedan atribuirse a un interesado sin utilizar información adicional, siempre que dicha información adicional figure por separado y esté sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos personales no se atribuyan a una persona física identificada o identificable*. La seudonimización implica la sustitución de la información personal identificable, como puede ser el nombre del interesado, por un identificador único que no está relacionado con su identidad, utilizando para ello diversas técnicas como la codificación o el *hashing*.

A) La seudonimización en el RGPD

La seudonimización se configura en el RGPD a través de tres vertientes:

En primer lugar, como una medida de seguridad que han de implantar los responsables del tratamiento con el objetivo de reducir los riesgos de identificación de los interesados (Artículo 32 RGPD). Así, por ejemplo, en la medida de lo posible, tanto el encargado como el responsable deberían seudonimizar los datos que están presentes tanto en las bases de datos como aquellos que se ha preparado para el proceso de aprendizaje de los modelos algorítmicos[639].

En segundo lugar, como una medida de garantía esencial que han de implementar los responsables del tratamiento cuando pretendan utilizar los datos con fines de archivo en interés público, fines de investigación científica o histórica o fines

[638] Disposición adicional decimoséptima apartado 2 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

[639] Las técnicas de seudonimización se han de implantar tanto en los datos de entrenamiento como en los datos que puedan estar contenidos en el modelo. Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.40. A modo de ejemplo, la doctrina ha señalado que la seudonimización de los datos resulta obligatoria en el contexto de los proyectos de Smart City. En: VELASCO RICO, C: "La ciudad inteligente: entre la transparencia y el control". *Revista General de Derecho Administrativo* 50, 2019, pág. 13.

estadísticos. (Artículo 89.1. RGPD). Este tipo de tratamientos se fomentan en el RGPD con la exención de algunas exigencias normativas, a cambio, el responsable debe desplegar toda una serie de garantías, entre estas[640], destaca la seudonimización. Así, por ejemplo, la LOPD de 2018 considera lícito el uso de datos personales seudonimizados con fines de investigación en salud y, en particular, biomédica[641]. Por tanto, la exención del consentimiento para estos tratamientos se compensa entre otras garantías con el uso de datos personales seudonimizados. También la PRAI ha considerado la seudonimización como una medida esencial para llevar a cabo los tratamientos de datos que tengan como objetivo detectar y corregir sesgos presentes en los sistemas de IA de alto riesgo[642].

En tercer lugar, la seudonimización también resulta una garantía básica a tener en cuenta a la hora de analizar la compatibilidad entre la finalidad inicial del tratamiento y la ulterior pretendida. (Artículo 6.4.e del RGPD). En estos casos, esta medida junto a otras ayudarán al responsable a justificar la compatibilidad de los fines secundarios de los datos pretendidos[643]. Tal y como ha señalado la AEPD, si no están previstas estas salvaguardas durante el análisis de compatibilidad, el tratamiento posterior pretendido difícilmente se habilitará[644]. Esto puede suceder cuando el responsable no ha valorado la seudonimización de datos o contemplándola, el tratamiento posterior necesariamente requiere la identificación de las personas para que se lleve a cabo[645].

En los tres supuestos se favorece tanto el principio de privacidad desde el diseño como el de minimización de datos tal y como indica el artículo 25 del RGPD[646].

---

[640] Se facilita el tratamiento ulterior de los datos para este tipo de finalidades y se restringen desminados derechos de los interesados. Artículo 89 del RGPD.

[641] Disposición adicional 17.2.d) de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

[642] Artículo 10.5 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

[643] Grupo del artículo 29. *Opinion 03/2013 on purpose limitation*. Texto adoptado el 2 de abril de 2013, págs.26.

[644] La AEPD, con relación a determinados tratamientos realizados por los colegios profesionales, consideró que no era compatible la recogida de datos personales para formar parte de la corporación profesional y tener los derechos y obligaciones que tal condición confiere y la pretensión posterior de tratar con los datos con fines de mercadotecnia. La AEPD señala que deberían haberse previsto medidas como la seudonimización de los datos. Agencia Española de Protección de Datos. Informe N/REF: 0014/2021, pág.14. Texto disponible en: https://www.aepd.es/es/documento/2021-0014.pdf

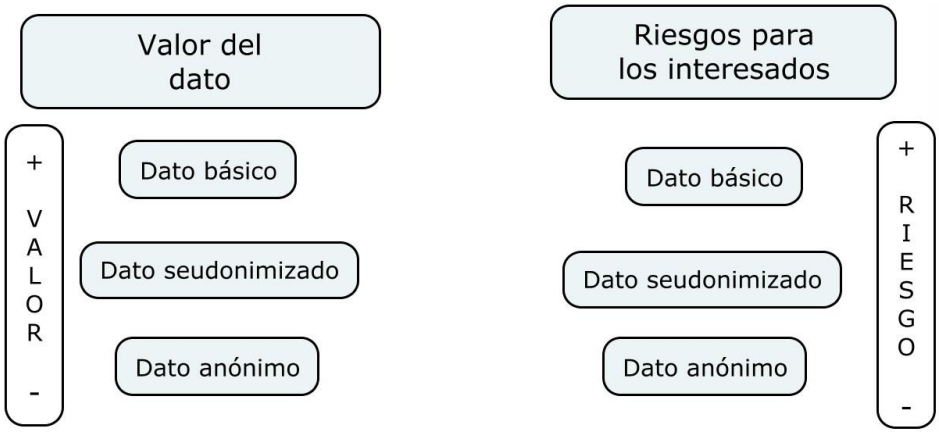
[645] Agencia Española de Protección de datos. Informe N/REF: 0089/2020. Pág.132. Texto disponible en: https://www.aepd.es/es/documento/2020-0089.pdf

[646] Véase en este sentido el estudio de minimización de datos al que hacemos referencia en el Capítulo IV, apartado III, punto 1 de esta tesis.

La configuración de la seudonimización prevista tanto en el artículo 89.1 como el 6.4.e) del RGPD resulta muy adecuada para los entornos de big data y la analítica masiva de datos personales. En este sentido, como ya hemos indicado, estas técnicas algorítmicas permiten la obtención de información oculta presente en las bases de datos. Las ventajas que ofrece la seudonimización son tanto para el responsable como para los interesados.

Por el lado del responsable, el uso de datos seudonimizados mejora la calidad de los resultados que genera la analítica masiva de datos personales sin necesidad de proceder a la anonimización de los mismos. Hay que tener en cuenta que muchos proyectos de analítica masiva de datos requieren del tratamiento de datos personales para que los mismos arrojen aportaciones relevantes[647]. La seudonimización permite mantener el valor presente en los datos personales sin que aquellos que los utilizan puedan identificar a los titulares de los mismos. Es importante recordar que los procesos de anonimización resultan complejos y costosos para los responsables y además, la pérdida de calidad de los datos puede resultar muy grave tras el proceso de anonimización. Además, la obtención del consentimiento para el tratamiento de datos con fines secundarios puede resultar en muchos casos difícil de obtener. La seudonimización suple estas carencias al permitir a las organizaciones poder tratar los datos personales y obtener la información que se infiere de los mismos sin necesidad del consentimiento del interesado, todo ello eso sí, bajo el paraguas normativo del RGPD y las leyes aplicables correspondientes. Junto a las ventajas del responsable, la seudonimización también beneficia a los interesados ya que los datos sometidos a esta medida siguen formando parte del campo de la protección de datos, algo que no ocurre cuando los datos se anonimizan.

[647] Así, tal y como se ha señalado, muchas investigaciones necesitan conocer además de los datos clínicos, ciertos datos demográficos de los pacientes para poder extraer resultados significativos. En: SOMOLINOS,CRISTÓBAL,R; MUÑOZ CARRERO,A; HERNANDO PÉREZ, M,E; PASCUAL CARRASCO,M; SÁNCHEZ DE MADARIAGA,R; MORENO GIL, O; FRAGUA MÉNDEZ,J,A; LÓPEZ RODRÍGUEZ,F;. HERNÁNDEZ SALVADOR,C: "Pseudonimización de información clínica para uso secundario. Aplicación en un caso práctico ISO/EN 13606", *XXXII Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, 2014.



Elaboración propia

B) La aplicación efectiva de la seudonimización y su legitimación como tratamiento de datos

La seudonimización de datos personales requiere el despliegue de toda una serie de medidas técnico organizativas para que esta resulte adecuadamente implementada. A su vez, y teniendo en cuenta que la propia seudonimización supone un tratamiento de datos, esta debe cumplir con el principio de licitud. Es turno de analizar estos elementos.

En primer lugar, por lo que se refiere a las medidas técnico organizativas, el considerando 28 del RGPD obliga a las organizaciones a establecer herramientas que mantenga por separado la información adicional para la atribución de los datos personales a una persona concreta. A su vez, se ha de indicar claramente quiénes son las personas que están autorizadas para acceder a esos registros. En este sentido, la LOPD de 2018 establece que el uso de datos personales seudonimizados con fines de investigación en salud pública y biomédica requerirá de: i) una separación técnica y funcional entre el equipo investigador y quienes realicen la seudonimización y conserven la información que posibilite la reidentificación y, ii) que los datos seudonimizados únicamente sean accesibles al equipo de investigación en circunstancias muy específicas[648]. Así, estas técnicas otorgan mayor protección a los interesados ya que el acceso está restringido a determinadas personas autorizadas, lo

[648] Los investigadores podrán acceder cuando: *i) Exista un compromiso expreso de confidencialidad y de no realizar ninguna actividad de reidentificación. ii) Se adopten medidas de seguridad específicas para evitar la reidentificación y el acceso de terceros no autorizados.* Disposición Adicional 17.2.d) de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales

cual reduce el riesgo en el tratamiento[649]. En cuanto a las técnicas informáticas en sentido estricto, el GT29 ha hecho referencia entre otras al cifrado con clave secreta, la función hash[650], función con clave almacenada y la descomposición en tokens[651].

De esta manera, y previo al proceso de seudonimización, resulta necesario responder a las preguntas de quién lo realizará, qué técnicas se aplicarán, quién tendrá acceso a los datos finales, cómo serán utilizados, cuál será el riesgo de reidentificación residual, etc[652]. El objetivo pretendido será evitar que la analítica masiva de datos de lugar a la identificación de las personas titulares de los mismos. En este sentido, si las medidas técnico organizativas no se implantan adecuadamente no puede considerarse efectivamente realizada la seudonimización[653].

En segundo lugar, con relación a la base de legitimación que habilita a la seudonimización hay que partir de que esta en sí misma es un tratamiento de datos, por lo tanto, es necesario que responsable del tratamiento tenga un mecanismo que le permita llevar a cabo el mencionado tratamiento y cumplir con el principio de licitud. Es turno de analizar algunas de las situaciones.

A) Cuando la seudonimización únicamente se implemente a efectos de establecer mayores medidas de seguridad sobre los datos que se están tratando y el uso de dichos datos no se proyecte para otras finalidades ulteriores, bastará con informar a los interesados durante la recopilación de los datos de que los mismos se pretenden seudonimizar. Entendemos que, el mecanismo de legitimación que se utilice junto con ese derecho de información serán más que suficientes para entender satisfecho el principio de licitud que habilita la seudonimización. Además, en el caso de que fuera necesario valorar la compatibilidad entre la finalidad inicial y la seudonimización, tal compatibilidad quedaría superada ya que el propio RGPD establece la necesidad de implementar la seudonimización como medida que favorece el principio de privacidad

[649] POLO ROCA,A: "Datos, Datos, Datos: El Dato Personal, El Dato No Personal, El Dato Personal Compuesto, La anonimización, La Pertenencia Del Dato Y Otras Cuestiones Sobre Datos". *Estudios De Deusto* 69 (1), 165-94, 2021, pág.228. Texto disponible: https://doi.org/10.18543/ed-69(1)-2021,

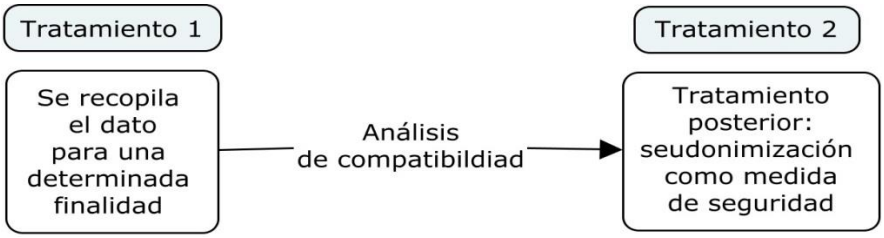
[650] Para más información sobre esta técnica de seudonimización véase: Agencia Española de Protección de Datos. *Introducción al hash como técnica de seudonimización de datos personales*. Octubre de 2019.

[651] Grupo del Artículo 29. Dictamen 05/2014 sobre técnicas de anonimización. Adoptado el 10 de abril de 2014, pág.22 y 23.

[652] ORTIZ UROZ,R: "El tratamiento de datos personales en los proyectos universitarios de investigación a la vista de la actual normativa de protección de datos personales". *LA LEY privacidad* , Nº 7, Sección El foro de la privacidad, Primer trimestre de 2021,.pág.10.

[653] Autoritat Catalana de Protecció de dades. Consulta CNS 6/2020, pág.15. Texto disponible en: https://apdcat.gencat.cat/web/.content/Resolucio/Resolucions_Cercador/Dictamens/2020/Documents/ca_cns_2020_006.pdf

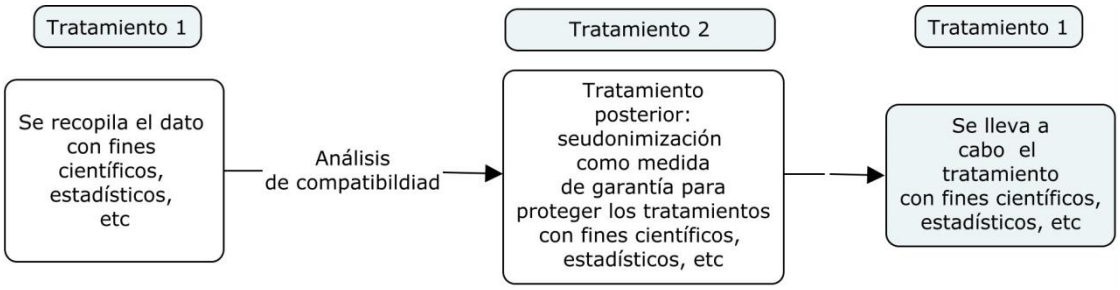
desde el diseño, el principio de seguridad y el de minimización de datos. Tanto el artículo 25 como el 32 del RGPD refuerzan esta idea.



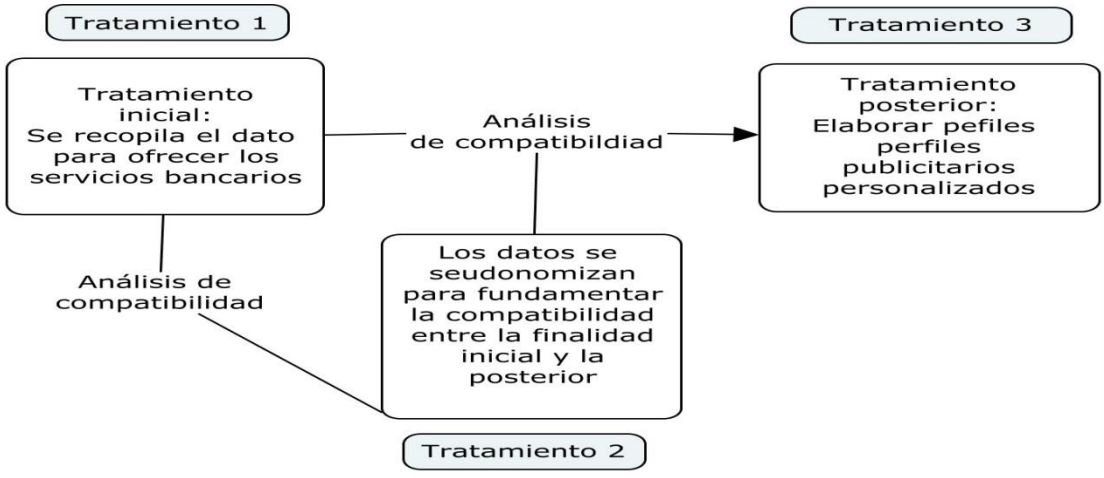
B) Cuando la seudonimización se despliegue como medida de garantía para facilitar los tratamientos que tengan como finalidad el archivo en interés público, fines de investigación científica o histórica o fines estadísticos, el responsable deberá informar sobre la seudonimización de esos datos en el momento de la recopilación de los mismos. Al igual que en el caso anterior, entendemos que el análisis de compatibilidad también se entiende superado ya que la seudonimización se estructura como una garantía necesaria para que se produzca una adecuada protección de los tratamientos descritos en el artículo 89.1 del RGPD. Durante la recopilación de los datos se deberá informar que los mismos se seudonimizarán. Por tanto, si una organización pretende recopilar datos de diversas personas con el objetivo de analizarlos a efectos de desarrollar modelos algorítmicos cuyas finalidades queden cubiertas por las descritas en el artículo 89.1 del RGPD, este habrá de indicar que dichos datos, además de servir para la investigación científica a través de técnicas de *machine learning*, previamente se seudonimizarán.

Cuando la finalidad posterior pretendida con el tratamiento corresponda con los fines indicados en el artículo 89.1, la compatibilidad se presume tal y como establece el artículo 5.1.b del RGPD. Por lo tanto, tal presunción también se extiende respecto de la seudonimización, la cual, recordemos, supone una medida que habrán de desplegar los responsables del tratamiento si desean tratar los datos con esa finalidad. Así, por ejemplo, el uso de la información del historial clínico inicialmente recopilado para prestar la asistencia sanitaria se entenderá compatible con el uso posterior para fines de investigación en salud siempre que dichos datos se seudonomicen[654].

[654] Autoritat Catalana de Protecció de dades. Consulta CNS 15/2019, pág.15. Texto disponible en: https://apdcat.gencat.cat/web/.content/Resolucio/Resolucions_Cercador/Dictamens/2019/Documents/ca_cns_2019_015.pdf



C) Cuando la seudonimización se despliegue como medida de garantía a tener en cuenta para valorar la compatibilidad o no del tratamiento inicial previsto de los datos respecto del tratamiento posterior pretendido, lo más relevante será valorar la compatibilidad entre el tratamiento inicial y el posterior de acuerdo a lo establecido en el artículo 6.4 del RGPD. Si ambos tratamientos resultan compatibles, teniendo en cuenta que la seudonimización se utiliza como medida que favorece dicha compatibilidad, la legitimación de la misma también se entenderá. Por ejemplo, una entidad bancaria ostenta una serie de datos de sus clientes que fueron recopilados con la finalidad de ofrecer sus servicios. Posteriormente, esta misma organización pretende utilizar esos datos para ofrecer publicidad personalizada y desarrollar perfiles. A la hora de valorar la compatibilidad entre el fin inicial y el posterior, la seudonimización de esos datos se valorará positivamente. Pues bien, si finalmente se considera que el tratamiento inicial y el posterior son compatibles, entendemos que la seudonimización de esos datos, que recordemos que también es otro tratamiento, también resultará compatible.



En definitiva, la seudonimización es una medida de responsabilidad activa y a su vez un tratamiento de datos. Como regla general, la seudonimización estará legitimada siempre que se utilice como una medida de seguridad o de garantía para proteger más

eficazmente los derechos de los interesados. Cuando la seudonimización se implante como medida para potenciar el análisis de compatibilidad descrito en el artículo 6.4 del RGPD, la seudonimización estará legitimada siempre que se llegue a la conclusión de la existencia de compatibilidad entre el fin inicialmente previsto y el ulterior. En todos estos supuestos, para que el proceso de seudonimización esté legitimado, el mismo ha de ser adecuadamente desplegado e implementado a través de distintas medidas técnico organizativas.

3 Nuevos escenarios para la analítica masiva de datos personales

Normalmente, las organizaciones, con el objetivo de evitar las exigencias que se derivan de la normativa de protección de datos acuden a los procesos de anonimización a la hora de analizar masivamente los datos y diseñar los modelos algorítmicos sobre los cuales se crearán sistemas automatizados. Existen distintos enfoques que ofrece la legislación para desarrollar modelos algorítmicos donde se hace uso del tratamiento de datos personales. Estos escenarios permiten, por un lado, una protección efectiva de las personas cuyos datos están siendo tratados ya que la normativa de protección de datos la respalda y a su vez, por otro lado, se consigue sacar más partido a los datos y al conocimiento oculto tras los mismos ya que estos últimos no pierden tanta utilidad tras los procesos de anonimización a los que se ven sometidos.

El primer enfoque que resulta sumamente interesante es que las organizaciones utilicen el tratamiento de datos seudonimizados a la hora de diseñar modelos algorítmicos. Como ya hemos indicado, el RGPD hace una apuesta clara por esta medida al configurarla como una herramienta que facilita el tratamiento de datos con diversas finalidades, sobre todo, las descritas en el artículo 89.1. En este sentido, muchos proyectos de *big data* requieren del uso de datos personales para que los mismos arrojen resultados concluyentes[655]. A su vez, los procesos de anonimización, además de resultar costosos, siempre mantienen un alto riesgo de reidentificación de las personas, riesgo que se ve acrecentado con la analítica masiva de datos a través de técnicas de inteligencia artificial[656]. Además, y a pesar de que la anonimización

[655] Comisión Europea. *Ethics and data protection*. Texto aprobado el 14 de noviembre de 2018, pág.7.
[656] Comisión Europea. *Libro Blanco sobre la inteligencia artificial. Un enfoque europeo orientado a la excelencia y la confianza*, Bruselas, 19.2.2020, pág.15.

resulta la medida más adecuada de protección para los titulares de los datos, los riesgos de reidentifiación continúan, sin embargo, la normativa de protección de datos no se aplica en estos casos. La seudonimización permite seguir tratando datos personales sin necesidad de identificar a esas personas, por tanto, la normativa de protección de datos mantiene su aplicación sobre dichos tratamientos[657]. Visto así, la seudonimización respecto de la anonimización ofrece ventajas tanto para los responsables como para los interesados. Esta última se muestra como una herramienta que en parte puede aliviar las tensiones existentes entre la normativa de protección de datos y la analítica masiva de los mismos. Así, uno de los ámbitos previamente comentado donde el uso de datos seudonimizados con fines de creación de sistemas algorítmicos ha encontrado un respaldo legal explícito se encuentra en el sector de la investigación sanitaria. La LOPD de 2018 ha desarrollado todo un marco legislativo en torno al uso de datos seudonimizados en el cual, se relajan algunos de los principios básicos de protección de datos como puede ser el de limitación de la finalidad o licitud o algunos de los derechos de los interesados pero a su vez[658], se compensan con la exigencia de cumplimiento de medidas férreas de responsabilidad activa. El fomento de este tipo de tratamientos encuentra su respaldo en el propio RGPD. Esta norma claramente apuesta por la investigación, la cual, representa un bien jurídico colectivo que se ha de respetar y respaldar[659]. En este ámbito, la seudonimización presenta ventajas muy relevantes respecto de la anonimización. Así, al mantenerse el vínculo entre el dato y la persona, es posible contrastar los resultados de la explotación de datos con, por ejemplo, la verdadera evolución de los pacientes. Ello permite valorar las posibles causalidades espurias, riesgo potencial presente en la

[657] TRONCOSO REIGADA,A: "Investigación, salud pública y asistencia sanitaria en el Reglamento General de Protección de Datos de la Unión Europea y en la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales". *Revista de derecho y genoma humano: genética, biotecnología y medicina avanzada*, nº49,2018, pág.210.

[658] Disposición adicional decimoséptima apartado 2, epígrafe d) de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales

[659] El considerando 4 del RGPD establece que: *El tratamiento de datos personales debe estar concebido para servir a la humanidad. El derecho a la protección de los datos personales no es un derecho absoluto sino que debe considerarse en relación con su función en la sociedad y mantener el equilibrio con otros derechos fundamentales, con arreglo al principio de proporcionalidad*. A su vez, el considerando 157 establece que: *Los resultados de investigaciones obtenidos de registros proporcionan conocimientos sólidos y de alta calidad que pueden servir de base para la concepción y ejecución de políticas basada en el conocimiento, mejorar la calidad de vida de numerosas personas y mejorar la eficiencia de los servicios sociales. Para facilitar la investigación científica, los datos personales pueden tratarse con fines científicos, a reserva de condiciones y garantías adecuadas establecidas en el Derecho de la Unión o de los Estados miembros.*

analítica masiva de datos y que en parte se soluciona con el uso de datos seudonimizados[660].

Este modelo normativo, con sus especificaciones propias del contexto que regula, puede ser replicado por las organizaciones en otros contextos donde también se pretenda la analítica masiva de datos y no se quiera perder parte de la utilidad de los mismos fruto de los procesos de anonimización. Sobre todo, la clave estará en permitir finalidades más genéricas durante el proceso de especificación de las mismas en las fases iniciales del desarrollo de estos sistemas. A su vez, el uso de datos seudonimizados ha de ser un elemento muy relevante a la hora de valorar la compatibilidad de los tratamientos cuando se pretendan fines secundarios. Todo ello claro está, con el diseño de otra serie de garantías que compensen la relajación de la normativa de protección de datos. En estos contextos, los responsables han de ser conscientes de que no existe una regulación expresa que les habilita a tratamientos más ventajosos de analítica de datos como los previstos para la investigación científica, estadística o de salud[661]. Es por ello que todo paso que busque replicar el modelo normativo comentado deberá hacerse con cautela y siempre bajo el paraguas general del RGPD y el resto de normativa de protección de datos. Las consultas que puedan realizar los responsables a las respectivas a autoridades de protección de datos cuando tengan dudas de las operaciones a realizar pueden ser también un buen instrumento para afianzar los tratamientos de datos pretendidos. Sería además recomendable que en estos contextos el legislador legitimara expresamente tratamientos relacionados con la analítica masiva de datos donde los riesgos para los particulares sean menores a los beneficios generales que dichos tratamientos pueden generar a los responsables del tratamiento. Se conseguiría así facilitar la analítica masiva de datos bajo un marco normativo garantista[662].

En segundo lugar, ligada a la seudonimización de los datos, otro escenario ya mencionado y que cada vez recibirá mayor atención es el relativo al papel de los intermediarios. Es decir, estas organizaciones pueden actuar como facilitadores de

[660] DE MONTALVO JÄÄSKELÄINEN, F: "Una reflexión desde la teoría de los derechos fundamentales sobre el uso secundario de los datos de salud en el marco del Big Data". *Revista De Derecho Político*, 1(106), 2019, pág.63. Texto disponible en: https://doi.org/10.5944/rdp.106.2019.26147

[661] Ya hemos comentado previamente este modelo normativo. Véase el artículo 89 del RGPD y la D.A.17 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

[662] Sobre esto último también hablamos en el Capítulo IV, apartado I, punto 6, epígrafe D).

entornos seguros entre aquellos que facilitan los datos y los usuarios de los mismos que pretenden analizarlos. Así, la Unión Europea, a través de la Propuesta de Reglamento relativo a la gobernanza europea de datos pretende implantar una serie de instrumentos que buscan facilitar la reutilización de datos especialmente protegidos como pueden ser los personales o los derivados de la propiedad intelectual. En este sentido, el primer mecanismo es referido a la creación de entornos seguros desarrollados por organismos públicos específicos que permitan la reutilización de datos anonimizados y seudonimizados en posesión de las Administraciones Públicas en favor de terceras organizaciones[663], de manera que ese organismo público actúa entre medias de los facilitadores de datos y aquellos que pretendan usarlos. El segundo mecanismo consiste en regular una nueva figura conocida como los proveedores de los servicios de intercambio de datos, estos actuarán como intermediarios de datos independientes tanto de los titulares de los datos como de los usuarios que pretendan analizarlos[664]. Precisamente, y con el objetivo de proteger los derechos que se derivan de la normativa de protección de datos, se crea una figura específica que tendrá como finalidad esencial facilitar el ejercicio de las facultades reconocidas en el RGPD por parte de los interesados en los casos en los que se pretendan usar sus datos[665]. Estos mecanismos pueden potenciar por tanto el análisis masivo de datos en un entorno respetuoso con la normativa de protección de datos. En estos entornos seguros, los datos seudonimizados o anonimizados pueden tratarse de forma adecuada y el riesgo de reidentifiación se reduce. Además, se convierte en una importante alternativa al tratamiento de datos en abiertos, los cuales, son anonimizados de tal manera que, en muchos casos pueden resultar poco valiosos para los analistas de datos.

En tercer lugar, proponemos otro escenario posible. Como ya hemos comentado en otras páginas, resulta conveniente que aquellas organizaciones que diseñen sistemas de toma de decisiones automatizadas tengan previstos desde las

[663] Considerandos 6, 11 y artículo 5 de la Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos). Resolución del 25 de noviembre de 2020.
[664] Considerando 22 y artículo 9 de la Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos). Resolución del 25 de noviembre de 2020.
[665] Se trata de los proveedores de servicios de intercambio de datos que ofrecen sus servicios a los interesados. Considerando 23 y articulo 9.1.b) de la Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos).

primeras fases de los mismos el cumplimiento normativo de protección de datos ya que una vez el sistema comience a desplegar sus efectos, muy posiblemente la normativa de protección de datos se aplicará y muchas de estas exigencias deberían estar implantadas durante la fase del diseño. A su vez, si una tercera organización, como puede ser una Administración Pública, requiere que ese sistema cumpla con la normativa de protección de datos, muy posiblemente exigirá previa a la adquisición del mismo el cumplimiento de la normativa de protección de datos. Es por ello que en nuestra opinión, e independientemente de que se traten datos o no personales durante el diseño de los sistemas automatizados, los principios del tratamiento reconocidos en los artículos 5 del RGPD se alineen en la medida de lo posible con los distintos principios y exigencias de robustez propias que han de estar presentes durante el desarrollo de estos sistemas y modelos algorítmicos. En las siguientes páginas se irá haciendo referencia a esa alineación entre estos principios durante la fase del diseño. Este enfoque, en nuestra opinión presentará ventajas tanto para los particulares que en un momento posterior puedan verse sometidos a decisiones automatizadas como a las organizaciones que desarrollan estos sistemas. Concretamente, los beneficios de estas últimas pasan tanto por la obtención de una mayor utilidad de los datos como por un mejor posicionamiento a la hora de ofrecer ese producto al mercado.

X. EL MONITOREO Y LA EVALUACIÓN DE LOS SISTEMAS AUTOMATIZADOS

Una de las medidas que necesariamente habrán de implantar los responsables del tratamiento que diseñen y/o implementen sistemas de toma de decisiones automatizadas son las referidas a las formas de validar los mismos. Esta medida de responsabilidad activa, aunque no esté expresamente indicada por el RGPD, es fundamental debido a la importancia que tiene la misma a la hora de evaluar si un determinado sistema que adoptará o está adoptando decisiones automatizadas basándose en datos persones actúa adecuadamente en el contexto donde se integrará o está integrado. En este sentido, tal y como indica la AEPD, establecer mecanismos que permitan la validación del sistema ayudará a los responsables del tratamiento no sólo a comprobar su robustez sino

también a demostrar el cumplimiento de la normativa de protección de datos[666]. Los chequeos o evaluaciones de los modelos algorítmicos pueden realizarse a lo largo de las distintas etapas que comprende el ciclo de vida estos sistemas, resultando necesario que al menos dichas pruebas de evaluación se realicen en tres momentos diferenciados. Estos son; la fase de desarrollo del modelo, la fase previa al despliegue del modelo y, finalmente, la fase de adopción de decisiones.

Por lo que se refiere a la *fase del diseño*, la validación habrá de realizarse al menos durante la etapa que comprende la elección de los distintos modelos generados. Como ya dijimos en páginas anteriores, una correcta validación del sistema facilitará su despliegue de una forma más segura y facilitará su posterior venta a terceros al dotar a dicho modelo algorítmico de índices objetivos de solvencia y robustez.

Por otro lado, en lo que respecta a *la fase previa a la puesta en marcha del modelo*, la validación del mismo se hace sumamente necesaria[667]. En este sentido, y dado que ya se conocerá el entorno en el que el sistema se integrará, será adecuado que el mismo pueda someterse en su caso a pruebas piloto que le permitan a la organización que pretenden implantarlo valorar mínimamente su funcionamiento y valorar si el sistema se adecúa a dicho contexto[668]. En nuestra opinión, cuando la organización que diseñó el sistema sea diferente a la que lo pretende implementar, el responsable del tratamiento que quiera utilizar dicho sistema debería evaluarlo antes de ponerlo en marcha para cerciorarse que el mismo cumple con la normativa de protección de datos. Todo ello en virtud del deber de diligencia que concierne al responsable a la hora de elegir encargados y productos algorítmicos que respeten la normativa de protección de datos[669].

Finalmente, durante la fase de *seguimiento y la evolución del algoritmo* también será necesario el monitoreo de los resultados. Son diversas las razones que justifican ello: En primer lugar, la propia validación del modelo a la hora de comprobar que este

---

[666] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.15. Véase también: Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág.26.

[667] Consejo de Europa. Consultative committee of the convention for the protection of individuals with regard to automatic processing of personal data. (Convention 108). *Artificial Intelligence and Data Protection: Challenges and Possible Remedies*. Informe adoptado el 25 de enero de 2019, pág.10.

[668] Como ha señalado la doctrina, es habitual que los sistemas automatizados se pongan en funcionamiento sin la realización previa de pruebas piloto. En: SPIELKAMP,M;KAYSER-BRIL,N: "Experimentation without a plan: automated decision-making in European public services journalist". *European Public Mosaic* , 2019, pág.8.

[669] Sobre el deber de diligencia del responsable en este contexto véase el Capítulo III, apartado III de esta tesis.

mantiene el mismo o un índice aproximado de errores que previamente se indicó o se detectó en la fase previa al despliegue[670]. En segundo lugar, a través del seguimiento de los modelos se pueden llegar a detectar errores que no se habían previsto durante la fase de diseño del mismo. En este sentido, los sistemas automatizados pueden ser una herramienta muy adecuada para detectar sesgos presentes en los procesos rutinarios que hasta la fecha no se habían detectado[671]. Así, tal y como ha señalado la autoridad de protección de datos italiana, el responsable del tratamiento está obligado a verificar periódicamente la corrección y exactitud de los resultados de los sistemas algorítmicos con el fin de asegurar que el riesgo de errores se minimice[672]. En tercer lugar, los testeos también pueden servir como herramienta para valorar si efectivamente el uso de sistemas automatizados en un entorno resulta o no adecuado teniendo en cuenta los objetivos marcados inicialmente con la implementación de los mismos y los objetivos conseguidos a medio o largo plazo. Es decir, el seguimiento puede servir como mecanismo que justifique la integración del sistema o en cambio poner en entredicho su uso en ese contexto específico al no ofrecer los resultados esperados[673]. En cuarto lugar, y teniendo en cuenta que muchos de estos sistemas se implantan en entornos altamente adaptativos o variables, la posibilidad de que la precisión inicial de estos modelos comience a descender aumenta drásticamente[674]. Es por ello que el seguimiento y monitoreo sea también fundamental. Además, incluso en entornos poco variables, la validación también será necesaria ya que esos contextos también pueden ir cambiando con el paso del tiempo lo que puede requerir una rediseño de esos modelos, los cuales,

---

[670] Como se ha indicado, la racionalidad perfecta de un sistema, es decir, la toma de decisiones correctas nunca será posible o difícilmente se logrará en entornos sumamente complejos dada la demanda computacional que ello exige. Por tanto, siempre habrá se ha de contar con un cierto margen de error. En: RUSSELL, S Y NORVING, P: *Inteligencia Artificial. Un enfoque moderno*, op.cit., pág. 6.

[671] SORIANO, ARNANZ,A: "Decisiones automatizadas y discriminación: aproximación y propuestas generales".op.cit.,pág.29.
También lo ha señalado Red iberoamericana de protección de datos personales. *Orientaciones Específicas para el Cumplimiento de los Principios y Derechos que Rigen la Protección de los Datos Personales en los Proyectos de Inteligencia Artificial.*2019, pág.16.

[672] Garante per la Protezione dei Dati Personali. Resolución nº 234 de 10 de junio de 2021. Apartado 3.3.6. Resolución disponible en:
https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9675440

[673] Se demostró que un sistema de predicción utilizado por la policía de Chicago no había logrado reducir la violencia con armas de fuego ni reducir la probabilidad de victimización. La inclusión de este sistema la inclusión sólo tuvo un efecto directo en los arrestos. RICHARDSON, R; SCHULTZ, J Y CRAWFORD, K.: "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice", op.cit., pág.18.

[674] Los algoritmos de aprendizaje automático se ajustan en función de la nueva retroalimentación y eventualmente toman decisiones basadas en criterios que pueden no haber sido elegidos explícitamente por los programadores humanos. En: Center for Democracy and Technology. *Digital Decisions,2017*, pág.4.

pueden quedar desvirtuados de la realidad legal y contextual en la que se ingresaron previamente. En todos estos supuestos, la continua monitorización también permitirá conocer las causas principales que han generado los errores o fallos en el sistema lo que facilitará la correspondiente atribución de responsabilidades por los daños que se produzcan[675].

En definitiva, la importancia de realizar testeos en las distintas etapas que comprende el uso de sistemas automatizados se muestra esencial. Es momento de centrarnos específicamente en determinadas métricas que resultan esenciales para realizar estos procesos de validación.

1. El análisis de los errores de los sistemas en el contexto específico de la toma de decisiones

Resulta esencial utilizar todo tipo de herramientas que permitan a las organizaciones evaluar el rendimiento de los modelos que en un momento posterior comenzarán a adoptar decisiones. Pues bien, como ya se indicó en el capítulo inicial de esta tesis, a través de distintas métricas puede llegar a evaluarse un determinado modelo algorítmico[676]. Ahora, analizaremos dichas métricas aplicadas a un caso específico utilizando como referencia los trabajos realizados por la doctrina estadounidense sobre el algoritmo de YouTube de retirada de contenido ilícito[677]. Se estudiará este algoritmo y los posibles errores del mismo teniendo en cuenta la normativa de derechos de autor de la Unión Europea que habilita a las plataformas a retirar contenido ilícito que se difunde a través de las mismas[678].

Así, YouTube cuenta con una base de datos que contiene multitud de archivos de audio y videos proporcionada por los titulares de los mismos con el fin de que sus algoritmos automáticos puedan detectar aquel contenido que se suba por otros usuarios y que atente contra la normativa de derechos de autor. De esta manera, el denominado

[675] Expert Group on Liability for New Technologies. Comisión Europea. *Liability for Artificial Intelligence and other emerging digital technologies*, 2019, pág.49.
[676] Capítulo I, aparatado II, punto 2, epígrafe E) de esta tesis.
[677]Un estudio muy interesante sobre el algoritmo de YouTube y la valoración de las consecuencias jurídicas en la normativa sobre derechos de autor en Estados Unidos por los errores generados por dicho sistema puede encontrarse en: LESTER, T; PACHAMANOVA,D: "The Dilemma of False Positives: Making Content Id Algorithms More Conducive to Fostering Innovative Fair Use in Music Creation". *UCLA Entertainment Law Review*, Vol. 24, No. 51, 2017.
Texto disponible en: https://escholarship.org/uc/item/1x38s0hj
[678] 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.

algoritmo Content ID compara de forma automática los videos subidos por los usuarios de esta plataforma y trata de buscar coincidencias con la base de datos en la que se encuentran albergados los contenidos de los titulares de los mismos de manera que, si dicho contenido coincide, notifica al titular del mismo y este decide si: i) el video se retira, ii) obtiene los ingresos que se derivan del mismo, iii) comparte esos ingresos con el usuario que subió el contenido o, iv) decide mantener el video con fines de medición de audiencia[679]. Dicho bloqueo de contenido podrá ser automático cuando así haya sido previsto en los acuerdos de licencia entre los titulares del contenido y la plataforma YouTube. Pues bien, independientemente de que el bloqueo del contenido sea o no automático, lo que está claro es que esta plataforma utiliza un algoritmo de reconocimiento automático para valorar la eliminación o no de contenido[680]. Es turno de analizar las distintas métricas de este modelo.

Así, cuando un algoritmo como el de YouTube recomienda clasificar contenido subido como infractor o no infractor de la normativa de derechos de autor se pueden originar cuatro posibles resultados: verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos[681]. En primer lugar, estaremos ante un verdadero positivo cuando el algoritmo identifique un video como infractor de la normativa y dicho contenido sea infractor. En segundo lugar, será un falso positivo cuando el algoritmo indique que una canción infringe la normativa de derechos de autor pero en realidad no la infringe. En tercer lugar, nos encontraremos ante un verdadero negativo cuando el algoritmo detecte que un video no infringe la normativa de derechos de autor y efectivamente dicho contenido no la infringe. Finalmente, cuando el algoritmo no considere como infractora una obra pero realmente esta sí lo sea, nos encontraremos ante un falso negativo.

---

[679] Conclusiones del Abogado General Sr. Henrik presentadas el 16 de julio de 2020. Asuntos acumulados C-682/18 y C-683/18. Frank Peterson contra Google LLC. Apartado 22. Disponible en: https://curia.europa.eu/juris/document/document.jsf;jsessionid=8667950A1A86F73CCDE040C63B998688?text=&docid=228712&pageIndex=0&doclang=ES&mode=req&dir=&occ=first&part=1&cid=1884929

[680] Algoritmo que como acertadamente ha indicado la doctrina necesariamente deberán implantar las plataformas como YouTube dadas las nuevas obligaciones de control del contenido que establece la Directiva de derechos de autor a las mismas. En: VEGA GARCÍA, P: "El algoritmo de YouTube, el artículo 17 de la Directiva 2019/790 y la protección de los derechos de autor ". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed. Aranzadi, Navarra, 2020, págs. 336 y 340.

[681] LESTER, T; PACHAMANOVA,D: "The Dilemma of False Positives: Making Content Id Algorithms More Conducive to Fostering Innovative Fair Use in Music Creation", op.cit.,pág.65.

	Verdadero positivo	**Falso positivo**
El sistema dice	Contenido infractor	Contenido infractor
La realidad es	Contenido infractor	No contenido infractor
	Falso negativo	**Verdadero negativo**
El sistema dice	No contenido infractor	No contenido infractor
La realidad es	Contenido infractor	No contenido infractor

2. Métricas utilizadas para calcular los valores derivados de las decisiones

Una vez se conocen estos valores, lo ideal es que los mismos se trasladen al contexto específico donde el sistema irradiará sus efectos y evaluar quiénes son los principales actores afectados por estas decisiones y los derechos e intereses en juego presentes. Así, por ejemplo, un falso positivo perjudicaría a los usuarios que suben contenido a la plataforma YouTube, quedando afectado su derecho a la libertad de expresión o artística ya que se bloquearía un archivo que no infringe la normativa. En cambio, un falso negativo afectaría a los titulares de aquellos contenidos que han de ser protegidos por la plataforma, afectando en estos supuestos al derecho a la propiedad, concretamente a la propiedad intelectual, ya que se mantiene en la red un archivo que infringe la normativa sobre derechos de autor.

	CONSECUENCIAS/DERECHOS AFECTADOS
FP	Se bloquea un contenido legal El derecho a la libertad de expresión y libertad artística aparecen afectados
FN	No se bloquea un contenido ilegal El derecho a la propiedad intelectual resulta afectado
VP	Se bloquea un contenido ilegal El derecho a la libertad intelectual se protege sin afectar a la libertad de expresión
VN	No se bloquea contenido que no es ilegal Se respeta la liberta de expresión y la artística sin afectar a la propiedad intelectual

Resulta por tanto esencial evaluar cada uno de estos valores a través de distintas métricas. Concretamente, estas métricas a las que ya nos hemos referido en otro momento del estudio son[682]; la tasa de acierto o *accuracy*, la tasa de verdaderos

[682] Estas métricas ya han sido analizada en el capítulo inicial de este trabajo, nos remitimos a las explicaciones indicadas anteriormente. Capítulo I, apartado IV, punto 2, epígrafe D, sub epígrafe d.2).

positivos o sensibilidad, la tasa de verdaderos negativos o especificidad, la tasa de verdaderos negativos o especificidad, el valor predictivo positivo o precisión y el valor predictivo negativo. Cada una de estas métricas, como ahora comprobaremos, presentan distintas implicaciones y pueden ser más o menos relevantes dependiendo del contexto en el cual se utilicen.

En primer lugar, por lo que se refiere a la tasa de acierto o *accuracy*, nos estamos refiriendo a la precisión global de un algoritmo de clasificación. De esta manera, si por ejemplo se indica que el algoritmo ostenta una tasa de acierto del 95%, quiere decir que el algoritmo identifica correctamente el 95% de los contenidos, ya sea infractor o no infractor el contenido, el 5% restante los clasifica incorrectamente. La tasa de acierto es una métrica relevante pero no diferencia el porcentaje correspondiente de falsos negativos y positivos. Se desconoce por tanto cómo se distribuyen los errores.

En segundo lugar, la *tasa de verdaderos positivos o sensibilidad* refleja la probabilidad de que el sistema algorítmico encuentre infracciones de la normativa de derechos de autor en los contenidos. Es decir, del conjunto obras ilícitas que se sube a la red, qué porcentaje de acierto tiene el sistema para detectarlas como ilícitas. Esta métrica resulta relevante para los titulares de las obras cuyos derechos son protegidos a través del algoritmo.

En tercer lugar, la tasa de *verdaderos negativos o especificidad*. Esta tasa refleja la probabilidad de que el sistema algorítmico excluya correctamente las obras que se suben a la plataforma como no infractoras. Es decir, del conjunto de obras lícitas que se suben a la plataforma, qué porcentaje de acierto tiene el sistema para detectarlas como lícitas. Esta métrica resulta relevante para los usuarios que suben contenido a la plataforma.

En cuarto lugar, en la métrica *de la precisión o valor predictivo positivo* se refleja la probabilidad de que un contenido clasificado como infractor lo sea realmente. Es decir, del conjunto de instancias catalogadas por el algoritmo como infractoras de la norma, cuales son realmente infractoras.

En quinto lugar, el *valor predictivo negativo* refleja la probabilidad de que un contenido clasificado como no infractor lo sea realmente. Es decir, del conjunto de instancias catalogadas por el algoritmo como no infractoras de la norma, cuales son realmente infractoras.

3. La importancia de las métricas

La información obtenida a partir de las métricas sitúa al responsable del tratamiento en una mejor posición para tomar una serie de decisiones relacionadas con el modelo algorítmico que se pretende implantar[683]. Estas decisiones pueden ser:

En primer lugar, estas métricas pueden utilizarse como referencia para evaluar el riesgo de que se materialice una determinada amenaza, en este caso, una decisión incorrecta. En este sentido, dado que ya se conocen los porcentajes específicos de los distintos resultados analizados, el responsable puede estimar adecuadamente la probabilidad de que el sistema adopte una decisión incorrecta que afecte a los particulares, es decir, puede facilitar la valoración del riesgo cuando se esté elaborando la evaluación de impacto en materia de protección de datos. Recordemos que una de las variables que se utilizaba para analizar y evaluar los riesgos era la probabilidad de que se materializará esa potencial amenaza, pues bien, al disponer de esas métricas, la probabilidad ya se puede cuantificar. En el caso de YouTube, una potencial amenaza podría ser que el sistema identifique incorrectamente el contenido ilícito subido por un usuario cuando dicho contenido realmente no lo era. La probabilidad de que suceda esa amenaza ya es cuantificable.

En segundo lugar, estas métricas también pueden ayudar a establecer las medidas que precisamente puedan reducir o mitigar esos riesgos potenciales que se deriven de incorrectas decisiones que afecten a las personas[684]. Concretamente, si YouTube apuesta por potenciar un sistema que refleja un valor predictivo positivo alto en defecto de un valor predictivo negativo bajo lo que está diseñando es un entorno en el cual se tratarán de detectar el mayor número de posibles contenidos ilícitos pero en cambio, no existirá alta precisión para detectar contenidos que no son ilícitos. Este modo de proceder puede afectar gravemente a los derechos de libertad de expresión de los usuarios. Para compensar esta situación de la cual YouTube es consciente, es decir, conoce la amenaza y además la probabilidad de que se materialice, este puede establecer

[683] Teniendo en cuenta que siempre existirá la probabilidad de que un sistema genera errores, resulta muy relevante que se puedan detectar, sobre todo, en situaciones en las que el sistema de IA afecte de manera directa a la vida de las personas. En: Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. Directrices éticas para una IA fiable.2019, pág.21.

[684] De esta manera, si las consecuencias de los falsos positivos son más graves, el nivel de garantías ha de aumentar. En :BERMAN, E:"A government of laws and not of machines". *B.U. L. RE*, op.cit., pág.48.
Esta idea no deja de representar el enfoque del riesgo sobre el que se asienta el principio de responsabilidad activa ya mencionado.

aquellas medidas que complemente tal situación. Entre las medidas podemos destacar; i) la creación de mecanismos ágiles de revisión de las impugnaciones presentadas por los usuarios ante los bloqueos de los contenidos, ii) información clara sobre los elementos que tiene en cuenta el algoritmo a la hora de bloquear el contenido, iii) el replanteamiento a la hora de considerar la automatización o no plena de todo ese procedimiento, iv) la formación adecuada de los trabajadores a los que les corresponde evaluar esas impugnaciones, etc. Estas medidas se han de actualizar si existe un cambio en los filtros de precisión. Así, por ejemplo, durante los meses más duros de la pandemia provocada por la Covid, la plataforma YouTube redujo parte de la plantilla de sus trabajadores que se dedicaban a evaluar el contenido que se vierte en sus plataformas, ello llevó esta plataforma a expandir aún más el uso de filtros automatizados. Pues bien, en el segundo trimestre del año 2020 se duplicó el número de videos que fueron eliminados por esta plataforma, la razón fue que esta plataforma decidió aumentar el umbral de precisión de sus sistemas automatizados para detectar posibles contenidos ilícitos[685]. Las consecuencias fueron que, automáticamente, también aumentaron el número de impugnaciones por parte de los usuarios ante la retirada masiva de contenido, el cual, se consideraba por estas personas como lícito. Este tipo de decisiones como puede comprobarse tiene efectos automáticos en la esfera de los particulares, concretamente y una vez más, en el derecho a la libertad de expresión de aquellos que suben el contenido y en el derecho a la libertad de información del resto de usuarios que pretenden visualizar los contenidos retirados por el algoritmo. Por tanto, ante estos cambios, también se han de prever medidas que refuercen esa afectación de derechos. En este caso lo lógico es que a la vez que se aumentaron los filtros de control de contenido, se facilitaran aún más los mecanismos para recurrir la retirada de dicho contenido.

En tercer lugar, estas métricas también pueden ayudar a justificar la incorporación o no de ese modelo algorítmico en ese concreto contexto. Recordemos, que una de las fases elementales de la evaluación de impacto era la valoración de la proporcionalidad y necesidad del tratamiento que se pretendía implantar. Los resultados derivados de estas métricas pueden ayudar al responsable a justificar la incorporación de estos modelos algorítmicos en la medida que se adecúan a las exigencias establecidas

[685] Fuente de la noticia: LAPOWSKY,I: "After sending content moderators home, YouTube doubled its video removals". *Protocol*, 25/08/2020. Disponible en: https://www.protocol.com/youtube-content-moderation-covid-19

por la ley que habilita a implantar estos mecanismos. Como ya indicamos en el apartado de la EIPD, en la Directiva de Derechos de autor se apreciaba un interés por parte de legislador europeo por establecer un nivel alto de protección de los derechos de autor de los contenidos que se suben a las plataformas sociales como YouTube. Esta plataforma justificaría el uso de modelos que potencien la identificación de contenido ilícito en defecto de modelos que tengan una menor precisión a la hora de detectar el contenido lícito, y ello, a pesar de que la libertad de expresión o de información de los usuarios se vea perjudicada en mayor medida[686].

En cuarto lugar, estas métricas también pueden resultar útiles para aquellas organizaciones que pretenda adquirir un determinado sistema algorítmico. Así, una Administración Publica podría valorar las consecuencias de los resultados en el entorno específico donde se implantaría ese sistema y a partir de ahí decidir si lo adquiere o no. En estos supuestos la operatoria posiblemente sería a la inversa, es decir, en el pliego contractual se fijaría la horquilla porcentual de todos o algunas de estas métricas sobre las cuales resultaría asumible la contratación de modelos algorítmicos. En este sentido, la PRAI obliga a los desarrolladores de sistemas de alto riesgo a incorporar en las instrucciones del modelo algorítmico los niveles de precisión y las métricas pertinentes[687]. Esa información puede resultar muy útil para que los potenciales compradores valoren la precisión de los sistemas que pretenden implementar en sus procesos decisorios.

Finalmente, en quinto lugar, la publicación de las métricas y los resultados de las mismas también pueden ayudar a los particulares que se ven sometidos a estos sistemas automatizados a conocer el funcionamiento general de estos modelos. En estos casos, dicha publicidad ha de ser meridanamente clara y facilitada en un lenguaje sencillo y comprensible para un usuario medio. Así, es muy habitual que las organizaciones, ya sean públicas o privadas, alardeen de sus modelos algorítmicos haciendo referencia a métricas muy generalista como puede ser la *accuracy* o tasa de acierto[688]. Sin embargo,

---

[686] Para algunas organizaciones pro derechos humanos, en cualquier proceso de moderación de contenido realizada por un algoritmo debería apostarse por establecer tasas de falsos negativos más altas-errar y que el sistema permita mayor circulación de noticias- y una tasa de falsos positivos más baja. En: DUARTE,N Y LLANSÓ,E: "Mixed Messages? The Limits of Automated Social Media Content Analysis", *Center for Democracy and technology*, op.cit., pág.19.

[687] Artículo 15.2 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

[688] Así, el Ministerio de Interior presume que su herramienta Veripol, encargada de estimar la probabilidad de que una denuncia por robo con violencia e intimidación o tirón sea falsa, ostenta una precisión de más del 90%. Fuente: Ministerio del Interior. En:

como hemos indicado anteriormente, sin un análisis global de todas las métricas es posible que se camuflen resultados que reflejan un funcionamiento inadecuado o poco preciso en determinados valores altamente relevantes, tal y como ocurre con el número de falsos positivos o negativos. Nótese que esta misma información también puede ser importante para las organizaciones que pretendan adquirir un determinado modelo.

4. Seguimiento de los resultados

Hasta ahora se han analizado los resultados del modelo algorítmico previos a la puesta en funcionamiento del mismo. No obstante, una vez que el sistema es implementado en el entorno donde comenzará a adoptar decisiones la precisión puede verse alterada. Es por tanto necesario que, como medida de responsabilidad activa, las organizaciones implanten mecanismos adecuados para monitorizar los resultados que el sistema va generado con el paso del tiempo y comprobar que estos mantienen los niveles de aciertos que en un momento inicial se obtuvieron durante la fase de desarrollo. En este sentido, sería recomendable que, previa a la puesta en marcha del sistema en el entorno real, este se probara de forma parcial a través de simulacros o pruebas piloto en las cuales se pueda validar nuevamente esos resultados y comprobar que los mismos se mantienen en un nivel similar o muy parecido a los mostrados en la fase de diseño del sistema[689].

Para llevar un seguimiento adecuado de un sistema automatizado se requiere la creación de mecanismos que permitan a los titulares de estos algoritmos llevar un registro adecuado de los resultados que se van generando, esto es, es necesario que se puedan ir contabilizando los verdaderos y falsos positivos, así como en su caso, los verdaderos y falsos negativos.

En general, mantener un seguimiento del rendimiento de los resultados del sistema dependerá en gran medida de los métodos que ostente el titular para comparar dichos resultados con una verdad fundamental que corrobore los resultados el

---

http://www.interior.gob.es/ca/prensa/noticias/-/asset_publisher/GHU8Ap6ztgsg/content/id/9496864
También YouTube hace mención a su algoritmo indicando que Content ID tiene una precisión para detectar los archivos de audio del 99,7% .Fuente de la noticia: KARP,H: "Industry Out of Harmony With YouTube on Tracking of Copyrighted Music", *The wall Street Journal*. 28/0672016. Disponible en:https://www.wsj.com/articles/industry-out-of-harmony-with-youtube-on-tracking-of-copyrighted-music-1467106213?tesla=y
En ambos supuestos no quedan claras las métricas a las que se hace referencia.

[689] Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021,.pág.27.

algorítmico[690]. En algunos supuestos, esa verdad fundamental no podrá ser corroborada, sobre todo cuando lo que se esté tratando de predecir o clasificar con el modelo algorítmico sean hechos subjetivos como por ejemplo decidir si un determinado contenido es ilícito o no. En estos casos, las organizaciones deberán acudir en la medida de lo posible a elementos objetivos para valorar dicho rendimiento. Así, ante el bloqueo de un determinado video en dicha plataforma, YouTube debería establecer las herramientas necesarias para contabilizar el rendimiento del sistema.

	Seguimiento del modelo de retirada de contenido	
FP	Consecuencia	Se bloquea contenido legal
	Seguimiento	Se impugna el bloqueo de contenido y se reinserta en la plataforma tras una evaluación del recurso. Se impugna el bloqueo del contenido, no se reinserta en la plataforma tras la primera evaluación pero, finalmente un juez obliga a su reinserción.
FN	Consecuencia	No se bloquea un contenido ilegal.
	Seguimiento	Se denuncia el contenido por terceros al considerar que es ilícito, ya sea la sociedad o los titulares potenciales de derechos de ese contenido y se decide bloquear el contenido
VP	Consecuencia	Se bloquea un contenido ilegal.
	Seguimiento	No se impugna el bloque del contenido o se impugna el bloqueo del contenido pero tras una evaluación del recurso no se reinserta el contenido
VN	Consecuencia	No se bloquea contenido que no es ilegal.
	Seguimiento	No se recurre el contenido por parte de terceros, ya sea la sociedad o los titulares de potenciales derechos de ese contenido

La forma por tanto de medir el rendimiento variará en función de multitud de factores, resultando muy relevante el contexto y la relación que mantenga la organización con la decisión que adopta el sistema. En este sentido, conforme este sujeto se aleja de la esfera de control de la organización que adoptó la decisión, más complejo resultará valorar si el sistema fue preciso en determinados valores. Vamos a representar un nuevo ejemplo que manifiesta lo indicado, concretamente, en este caso se

[690] Tal y como ha indicado la Information Commisioner´s Office británica. Información disponible en: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/what-do-we-need-to-do-to-ensure-lawfulness-fairness-and-transparency-in-ai-systems/

hace mención a un sistema automatizado que utiliza una entidad bancaria para decidir si conceder o no un préstamo.

	Seguimiento del modelo de concesión de préstamos	
FP	Consecuencia	Se concede un préstamo a una persona que no lo puede devolver
	Seguimiento	Se produce un impago por parte del particular
FN	Consecuencia	No se concede un préstamo a una persona que puede devolverlo
	Seguimiento	Es difícil realizar un seguimiento, salvo que se altere la decisión una vez que el particular la recurra
VP	Consecuencia	Se concede un préstamo a una persona que lo puede devolver
	Seguimiento	La persona devuelve el préstamo en el plazo fijado
VN	Consecuencia	No se concede el préstamo a una persona que no lo puede devolver
	Seguimiento	Difícil mantener el seguimiento

La complejidad de realizar un seguimiento de algunas métricas una vez el sistema se pone en marcha puede ser un factor a tener en cuenta por las organizaciones a la hora de potenciar la precisión de aquellas métricas en las que existe precisamente esa dificultad. Así, por ejemplo, si en la tabla anterior resulta difícil mantener una evaluación de los falsos negativos -denegación de un crédito a una persona solvente-, esa organización debería tratar de potenciar esa métrica para que sea lo más precisa posible en defecto de otras métricas que son más fáciles de evaluar una vez el sistema se pone en marcha. Y es que, en estos supuestos, el castigo para estas personas sería doble. Primero se ven perjudicados por la denegación de un crédito que se merecen y segundo, dado que existen más dificultades para medir esa métrica, el sistema nunca se corrige.

En definitiva, en base al principio de responsabilidad activa, el responsable del tratamiento ha de establecer cauces que permitan al mismo contabilizar todas las posibles alteraciones relacionadas con la precisión de estos sistemas una vez que estos se despliegan en el entorno y comienzan a irradiar sus efectos, pudiendo detectar aquellos ámbitos específicos donde el sistema puede arrojar mayores errores. En este sentido, el seguimiento específico de determinadas métricas puede alumbrar errores

muy relevantes presentes en el modelo algorítmico no detectado durante la fase de diseño. Así, por ejemplo, el algoritmo de YouTube puede presentar dificultades a la hora de diferenciar entre un contenido que vulnera la normativa de derechos de autor y un contenido que queda exceptuado de la misma. Ello ocurre por ejemplo cuando el usuario utiliza una canción que está protegido por la normativa de derechos de autor para realizar una parodia de la misma[691]. Así, la cita, la crítica, la sátira o la parodia son contenidos a los que no se les aplica las exigencias de esta normativa por considerarse límites legítimos a los derechos de autor[692]. Las organizaciones han de prestar especial atención a estos elementos y potenciar los mecanismos que puedan compensar estos problemas ya que la eliminación de estos contenidos supondrá una vulneración, y no una restricción, de sus derechos.

XI. LAS AUDITORÍAS EN PROTECCIÓN DE DATOS

La auditoría en materia de protección de datos de sistemas de toma de decisiones automatizadas se convierte en un mecanismo mediante el cual, un equipo auditor emite una opinión independiente, objetiva y documentado sobre el grado de cumplimiento de dicho sistema en relación con los principios y la normativa en materia de protección de datos[693], permitiendo que los responsables del tratamiento puedan en su caso tomar las medidas oportunas para subsanar aquellas deficiencias identificadas y atender en su caso a las observaciones que haya planteado el equipo auditor[694]. Además, estas auditorías también servirán como mecanismo que sustenta la confianza del sistema para terceros potenciales involucrados, ya sean tanto para particulares sometidos a las

[691] LESTER, T; PACHAMANOVA,D: "The Dilemma of False Positives: Making Content Id Algorithms More Conducive to Fostering Innovative Fair Use in Music Creation", op.cit.,pág.64.
También ha ocurrido este problema cuando el sistema ha identificado un contenido como contrario a los términos y condiciones de la plataforma y ello no es realmente así. Por ejemplo, el algoritmo de Facebook ha bloqueado reiteradamente fotos de grupos indígenas considerando que muestran desnudos. Dado que la publicación de fotos de personas desnudas es contraria a los términos y condiciones de esta plataforma, el algoritmo a veces no es capaza de distinguir. Visto en:
https://www.theguardian.com/technology/2021/may/28/facebook-accused-of-discriminatory-and-racist-behaviour-after-removing-historical-png-images

[692] Véase el Considerando 70 de 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.

[693] Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, págs.11 y 12.

[694] Véase el objeto de una auditoría en materia de seguridad de la información. En: Centro Criptológico Nacional. *Guía de auditoría. Guía de Seguridad de las TIC CCN-STIC 802*, apartado 9, pág.6.

decisiones como a potenciales adquirentes de estos modelos algorítmicos, pudiendo entrar en juego aquí las posibles certificaciones derivadas de procesos auditables.

Al igual que ocurre con otras medidas de responsabilidad activa, la auditoría de un sistema de toma de decisiones automatizadas generalmente requerirá un mayor o menor nivel de exhaustividad y control en función del riesgo que dicho sistema pueda generar a los derechos y libertades de los interesados[695]. Junto al riesgo, el alcance y la finalidad que se pretenda con la auditoría también serán determinantes para definir los elementos auditables de estos modelos algorítmicos. Por lo que se refiere al alcance, será necesario indicar las fases que comprenderán dicha auditoría, esto es, todas las etapas del ciclo de vida del sistema de toma de decisiones o sólo algunas. Es decir, se puede auditar más o menos tratamientos de datos. En los supuestos en los que la organización que diseñó el sistema sea distinta a la que lo quiere desplegar, el responsable del tratamiento que pretenda utilizar dicho sistema debería auditarlo antes de ponerlo en marcha para cerciorarse que el mismo cumple con la normativa de protección de datos. Todo ello en virtud del deber de diligencia que atañe al responsable a la hora de elegir encargados y productos algorítmicos que respeten la normativa de protección de datos[696].

Respecto a quién está obligado a realizar las auditorías, el RGPD guarda silencio. Únicamente el artículo 39.1.a) de dicha norma establece que corresponde al delegado de protección de datos la supervisión de dichas auditorías. Es por ello que en principio, estas auditorías puedan realizarse tanto por personal interno de la organización como por personas ajenas a la misma[697]. En este sentido, resulta conveniente que estas auditorías sean realizadas por personal externo de la organización ya que garantiza en mayor medida la independencia que caracteriza a este mecanismo. Así, las auditorías realizadas por organismos independientes han sido defendidas como instrumento de garantía que debería estar presente en todos los sistemas automatizados cuyas decisiones afecten a derechos fundamentales[698], resultando muy indicadas para la doctrina cuando se lleven a cabo los tratamientos de datos reconocidos por el artículo 22

[695] Artículo 24 del RGPD.
[696] Véase el Capítulo III, apartado III de esta tesis.
[697] Center for Democracy and Technology. *Algorithmic Systems in Education: Incorporating Equity and Fairness When Using Student Data*, 2019,pág.23.
[698] Grupo Independiente de Expertos de Alto nivel sobre *Inteligencia Artificial creado por la Comisión Europea. Directrices éticas para una IA fiable*.2019, pág.24.

del RGPD[699]. En este mismo sentido se ha pronunciado el Tribunal de la Haya cuando analizó el sistema SyRI. Este órgano judicial consideró como garantía mínima del funcionamiento de estos sistemas la realización de auditorías independientes[700]. En nuestra opinión, no existe una obligación a realizar auditorías por parte de un responsable del tratamiento. Sin embargo, teniendo en cuenta el enfoque del riesgo sobre el que se asientan todos los mecanismos de responsabilidad activa y las amenazas ya constadas que comportan el uso de sistemas automatizados, muy posiblemente las auditorias resulten una de las herramientas necesarias para mitigar los riesgos potenciales que se derivan de estos sistemas. Así, existen toda una serie de consecuencias negativas ligadas a la toma de decisiones automatizadas que resultan en muchos supuestos difíciles de detectar por parte de los particulares que se someten a dichas decisiones como puede ser la existencia de sesgos inherentes al sistema o alteraciones inesperadas del algoritmo. En estos casos, si el responsable no cuenta con medios adecuados para evitar o detectar estas posibles amenazas y reducir el riesgo de las mismas a un nivel aceptable, la auditoría se muestra como un mecanismo necesario para suplir esas carencias.

XII. LA FORMACIÓN DEL PERSONAL INTERNO DE LA ORGANIZACIÓN

Aunque a través de la incorporación de los procesos automatizados se tienda a la sustitución de la máquina por el hombre, lo cierto es que durante todo el ciclo de vida de los modelos algorítmicos el papel del personal de las organizaciones resulta sumamente relevante. Es tal ese protagonismo que muchas de las amenazas presentes en estos sistemas derivan del papel de estos agentes. Es por ello que entre las medidas de responsabilidad activa que traten de demostrar el cumplimiento de la normativa de protección de datos y reducir los potenciales riesgos se encuentre la formación adecuada del personal.

[699] Así se ha indicado con relación a las medidas y salvaguardas mínimas que han de estar presentes en este tipo de tratamiento tal y como indica el artículo 22.3 del RGPD. En: KAMINSKI,M; MALGIERI,G: "Algorithmic impact assessments under the GDPR: producing multi-layered explanations", *International Data Privacy Law*, 2020, ipaa020, pág.9. Disponible: https://doi.org/10.1093/idpl/ipaa020

[700] Sentencia del Tribunal de la Haya de 5 de febrero de 2020. Apartados. 6.97, 6.99 y en especial el 6.106. Resolución disponible en: :https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878

Así, por lo que se refiere a la *fase del diseño*, le corresponde al equipo que se encargará de elaborar el modelo definir el problema que se pretende resolver, indicar las variables que alimentarán al sistema, el peso de estas variables, las bases de datos elegidas, etc[701]. Todas estas decisiones deberían realizarse por personal debidamente cualificado. Dicho personal no ha de quedar limitado a aquellos perfiles que posean competencias técnicas e informáticas. En estos supuestos, la existencia de comités de ética puede ayudar a compensar esas carencias presentes en los grupos encargados de desarrollar los sistemas algorítmicos. A su vez, el delegado de protección de datos se puede convertir en el personal de la organización que aporte la experiencia jurídica en materia de protección de datos aplicada a los tratamientos que se pretenden o se estén llevando a cabo durante la elaboración del modelo algorítmico.

Por otro lado, en lo que respecta a la *fase de despliegue* del sistema, el personal encargado de analizar los resultados emitidos por el algoritmo ha de ostentar las competencias necesarias para interpretar y explicar las salidas. Las explicaciones que emita el personal de la organización a los particulares afectados por las decisiones han de ser personalizadas. Por ejemplo, este personal ha de ser consciente de las principales variables que potencialmente hayan intervenido en la formación de esa concreta decisión y adecuarse al contexto en el cual se haya adoptado la misma. Así, el artículo 22 reconoce a los particulares la intervención del personal de la organización para que le informe de la decisión adoptada[702], a la hora de facilitar esa explicación, esta ha de realizarse en un lenguaje adecuado al tipo de usuario sobre la cual se ha adoptado la decisión.

La formación debe también mitigar los principales riesgos que se derivan cuando el personal interactúa con los sistemas automatizados a la hora de interpretar los resultados. Entre los riesgos que se han detectado destacamos los siguientes:

En primer lugar el llamado sesgo de automatización o principio de autoridad de la máquina, es decir, el personal de la organización termina acatando frecuentemente los resultados que vierte el sistema sin espíritu crítico. Es decir, se genera una confianza

[701] Center for Democracy and Technology. *Algorithmic Systems in Education: Incorporating Equity and Fairness When Using Student Data*, 2019, pág.7.

[702] El artículo 22.3 del RGPD reconoce a los particulares el derecho a obtener intervención humana por parte del responsable del tratamiento tras la decisión adoptada por el sistema automatizado.

total hacia las salidas que emite el algoritmo[703]. Ello suele suceder en aquellos casos en los que los resultados del sistema confirman las creencias o hipótesis que puede tener ese trabajador sobre el caso específico sobre el que se toma la decisión. En este sentido, aquellos a los que se les asigna la función de interpretación de estos resultados han de partir siempre de una premisa básica, esto es, los sistemas de toma de decisiones no son infalibles[704]. Esta premisa debe reforzarse sobre todo en los supuestos en los que la decisión es impugnada por un particular. En este sentido, la PRAI prevé expresamente que en los sistemas de alto riesgo se garantice que las personas de la organización encargadas de supervisar estos sistemas sean conscientes de este sesgo[705].

En segundo lugar encontraríamos el efecto contrario, esto es, el personal de la organización muestra un alto recelo a los resultados que emite el modelo algorítmico[706]. En estos supuestos una total desconfianza a estos sistemas puede generar importantes riesgos para la esfera de los particulares teniendo en cuenta que en muchos contextos ya indicados, la ecuación máquina más humano se muestra como la ideal. En este sentido, la Carta de Derechos Digitales española establece que los particulares tendrán derecho a que se motive la decisión administrativa cuando esta se separe del criterio propuesto por un sistema automatizado o inteligente[707].

En definitiva, un adecuado plan formativo en favor de las personas que trabajan para las organizaciones que diseñan e implementan sistemas automatizados ayudará a prevenir algunos de los riesgos potenciales presentes en tales algoritmos. Ello a su vez beneficiará a los particulares sometidos a dichas decisiones automatizadas y favorecerá el cumplimiento de la normativa de protección de datos personales.

---

[703] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020.pág.8

[704] KEAT,CITRON,D: "Technological Due Process". *Washington University Law Review*, volume 85, issue 6, 2008, pág.1306. Disponible en: https://openscholarship.wustl.edu/law_lawreview/vol85/iss6/2/

[705] Artículo 14.4.b) de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la unión. Resolución de 21 de abril de 2021.

[706] Information Commissioner's Office y The Alan Turing Institute. *Explaining decisions made with AI*. 2020, págs. 78 y 79.

[707] Artículo 18.6.c) de la Carta de Derechos Digitales española.

CAPÍTULO IV. LOS PRINCIPIOS DE TRATAMIENTO DE DATOS PERSONALES EN EL USO DE SISTEMAS DE TOMA DE DECISIONES AUTOMATIZADAS

Los principios reconocidos en el artículo 5 del RGPD son un conjunto de reglas mínimas que han de estar presentes en cualquier tratamiento de datos personales. Junto a los derechos que se reconocen a los titulares de los datos y las medidas de responsabilidad activa que se imponen a los responsables del tratamiento, los principios son el tercer pilar sobre el que se estructura el derecho fundamental a la protección de datos[708]. Como ahora se comprobará, muchas de las exigencias que se derivan de estos principios coinciden en parte con los requisitos de robustez de los sistemas de toma de decisiones automatizados. La idea es que, a través del análisis de estos principios no sólo se busque una interpretación que proteja a los individuos frente a los sistemas de toma de decisiones automatizada sino que a la vez y en la medida de lo posible, se desarrollen modelos algorítmicos adecuados desde un punto de vista técnico.

I. EL PRINCIPIO DE LICITUD

El principio de licitud exige que todo tratamiento de datos que lleve a cabo el responsable esté amparado al menos por un mecanismo o base legítima que permita el mentado tratamiento. Este principio se reconoce en el artículo 5.1.a) y posteriormente se desarrolla en el artículo 6 del RGPD donde se establecen las seis bases de legitimación que habilitan a tratar cualquier dato personal.

A lo largo del ciclo de vida del desarrollo de los sistemas de decisiones automatizadas se llevan a cabo distintas operaciones de tratamiento de datos. Para la AEPD, cada una de las fases que engloba este proceso puede requerir de una base de legitimación distinta, esto es, recopilación de datos, entrenamiento, validación, etc[709].

[708] Una aproximación a los principios del derecho a la protección de datos puede encontrarse en: PUYOL MONTERO, F, J: "Los principios del derecho a la protección de datos". En PIÑAR MAÑAS,J,L (dir): *Reglamento Europeo de Protección de Datos. Hacia un nuevo modelo europeo de privacidad*. Ed. Reus, Madrid, 2016, págs.135 a 150. A su vez, véase: PUENTE ESCOBAR,A: "Principios y licitud del tratamiento". En RALLO LOMBARTE, A (dir): Tratado de Protección de Datos. Ed. Tirant lo Blanch, Valencia, 2019, págs. 115 a 169. También puede verse un análisis de los principios del tratamiento en las Administraciones Públicas en: TRONCOSO, REIGADA,A: *La Protección de Datos Personales .En Busca del Equilibrio*. Ed. Tirant lo Blanch, Valencia, 2011, págs 393 y ss.

[709] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.20.

En este apartado, nosotros nuevamente nos centraremos en la división que venimos realizando a lo largo de todo este estudio, es decir, analizaremos cada una de las bases de legitimación adecuadas para las dos grandes fases que engloba el uso de sistemas de decisiones automatizados, esto es, la etapa de diseño y la de despliegue[710].

Es turno por tanto de ir analizando una a una las distintas bases de legitimación que configura el RGPD para tratar los datos personales cuando se utilizan sistemas de decisiones automatizadas.

1. El consentimiento del interesado en los tratamientos de datos presentes durante el ciclo de vida de los sistemas de toma de decisiones automatizadas

El artículo 4, apartado 11, del RGPD define el consentimiento como: *«toda manifestación de voluntad libre, específica, informada e inequívoca por la que el interesado acepta, ya sea mediante una declaración o una clara acción afirmativa, el tratamiento de datos personales que le conciernen»*.

De esta manera, el consentimiento, de acuerdo a lo exigido por este precepto ha de otorgar a los interesados el poder efectivo para decidir sobre si sus datos personales son o no tratados. Si no es así, el control del interesado no puede considerarse real y como es lógico, la base del consentimiento en la que se sustenta el tratamiento carecerá de respaldo normativo[711]. Es turno de analizar este mecanismo de legitimación poniendo el foco en las principales implicaciones relacionadas con el ciclo de vida de los sistemas automatizados.

A) Consentimiento libre

El primer requisito referido al consentimiento es que este sea libre. Es decir, este consentimiento será válidamente emitido cuando el interesado que emite la declaración de voluntad pueda de manera efectiva ejercer una opción sobre la concesión del mismo sin que ello genere consecuencias negativas en la esfera del titular de los datos cuando

[710] Este último enfoque es el que propone la ICO a la hora de afrontar el principio de licitud en los tratamientos que engloban técnicas basadas en inteligencia artificial. Information Commissioner's Office: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/what-do-we-need-to-do-to-ensure-lawfulness-fairness-and-transparency-in-ai-systems/

[711] Comité Europeo de Protección de Datos. *Directrices 5/2020 sobre el consentimiento en el sentido del Reglamento (UE) 2016/679*, versión 1.1. Directrices adoptadas el 4 de mayo de 2020, págs. 7 y ss.

este último se niegue a otorgarlo. Así, el Considerando 43 del RGPD establece varias presunciones en las que no se considera que el consentimiento ha sido libremente concedido. Estas son:

En primer lugar, el consentimiento se presume que no es válido cuando el mismo se haya otorgado en una situación donde exista un desequilibro claro entre el interesado y el responsable del tratamiento. Ello puede suceder cuando el responsable sea una autoridad pública o un empleador. En este último caso, dada la jerarquía existente entre el empleador y el empleado, el consentimiento otorgado por el trabajador se considera improbable que sea libre ya que este último siempre tendrá el temor o el riesgo de que su negativa a conceder tal consentimiento le produzca efectos perjudiciales[712].

En segundo lugar, se presume que el consentimiento no ha sido libremente otorgado cuando el mismo no permita autorizar por separado las distintas operaciones de tratamiento de datos personales sobre las cuales se pretenden autorizar los tratamientos. Es decir, en aquellos supuestos en los que el responsable pretenda tratar datos para distintas finalidades, el responsable, si pretende legitimar dichos tratamientos por la vía del consentimiento deberá establecer de forma individualizada o granular la solicitud del consentimiento para cada una de estas finalidades. El objetivo pretendido es que sean los particulares los que decidan para qué finalidad autorizan el tratamiento de sus datos y para cuáles no. Esta exigencia resulta sumamente relevante por ejemplo a la hora de autorizar el uso de cookies. Normalmente, las cookies pueden ser recopiladas para muchas finalidades, muchas de estas finalidades tiene una fuerte incidencia a la hora de desarrollar e implementar sistemas de toma de decisiones automatizadas además

[712] Comité Europeo de Protección de Datos. *Directrices 5/2020 sobre el consentimiento en el sentido del Reglamento (UE) 2016/679*, versión 1.1. Directrices adoptadas el 4 de mayo de 2020, págs. 8 y 9.
Así, cabe destacar que en el Estado de Illinois se permite a las empresas el uso de técnicas de reconocimiento facial en los procesos de selección de potenciales aspirantes siempre que, se informe inicialmente al candidato de las principales características del sistema algorítmico y además, este candidato preste su consentimiento, en el supuesto de que este último se niegue a prestarlo, el empresario no podrá hacer uso de esta herramienta. Artificial Intelligence Video Interview Act, Public Act 101-0260,1/1/2020. Texto accesible en: https://www.ilga.gov/legislation/publicacts/fulltext.asp?Name=101-0260 .En Europa, el consentimiento podría ser un mecanismo válido siempre que se ofrezca una vía alternativa a la entrevista realizada por la máquina , no obstante, dada la presunción que existe sobre la invalidez del consentimiento prestado, puede resultar complejo acudir a esta base de legitimación para legitimar este tratamiento. Y es que, si bien es cierto que en este supuesto aún no existe una relación laboral como tal, la necesidad por obtener ese trabajo puede empujar a dicho aspirante a aceptar la entrevista automatizada por temor a no ser contratado u ofrecer una imagen negativa.

de para la elaboración de perfiles, como es lógico, cada una de estas finalidades ha de presentarse de forma separada.

Finalidad de las cookies	Fase
Desarrollar y mejorar productos Mejorar la experiencia del usuario Utilizar estudios de mercado a fin de generar información sobre el público Almacenamiento y acceso a información de geolocalización para realizar estudios de mercado	Fase de diseño de sistemas automatizados
Crear un perfil para la personalización de contenidos Seleccionar anuncios personalizados Crear un perfil publicitario personalizado Seleccionar anuncios básicos Seleccionar contenido personalizado Enriquecer el perfil con información de terceros	Fase despliegue de sistemas automatizados

Elaboración propia a partir del análisis de las principales cookies utilizadas en los portales webs.

La granularidad del consentimiento está estrechamente relacionada con la necesidad de que se explicite específicamente la finalidad para la cual se otorga el mismo. En este sentido, es frecuente que en los proyectos de diseño de sistemas de aprendizaje no supervisado se desconozca la finalidad inicial del proyecto. Pese a dichas dificultades, el responsable ha de establecer al menos la finalidad general que se pretende con ese proyecto. Así, en el ámbito de la investigación científica, el considerando 33 del RGPD permite cierta flexibilidad a la hora de establecer la finalidad específica del proyecto de investigación que se pretenden llevar a cabo. Permitiéndose por tanto un consentimiento más amplio[713]. Para compensar esta restricción de los principios de licitud y limitación de la finalidad se habrán de establecer suficientes medidas de garantía las cuales serán analizadas en el apartado dedicado al principio de limitación de la finalidad.

En *tercer lugar*, el artículo 7.4 del RGPD, referido a las condiciones del consentimiento establece que, a la hora de evaluar si el consentimiento ha sido prestado libremente se tendrá en cuenta en la mayor medida posible el hecho de si, entre otras cosas, la ejecución de un contrato, incluida la prestación de un servicio, se supedita al

[713] También la norma española permite un consentimiento amplio para tratar los datos de salud. Disposición adicional decimoséptima, apartado 2, letra a) Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. La propia AEPD ya lo ha declarado también con relación al tratamiento de datos con fines científicos y en concreto biomédicos. Informe AEPD Nº: 073667/2018, pág.8. Disponible en: https://www.aepd.es/sites/default/files/2019-09/2018-0046-investigacion-biomedica.pdf

consentimiento del tratamiento de datos personales que no son necesarios para la ejecución de dicho contrato. Tal y como indica el CEPD, del considerando 43 y del artículo 7.4 del RGPD se extrae la conclusión de que[714], supeditar el cumplimiento de un contrato o la prestación de un servicio a una solicitud de consentimiento para el tratamiento de datos personales que no son necesarios para la realización de dicho contrato o servicio resulta muy inapropiado[715]. Ello es así porque indirectamente se estaría obligando al titular a permitir el tratamiento de datos que realmente no son necesarios para la prestación del servicio requerido por el interesado. En este sentido, se corre el riesgo de que dicho interesado acabe "aceptando", esto es, consintiendo el tratamiento de sus datos personales independientemente de que sean o no necesarios con el objetivo único de poder acceder al servicio o concertar el contrato.

A la hora de examinar el carácter necesario de esos datos respecto de ese contrato o servicio el responsable ha de analizar si existe un vínculo directo y objetivo entre el tratamiento de los datos y la finalidad de la ejecución del contrato[716]. Las actividades de tratamiento que sean necesarias para la ejecución del contrato o prestación del servicio utilizarán generalmente la base de legitimación prevista en el artículo 6.1.b) del RGPD. Por otro lado, si el responsable pretende acudir al consentimiento como mecanismo que autoriza el tratamiento de datos para otras finalidades distintas a la ejecución del contrato o la prestación de servicio ha de saber que, si condiciona dicho contrato o servicio al consentimiento, muy probablemente se

---

[714] La propuesta inicial que presentó la Comisión Europea sobre el RGPD no contemplaba este precepto. Sin embargo, este se incorporó tras su paso por el Parlamento Europeo. Concretamente, este precepto establecía que: *La ejecución de un contrato o la prestación de un servicio no podrán supeditarse al consentimiento respecto de un* tratamiento *de datos que no sea necesario para dicha ejecución o prestación con arreglo a lo dispuesto en el artículo 6, apartado 1, letra b)*. Véase la enmienda 101 del Proyecto de resolución legislativa del Parlamento Europeo sobre la propuesta de Reglamento del Parlamento Europeo y del Consejo 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 (Reglamento general de protección de datos) (COM(2012)0011 – C7-0025/2012 – 2012/0011(COD)). Disponible en: https://www.europarl.europa.eu/doceo/document/A-7-2013-0402_ES.html?redirect#title2

[715] Comité Europeo de Protección de Datos. *Directrices 5/2020 sobre el consentimiento en el sentido del Reglamento (UE) 2016/679*, versión 1.1. Directrices adoptadas el 4 de mayo de 2020, apartado 26,pág.9. El TJUE tuvo una gran oportunidad para analizar este precepto en un litigio que versaba sobre el uso de cookies y el consentimiento prestado. Sin embargo, decidió no pronunciarse ya que en las cuestiones prejudiciales que se plantearon no se aludió a este precepto. En: SENTENCIA DEL TRIBUNAL DE JUSTICIA (Gran Sala) de 1 de octubre de 2019, asunto C-673/17, caso Planet49, FJº 64. Resolución disponible en: https://curia.europa.eu/juris/document/document.jsf;jsessionid=43E617ABD8B5CB0E1051B9BCDD70A654?text=&docid=218462&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=24967676

[716] Comité Europeo de Protección de Datos. *Directrices 5/2020 sobre el consentimiento en el sentido del Reglamento (UE) 2016/679*, versión 1.1. Directrices adoptadas el 4 de mayo de 2020, apartado 30, pág.10.

considerará que dicho consentimiento no ha sido libremente prestado. Es por ello que en estos supuestos donde esté presente esa condicionalidad el consentimiento no sea un mecanismo idóneo. Esta conclusión resulta muy relevante ya que cada vez más organizaciones, siendo conscientes del valor que esconde los datos, recopilan estos últimos no sólo con la finalidad de prestar el servicio principal sino con otra serie de finalidades adicionales. Entre esas finalidades secundarias se encuentra el desarrollo de modelos algorítmicos o la mercadotecnia automatizada personalizada. En estos casos es habitual que la prestación del servicio o la ejecución del contrato queden condicionadas al otorgamiento del consentimiento para esas otras actividades. En estos supuestos, dicho consentimiento como norma general no sería válido[717]. Y es que, no podemos olvidar que en principio, las razones que empujan a un interesado a solicitar la prestación de un servicio no son precisamente recibir publicidad o fomentar el desarrollo de modelos algorítmicos sino recibir el servicio solicitado. Que consienta o no la mercadotecnia directa ha de ser una decisión libre y no quedar condicionada a la prestación de ese servicio. Esta problemática aumenta cuando la contraprestación del servicio o el contrato son los propios datos personales del interesado[718]. Así, la Directiva de suministros digitales considera legítimo el suministro de contenidos o servicios digitales por parte de un empresario a cambio de facilitar los datos personales del consumidor[719]. En estos contratos, los datos personales que se suministran para la

[717] Totalmente contrario a nuestra opinión se manifiesta GARCÍA-RIPOLL MONTIJANO,M: "El consentimiento al tratamiento de datos personales". En: GONZÁLEZ PACANOWSKA,I (Coord): *Protección de datos personales*. Ed. Tirant lo Blanch, Valencia, 2020, pág.124.

[718] Es decir, el dato personal se convierte en una especie de moneda de cambio. Así, a modo de ejemple, en 2019 Google, fruto de los problemas que tenía sus sistema de reconocimiento fácil de las persona de raza negra decidió contratar a una empresa para que se encargara de la recopilación de estos datos a efectos de mejorar el entrenamiento de sus sistemas. Sin embargo, el proceso de recopilación de estos datos por parte de dicha empresa fue muy polémico Así, se demostró que en muchos casos, la recopilación de rostros faciales se basó en el pago de cantidades inferiros a 5 dólares a colectivos como jóvenes, estudiantes o personas sin hogar. En otros supuestos, la recopilación se basaba en el engaño, por ejemplo, que su rostro era necesario recopilarse para una encuesta o como parte de un juego donde estas personas debían hacerse un *selfie*. Fuente de la noticia; CARRIE WONG,J: "Google reportedly targeted people with 'dark skin' to improve facial recognition". *The Guardian*. 03/10/2019. Disponible en: https://www.theguardian.com/technology/2019/oct/03/google-data-harvesting-facial-recognition-people-of-color

Estas técnicas sería contrarias a la normativa de protección de datos de acuerdo a lo indicado por el GT29 con relación al uso de sistemas biométricos. Así, este órgano indica que los datos personales no son bienes que puedan intercambiarse por un servicio. Por tanto, aquellos contratos que prevean o que ofrezcan un servicio bajo la condición de que una persona consienta el tratamiento de sus datos biométricos para otro servicio no puede servir de base jurídica para dicho tratamiento. En: Grupo del Artículo 29. *Dictamen 3/2012 sobre la evolución de las tecnologías biométricas.* Resolución adoptada el 27 de abril de 2012. Pág.12.

[719] Artículo 3.1 párrafo segundo de la DIRECTIVA (UE) 2019/770 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 20 de mayo de 2019 relativa a determinados aspectos de los contratos de suministro de contenidos y servicios digitales.

prestación del servicio digital se utilizan en gran parte para finalidades distintas a la prestación del mencionado servicio o contenido digital[720], entre las cuales, suelen estar presentes el desarrollo de modelos algorítmicos o la analítica masiva de datos. Pues bien, teniendo en cuenta lo indicado hasta ahora, si el suministro del contenido o el servicio digital queda condicionado al consentimiento otorgado por el particular para tratar esos datos con finalidades distintas a las propias de esos servicios o contenidos digitales, algo que resultará normalmente habitual, es muy probable que el consentimiento no se considere libremente prestado[721]. En este sentido, el artículo 6.3 de la LOPD de 2018 es bastante claro ya que no permite que la ejecución del contrato quede supeditada al consentimiento que otorgue el interesado para el tratamiento de datos personales para finalidades que no guarden relación con el mantenimiento, desarrollo o control de la relación contractual. Esto supone que, además de entrar en una probable contradicción con la norma europea previamente comentada[722], el legislador español directamente prohíbe a los responsable del tratamiento acudir al consentimiento como base de legitimación cuando estos pretendan tratar los datos para finalidades distintas de la que se derivan del contrato si este último queda condicionado al consentimiento de esas otras finalidades. Esto supone ir un paso más allá de las

[720] El Considerando 24 de la Directiva de suministros digitales establece expresamente*: la presente Directiva debe aplicarse en aquellos casos en que el consumidor abre una cuenta en una red social y facilita un nombre y una dirección de correo electrónico, y estos se utilizan para fines que no sean exclusivamente el suministro de los contenidos o servicios digitales, o distintos del cumplimiento de los requisitos legales. También debe aplicarse en aquellos casos en que el consumidor dé su consentimiento para que cualquier material que constituya datos personales, como fotografías o mensajes que cargue, sea tratado por el empresario con fines comerciales*.

[721] Tal y como indica la doctrina, el consentimiento no resulta el mejor mecanismo cuando los datos facilitados por el consumidor son tratados por el empresario con otras finalidades a cambio de acceder a contenidos o servicios digitales. En: GARCÍA PÉREZ, RM: "Bases jurídicas relevantes del tratamiento de datos personales en la contratación de contenidos y servicios digitales". *Cuadernos de Derecho Transnacional*, marzo 2020, Vol. 12, Nº 1, pág 20.
El propio CEPD ya advirtió también que al reconocerse por esta legislación los contratos de suministro digitales a cambio de datos personales se podría inducir a error a los proveedores de servicios ya que estos podrán llegar a creer que el procesamiento de datos basado en el consentimiento en este contexto es legalmente compatible en todos los casos, incluso cuando no se cumplen las condiciones para el consentimiento válido establecidas en el RGPD. Afectando con ello a la seguridad jurídica. Comité Europeo de Protección de Datos. *Opinion 8/2018 on the legislative package "A New Deal for Consumers"*. Resolución adoptada el 5 de octubre de 2018. Apartado 52, pág.17.

[722] No pretendemos entrar a analizar esa posible contradicción de la norma estatal con la europea, si bien cabe decir que la Directiva menciona al consentimiento como una de las posibles formas de legitimar los tratamientos de datos que regula. Entre esos tratamientos de datos cabe destacar aquellos se llevan a cabo en los contratos de suministros digitales a cambios de datos que pueden utilizarse para finalidades distintas a las que se derivan de esos suministros, la contradicción parece clara.
Una interpretación alternativa e interesante a estos preceptos que pretende armonizar el conflicto presente entre el dato como valor económico y el respeto del derecho fundamental a la protección de datos lo encontramos en: CASTILLO PARRILLA, J,A: "Los datos personales como contraprestación en la reforma del TRLGDCU y las tensiones normativas entre la economía de los datos y la interpretación garantista del RGPD". *La Ley mercantil*, Nº 82, Sección Consumo/ Doctrina, Julio 2021, págs 12 y ss.

previsiones contenidas en el RGPD. Así, mientras que la norma europea atribuye a esta conducta la presunción de consentimiento inválido, la norma española directamente lo prohíbe. Para salvar este escollo de la condicionalidad y permitir que el responsable demuestre que esta no existe, el CEPD establece que, se considerará libremente concedido el consentimiento cuando el responsable ofrezca a los interesados la posibilidad real de escoger entre un servicio que incluya el consentimiento para el uso de datos personales con fines adicionales y un servicio equivalente ofrecido por el mismo responsable que no suponga prestar el consentimiento para el uso de datos con fines adicionales[723]. Ese servicio alternativo no sólo se ha de ofrecer durante el inicio del contrato o la prestación del servicio sino que el mismo ha de mantenerse durante todo el tiempo que perdure. En nuestra opinión, si el servicio alternativo deja de ofrecerse, el nuevo consentimiento que se solicite será inválido ya que el mismo queda supeditado al mencionado contrato o servicio. En estos supuestos la presunción de invalidez es aún más patente. Y es que, dado que el particular ya venía recibiendo ese servicio, resulta más probable que este acabe consintiendo el tratamiento para otras finalidades distintas a las estrictamente necesarias con el objetivo de mantener el servicio o el contrato. En este sentido, la App de mensajería instantánea WhatsApp actualizó su política de privacidad. Aquellas personas que no actualizaron estas políticas vieron reducidas las funcionalidades que esta aplicación ofrecía hasta que la misma quedara casi impracticable[724]. El Comisionado de Hamburgo para la Protección de Datos y la Libertad de Información anunció el 11 de mayo de 2021 la emisión de una orden en base al artículo 66 del RGPD que prohibió provisionalmente a Facebook seguir procesando los datos de los usuarios de WhatsApp para sus propios fines por entender que el consentimiento que prestan los ciudadanos a las nuevas políticas de

[723] Comité Europeo de Protección de Datos. *Directrices 5/2020 sobre el consentimiento en el sentido del Reglamento (UE) 2016/679*, versión 1.1. Directrices adoptadas el 4 de mayo de 2020, apartado 37, pág.11.
La doctrina ha propuesto otras fórmulas para demostrar la no condicionalidad, así se viene a decir que, si el responsable *quisiera obtener en el mismo acto o negocio «otros datos» debería: (i) solicitarse de forma separada, distinguiendo bien los datos precisos para la ejecución de un contrato y los que no son necesarios [ex. art. 7.2 RGPD] (ii) identificarse la otra finalidad, que obviamente no será la de ejecución del contrato (iii) identificar en este «otro caso» la base jurídica que habilitaría el tratamiento, que sería el consentimiento del afectado [el art. 13.1.c) RGPD indica que debe informarse al afectado, entre otros extremos, de la base jurídica del tratamiento]*. En VILASAU SOLANA,M: "El consentimiento general y de menores". En RALLO LOMBARTE, A (dir): Tratado de Protección de Datos. Ed. Tirant lo Blanch, Valencia, 2019, págs. 204 y 205.

[724] Conforme a lo indicado por esta compañía, las limitaciones en las funciones disponibles para el usuario irán en aumento a medida que pasen las semanas sin aceptar dichas condiciones. En: https://faq.whatsapp.com/general/security-and-privacy/about-new-business-features-and-whatsapps-privacy-policy-update

privacidad no era libre[725]. A falta de que el asunto acabe resolviéndose, lo que demuestran este tipo de asuntos es que el consentimiento no siempre será la base de legitimación ideal para este tipo de tratamientos.

En definitiva, se presume que un consentimiento no ha sido libremente otorgado cuando un responsable establece como condición para concertar un contrato o prestar un servicio el tratamiento de datos personales que no son estrictamente necesarios para la ejecución de dicho contrato o servicio[726]. Esta presunción no implica a priori una prohibición absoluta para utilizar el consentimiento como base de legitimación en estos supuestos[727]. Ahora bien, si el responsable acude a este mecanismo deberá justificar que no existe esa condicionalidad, o existiendo, habrá de probar que esta última no invalida el consentimiento. Para justificar que no existe esa condicionalidad, entre otros cauces, junto al servicio o contrato que pretende obtener esos datos que no son estrictamente necesarios, el responsable debería ofrecer a los interesados un servicio alternativo que no requiera del tratamiento de esos datos de más. El objetivo final es evitar que el particular se vea abocado a consentir determinados datos que ni siquiera son necesarios para el servicio principal por el cual, el titular de los datos se interesó. Por supuesto, siempre quedarán para el responsable otros mecanismos de legitimación como puede ser el interés legítimo regulado en los apartados f) del artículo 6.1 del RGPD para legitimar las otras finalidades al servicio o al contrato[728]. Finalidades alternativas que en muchos casos tendrán como objetivo la recopilación de datos para desarrollar sistemas algorítmicos.

B) Manifestación de voluntad inequívoca

De acuerdo al considerando 32 del RGPD, el consentimiento debe otorgarse mediante un acto afirmativo claro que refleje una manifestación libre e inequívoca del interesado para que se traten sus datos personales. Es por ello que, el silencio, las

[725] Esta autoridad de control activó el procedimiento de urgencia previsto en el artículo 66 del RGPD por el cual, una autoridad de protección de datos puede intervenir y adoptar inmediatamente medidas provisionales destinadas a producir efectos jurídicos en su propio territorio, con un periodo de validez determinado que no podrá ser superior a tres meses con el objetivo de proteger los derechos y libertades de los interesados. Véase la Resolución del 11 de mayo de 2021. Resolución disponible en: https://datenschutz-hamburg.de/assets/pdf/2021-05-11-press-release-facebook.pdf

[726] Se trata de una presunción, es por ello que nunca puede entenderse como una prohibición absoluta.

[727] Conclusiones del Abogado General Sr. Maciej Szpunar presentadas el 21 de marzo de 2019, Asunto C-673/17, Planet49 GmbH contra Bundesverband. Apartado 98. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:62017CC0673&from=EN

[728] Véase el Capítulo IV, aparatado I, punto 6 de esta tesis.

casillas marcadas por defecto, los sistemas de exclusión voluntaria[729] o la inacción no puedan englobarse en esa actitud activa[730]. De esta manera, tal y como ha indicado el CEPD, navegar por un sitio web no puede considerarse una conducta sobre la que se pueda inferir una indicación de que el interesado desee manifestar su acuerdo con respecto a una operación de tratamiento propuesta[731]. En este sentido, cabe destacar que este órgano se ha preocupado especialmente en elevar las exigencias mínimas de validez del consentimiento durante el proceso de aceptación de las cookies en los entornos digitales. Así, la interacción de los usuarios con los muros de cookies o la no validez del consentimiento a través de la acción de *scroll* fueron expresamente incluidas en la actualización de las directrices de consentimiento para indicar que esas acciones nunca podían considerarse acciones afirmativas[732]. Esta aclaración resulta sumamente relevante ya que en muchas ocasiones la puerta de entrada de legitimación de grandes cantidades de datos tanto para el diseño de modelos algorítmicos como para la toma de decisiones automatizadas se deriva de la aceptación de las cookies previo al acceso a los sitios webs. Teniendo en cuenta los numerosos tratamientos de datos y las diversas finalidades que se pueden derivar de la aceptación de las cookies, asimilar el consentimiento al mero hecho de descender o interactuar en el sitio de una página web resultaba cuanto menos problemático.

[729] No puede considerarse un consentimiento libremente prestado cuando el mismo se otorgue a través de cláusulas de exclusión voluntaria en las que el interesado se ha de oponer expresamente al intercambio de sus datos personales. En Resolución AEPD Nº: PS/00187/2019, pág.7. Resolución disponible en: https://www.aepd.es/es/documento/ps-00187-2019.pdf

[730] Así, el TJUE consideró que el consentimiento dado mediante una casilla marcada por defecto no implica un comportamiento activo por parte del usuario de un sitio de Internet. SENTENCIA DEL TRIBUNAL DE JUSTICIA (Gran Sala) de 1 de octubre de 2019, asunto C-673/17, caso Planet49**,** FJº 52 y 57. Resolución disponible: https://curia.europa.eu/juris/document/document.jsf?text=&docid=218462&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=24990078

[731] Comité Europeo de Protección de Datos. *Directrices 5/2020 sobre el consentimiento en el sentido del Reglamento (UE) 2016/679*, versión 1.1. Directrices adoptadas el 4 de mayo de 2020, apartados 84 y 86, pág.18.

[732] El 10 de abril de 2018, el GT29 adoptó sus Directrices sobre el consentimiento en el sentido del Reglamento 2016/679 (WP259.01). Estas directrices fuero ligeramente modificadas por el Comité Europeo de Protección de Datos el 4 de mayo de 2020. Estas alteraciones se focalizaron esencialmente en los aspectos relativos a la validez del consentimiento facilitado por el interesado al interactuar con las denominadas barreras de cookies o *cookie walls*. Estas modificaciones llevaron a la propia AEPD a cambiar su criterio ya que hasta ese momento se consideraba como consentimiento válido continuar utilizando la página web o la aplicación. Véase por un lado la Guía sobre el uso de las cookies de noviembre de 2019 de la AEPD, pág.21, la cual autoriza esta conducta y por otro lado, la Guía sobre el uso de las cookies de julio de 2020 también de la AEPD, pág.30, la cual se adecúa a los nuevos criterios establecidos por el CEPD.

C) El consentimiento explícito

El consentimiento explícito se requiere en determinadas situaciones en las que el legislador europeo ha considerado que existe un grave riesgo para los derechos y libertades del particular. Para ello, cuando el responsable pretenda acudir a la vía del consentimiento para legitimar el tratamiento de datos, en determinados supuestos dicho consentimiento habrá de ser explícito ya que se entiende que es necesario conceder un nivel elevado de control en favor del interesado en relación con sus datos personales. Concretamente, y por lo que se refiere al ciclo de vida de los sistemas automatizados, el consentimiento explícito se exige en dos situaciones concretas; por un lado, cuando el responsable del tratamiento pretenda tratar datos de categoría especial de los contemplados en el artículo 9 del RGPD, ya sea en la fase de diseño o de despliegue, y, por otro lado, cuando pretenda llevar a cabo los tratamientos descritos en el artículo 22 del RGPD, es decir, las decisiones plenamente automatizadas relevantes, incluidas la elaboración de perfiles que se materialicen en dichas decisiones. En estos casos, el responsable del tratamiento puede seguir acudiendo a otras bases de legitimación, sin embargo, si pretende habilitar el tratamiento de datos bajo el consentimiento del interesado, este ha de ser explícito.

Al tildar de explícito el consentimiento para determinados tratamientos, el legislador europeo consideró que era necesario que el responsable llevara a cabo unos esfuerzos adicionales que probaran que este mecanismo era real. Es por ello que la obtención del mismo ha de ser reforzada respecto del consentimiento normal o estándar[733]. En este sentido, la AEPD ha considerado que dicho consentimiento explícito en estos contextos obliga al responsable a ofrecer alternativas equivalentes y viables a la decisión automatizada y además a garantizar que si el particular elige no ser objeto a la decisión automatizada, ello no le afecte[734].

Por su parte, el CEPD ha indicado que el término explícito se refiere a la manera en que el interesado expresa el consentimiento. Es decir, significa que el interesado debe realizar una declaración expresa de consentimiento. Entre las formas consideradas bajo esta fórmula encontramos; las declaraciones escritas, cumplimiento de impresos

[733] Así lo ha indicado la AEPD con relación al consentimiento explícito para la toma de decisiones plenamente automatizadas previstas en el artículo 22 del RGPD. Resolución AEPD Nº: PS/00037/2020, pág. 112. Resolución disponible: https://www.aepd.es/es/documento/ps-00037-2020.pdf

[734] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.28.

electrónico, documentos escaneados con firma, confirmaciones verbales que sean claras o incluso la verificación del consentimiento en dos fases. Estas técnicas pueden suponer importantes escollos cuando el responsable del tratamiento pretenda reutilizar los datos con otros fines distintos a los iniciales y pretenda utilizar como base del tratamiento el consentimiento. Para ello, la Propuesta de Reglamento de Gobernanza de datos prevé la posibilidad de que los organismos públicos ayuden a las organizaciones que pretenden reutilizar los datos personales que ostentan los poderes públicos facilitando la obtención del consentimiento de los titulares de esos datos para esas finalidades de reutilización[735].

D) La retirada del consentimiento

En lo que concierne a la retirada del consentimiento, el artículo 7.3 del RGPD establece que el responsable del tratamiento debe garantizar que el interesado pueda retirar su consentimiento en cualquier momento y además, dicha retirada ha de ser tan fácil como cuando el consentimiento se concedió. Tal y como establece el CEPD, si el particular retira el consentimiento, todas las operaciones de tratamiento que se basaban en el consentimiento y que hayan tenido lugar antes de dicha retirada seguirán siendo lícitas. En cambio, los tratamientos futuros se deberán paralizar. La retirada del consentimiento puede tener incidencia en las distintas fases que comprende el ciclo de vida de los sistemas de toma de decisiones.

Así, por lo que se refiere a la fase de diseño, se ha indicado que la retirada del consentimiento puede tener fuertes implicaciones en el desarrollo de modelos algorítmicos ya que, la eliminación de los datos que entrenaron al modelo puede afectar gravemente al diseño de estos. Además, la exclusión de esos datos del modelo puede resultar compleja desde el punto de vista técnico[736]. Este inconveniente en parte se ha resuelto por la AEPD al indicar que la retirada del consentimiento no tiene un efecto retroactivo con relación a los resultados obtenidos en un tratamiento ya realizado. De

[735] Artículo 5.apartado 6 de la Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos). Texto aprobado el 25 de noviembre de 2020.

[736] Entre las soluciones que se propone para atajar esta problemática desde el punto de vista técnico se encuentran el de aislamiento o supresión de la línea de aprendizaje que incorporó los datos ahora no consensuados o el reentrenamiento de los modelos de IA existentes utilizando los conjuntos de datos modificados. MITROU,J: "Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?", 2018, pág.40. Texto disponible en: https://ssrn.com/abstract=3386914 or http://dx.doi.org/10.2139/ssrn.3386914

esta manera, la retirada del consentimiento no invalida la explotación del modelo[737]. Se prima por tanto en estos casos los intereses de la organización sobre los del particular debido a que la primera creo un modelo amparado en el consentimiento del segundo. Y es que, la mera presencia de datos personales en un modelo no equivale en modo alguno al ejercicio pleno e ilimitado de derechos sobre ellos[738]. Ahora bien, en nuestra opinión, si se realizan futuras actualizaciones del modelo o este se reentrena, la retirada del consentimiento sí que obligará a eliminar esos datos personales del mismo ya que este se ha alterado. Por otro lado, cuando los datos aún no hayan sido procesados o dichos datos continúen en los *data lakes* para futuros desarrollos de modelos algorítmicos, la retirada del consentimiento sí que afectará a dichos datos y los mismos se habrán de eliminar. Es decir, los modelos que se hayan creado no quedarán afectados por esa retirada del consentimiento pero los datos no procesados o los que se mantengan almacenados sí que se deberán suprimir. La supresión de los datos será analizada con mayor detenimiento en el epígrafe destinado al derecho de supresión[739].

La retirada del consentimiento también tendrá su incidencia en la fase de despliegue, así, esta revocación obligará al responsable a paralizar los perfiles que estuviera desarrollando sobre el particular.

E) Limitaciones e inconvenientes del consentimiento

A pesar de que el consentimiento suponga a priori la base de legitimación que más capacidad de control otorga al titular sobre el uso de sus datos personales, hace ya algún tiempo que se duda sobre si este es el mecanismo más adecuado para habilitar el tratamiento de datos, sobre todo, en los contextos tecnológicos. Las razones son varias:

En primer lugar se indica que, dada la especial complejidad que comprenden muchos de los tratamientos de datos que hoy día se realizan y el número importante de obligaciones que en materia de transparencia se exigen al responsable sobre dichos tratamientos, los avisos de privacidad suelen estar compuestos por textos largos y difíciles de entender con un alto contenido técnico-jurídico. Esta situación lanza a los

737 Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.21.
738 WACHTER, S; MITTELSTADT, B: "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI". op.cit, pág.109.
739 Véase el Capítulo V, apartado II, punto 3.

particulares a aceptar sin más el tratamiento de datos sin llegar a ser conscientes realmente de los usos que están consintiendo[740].

En *segundo lugar* se ha señalado que el consentimiento en los entornos digitales se convierte en una especie de ritual o costumbre por el cual, los interesados que acuden a los sitios webs y aplicaciones tecnológicas aceptan sin más los avisos de privacidad como un paso previo al acceso a dichos servicios. Esa especie de tradición socava los cimientos sobre los descansan todos los requisitos establecidos por el RGPD en relación con el consentimiento previamente explicados.

En *tercer lugar*, la obtención de ese consentimiento suele enmarcarse en un entorno predefinido que empuja al interesado a aceptar el tratamiento de datos. Esta situación se da a través de los llamados *dark paterns* comúnmente presentes en los avisos de privacidad previos a la aceptación de las cookies. Los patrones oscuros están formados por herramientas e interfaces digitales cuidadosamente diseñadas para dirigir a los usuarios a que hagan aquello que la organización que los diseña pretenda[741]. En este caso, aceptar las políticas de privacidad. Son diversas las formas en las que se manifiestan estos patrones en los entornos digitales destacando el uso de lenguaje confuso, priorización de las opciones destinadas a la aceptación de las cookies sobre aquellas que habilitan a configurarlas o a rechazarlas, opciones preseleccionadas, etc. En el Estado de California los patrones oscuros están prohibidos expresamente por la *California Consumer Privacy Act of 2018*[742]. En Europa, aunque no existe tal prohibición, lo cierto es que, el uso de estas técnicas engañosas para obtener el consentimiento va en contra de los requisitos mínimos que este último ha de ostentar. A

---

[740] TRUJILLO CABRERA,C: "Aproximación a la regulación del consentimiento en el Reglamento General de Protección de Datos". *Anales de la Facultad de Derecho*, 34; septiembre 2017, pág.75.

[741] Comité de ética alemán. Gutachten der Datenethikkommission, 2019, pág.97.
Para más información sobre los patrones oscuros en: Fuente de la noticia: MATUSZEWSKA,K: "When design goes away – How dark patterns conflict with GDPR and CCPA". *Piwik*, 26/04/2021. Disponible en: https://piwik.pro/blog/how-dark-patterns-conflict-with-gdpr-ccpa/

[742] Concretamente el texto indica que:
Issuing regulations to govern how a business that has elected to comply with subdivision (b) of Section 1798.135 responds to the opt-out preference signal and provides consumers with the opportunity subsequently to consent to the sale or sharing of their personal information or the use and disclosure of their sensitive personal information for purposes in addition to those authorized by subdivision (a) of Section 1798.121. The regulations should….: Does not make use of any dark patterns.
Véase la California Consumer Privacy Act of 2018, 1798.185. 20, C) iii)
https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5
En este mismo sentido, a nivel federal destaca un proyecto de ley conocido como *Deceptive Experiences To Online Users Reduction Act* que pretende prohibir el uso de prácticas de explotación y engaños por parte de grandes operadores en línea al consumidor. Véase: S. 1084 - 116 ° Congreso: Ley de reducción de experiencias engañosas para usuarios en línea. 2019. Proyecto de ley disponible en: https://www.govtrack.us/congress/bills/116/s1084

ello hay que sumarle además que no sólo quedaría afectado el principio de licitud, sino también los de transparencia y lealtad en el tratamiento de los datos personales conforme al artículo 5.1.a) RGPD.

Pues bien, los problemas comentados previamente se ven acrecentados en el contexto del tratamiento y análisis masivo de datos. En este sentido, la complejidad y el uso transformador de los macrodatos limitan las posibilidades reales que tienen los interesados de comprender los posibles usos futuros de sus datos personales[743]. Así, por ejemplo, en la fase de diseño de los sistemas automatizados los particulares pueden consentir el tratamiento de sus datos para la elaboración de modelos algorítmicos. Estos modelos en innumerables ocasiones ni si quiera se aplicarán a aquellos individuos cuyos datos se utilizaron para entrenar y desarrollar el sistema sino a terceras personas que en su caso pueden presentar características similares o no a las del perfil de los sujetos cuyos datos fueron utilizados. Es decir, el consentimiento de unos para diseñar los modelos algoritmitos afecta a terceras personas cuando estos sistemas comienzan a adoptar decisiones sobre estas últimas. Por otro lado, en la fase de despliegue, esto es, la toma de decisiones, los interesados habitualmente desconocen qué supone para ellos la elaboración de un perfil personal o grupal y qué efectos o consecuencias se derivan del mismo. Siempre que exista una pretensión de personalización en el tratamiento de datos, el interesado debería ser consciente de qué supone dicha personalización, ya sea de anuncios, de contenidos digitales, de precios, etc.

De lo dicho hasta ahora queda claro por tanto que el consentimiento obtenido en el contexto digital, generalmente a través de las *cookies*, donde se autorizan multitud de tratamientos de datos con importantes repercusiones para el diseño y despliegue de sistemas de toma de decisiones automatizadas resulta cuanto menos complejo. Es por ello necesario establecer toda una serie de propuestas.

Por un lado, y como ya indicábamos anteriormente, los interesados han de ser conscientes de los procesos de personalización que sufren desde el momento que aceptan las políticas de privacidad. Ello ha de ser extensible no sólo a la fase inicial de recopilación sino durante todo el proceso que abarca la misma. Es decir, el particular ha de ser consciente de que el anuncio, el precio, la noticia o el contenido que recibe es personalizado y además ha de conocer los elementos básicos de esa personalización. En este sentido, la propuesta de Directiva de servicios digitales impone determinados

[743] MITROU,J: "Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?", op.cit.,pág.39.

deberes de trasparencia a las plataformas en relación con la publicidad personalizada. Así, por cada anuncio que se muestre, las plataformas están obligadas a señalar que: i) el contenido mostrado es un anuncio, así como la empresa que muestra el mismo ii) información significativa sobre los principales parámetros utilizados para determinar el destinatario al que se muestra el anuncio[744]. Esta previsión normativa es un paso importante en la transparencia sobre los cometidos personalizados y habría de extenderse a otros contextos. Resulta frecuente que los particulares desconozcan los efectos que se derivan de la aceptación de las cookies o las políticas de privacidad de las plataformas digitales. Mecanismos a posteriori como el previsto en esta propuesta legislativa permiten a los interesados conocer los efectos y consecuencias reales que suponen para estos particulares la aceptación de esas condiciones de privacidad. Junto a esos mecanismos de información se debería facilitar además un mecanismo ágil de retirada del consentimiento, de manera que, si el particular considera molesta o perjudicial la mentada personalización, este pueda revocar el consentimiento con la misma facilidad que tuvo para aceptar las cookies.

Por otro lado, si partimos de una visión negativa en la que consideramos que los avisos de *cookies* no se leen y se produce una aceptación del particular de las mismas sin que exista conciencia real sobre lo que se acepta, como mínimo, estas políticas de privacidad deberían permitir a los interesados tener la misma facilidad tanto para rechazarlas como para aceptarlas. Es decir, por defecto, los avisos de privacidad deberían permitir a los interesados en plano de igualdad elegir entre; i) la aceptación de las cookies, ii) el rechazo de las mismas o, iii) configurar una elección granulada de los tratamientos que se solicitan. También tendrían encaje aquellas que únicamente dan la opción inicial de aceptar o configurar las cookies, siempre que al aceptarlas, el interesado por defecto sólo permita las cookies estrictamente necesarias. En nuestra opinión, aquellas configuraciones que prioricen la aceptación de las cookies o castigue el rechazo de las mismas no pueden considerarse adecuadas para legitimar el tratamiento de datos por la vía del consentimiento. En este sentido, el TJUE ha indicado que no se considera demostrado que el consentimiento ha sido válidamente otorgado

[744] Artículo 24 de la 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. Resolución de 15 de diciembre de 2020.

cuando el responsable pone trabas burocráticas a los particulares que se niegan a consentir un determinado tratamiento de datos personales[745].

Forma de presentar las cookies. Elaboración propia

A su vez, y con el objetivo de facilitar un control sobre cómo los responsables recopilan los datos a través de las cookies, sería conveniente que estos últimos, en base al principio de responsabilidad activa guarden copias de las interfaces y métodos que han utilizado a la hora de recopilar los consentimientos de los interesados. De esta manera, las autoridades de control, cuando hayan de evaluar si un consentimiento ha sido otorgado conforme al RGPD puedan valorar si en el momento que se concedió dicho consentimiento se habían utilizado o no prácticas adecuadas para obtenerlo, lo que permitirá detectar posibles patrones oscuros o prácticas engañosas.

Finalmente cabe decir que el consentimiento no deja de ser uno de los mecanismos que habilita el tratamiento de datos personales, el responsable puede acudir al resto de bases comprendidas en el artículo 6 del RGPD para tratar de legitimar su tratamiento de datos. En las páginas siguientes analizamos estas bases.

2. La existencia de un contrato como base para legitimar la toma de decisiones automatizadas

De acuerdo al artículo 6.1.b) del RGPD el tratamiento será lícito *cuando sea necesario para la ejecución de un contrato en el que el interesado es parte o para la aplicación a petición de este de medidas precontractuales*. Son por tanto dos situaciones las que permite a los responsables legitimar su tratamiento de datos a través de este mecanismo. En primer lugar, esta disposición engloba aquellos supuestos en los que el tratamiento es necesario para la ejecución de un contrato en el que el interesado

[745] En este caso, el interesado que se negara a dar su consentimiento debía cumplimentar un formulario adicional en el que hiciera constar dicha negativa al tratamiento de sus datos. Esta forma de proceder llevaba a que muchos particulares acabaran consintiendo el tratamiento de datos por las trabas que ponía la empresa cuando un particular se negaba a consentir el tratamiento. Sentencia del TJUE (Sala Segunda) de 11 de noviembre de 2020, asunto C-61/19, caso Orange Romania SA / ANSPDCP. FJº 50 y 52. https://curia.europa.eu/juris/document/document.jsf?text=&docid=233544&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=13705373
La doctrina ha considerado incluso que no se deberían permitir opciones en los que los particulares puedan aceptar por defectos todas las cookies. En: SORIANO, ARNANZ, A: "Decisiones automatizadas: problemas y soluciones jurídicas. más allá de la protección de datos". op.cit.,pág.122.

es parte. Esto se da por ejemplo en las relaciones contractuales entre empleador y trabajador o empresario y cliente de un servicio[746]. Así, en el ámbito laboral, durante el transcurso de la relación de trabajo, el empresario puede adoptar decisiones automatizadas que afecten a los trabajadores con relación a derechos de ascenso, renovación o extinción de su contrato[747]. En segundo lugar, también quedan englobados en este mecanismo de legitimación aquellos tratamientos que sean necesarios para tratar datos personales antes de celebrar el contrato a fin de facilitar la propia celebración del mismo. Pueden quedar englobados aquí el uso de sistemas de evaluación de riesgos de solvencia previo a la concesión de un préstamo o el uso de sistemas automatizados para cribar a aspirantes a una oferta de trabajo en un proceso de selección de personal.

A) El carácter necesario

Uno de los elementos básicos contenidos en esta base de legitimación es el referido a cuándo se ha de considerar que el tratamiento es necesario para la ejecución del contrato o para las medidas precontractuales. El carácter necesario es un concepto jurídico indeterminado que ha de ser aclarado para evitar un uso inadecuado de esta base de legitimación[748]. En este sentido, el CEPD ha fijado una serie de pasos que se han de seguir a la hora de valorar dicho dicha necesidad del tratamiento. En primer lugar, se ha de comenzar por identificar la finalidad del tratamiento, para ello se ha de determinar la justificación exacta del contrato, es decir, su esencia y el objetivo fundamental[749]. Seguidamente, en segundo lugar, corresponde al responsable del tratamiento demostrar que el objeto principal del contrato o de la prestación del servicio no puede alcanzarse si no se lleva a cabo el tratamiento concreto de los datos personales

[746] En el contexto laboral, en la mayoría de las ocasiones el consentimiento se entiende que no es válido. Grupo del artículo 29. *Dictamen 2/2017 sobre el tratamiento de datos en el trabajo.* Adoptado el 8 de junio de 2017**.** Pág.6.

[747] Agencia Española de Protección de Datos. *La protección de datos en las relaciones laborales*. 2021, pág.27.

[748] Así lo indicó el Gobierno Neerlandés durante la preparación de la posición del Consejo sobre la evaluación y revisión del Reglamento General de Protección de datos. En: General Secretariat of the Council. *Preparation of the Council position on the evaluation and review of the General Data Protection Regulation (GDPR)*. Resolución del 9 de octubre de 2019, pág.50. Documento disponible en:https://data.consilium.europa.eu/doc/document/ST-12756-2019-REV-1/en/pdf

[749] Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, pág.20.

en cuestión que se pretende ejecutar[750]. Para ello, entre otros elementos, antes de proceder al tratamiento se ha de evaluar y comprobar que no existen otras alternativas realistas y menos intrusivas. De esta manera, en caso de existir otros medios efectivos y menos invasivos para lograr el mismo objetivo, el tratamiento no será considerado necesario[751]. Así, por ejemplo, resulta cada vez más habitual el uso de sistemas de reconocimiento facial automatizados basados en técnicas de IA en los procesos de reclutamiento de personal con el objetivo de encontrar al perfil idóneo teniendo como referencia las respuestas y el comportamiento de la persona en las entrevistas a través de su estudio del rostro[752]. Muy posiblemente, el uso de estas herramientas exceda de ese carácter necesario por existir medios menos invasivos para las personas. Es por ello que la mayor eficiencia de un modelo algorítmico respecto de una persona a la hora de adoptar decisiones no justifica por si sola el carácter necesario de ese tratamiento[753], se ha de demostrar que efectivamente el ingreso de ese modelo algorítmico es necesario para que se pueda llevar a cabo el objeto esencial del contrato o la prestación del servicio. Así, a modo de ejemplo, el GT29 ha indicado que sería necesario para el objeto del contrato establecer un sistema automatizado de filtrado de solicitudes en procesos selectivos en los que hayan solicitado dicho trabajo un número muy alto de potenciales candidatos[754]. Este filtro sería necesario para realizar un cribado inicial de las

---

[750] Comité Europeo de Protección de Datos. *Directrices 2/2019 sobre el tratamiento de datos personales en virtud del artículo 6, apartado 1, letra b), del RGPD en el contexto de la prestación de servicios en línea a los interesados*. Versión 2.0. Aprobadas el 8 de octubre de 2019. Apartado 30 *in fine*.. Págs. 10 y 11, Este mismo órgano ha señalado que: ***el** responsable debe estar en condiciones de justificar la necesidad del tratamiento concreto haciendo referencia al fin fundamental del contrato tal como lo entienden las dos partes del mismo. Esto no sólo depende del enfoque del responsable del tratamiento, sino también del punto de vista razonable que tuviera el interesado a la hora de celebrar el contrato y de si este aún puede considerarse «ejecutado» sin el tratamiento en cuestión.* Pág.11, apartado 32 de las mismas directrices.

[751] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*. Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.26

[752] En Corea del Sur cada vez se está implantando más el uso de estos sistemas automatizados en los procesos de reclutamiento de personal. Para ello, existen ya centros formativos especializados que preparan a los aspirantes con diversas estrategias para poder lidiar con un sistema automatizado y así tener más posibilidades de contratación. En: Fuente de la noticia: CHA,S: "Smile with your eyes': How to beat South Korea's AI hiring bots and land a job". *Thomson Reuter*. 13/01/2020. Disponible en: https://www.reuters.com/article/us-southkorea-artificial-intelligence-jo/smile-with-your-eyes-how-to-beat-south-koreas-ai-hiring-bots-and-land-a-job-idUSKBN1ZC022

[753]Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*. Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.14. La doctrina considera por su parte considera que sólo cuando el sistema algorítmico sea como mínimo tan preciso como un humano este estará justificado. En: BERMAN, E:"A government of laws and not of machines". *B.U. L. RE*, op.cit., pág.22

[754] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*. Adoptadas el 3 de octubre de 2017 y revisadas el 6 de

solicitudes irrelevantes. En términos similares, se ha considerado también necesario para un contrato de seguro el análisis de los hábitos de conducción de un cliente a la hora de establecer la prima cuando el precio de la misma dependa del comportamiento de esa persona en el vehículo[755]. Recordemos que la valoración de la necesidad de las operaciones del tratamiento respecto de la finalidad del mismo ya había sido evaluada a la hora de analizar la EIPD, es por ello que lo indicado en esas líneas pueda resultar útil para justificar el carácter necesario del contrato[756].

Resulta por tanto fundamental acreditar el carácter necesario del tratamiento para poder acudir a la base de legitimación del artículo 6.1.b). Cuando el responsable tenga dudas, habrá de valorar los posibles mecanismos de licitud que le brinda el artículo 6 y optar por aquel que sea más acorde a la situación. Es turno de analizar algunos ejemplos donde se han presentado dudas sobre el uso del mecanismo de este mecanismo de legitimación.

Así, en *primer lugar* y por lo que se refiere a los perfilados que puede realizar una entidad bancaria. La AEPD, en uno de sus informes jurídicos deja clara la diferencia entre el uso de la base de legitimación de la ejecución del contrato (Artículo 6.1.b) y el interés legítimo (Artículo 6.1.f). De esta manera, cuando un particular solicite un producto financiero a una entidad bancaria, esta legitimará el análisis de la solvencia financiera de esa persona a través del mecanismo previsto en el artículo 6.1.b) ya que existe una disposición legal que obliga a este responsable a realizar tal evaluación previa a la concesión de ese préstamo o crédito, desprendiéndose de ello el carácter necesario de ese tratamiento en la formalización del contrato. Sin embargo, cuando esa misma entidad quiera utilizar esos datos para realizar perfilados personalizados para ofrecer nuevos productos financieros a esos clientes, dado que el particular no es el que ha iniciado la solicitud del producto, el mecanismo de legitimación sobre el cual descansará dicho tratamiento será el consentimiento o el interés legítimo ya que dicho tratamiento no es necesario para el correcto funcionamiento de la ejecución del

---

febrero de 2018, pág.26. También lo ha señalado recientemente la AEPD en: Agencia Española de Protección de Datos. *La protección de datos en las relaciones laborales*. 2021, pág.24.

[755] Los contratos conocidos como *Pay As You Drive* basan el coste de la prima en función de cómo esa persona conduce. Es por ello que, dado que dicho tratamiento es totalmente necesarios para que se pueda ejecutar el contrato, los mismos resultan plenamente encajables en el apartado 6.1.b).En: Comité Europeo de Protección de Datos. *Guidelines 1/2020 on processing personal data in the context of connected vehicles and mobility related applications*. Versión 1.0 Directrices adoptadas el 28 de enero de 2020. Apartado 103, págs 21 y 22.

[756] Capítulo III, apartado II, punto 5 de esta tesis.

contrato[757]. Pese a la rotundidad de la AEPD mostrada en esta resolución, el GT29 ha señalado que cuando exista una obligación legal hacia el responsable para llevar a cabo un tratamiento, la base de legitimación ideal habría de ser la prevista en el artículo 6.1.c (obligación legal del responsable) y no la referida a la ejecución del contrato[758]. Existen por tanto dudas sobre el mecanismos de legitimación. Base de legitimación que siempre se deberá justificar.

Por otro lado, *en segundo lugar*, no existen tantas dudas con relación al carácter o no necesario de la publicidad compartimental en los servicios digitales. Así, muy difícilmente se podrá acreditar ese carácter necesario de la misma, y es que, si bien es cierto que la publicidad puede ser en muchos casos el sustento en el que se basan un número importante de contratos y servicios digitales actuales, en general, esta no es necesaria para el objeto del contrato o la prestación del servicio que se ofrece. En estos supuestos será generalmente el consentimiento o el interés legítimo las bases legales adecuadas.

Por último, *en tercer lugar*, quedan en duda el carácter o no necesario de la personalización de los contenidos que ofrecen los servicios digitales, tales como sistemas de recomendación de noticias, videos, imágenes, etc. Así, el CEPD viene a indicar que es posible que la personalización de contenidos forme parte del carácter necesario del objeto del contrato. Cuando la personalización del contenido no sea necesaria desde el punto de vista objetivo para los fines del contrato subyacente (por ejemplo, cuando el contenido personalizado ofrecido tenga por objeto incrementar el uso del servicio por el usuario pero no forme parte esencial del uso del servicio), los responsables del tratamiento de los datos deberán examinar la posibilidad de usar un fundamento jurídico alternativo[759]. El elemento clave girará en valorar si la personalización de contenidos realizada por algoritmos que se ofrece es o no esencial para ese servicio. Así, en nuestra opinión, si un periódico digital ofrece noticias personalizadas a los suscriptores que han contratado el mentado servicio, la

[757] Agencia Española de Protección de Datos. Informe jurídico 0195/2017. Págs.10 a 12. Resolución disponible en: https://www.aepd.es/es/documento/2017-0195.pdf

[758] Así lo ha indicado el GT29 tanto para las evaluaciones de solvencia de crédito como para el cálculo de las pólizas del seguro. Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, pág.22.

[759] Comité Europeo de Protección de Datos. *Directrices 2/2019 sobre el tratamiento de datos personales en virtud del artículo 6, apartado 1, letra b), del RGPD en el contexto de la prestación de servicios en línea a los interesados*. Versión 2.0. Aprobadas el 8 de octubre de 2019, apartado 57, pág.17.

personalización hay que considerarla esencial[760]. Sin embargo, no está tan claro ese carácter necesario cuando la personalización de noticias o contenidos se realiza por plataformas digitales como Facebook ya que el servicio principal de estas no es el de ofrecer o publicar noticias. Para estos últimos tratamientos, muy posiblemente el interés legítimo o el consentimiento resulten las vías más adecuadas, salvo que se demuestre tal carácter esencial, algo que debería quedar debidamente justificado.

B) El carácter necesario en las decisiones plenamente automatizadas relevantes

Tal y como ya se ha explicado, el artículo 22 del RGPD hace referencia a las decisiones plenamente automatizadas relevantes[761]. Estas, aunque a priori estén prohibidas, seguidamente, son autorizadas cuando entren en juego algunas de las excepciones previstas en el artículo 22.2. Así, el apartado a) de este precepto permite el uso de sistemas de toma de decisiones automatizadas relevantes cuando dicha decisión sea necesaria para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento[762].

El carácter necesario al que alude este precepto ha de ser interpretado en los mismos términos que previamente se han explicado. No obstante, el tipo de decisiones que caracterizan estos tratamientos, estas son, plenamente automatizadas y relevantes, obligan al responsable del tratamiento a justificar en mayor grado la necesariedad del despliegue de estos sistemas. En este sentido, la doctrina ha considerado que el carácter necesario no queda justificado cuando una entidad bancaria concede un préstamo bancario de forma plenamente automatizada ya que existen otros medios efectivos menos invasivos para conseguir los mismos objetivos[763]. En nuestra opinión, la automatización se deberá tener en cuenta a la hora de analizar la proporcionalidad de la medida, la cual, por sus características, supone una mayor afectación en los derechos de los particulares. Aquí, un elemento interesante será analizar las garantías que en su caso despliegue el responsable para compensar el uso de ese tipo de decisiones, garantías que han de ir más allá de las que expresamente establece la norma para este tipo de

[760] Hay quien incluso considera que la personalización de noticias en los periódicos digitales tampoco puede presentar ese carácter necesario ya que los contenidos personalizados nunca formaron parte del servicio original. En: ESKENS,S: "A right to reset your user profile and more: GDPR-rights for personalized news consumers", op.cit., pág.11

[761] Capítulo II, aparatado I, punto 1, epígrafe A) de la tesis.

[762] Véase el Capítulo V, aparatado III, punto 2 de esta tesis.

[763] MAS BADIA, M,D: *Sistemas privados de información crediticia. Nueva regulación entre la protección de datos y el crédito responsable*. Ed. Tirant lo Blanch, Valencia, 2021,págs..372 y 373.

tratamientos[764]. Dichas garantías, tanto las explícitamente reconocidas por el RGPD como aquellas que el propio responsable ha de incorporar se analizarán en el capítulo V de esta tesis[765].

3. Cumplimiento de una obligación legal o autorización del tratamiento a través de una norma

A) Cumplimiento de una obligación legal

El artículo 6.1.c) establece como mecanismo de legitimación aquellas situaciones en las que el tratamiento sea necesario para el cumplimiento de una obligación legal aplicable al responsable del tratamiento. Dicho tratamiento por tanto debe basarse en una norma de derecho de la Unión Europea o del derecho de los Estados Miembros (Artículo 6.3 RGPD). En el caso de España una norma con rango de ley, tal y como establece la LOPD de 2018[766], la cual, viene a secundar la reserva de ley impuesta por la Constitución Española[767]. De esa norma con rango de ley se debe desprender una obligación real por la cual el responsable ha de llevar a cabo el tratamiento si quiere cumplir con la mentada obligación normativa[768]. Es decir, el responsable no puede improvisar la obligación legal sobre la que basa el tratamiento de datos, ello le obliga a justificar y documentar el precepto y la obligación que se desprende del mismo sobre la que pretende vertebrar el tratamiento de datos. Tal y como señala el GT29, el mencionado tratamiento ha de basarse en las disposiciones jurídicas que hacen referencia explícitamente a la naturaleza y al objeto del tratamiento. Es decir, el responsable del tratamiento no deberá tener un grado indebido de discreción sobre cómo cumplir con dicha obligación jurídica[769].

---

[764] El artículo 22.3 establece que cuando se lleven a cabo este tipo de tratamientos el responsable recocerá a los interesados *el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión*.
Estos derechos además han de ser complementados con otras medidas de garantía, medias que como hemos señalado serán esenciales a la hora de valorar ese carácter necesario de la decisión.

[765] Véase el Capítulo V, apartado III, punto 3 de esta tesis.

[766] Artículo 8.1 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

[767] El artículo 53.1 de la Constitución Española reconoce la reserva de ley para aquellos supuestos en los que se restringe un derecho fundamental.

[768] SENTENCIA DEL TRIBUNAL DE JUSTICIA (Sala Tercera) de 30 de mayo de 2013, asunto C-342/12, caso Worten. Véase FJ 45. Resolución disponible en:
https://curia.europa.eu/juris/document/document.jsf;jsessionid=10D1650CB9972325A147A45EE3367A8C?text=&docid=137824&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=500150

[769] Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, pág.24.

De lo dicho anteriormente, surge la duda sobre los posibles tratamientos de datos que pueden estar amparados por esta base de legitimación cuando se pretenda implantar un sistema de toma de decisiones automatizadas, sobre todo, cuando la norma que habilita a ese hipotético tratamiento no hace referencia expresa al uso de estos modelos algorítmicos. Es decir, lo que vamos a analizar es cómo, a través del artículo 6.1.c.), un responsable puede legitimar el usos de sistemas automatizados. Ello nos exige estudiar distintos supuestos atendiendo al contenido de diversas normas jurídicas y al grado de mayor a menor explicitud del uso de dichos sistemas en estos textos.

Así, en *primer lugar*, el uso de sistemas automatizados como tratamiento de datos estará plenamente legitimado cuando la propia norma expresamente haga mención al sistema algorítmico y, además, dicho modelo se utilice como herramienta para cumplir la obligación legal que se deriva de esa norma. A modo de ejemplo, la Ley valenciana de 22/2018 regula el sistema de alertas tempranas para la prevención de malas prácticas en la Generalitat[770]. En esta norma se describe el mencionado sistema y además se establecen las obligaciones que corresponden al responsable del tratamiento en relación con su uso. En términos similares, la norma neerlandesa que reconoce legalmente el uso del sistema SyRI en Holanda ya comentado para combatir el fraude también presenta las mismas características[771]. A su vez, en la Unión Europea podemos destacar el Reglamento contra los abusos sexuales de menores en línea, esta norma obliga a las plataformas a detectar y retirar contenido pedófilo por medio de sistemas de toma de decisiones total o parcialmente automatizadas. Aunque este texto no hace referencia a un sistema concreto, sí que menciona el uso de tecnologías específicas de control automatizado del contenido basadas en inteligencia artificial[772]. En todos estos supuestos se hace mención a la tecnología o sistema algorítmico, se hace referencia a las obligaciones legales que ha de afrontar el responsable y además se indica que, para cumplir con dichas obligaciones legales, entre otros medios, el responsable debe utilizar

[770] Artículos 17 y ss. de la Ley 22/2018, de 6 de noviembre, de Inspección General de Servicios y del sistema de alertas para la prevención de malas prácticas en la Administración de la Generalitat y su sector público instrumental.

[771] Decreto del 1 de septiembre de 2014 para enmendar el Decreto SUWI en relación con las reglas para combatir el fraude a través del intercambio de datos y el uso efectivo de los datos conocidos por el gobierno con el uso de SyRI. Norma disponible en el Boletín Oficial del Reino de los Países Bajos: **https://zoek.officielebekendmakingen.nl/stb-2014-320.html**

[772] Considerando 7 y artículos 1 y 3 del REGLAMENTO (UE) 2021/1232 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 14 de julio de 2021 por el que se establece una excepción temporal a determinadas disposiciones de la Directiva 2002/58/CE en lo que respecta al uso de tecnologías por proveedores de servicios de comunicaciones interpersonales independientes de la numeración para el tratamiento de datos personales y de otro tipo con fines de lucha contra los abusos sexuales de menores en línea.

dicho sistema o tecnología. El mecanismo de legitimación previsto en el artículo 6.1.c del RGPD es evidente y adecuado.

En *segundo lugar* encontraríamos todas aquellas disposiciones legislativas que si bien imponen una obligación al responsable para llevar a cabo el tratamiento de datos, en dicha norma no se contempla el uso de sistemas automatizados para cumplir con tal obligación. Así por ejemplo, el artículo 29.1 de Ley de Economía Sostenible establece que las entidades de crédito están obligadas a evaluar la solvencia del potencial cliente antes de celebrar un contrato de crédito o préstamo[773]. Para cumplir con esta obligación, el responsable del tratamiento ha de tratar los datos para evaluar el posible riesgo de impago. Sin embargo, no se indica cómo ha realizar esa evaluación de riesgos. Al no establecerse de forma expresa el uso de sistemas automatizados para evaluar el riesgo se podría llegar a considerar que tal tratamiento de datos no podría ampararse en el artículo 6.1.c). En nuestra opinión, a pesar de que no se prevea de forma expresa, el responsable podrá acreditar este mecanismo siempre que cumpla varios requisitos, estos son: i) la norma ha de imponer al responsable una obligación clara, a pesar de que no se indique cómo ha de ejecutar dicha obligación, ii) se ha de realizar un análisis del carácter necesario de utilizar ese sistema automatizado para cumplir con esa obligación. Ello obligará a analizar la proporcionalidad de incorporar este sistema, sobre todo, con relación a la necesidad y a la proporcionalidad en sentido estricto del sistema[774]. iii) Además, dicho modelo algorítmico deberá reflejar los objetivos que se pretenden conseguir con las obligaciones impuestas por la norma. Estos elementos deberán quedar registrados en la EIPD en los términos que ya se explicaron a la hora de analizar esta herramienta[775]. iv) Por último, se deberán establecer suficientes medidas de garantía.

En *tercer lugar*, se plantean más dudas para aquellas normas que si bien reconocen el posible uso de sistemas automatizados, las obligaciones impuestas a través de esos textos legales no están del todo claras. Así, la Unión Europea se encuentra actualmente en un periodo de transición legislativa que pretende imponer mayores obligaciones a las grandes plataformas digitales para que controlen el contenido que se vierte en el interior de las mismas. En estas propuestas legislativas se establece que

[773] Artículo 29.1 de la Ley 2/2011, de 4 de marzo, de Economía Sostenible y artículo 18 de la Orden EHA/2899/2011, de 28 de octubre, de transparencia y protección del cliente de servicios bancarios.
[774] Sobre el juicio de proporcionalidad en estos contextos véase el Capítulo III, apartado II, punto 5 de esta tesis.
[775] Capítulo III, apartado II de esta tesis.

estas plataformas no tienen la obligación general de supervisar ese contenido[776]. Sin embargo, algunos de estos preceptos incentivan a las mismas a que realicen actuaciones proactivas de control, ese control, dado el número tan alto de información que circula en estas plataformas acabará en parte automatizándose. Es por ello que a priori y a pesar de que muchas de las exigencias previstas por estas normas hayan de ejercitarse a través de medios automáticos[777], dado que no se desprende una obligación de supervisión general para los responsable del tratamiento, hay que entender que no existe dicha obligación y por tanto no podría acudirse para legitimar el control automatizado de contenido a la vía del artículo 6.1.c) del RGPD. Ello queda aún más patente cuando la propia norma expresamente indica que el uso de medios automatizados para cumplir con las obligaciones legales de la norma no es necesario[778]. En este sentido, cabe señalar que hasta la fecha, el TJUE también se ha inclinado por esta interpretación ya que ha considerado que los proveedores de servicios no tienen una obligación general de supervisión sobre los contenidos y comunicaciones que se vierten en sus plataformas, y por tanto, medidas que impongan el deber de establecer sistemas de filtrado y bloqueo automático va en contra de la normativa europea ya que afecta gravemente a los derechos de libertad de empresa, protección de datos y libertad de información[779]. En estos supuestos entendemos que el responsable deberá acudir a otros mecanismos para legitimar el tratamiento como el contemplado en el artículo 6.1.f), es decir, intereses legítimos. Ahora bien, en nuestra opinión, aunque compartimos los argumentos del TJUE, muy probablemente con el paso del tiempo todas estas actuaciones acaben

[776] Véase el artículo 7 de la Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a un mercado único de servicios digitales. También en el mismo sentido el Artículo 17.8 de 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.

[777] Debido a la enorme cantidad de trabajo requerido para moderar el contenido, las plataformas de Internet han comenzado a desarrollar e implantar sistemas de inteligencia artificial para automatizar la toma de decisiones sobre las solicitudes de eliminación de contenido. En: DIAS OLIVE,T: "Content Moderation Technologies: Applying Human Rights Standards to Protect Freedom of Expression". *Human Rights Law Review*, 2020, pág.609.
Texto disponible en: https://academic.oup.com/hrlr/article/20/4/607/6023108

[778] Tal y como establece el artículo 5.8. párrafo segundo 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.

[779] Sentencia del Tribunal de Justicia (Sala Tercera) de 24 de noviembre de 2011. Asunto C-70/10, caso Scarlet Extended SA. FJ 47 y ss. Resolución disponible en::
https://curia.europa.eu/juris/document/document.jsf?text=&docid=115202&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=502470
Sentencia del Tribunal de Justicia (Sala Tercera) de 16 de febrero de 2012, Asunto C-360/10, (SABAM, FJ 45 y ss.) Resolución disponible:
https://curia.europa.eu/juris/document/document.jsf?text=&docid=119512&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=502910

automatizándose fruto del aumento de responsabilidades que cada vez se exigirán a estas plataformas y la imposibilidad de que las mismas las realice un humano, cuando ello ocurra, la posibilidad de acudir a la vía del artículo 6.1.c) será más clara.

B) Autorización o habilitación del tratamiento basado en la toma de decisiones automatizadas a través de una norma. Especial referencia al sector público

Resulta habitual, sobre todo en el contexto de las Administraciones Publicas, que las normas no sólo obliguen a los responsables a llevar a cabo un tratamiento de datos sino que más bien, estas normas autorizan o habilitan a dichos responsables a través de las mismas a desplegar ese tratamiento. Es decir, la norma en sí se convierte en un mecanismo de legitimación del tratamiento. De manera que el responsable evita con ello tener que acudir a otras bases de legitimación como el consentimiento o el interés legítimo[780]. Es turno de analizar los requisitos tanto formales como materiales que ha de contener esa norma.

b.1) Los requisitos formales de la norma

Pues bien, el artículo 23 del RGPD permite que el derecho interno o el de la UE reconozcan o autoricen tratamientos de datos en favor de los responsables. Este precepto no establece cuál ha de ser el instrumento legal adecuado para autorizar a estos tratamientos. Sin embargo, en nuestra opinión, teniendo en cuenta los riesgos que pueden comportar para los particulares la elaboración de perfiles y la toma de decisiones significativas total o parcialmente automatizadas, la norma que autorice estos tratamientos, al menos en el sector público, deberá ostentar como mínimo el rango de ley. En este sentido, el TC ha argumentado en varias ocasiones la necesidad de regular por ley los límites al derecho fundamental a la protección de datos[781], tal y como ocurre cuando una norma pretende autorizar el uso de un sistema de toma de decisiones automatizadas. Así lo ha indicado también la AEPD en diversas resoluciones donde se ha pretendido legitimar el uso de sistemas de reconocimiento facial automatizados[782].

[780] A modo de ejemplo, los artículos 22 a 27 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales legitiman toda una serie de tratamientos de datos personales. Por ejemplo, videovigilancia, denuncias internas, función estadística, archivo de interés público, etc.

[781] STC 17/2013 de 31 de enero de 2013, FJ 4º. Texto disponible en: http://hj.tribunalconstitucional.es/es-ES/Resolucion/Show/23272
Véase también: Véase la STC Sentencia 76/2019, de 22 de mayo de 2019. FJº 5 Y 8. Texto disponible en: https://hj.tribunalconstitucional.es/HJ/docs/BOE/BOE-A-2019-9548.pdf

[782] Agencia Española de Protección de Datos. Informe N/REF: 0047/2021, pág.14. Texto disponible en:

En estos supuestos, los especiales riesgos que comporta el ingreso de esa tecnología en la esfera de los particulares requieren para su legitimación de una autorización a través de una norma con rango de ley. En nuestra opinión, la exigencia legal que se impone para el uso de los sistemas de reconocimiento facial entendemos que también es trasladable al contexto de la toma de decisiones automatizadas, sobre todo, en el ámbito público. Y ello, independientemente de si se tratan o no datos de categoría especial o el sistema adopta decisiones total o parcialmente automatizadas. El carácter o no automatizados del proceso algorítmico o la existencia o no de categorías especiales de datos serán factores a tener en cuenta a la hora de imponer mayores o menos garantías, pero no vincularán con relación al tipo de instrumento legal que legitime el tratamiento de datos, el cual, ha de ser una norma con rango de ley.

Sector Público	
Rango de la norma mínimo	**Tipo de tratamiento de datos**[783]
Norma con rango de ley	Decisiones plenamente automatizadas relevantes con o sin elaboración de perfiles Decisiones parcialmente automatizadas relevantes con o sin elaboración de perfiles
Reglamento administrativo	Decisiones plenamente automatizadas no relevantes Decisiones parcialmente automatizadas no relevantes

b.2) Los requisitos materiales de la norma

Junto a la reserva de ley, también se ha de valorar el contenido de esa norma jurídica. Es decir, como ha señalado el TC, la ley no sólo ha de habilitar la medida restrictiva de derechos fundamentales sino también las exigencias que se derivan de la calidad de la ley, esto es, en dicha regulación el legislador está obligado a ponderar los derechos o intereses en pugna, predeterminar los supuestos del tratamiento y fijar las condiciones y las garantías en que procede la adopción de medidas restrictivas de derechos fundamentales[784]. Así, el artículo 23 del RGPD, en referencia a los requisitos

https://www.aepd.es/es/documento/2021-0047.pdf Véase también: Resolución Nº: PS/00120/2021, pág.48. En el mismo sentido se ha pronunciado la Red DAIA (Derecho Administrativo e Inteligencia Artificia. En: Declaración de Valencia, Red DAIA (Derecho Administrativo e Inteligencia Artificia), octubre 2019, apartado 4. Información disponible en: http://reddaia.org/documentos/

[783] En esta tesis se hace una distinción entre tipos de decisiones automatizadas y en su caso elaboración de perfiles. Véase el Capítulo II, apartado I de esta tesis.

[784] STC Sentencia 76/2019, de 22 de mayo de 2019. FJº 8. Texto disponible en: https://hj.tribunalconstitucional.es/HJ/docs/BOE/BOE-A-2019-9548.pdf
Véase también: COTINO HUESO, L: "Exigencias constitucionales para las comunicaciones de datos entre Administraciones y la problemática de la procedencia de los datos del censo para la celebración de la consulta catalana. Posibles soluciones legales", *Boletín APEP Informa 2014*, nº 10, Asociación

que ha de contener la norma que restrinja el derecho fundamental a la protección de datos establece que dicha norma ha de incluir entre otros elementos: la finalidad del tratamiento o de las categorías de tratamiento, las categorías de datos personales de que se trate, el alcance de las limitaciones establecidas, las garantías suficientes, los derechos de los interesados, plazos de conservación, etc[785].

Trasladado a nuestro contexto, tomando como referencia la doctrina constitucional comentada y las reglas previstas por el RGPD, la norma de rango de ley que autorice a implantar sistemas de toma de decisiones automatizadas ha de contener al menos: i) los tratamientos de datos que se pretenden autorizar con el ingreso de ese algoritmo, ii) la finalidad o finalidades específicas que se esperan acometer, iii) los datos personales a tratar y, iv) las garantías y salvaguardas específicas que requiere ese tratamiento.

De esta manera, preceptos legislativos que genéricamente hagan alusión a la actuación automatizada de un tratamiento sin especificar las finalidades concretas del mismo y además únicamente prevean una serie de garantías mínimas no pensamos que puedan ser mecanismos adecuados para legitimar un tratamiento de datos basado en las técnicas algorítmicas que estamos analizando. A modo de ejemplo, en el ámbito del sector público, el artículo 41 de la Ley 40/2015 de Régimen Jurídico del Sector Público establece una serie de reglas y garantías generales que han de incorporar las administraciones cuando pretendan llevar a cabo una actuación administrativa automatizada[786]. En nuestra opinión, este precepto no supondría un mecanismo de legitimación suficiente para habilitar los tratamientos de datos que venimos

---

Profesional Española de Privacidad, abril de 2014, págs 11 y 12. A su vez: REBOLLO DELGADO, L: "Big data, inteligencia artificial y derechos fundamentales: problemas jurídicos". En: *El derecho a la protección de datos personales en Europa y en América: diferentes visiones para una misma realidad*. Ed. Tirant lo Blanch. Valencia, 2021, pág.426.

[785] Véase el artículo 23.2 del RGPD.

[786] El artículo 41 denominado, "Actuación administrativa automatizada", establece que: 1. Se entiende por actuación administrativa automatizada, cualquier acto o actuación realizada íntegramente a través de medios electrónicos por una Administración Pública en el marco de un procedimiento administrativo y en la que no haya intervenido de forma directa un empleado público. 2. En caso de actuación administrativa automatizada deberá establecerse previamente el órgano u órganos competentes, según los casos, para la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente. Asimismo, se indicará el órgano que debe ser considerado responsable a efectos de impugnación. Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público. Sobre el uso de la inteligencia artificial en el contexto de las Administraciones Públicas véase: HUERGO LORA, A,J: "*Regular la inteligencia artificial (en Derecho administrativo"). Blog de la Revista de Derecho Público*. 08/03/2021. Texto disponible en: http://blogrdp.revistasmarcialpons.es/blog/regular-la-inteligencia-artificial-en-derecho-administrativo-por-alejandro-huergo-lora/

describiendo a lo largo de la tesis[787]. Esto es, la toma de decisiones parcial o totalmente automatizadas relevantes y la elaboración de perfiles relevantes. Ello es así porque este precepto en ningún momento hace referencia a las finalidades potenciales que habilitarían al tratamiento de esos datos, ni a los datos o categorías de datos generales que pudieran estar afectados y además, las garantías específicas recogidas en el mismo resultan cuanto menos insuficientes. Una interpretación contraria a la sugerida previamente dejaría en manos de la Administración por la vía del reglamento administrativo u otros instrumentos más opacos la legitimación de cualquiera de los tratamientos de datos sin las debidas garantías y exigencias que han de estar previstas por parte del legislador y que se derivan de la doctrina constitucional[788]. En este sentido, somos conscientes que este tipo de preceptos se incardinan en una legislación básica que establece las reglas generales del funcionamiento de las Administraciones Públicas, sin embargo, la inconcreción en relación con las finalidades a la hora de hacer mención a posibles elaboraciones de perfiles o uso de sistemas algorítmicos y la escasez de garantía suficientes nos lleva a la conclusión de que este precepto, a efectos de la normativa de protección de datos[789], no puede ser un instrumento suficiente para legitimar un determinado tratamiento de datos por parte de una Administración Pública

[787] Contrariamente a lo que defendemos, la doctrina en términos generales ha considerado que el artículo 41 de la Ley 40/2015 sí que es un mecanismo de legitimación suficiente a efectos de la normativa de protección de datos. En: BOIX PALOP,A: "Los algoritmos son reglamentos: La necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones". op.cit., pág. 246. También en el mismo sentido: CERRILLO MARTÍNEZ,A: "¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?". *European review of digital administration & law*, Vol. 1, Nº. 1-2, 2020, pág.27. Texto disponible en: http://www.aracneeditrice.it/pdf2/97888255389603.pdf

[788] Tal y como señaló el ex magistrado del TC Pablo Pérez Tremps en un voto particular en referencia a la calidad de la ley: *La garantía de la previsión legal es sólo uno de los condicionantes de la constitucionalidad de la limitación de este derecho fundamental. Sin embargo, esta garantía no queda limitada al hecho de que una ley habilite la medida, sino que es preciso, conforme a exigencias mínimas de calidad de la ley y de respeto al contenido esencial del derecho —como mandato dirigido al legislador de los derechos fundamentales— (art. 53.1 CE), que en esa regulación el legislador predetermine, como primer obligado a realizar la ponderación de derechos o intereses en pugna, los supuestos, las condiciones y las garantías en que procede la adopción de medidas restrictivas de derechos fundamentales. La posibilidad de que ese mandato de predeterminación respecto de elementos esenciales vinculados a la proporcionalidad de la limitación del derecho fundamental quede deferido a un eventual desarrollo reglamentario no sólo no satisface las elementales exigencias de la reserva de ley sino que, además, es especialmente grave en supuestos como el presente en el que las medidas, sin perjuicio de un eventual control judicial posterior, son de aplicación directa por parte de la Administración pública*. En: Voto Particular **o** don Pablo Pérez Tremps, apartado 2. STC Sentencia 17/2013, de 31 de enero. Resolución disponible en: https://hj.tribunalconstitucional.es/es-ES/Resolucion/Show/23272

[789] Tal y como indicó en su momento el TC sobre la limitación del derecho fundamental a la protección de datos., *la falta de precisión de la ley en los presupuestos materiales de la limitación de un derecho fundamental es susceptible de generar una indeterminación sobre los casos a los que se aplica tal restricción*. STC Sentencia 292/2000, de 30 de noviembre de 2000. FJº 15. Texto disponible en: https://www.boe.es/buscar/doc.php?id=BOE-T-2001-332

que en el marco de sus competencias decide desplegar un sistema de toma de decisiones automatizadas[790]. Y ello, pese a que actualmente sea la dinámica actual de nuestras Administraciones Publicas[791].

Señaladas las carencias actuales, son varias las propuestas que planteamos sobre el contenido de la norma que ha de legitimar estos tratamientos de datos:

En *primer lugar*, para preceptos generales como el previsto en el artículo 41 de la ley 40/2015 que pretenda legitimar un importante número de tipos tratamientos de datos e irradiar sus efectos en diferentes sectores. Sería necesario que esa norma haga mención expresa a la elaboración de perfiles o en su caso a las decisiones algorítmicas y además se regulen necesariamente las garantías específicas para ese concreto ámbito. Como hoja de ruta para regular un marco jurídico general adecuado pueden extraerse soluciones de algunos textos como la Carta de derechos digitales española o la PRAI[792]. Una vez reconocidas esas salvaguardas adecuadas y la mención expresa al uso de algoritmos en la norma de rango de ley, por vía reglamentaria sí que en su caso se podrían establecer las finalidades específicas que se pretenda para cada sistema con unas debidas nociones de publicidad de los mismas y siempre, siguiendo las pautas marcadas por la ley sobre la que desarrolla la norma, ley que deberá ser suficientemente garantista. Por vía reglamentaria también se podrán establecer las especificaciones técnicas del sistema algorítmico.

En *segundo lugar*, resultará siempre más adecuado que directamente sea la norma con rango de ley sectorial la que reconozca específicamente el tratamiento de datos que autorice al uso de estas tecnologías y dicha ley prevea las finalidades y garantías específicas de ese concreto ámbito donde se adoptarán las decisiones automatizadas. Entre los ejemplos que ya hemos comentado encontramos por un lado la

[790] Recientemente, el legislador español ha tenido una nueva oportunidad para regular de forma más garantistas las actuaciones automatizadas de la Administración pero tampoco ha enmendado la merma de garantías ya comentada. Así, el artículo 13 del Real Decreto 203/2021, de 30 de marzo sobre el funcionamiento del sector público por medios electrónicos viene prácticamente a calcar lo ya previsto y comentado por el artículo 41 de la Ley 40/2015.

[791] No es frecuente que las Administraciones acudan al instrumento de la ley para autorizar estos tratamientos de datos. Tal y como ha señalado la doctrina, *la vía preferente para acordar el uso de estas herramientas es la resolución del órgano administrativo competente, en la que se informa de que existe una aplicación informática que interviene o incide en procedimientos administrativos específicos, indicando su finalidad, así como el órgano competente a efectos de impugnación*. En: CAPDEFERRO, O: "La inteligencia artificial del sector público: desarrollo y regulación de la actuación administrativa inteligente en la cuarta revolución industrial", *IDP. Revista de Internet, Derecho y Política*. N. º 30, págs. 5 y 6. Texto disponible en: https://raco.cat/index.php/IDP/article/view/373603/467216

[792] Artículos 18 y 25 de la Carta de Derechos Digitales. También véase la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión.

Ley valenciana de 22/2018 que regula el sistema automatizado de alertas tempranas para la prevención de malas prácticas en la Generalitat y por otro lado, la ley francesa que legitima el análisis masivo de datos personales de los ciudadanos disponibles públicamente en las plataformas virtuales como método para luchar contra el fraude fiscal[793].

En *tercer lugar*, fuera del sector público, a pesar de que las exigencias pueden ser menores a las previstas para las administraciones públicas, seguimos considerando que la ley que decida legitimar los tratamientos de datos también deberá indicar claramente las finalidades y prever suficientes medias de garantía. En este sentido, la APED consideró que un sistema de reconocimiento facial automatizado no podía considerarse legitimado a través del artículo 22 de la LOPD de 2018. Este precepto legitima los tratamientos de videovigilancia pero no los de reconocimiento facial[794], los cuales generan mayores riesgos para los particulares.

Para finalizar, cabe mencionar un último apunte en relación con la PRAI, la cual, como ya hemos explicado en otro momento tiene como objetivo establecer normas armonizadas para la comercialización, la puesta en servicio y el uso de sistemas de inteligencia artificial tanto en el sector público como en el privado. Concretamente, esta norma establece requisitos específicos para los sistemas de IA de alto riesgo y obligaciones para los operadores de dichos sistemas[795]. En muchos casos, estos sistemas tratarán datos personales una vez que comiencen a usarse. De esta manera, si esos sistemas automatizados cumplen los requisitos exigidos por esta norma de derecho europeo, los mismos podrán venderse en el mercado europeo o utilizarse en la toma de decisiones automatizadas. Ahora bien, tal y como señala la PRAI en su considerando 41, el hecho de que un sistema de IA esté clasificado como de alto riesgo en virtud de esa norma no debe interpretarse como una indicación de que el uso del sistema es necesariamente lícito en virtud de otros actos del Derecho de la Unión o del Derecho nacional. De aprobarse la PRAI, el cumplimiento de los requisitos exigidos por la misma por parte de un algoritmo no deberá entenderse *per se* como un fundamento

[793] del artículo 154 de la Loi de finances pour 2020, sous le n° 2019-796 DC, le 20 décembre 2019 que legitima el análisis masivo de datos personales de los ciudadanos disponibles públicamente en las plataformas virtuales como método para luchar contra el fraude fiscal. Texto disponible en: https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000039683923

[794] Agencia Española de Protección de Datos. Resolución Nº: PS/00120/2021, págs.72 y 73.

[795] COTINO HUESO, L; CASTILLO PARRILLA, J,A; SLAZAR,I; BENJAMINS,R; CUMBRERAS,M; ESTEBAN,A,M: "Un análisis crítico constructivo de la Propuesta de Reglamento de la Unión Europea por el que se establecen normas armonizadas sobre la Inteligencia Artificial", op.cit., pág.3.

jurídico para el tratamiento de datos personales[796]. Esta interpretación tiene mucho sentido, así, cuando un fabricante desarrolla un reloj inteligente o un coche autónomo ha de ser consciente de que, además de cumplir con las exigencias técnicas que la normativa específica prevea para que ese producto pueda distribuirse o utilizarse en el mercado, dicho producto también ha de cumplir la normativa de protección de datos y en concreto, el principio de licitud, el cual requiere que el responsable tenga una base o mecanismo de legitimación para llevar a cabo el tratamiento de datos pretendido independientemente de que el producto haya pasado los controles de seguridad o de funcionamiento.

4. La misión de interés público o el ejercicio de poderes públicos en la toma de decisiones automatizadas

El artículo 6.1.e) proporciona un fundamento jurídico válido en situaciones en las que *el tratamiento es necesario para el cumplimiento de una misión realizada en interés público o en el ejercicio de poderes públicos conferidos al responsable del tratamiento*[797]. Al igual que ocurre con la base de legitimación prevista en el artículo 6.1.c), el tratamiento debe tener una base en el Derecho de la Unión o de los Estados miembros. Según el GT29, este mecanismo de legitimación tiene un ámbito muy amplio de aplicación que requiere una interpretación estricta y una clara identificación caso por caso del interés público en juego y de la potestad oficial que justifica el tratamiento[798]. Lo cierto es que este mecanismo de legitimación comúnmente se alterna y se confunde con el previsto en el artículo 6.1.c), esto es, obligación legal, sobre todo cuando el responsable es una autoridad pública, ya que estas son las dos bases de legitimación que resultan más adecuadas para los organismos públicos cuando pretenden legitimar sus

[796] Véase el considerando 41 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. En el mismo sentido se ha pronunciado el CEPD y el SEPD. En: En: Comité Europeo de Protección de Datos. EDPB-EDPS. Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act). Resolución de 18 de junio de 2021. Apartado 22, pág.9.

[797] Como regla general, esta será una de las bases de legitimación habituales que utilizarán los poderes públicos para legitimar por ejemplo la elaboración de perfiles por parte de las Administraciones Públicas. En: VELASCO RICO, C: "Personalización, proactividad e inteligencia artificial. ¿Un nuevo paradigma para la prestación electrónica de servicios públicos?", op.cit., pág.11.

[798] Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, pág.27.

tratamientos de datos[799]. En nuestra opinión, la diferencia principal entre uno y otro mecanismo es que el responsable que pretenda acudir al artículo 6.1.c) ha de encontrar la base legal que le obliga implantar ese tratamiento mientras que el artículo 6.1.e) obliga al responsable a fijar la competencia o el interés jurídico sobre el que ampara su tratamiento[800]. El TJUE ha tenido la oportunidad de analizar esta base jurídica en varias ocasiones y en esencia, a la hora de analizar el carácter necesario del tratamiento ha valorado que: i) exista una norma que confiera al responsable la misión o el poder público, ii) se ha de evaluar la necesidad e idoneidad del tratamiento y, iii) se ha de valorar si dicho tratamiento cumple con el resto de normativa de protección de datos[801]. De esta manera, y centrándonos en el uso de sistemas automatizados basados en este mecanismo de legitimación, una vez que el responsable haya detectado la norma con rango de ley que otorga el poder o competencia pública, este ha de valorar la necesidad de utilizar dicho sistema automatizado en los términos que hemos indicado en el apartado referido a la EIPD. Es decir, idoneidad, necesidad y proporcionalidad en sentido estricto[802]. Ese tratamiento además ha de cumplir con las exigencias en materia de protección de datos.

---

[799] Tal y como ha señalado la AEPD: *con carácter general, la base jurídica del tratamiento en las relaciones con la Administración, en aquellos supuestos en que existe una relación en la que no puede razonablemente predicarse que exista una situación de equilibrio entre el responsable del tratamiento (la Administración), y el interesado (el administrado) no sería el consentimiento (art. 6.1.a) RGPD), sino, según los casos, el cumplimiento de una obligación legal (art. 6.1.c) RGPD) o el cumplimiento de una misión de interés público o en el ejercicio de poderes públicos (art. 6.1.e) RGPD).* En: Agencia Española de Protección de Datos. Informe Nº 175/2018. Pág.4.

[800] La antigua ley orgánica de protección de datos no establecía el instrumento normativo sobre el que se fundaba la misión o interés público para basar el tratamiento de datos. Artículo 6.2 de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal. No obstante, el Real Decreto que desarrolla esta norma indica expresamente la necesidad de que esta base de legitimación quede encomendada a la ley. Artículo 10.3.a) Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal. En: PUENTE ESCOBAR, A: "Legitimación para el tratamiento". En: MARTÍNEZ MARTÍNEZ,R (coord.): *Protección de datos. Comentarios al Reglamento de Desarrollo de la LOPD*. Ed. Tirant lo Blanch. Valencia, 2008, págs.32 y 33.

[801] Sentencia del Tribunal de Justicia (Sala Segunda) de 27 de septiembre de 2017, asunto C-73/16, caso Peter Puškár, FJº 117. Resolución disponible:
https://curia.europa.eu/juris/document/document.jsf?text=&docid=195046&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=5339811
Véase también: Sentencia del Tribunal de Justicia de la Unión Europea, (Sala Tercera) de 30 Mayo de 2013,asunto C- 342/2012, caso Worten, FJº45. Resolución disponible en:
https://eur-lex.europa.eu/legal-content/es/TXT/?uri=CELEX:62012CJ0342
A su vez: Sentencia del Tribunal de Justicia de la Unión Europea, Sentencia de 20 Mayo de 2003, asuntos acumulados C-465/00, C-138/01 y C-139/01, caso Rechnungshof, FJº 94. Resolución disponible en:
https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A62000CJ0465

[802] En el asunto C-524/06, caso HUBER se consideró que un tratamiento era necesario para cumplir con esa misión de interés público si el responsable sólo recopilaba aquellos datos estrictamente necesarios y el tratamiento permitía una aplicación más eficaz de la normativa en la que se basaba. SENTENCIA DEL

Requisitos del mecanismo de legitimación previsto en el artículo 6.1.e) RGPD
-Norma con rango de ley ha de atribuir la misión o el poder público. (atribuye competencias) -El tratamiento ha de ser necesario para el cumplimiento de una misión realizada en interés público o para el ejercicio de poderes públicos. -Cumplimiento del resto de normativa de protección de datos.

Es turno de analizar más profusamente estos tres requisitos.

En *primer lugar*, por lo que se refiere al instrumento jurídico sobre el que se basa el tratamiento de datos, la LOPD de 2018 ya ha señalado que el interés público o el ejercicio del poder público han de basarse en una competencia atribuida al responsable a través de una norma con rango de ley[803]. No hace mención este precepto al grado de concreción de esa competencia establecido por la norma. Es decir, si bien es cierto que el tratamiento de datos se ha de basar en una competencia atribuida por una norma con rango de ley, no queda claro si la ley ha de definir de forma más o menos explícita la relación entre esa competencia y el tratamiento de datos sobre el que se basa el ejercicio o poder público. Por tanto, lo siguiente que tenemos que preguntarnos es si resulta suficiente para legitimar un tratamiento de datos el hecho de que una norma atribuya competencias genéricas sin más a una determinada Administración para implementar un tratamiento de datos relacionado con la toma de decisiones automatizadas o se requiere en cambio que esa norma que atribuye la competencia haga mención más o menos explícita a dicho tratamiento. Hasta la fecha, podemos decir que no existe debate en este sentido ya que tanto la doctrina[804], como los tribunales[805], así como la AEPD se han

---

TRIBUNAL DE JUSTICIA (Gran Sala) de 16 de diciembre de 2008, asunto C-524/06, caso Heinz Huber, FJ 66. Resolución disponible: https://curia.europa.eu/juris/document/document.jsf;jsessionid=079A4EF056BB8BF122AA32D3E58A253F?text=&docid=76077&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=5337828

[803] Artículo 8.2 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.
La antigua ley orgánica de protección de datos de 1999 no establecía el instrumento normativo sobre el que se fundaba la misión o interés público para basar el tratamiento de datos. Artículo 6.2 de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal. No obstante, el reglamento que desarrolló esta norma indicó expresamente la necesidad de que esta base de legitimación quedará encomendada a la ley. Artículo 10.3.a) Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal. En: PUENTE ESCOBAR, A: "Legitimación para el tratamiento". En: MARTÍNEZ MARTÍNEZ,R (coord.): *Protección de datos. Comentarios al Reglamento de Desarrollo de la LOPD*. Ed. Tirant lo Blanch. Valencia, 2008, págs.32 y 33.

[804] Para Toncoso Reigada, *toda actividad administrativa en el ámbito de sus competencias está encaminada a servir con objetividad los intereses generales -artículo 103.1 CE- y puede justificar el tratamiento de datos personales sin el consentimiento pero no la excepción de los derechos de acceso, rectificación y cancelación*. TRONCOSO, REIGADA, A: *La Protección de Datos Personales .En Busca del Equilibrio*, op.cit., pág.200. En el mismo sentido se ha pronunciado Fernández Salmerón al señalar que

declinado por la primera interpretación, esto es, una Administración Pública puede legitimar su tratamiento de datos aludiendo a una competencia que le ha atribuido una norma de rango de ley sin necesidad de que la norma explícitamente haga mención al tratamiento en cuestión. Esta interpretación tiene en parte su justificación, así, las Administraciones Públicas tienen como objetivo fundamental servir a los intereses generales[806], para ello, las normas les atribuyen toda una serie de potestades y competencias con el objetivo de que estas puedan posteriormente ejercer sus funciones en el mayor número de supuestos. Pues bien, en nuestra opinión, a pesar de lo indicado previamente, teniendo en cuenta los riesgos que presentan los tratamientos de datos objeto de esta tesis[807], cuando una organización pretenda acudir a este mecanismo de legitimación para justificar los tratamientos de datos indicados, estas organizaciones deberían acudir a normas con rango de ley que les atribuyan competencias que en cierto grado definan los posibles tratamientos, no resultando adecuadas aquellas que hacen referencia o atribuyen competencias excesivamente genéricas[808]. A modo de ejemplo, imaginemos que un ayuntamiento pretende implantar un sistema de gestión de residuos inteligentes sobre el cual se elabora un perfil de los usuarios de manera que se premia a aquellas personas que reciclan más. Dicho perfil lo realiza un algoritmo en función de toda una serie de datos personales. A priori, entendemos que no sería suficiente para justificar este tratamiento acudir al artículo 25.2.b) de la Ley de Bases de régimen Local, este precepto, de forma genérica atribuye a los municipios la gestión de los residuos sólidos urbanos[809]. Sin embargo, sí que podría justificarse este tratamiento respecto de otras normas que atribuyen de forma más clara la competencia. Un ejemplo sería el Texto refundido de la Ley reguladora de los residuos de Cataluña, el cual, no

---

esta excepción (misión o interés público) aplica cuando el tratamiento lo realice alguna Administración Pública y la misma actúe dentro del llamado giro o tráfico administrativos en cuanto actividad típica de la Administración Pública. En: FERNÁNDEZ SALMERÓN, M: *La protección de los datos personales en las Administraciones Públicas*. Ed. Thomson, Civitas. 2003, pág.281.

[805] Sentencia de la Audiencia Nacional AN (Sala de lo Contencioso-Administrativo, Sección 1ª) Sentencia de 23 noviembre 2012, FJº 3. En el mismo sentido, Sentencia del Tribunal Supremo, Sala Tercera, de lo Contencioso-administrativo, Sección 6ª, Sentencia de 12 Marzo de 2012, Rec. 2453/2009, FJº2 *in fine*.

[806] Artículo 103.1 de la Constitución Española.

[807] Toma de decisiones total o parcialmente automatizadas relevantes o elaboración de perfiles relevantes.

[808] La AEPD ya ha advertido de los graves riesgos que se derivan del automatismo de los tratamientos de datos basados en inteligencia artificial. En: Agencia Española de Protección de Datos. Resolución Nº: PS/00120/2021, pág.94.

[809] El artículo 25.2.b) de Ley 7/1985, de 2 de abril, Reguladora de las Bases del Régimen Local establece que: ***el Municipio ejercerá en todo caso como competencias propias****, en los términos de la legislación del Estado y de las Comunidades Autónomas, en las siguientes materias: b) Medio ambiente urbano: en particular, parques y jardines públicos,* ***gestión de los residuos sólidos urbanos*** *y protección contra la contaminación acústica, lumínica y atmosférica en las zonas urbanas.* (La negrita y la cursiva son nuestras).

sólo hace referencia a la competencia que se atribuye al municipio, sino también al objetivo que se pretende con la atribución de la misma, atribución de la que se puede desprender o relacionar el uso del sistema de inteligencia artificial previamente aludido[810]. Lo que proponemos es que al menos, el responsable del tratamiento, el cual generalmente será una Administración Pública, justifique adecuadamente la competencia sobre la que basa ese poder o interés público y que lo relacione adecuadamente con el tratamiento automatizado que pretende implantar, en este caso, la toma de decisiones automatizadas. Antes de continuar con el siguiente requisito exigido para esta base de legitimación, es importante destacar que en virtud del artículo 22 del RGPD, aquellos tratamientos de datos que supongan la toma de decisiones relevantes plenamente automatizadas no podrán basarse en el mecanismo de legitimación que estamos analizando, este tipo de tratamientos requieren de unas concretas bases de legitimación entre las cuales no está la descrita en el artículo 6.1.e) del RGPD[811].

En segundo lugar, de acuerdo al Artículo 6.3 del RGPD, el tratamiento de datos pretendido ha de ser necesario para el cumplimiento de una misión realizada en interés público o en el ejercicio de poderes públicos. Por ello, si un determinado tratamiento no es necesario, este no quedará amparado por el apartado e) del artículo 6.1 del RGPD. En este sentido, la AEPD ha considerado que el carácter necesario ha de interpretarse en sentido amplio. Es decir, un tratamiento es necesario de manera que sin existir el mismo, el responsable no podría llevar a cabo el ejercicio de la competencia o potestad atribuida a la Administración. A sensu contrario, si a través de otro tratamiento se puede ejercer la competencia, entonces el tratamiento pretendido se considerará superfluo y contrario tanto al principio de licitud como al de minimización de datos[812]. Esta interpretación restringe en parte la posible legitimación de los tratamientos de datos de

[810] Concretamente, el artículo 53.1 Decreto Legislativo 1/2009, de 21 de julio, por el que se aprueba el Texto refundido de la Ley reguladora de los residuos de Cataluña establece que: *Con el **objetivo de favorecer el reciclaje** y la valorización material de los residuos municipales, todos los municipios prestarán el servicio de recogida selectiva de las diversas fracciones de residuos. Los **municipios prestarán el servicio de recogida selectiva utilizando los sistemas de separación y recogida que se hayan mostrado más eficientes** y que sean más adecuados a las características de su ámbito territorial*. (La negrita y la cursiva son nuestras). El ejemplo planteado en este trabajo se ha inspirado en un supuesto planteado por un ayuntamiento de Cataluña a la Autoritat Catalana de Protecció de dades. Véase la Consulta nº CNS 36/2020. Texto disponible en: https://apdcat.gencat.cat/web/.content/Resolucio/Resolucions_Cercador/Dictamens/2020/Documents/ca_cns_2020_036.pdf

[811] Como ahora veremos, las bases de legitimación para estos tratamientos de acuerdo al artículo 22.2.b) del RGPD son: i) el consentimiento, ii) la existencia de una norma estatal o derecho de la UE que expresamente regule el tratamiento o, iii) que dicho tratamiento sea necesario para la ejecución de un contrato. Véase el Capítulo V, apartado III, punto 4, epígrafe A) de esta tesis.

[812] Agencia Española de Protección de Datos. Informe Nº 175/2018. Págs.7 y 11.

las Administraciones públicas cuando pretendan desplegar sistemas de toma de decisiones automatizadas ya que normalmente existirán otras actividades o tratamientos de datos que resultando menos invasivos, lograrán también ejercer adecuadamente la competencia atribuida. Así, los tribunales sí que han sido más restrictivos a la hora de justificar el carácter necesario de la atribución de la competencia valorando la adecuación y pertinencia de los datos tratados por ese tratamiento[813], sobre todo, con relación a la publicación de datos personales excesivos en los portales de transparencia amparados en las funciones públicas atribuidas por estas norma[814]. Es decir, corresponde a las organizaciones que quieran desplegar sistemas de toma de decisiones automatizadas justificar adecuadamente el carácter necesario de los tratamientos de datos que se pretenda llevar a cabo, lo que requerirá un análisis estricto de la proporcionalidad[815]. Recordemos que en este análisis se deberán valorar las garantías específicas que en su caso se implementarán para justificar estos tratamientos de datos.

En tercer lugar, para finalizar, estos tratamientos deberán respetar la normativa de protección de datos por lo que las organizaciones habrán de incorporar todas aquellas medidas y garantías que a lo largo de esta tesis se están señalando cuando pretendan desplegar los sistemas de toma de decisiones automatizadas. Entre estas exigencias resulta conveniente indicar que el responsable ha de ser consiente que los particulares sometidos a estos tratamientos bajo el mecanismo de legitimación previsto en el artículo 6.1.e) podrán oponerse al mismo en virtud del artículo 21.1 del RGPD[816].

5. Intereses vitales

El considerando 46 del RGPD establece que el tratamiento de datos personales también debe considerarse lícito cuando sea necesario para proteger un interés esencial para la vida del interesado o la de otra persona física. Este mecanismo de legitimación

[813] Agencia Española de Protección de Datos. Resolución: Nº AP/00075/2015, pág.9.

[814] Así, tal y como se ha señalado: *la difusión de datos personales en abierto, es decir, con acceso ilimitado, en una página web de la Administración Pública con ocasión de la tramitación de un procedimiento de contratación administrativa está sometida al cumplimiento de las obligaciones establecidas en la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal, y, específicamente, al deber jurídico de tener que recabar el consentimiento de los afectados sobre la recogida y tratamiento de datos cuanto no se revelen imprescindibles, necesarios o pertinentes para el adecuado y regular ejercicio de las funciones públicas, al no poder interpretar de forma expansiva el supuesto de excepción previsto en el artículo 6.2 de la citada Ley Orgánica*. En: Sentencia del Tribunal Supremo (Sala de lo Contencioso-Administrativo, Sección 3ª) Sentencia num. 1007/2019 de 8 julio, FJ 3º.

[815] El principio de proporcionalidad exige la superación del triple test, esto es: el juicio de idoneidad, necesidad y proporcionalidad en sentido estricto. Véase este análisis con relación al uso de sistemas automatizados en el Capítulo III, apartado II, punto 5 de esta tesis.

[816] Véase el Capítulo V, aparatado II, epígrafe A) de esta tesis.

se ha de utilizar como último recurso, por lo tanto, siempre que se pueda, el responsable ha de valerse de otras bases de legitimación.[817]. Como ejemplo se podría acudir a este mecanismo en casos de prevención y control de epidemias, respuestas a situaciones humanitarias, catástrofes, etc. Para todos estos supuestos, el uso de sistemas automatizados puede resultar altamente recomendable [818]. Como es lógico, este mecanismo de legitimación se justifica por la necesidad de tratar datos de forma urgente e inmediata. Si se pretende mantener en el tiempo dichas operaciones, será necesario posteriormente acudir a otras bases de legitimación del tratamiento.

6. El interés legítimo en los tratamientos de datos basados en sistemas de toma de decisiones automatizadas

El interés legítimo constituye una de las bases de legitimación que habilitan al tratamiento de datos personales. Este mecanismo exige tres requisitos o pruebas para que pueda operar: i) que exista un interés legítimo del responsable del tratamiento o de un tercero, ii) que el tratamiento sea necesario para la satisfacción del interés legítimo perseguido y, iii) que en la ponderación del interés legítimo del responsable o de terceros por una parte y del impacto que sobre los derechos y libertades del interesado provoca dicho tratamiento que se pretende efectuar por otra parte, prevalezca el interés del responsable o del tercero. La ponderación realizada por el responsable del tratamiento es la prueba esencial que determinará en gran medida que el tratamiento de datos bajo el amparo de esta base de legitimación es lícito[819]. Como ahora se

[817] Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, págs. 24 y 25.

[818] Se ha justificado a través de este mecanismo de legitimación el uso de sistemas automatizados para luchar contra la Covid. En: COTINO HUESO, L: "Inteligencia artificial, *big data* y aplicaciones contra la COVID-19: privacidad y protección de datos". *IDP. Internet, Derecho y Política*, núm. 31, pág.5. Texto disponible en: http://dx.doi.org/10.7238/idp.v0i31.3244 También en el mismo sentido: VALERO TORRIJOS, J ; CERDÁ MESEGUER, J,L: "Transparencia, acceso y reutilización de la información ante la transformación digital del sector público: enseñanzas y desafíos en tiempos del COVID-19". *EUNOMÍA. Revista En Cultura De La Legalidad*, (19), 2020, pág. 119. Texto disponible en: https://doi.org/10.20318/eunomia.2020.5705

[819] Tal y como señala el grupo del artículo 29. En el resto de bases de legitimación diferentes al interés legítimo los tratamientos de dato se consideran a priori lícitos ya que existe una presunción de que, cumplidas las exigencias previstas para dichos mecanismos, se alcanza el equilibrio entre los diferentes derechos e intereses en juego. En la base del interés legítimo ese equilibrio ha de ser ponderado por el responsable del tratamiento. Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, pág.11. Un estudio sobre cómo nuestros tribunales han interpretado el interés legítimo puede encontrarse en: FERNÁNDEZ-SAMANIEGOJ ; FERNÁNDEZ-LONGORIA,P: "El interés legítimo como principio para legitimar el tratamiento de datos". En RALLO LOMBARTE,A; GARCÍA

argumentará, el uso adecuado de esta base de legitimación puede resultar ideal en algunas de las fases que comprende la utilización de sistemas de decisiones automatizadas. Sin embargo, los especiales riesgos que comportan el proceso algorítmico en estos contextos exigen de unas cautelas especiales. Cabe mencionar que tal y como ha señalado la AEPD, el interés legítimo no es una base adecuada para justificar los tratamientos de datos que llevan a cabo las Administraciones Públicas, quedando reservado este mecanismo al sector privado[820].

A) Existencia de un interés legítimo

De esta manera, lo primero que ha de alegar la organización es el interés legítimo. Por interés hay que considerar el beneficio que el responsable del tratamiento, el tercero o la sociedad puedan obtener de dicho tratamiento. Dicho interés debe cumplir varios requisitos, en primer lugar, ese interés ha de ser lícito, es decir, cumplir con la normativa de protección de datos en particular y el resto del ordenamiento jurídico en general. En segundo lugar, ese tratamiento debe haberse identificado por parte del responsable con la suficiente claridad para que posteriormente pueda realizarse una ponderación adecuada de los intereses y derechos en juego, esto es, los del responsable o tercero por un lado y los de los titulares de los datos por otro. En tercer lugar, el interés alegado ha de ser real, no puede ser especulativo[821].

MAHAMUT,R (eds.): *Hacia un nuevo derecho europeo de protección de datos*. Ed. Tirant lo Blanch, Valencia, 2015, págs. 411 a 462. A nivel general existen otros trabajos que estudian el interés legítimo desde una perspectiva internacional. En: MARTÍNEZ CRUZ, J: "El interés legítimo en el marco de la protección de datos personales". En: MURGA FERNÁNDEZ, J,P ; FERNÁNDEZ SCAGLIUSI, M,A; ESPEJO LERDO DE TEJADA,M (dirs.): *Cuestiones actuales sobre protección de datos en España y México*. Ed. Tirant lo Blanch, Valencia, 2021, págs. 245 y ss.

[820] En palabras de la AEPD: *la Administración no puede utilizar como base jurídica del tratamiento el interés legítimo del apartado f) del párrafo 1 del artículo 6, siempre que se entienda que el apartado e) "misión de interés público" habrá de interpretarse en un sentido amplio de forma que permita a las Administraciones, incluso en el ámbito del Derecho Privado, los tratamientos de datos personales necesarios para las finalidades legítimas que el ordenamiento les concede o permite*. En: Agencia Española de Protección de Datos. Informe Nº 175/2018, pág.10.

[821] Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, pág.30.

Por ejemplo, el TJUE ha considerado intereses legítimos de un particular que instala un sistema de videovigilancia la protección de sus bienes, la salud y la vida de él mismo y la de su familia. SENTENCIA DEL TRIBUNAL DE JUSTICIA (Sala Cuarta) de 11 de diciembre de 2014, asunto C-212/13, caso Ryneš**,**, FJº 34. Resolución disponible en: https://curia.europa.eu/juris/document/document.jsf?docid=160561&doclang=ES

Por tanto, en la *fase del diseño* de los sistemas automatizados, las organizaciones que pretendan acudir a la vía del interés legítimo han de identificar el beneficio esperado que pretenden por ejemplo con la analítica de datos o la creación de un determinado modelo algorítmico. Identificado el beneficio, se ha de valorar si es lícito, claro y real. Por ejemplo, una entidad bancaria puede llevar a cabo una analítica masiva de datos de antiguos clientes con el objetivo de crear un modelo algorítmico que detecte el fraude en el uso de sus tarjetas de crédito o la evaluación de la solvencia financiera[822]. Los beneficios esperados de este tratamiento son tanto para la entidad bancaria como en su caso para la sociedad. Por lo que se refiere a la entidad bancaria, esta tiene interés en evitar el uso fraudulento de sus tarjetas de crédito con los correspondientes costes que ello acarrea a estas empresas, además, desde el punto de vista de la sociedad, una correcta asignación del crédito responsable también permite un sistema bancario más saneado[823].

Por otro lado, en *la fase de toma de decisiones*, una plataforma social puede ostentar un interés legítimo en implantar un sistema automatizado para analizar si los contenidos que se comunican en su red cumplen con sus políticas internas de comportamiento. Estas políticas pueden prohibir la emisión de contenidos ilícitos tales como aquellos que inciten al odio, divulgación de imágenes pedófilas, desnudos, posibles fraudes o estafas, etc[824]. El interés o beneficio en este caso nuevamente es doble. Por una parte la organización consigue que se cumplan sus reglas internas y por otra se protegen algunas de las afectaciones a derechos más relevantes que se pueden generar en dicha plataforma respecto de los usuarios de la misma al prevenir la comisión de actuaciones delictivas.

B) El tratamiento es necesario para el interés legítimo del responsable

Una vez que se ha identificado ese interés legítimo, el siguiente paso a valorar es si el tratamiento de datos personales que se pretende llevar a cabo es necesario para la

[822] El considerando 47 del RGPD establece que el tratamiento de datos de carácter personal estrictamente necesario para la prevención del fraude constituye un interés legítimo del responsable del tratamiento.
[823] Sobre los intereses generales del crédito responsable véase: MAS BADIA, M,D: *Sistemas privados de información crediticia. Nueva regulación entre la protección de datos y el crédito responsable*, op.cit., págs. 91 y ss.
[824] Por ejemplo, la red social Facebook establece toda una serie de contenidos que están prohibidos que se publiquen en su plataforma. Entre otros, contenidos que incitan a la violencia, al odio, a la delincuencia, contenido que fomenten la explotación sexual, acoso de menores, etc. Véase: https://www.facebook.com/communitystandards/introduction

satisfacción del interés legítimo perseguido por el responsable del tratamiento y previamente identificado[825]. Una vez más, se habrá de valorar si la organización dispone de otros medios menos invasivos para servir al mismo fin Así, a modo de ejemplo, el Reglamento Europeo sobre la lucha contra la difusión de contenidos terroristas en línea impone a los prestadores de servicios de alojamiento de datos el deber de eliminar y controlar dicho contenido[826]. Teniendo en cuenta el tráfico incesante de información que potencialmente puede considerarse terrorista que circula por estas plataformas, resultará adecuado acudir a medios automatizados, esto es, algoritmos para controlar dicho contenido[827].

C) Ponderación entre los intereses legítimos y los derechos en juego

Por último**,** el artículo 6.1.f) exige al responsable que dicho interés legítimo y necesario sea ponderado con los intereses y derechos de los particulares. La prueba de ponderación obliga al responsable a; i) evaluar el interés legítimo que alega, ii) valorar el impacto que puede generar el tratamiento proyectado sobre los particulares y , iii) valorar el equilibrio resultante entre el interés legítimo alegado y los impactos que genera a los particulares.

En el supuesto en el que el responsable del tratamiento siga considerando que los derechos de los particulares han de preponderar respecto del interés legítimo que alega el responsable, este último debe compensar la balanza previendo mayores medidas de garantía y, en los casos en los que dichas medidas de garantía no logren balancear lo suficiente el juicio en su favor, abstenerse de ejecutar ese tratamiento, al menos, sobre esa base de legitimación.

La prueba de ponderación pondrá sobre la balanza por un lado los intereses legítimos, los cuales pueden ser; insignificantes, poco importantes, importantes o

[825] Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, pág.34.
[826] Artículo 5 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.
[827] Es importante destacar que el Reglamento no obliga a las plataformas a utilizar sistemas automatizados de control del contenido aunque sí que lo permite cuando estas lo consideran adecuado. Considerando 25 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.

imperiosos y por otro lado[828], el impacto o la repercusión hacia los interesados, el cual puede ir desde despreciable hasta máximo.

Interés legítimo del responsable, tercero, sociedad	
1. Interés insignificante	3. Interés importante.
2. Interés poco importante	4. Interés imperioso

Grado de impacto sobre los interesados	
1. Impacto despreciable	3. Impacto significativo
2. Impacto limitado	4. Impacto máximo

De esta manera, cuando se llegue a la conclusión de que el interés perseguido es insignificante o poco importante, salvo que el impacto en los particulares se considere despreciable o limitado, no se podrá acudir a la base de legitimación del interés legítimo. A su vez, tal y como establece el GT29, un interés legítimo importante o imperioso puede en algunos casos y sujeto a garantías y medidas adecuadas justificar incluso un impacto significativo en los intereses o derechos de los interesados[829]. Por ejemplo, continuando con el análisis del Reglamento Europeo sobre la lucha contra la difusión de contenidos terroristas en línea, el despliegue de un sistema de toma de decisiones automatizadas que controla el contenido de la información que se vierte en esas plataformas presenta un interés imperioso tanto para las plataformas como para el conjunto de usuarios y la sociedad[830]. A su vez, el impacto que puede generar la irrupción de estos sistemas en los usuarios que ven eliminado o bloqueado la información que suben también presenta un impacto significativo o máximo. Para situar

[828] El TJUE ya se ha encargo en alguna ocasión de hacer alusión a la importancia de los intereses del responsable del tratamiento en tratar los datos como elemento relevante para legitimar el mecanismo del interés legítimo. En: SENTENCIA DEL TRIBUNAL DE JUSTICIA (Sala Cuarta) de 14 de septiembre de 2000, asunto C-369/98, caso Fisher, FJº 23 y ss, en especial el 28. Resolución disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:61998CJ0369&from=SV

[829] Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, pág.36 *in fine*.

[830] La presencia de contenidos terroristas en línea ha demostrado ser un catalizador para la radicalización de individuos que puede conducir a la comisión de actos terroristas y, por tanto, tiene graves consecuencias negativas para los usuarios, los ciudadanos y la sociedad en general, así como para los prestadores de servicios en línea que alojan esos contenidos, dado que menoscaba la confianza de sus usuarios y daña sus modelos de negocio. Considerando 5 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.
La doctrina ha diferenciado distintas técnicas que utilizan las grandes plataformas para detectar a través de algoritmos el contenido terrorista. En: VEDASCHI, A: "Intelligenza artificiale e misure antiterrorismo alla prova del diritto costituzionale". En VV.AA : *Liber Amicorum per Pasquale Costanzo – Diritto costituzionale in trasformazione Vol. I – Costituzionalismo, Reti e Intelligenza artificiale. Ed.* Associazione giuridica scientifico-culturale. Génova, 2020, pág.499. Texto disponible en: https://www.giurcost.org/studi/COLLANA/Liber_Pasquale_Costanzo_Tomo_I.pdf

la balanza en favor del interés legítimo del responsable, este último deberá además prever medidas de garantía suficientes.

c.1) Factores a sopesar a la hora de evaluar el interés del responsable y el impacto de la medida en los derechos de los interesados

Son diversos los factores que pueden ser relevante para evaluar tanto el interés legítimo del responsable o de terceros y los impactos en la esfera de los derechos e intereses de los particulares y equilibrar la balanza en uno u otro sentido. Tomaremos como base aquellos elementos que han sido indicados por el Grupo del Artículo 29.

Por lo que se refiere al *interés legítimo alegado por el responsable*, se han nombrado distintos factores que pueden ser muy relevantes. En primer lugar, cuando el interés identificado por el responsable no sólo beneficie al mismo sino a terceros o a la sociedad en general, ello se considerará favorable en la ponderación[831]. Así, por ejemplo, el algoritmo de YouTube que se encarga de bloquear el contenido que se difunde en esta red social potencialmente contrario a la normativa de derechos de autor, no sólo beneficia a esta plataforma para evitar que sobre la mismas recaigan responsabilidades por no evitar esas conductas fraudulentas[832], sino que también favorece a los titulares de aquellas obras que están protegidas por las normas sobre propiedad intelectual y que pueden verse afectados por un incorrecto uso de tales contenidos en las mentadas plataformas. Otro ejemplo pueden ser los algoritmos que utiliza Amazon para detectar que las reseñas que realizan los usuarios son realmente

[831] El TJUE considera que el interés económico que ostenta un gestor de motor de búsqueda cuando lleva a cabo los tratamientos de indexación de enlaces no es suficiente para superar el grado de afectación de los derechos de los particulares que se ven sometidos a tal tratamiento. Sin embargo, si junto a ese interés económico del responsable del tratamiento le sumamos el interés de los internautas por conocer la información que se deriva de ese tratamiento, en determinadas circunstancias, esa ponderación puede resultar favorable para el lado del responsable del tratamiento respecto del titular de los datos que se tratan. En: SENTENCIA DEL TRIBUNAL DE JUSTICIA (Gran Sala) de 13 de mayo de 2014, asunto C-131/12, caso Mario Costeja y Google Spain, S.L, FJ 81º. Resolución disponible en: https://curia.europa.eu/juris/document/document.jsf;jsessionid=3822C428505144A215380795155ECD14?text=&docid=152065&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=729025
En términos aún más restrictivos se ha mostrado la AEPD con relación al beneficio o interés económico alegado por el responsable como interés legítimo. Así, esta autoridad de control ha indicado que la obtención de un beneficio económico es un interés legítimo. Sin embargo, este último en ningún caso podrá prevalecer sobre el derecho fundamental a la protección de datos de las personas afectadas. En: Agencia Española de Protección de Datos. Informe Nº: PS/00070/2019, pág.128. Resolución disponible en: https://www.aepd.es/es/documento/ps-00070-2019.pdf

[832] BURK,L,D: "Algorithmic Fair Use". *University of Chicago Law Review*, 283 (2019), pág.289. Texto disponible en: https://ssrn.com/abstract=3076139.

verdaderas[833]. En estos supuestos, no sólo se beneficia esta organización sino también los potenciales consumidores que acuden a dichas reseñas a la hora de obtener esos productos. En segundo lugar, también resultará positivo para el interés del responsable que dicho interés, ya sea público o privado, se desprenda o esté latente en el ordenamiento jurídico. En este sentido, conforme el interés legítimo alegado resulte más explícitamente reconocido en una norma o instrumento jurídico, más peso tendrá el mismo en la balanza[834]. Es importante destacar que no siempre habrá que buscar el sustento en una norma vinculante. Cabe la posibilidad de acudir a directrices, guías e incluso[835] en determinados casos a las tradiciones o costumbres de una determinada comunidad[836]. Existen diversas normas, las cuales ya han sido analizadas, que directa o indirectamente respaldan el uso de sistemas automatizados como mecanismo necesario para proteger intereses públicos y privados.

Norma que implícita o explícitamente promueven uso de sistemas algorítmicos	**Interés público o privado presente**
Directiva 2019/790 del Parlamento Europeo y del Consejo de 17 de abril de 2019 sobre los derechos de autor.	-Buen funcionamiento del mercado interior. -Protección de los derechos de autor.
Reglamento 2021/784 del Parlamento Europeo y del Consejo de 29 de abril de 2021 sobre la lucha contra la difusión de	-Lucha contra el terrorismo. -Confianza en las plataformas. -Garantizar el correcto funcionamiento del

[833] El uso de algoritmos basados en *machine learning* por parte de Amazon tiene como objetivo encontrar aquellas reseñas fraudulentas que únicamente tienen como objetivo obtener un pago de dinero por parte del dueño del producto a cambio de una reseña positiva. Fuente de la noticia: RUS,C: "Una base de datos filtrada desvela un esquema con cientos de miles de personas implicadas en reseñas falsas en Amazon", *Xataka*, 10/05/2021. Disponible en: https://www.xataka.com/empresas-y-economia/base-datos-filtrada-desvela-esquema-cientos-miles-personas-implicadas-resenas-falsas-amazon

[834] Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, pág.43.

[835] Por ejemplo, podría encontrarse ese interés en documentos elaborados por representantes de un determinado colectivo. No obstante, estos documentos deben interpretarse con recelo ya que normalmente pueden estar sólo pensando en los intereses del colectivo que representa. Aquí lo ideal es que en su caso, una determinada autoridad de control o ratifique. A través de códigos de conducta aprobados por la autoridad de control se puede conseguir esa mayor legitimidad. En el sector de los seguros, se considera interés legítimo la publicidad de marketing de sus propios clientes relacionado con los productos aseguradores que se envíe a personas que se adecúe al perfil. En: "Guía para el tratamiento de los datos personales por las entidades aseguradoras", Asociación empresarial del seguro, febrero 2019, pág.26.

[836] Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, pág.43

contenidos terroristas en línea.	mercado único digital en una sociedad abierta y democrática.
Propuesta del Parlamento Europeo y del Consejo sobre un Mercado Único de Servicios Digitales.	-Lucha contra la desinformación en la red. -Salvaguardar y mejorar el funcionamiento del mercado interior.

Hay que recordar no obstante que para estos supuestos en concreto es posible que el responsable pueda acudir a otros mecanismo de legitimación como los descritos en el artículo 6.1.c) (obligación legal) o en el artículo 6.1.e) (misión o interés público).

Por otro lado, en la otra parte de la balanza, *los impactos en la esfera de los derechos e intereses de los particulares* también se han de evaluar. Aquí puede venir bien la metodología que se expuso durante los apartados referidos al análisis de riesgos y la EIPD a la hora de valorar los riesgos potenciales, teniendo en cuenta el impacto que genera una hipotética amenaza a los particulares y la probabilidad de que esas amenazas se materialicen[837]. En primer lugar, un elemento importante es la naturaleza de los datos. Así, cuanto más sensible sea la información que se pretende tratar, mayores consecuencias puede generar en la esfera de los mismos[838]. Una vez más, se han de tener en cuenta la inferencias que se pueden derivar de los resultados vertidos por el algoritmo. Por otro lado, con relación al uso de datos obtenidos de fuentes públicas[839], en principio, el tratamiento de los mismos tendrá menos incidencia en la esfera de los particulares[840], sobre todo, si estos últimos los han hecho manifiestamente públicos. Ahora bien, el hecho de que se hayan puesto a disposición del público esos datos para una finalidad determinada no supone una carta blanca para su reutilización

[837] Véase el Capítulo III, apartado I (análisis de riesgos) y Capítulo III, apartado II (EIPD).

[838] Por ejemplo, la ICO ha indicado que la divulgación de los datos referidos a la vida profesional de una persona representa una intrusión menor en la esfera de esta respecto de la divulgación de los datos correspondientes a su vida privada. Véase la resolución IC-45844-V2X8 , apartados 68 y 77. Disponible en:https://ico.org.uk/media/action-weve-taken/decision-notices/2020/2618241/ic-45844-v2x8.pdf

[839] Sobre los datos obtenidos de fuentes accesibles al público puede verse un resumen en: Fuente de la noticia: GARCÍA HERRERO,J: "Fuentes accesibles al público y RGPD (OjoAlDato)". *Blog de Jorge García Herrero*. 07/04/2021. Información disponible en: https://jorgegarciaherrero.com/fuentes-accesibles-al-publico-y-rgpd-ojoaldato/

[840] El TJUE ha considerado que el impacto en los derechos de los individuos es mayor en aquellos tratamientos de datos que figuran en fuentes no accesibles al público que cuando se tratan datos de fuentes de acceso público. Este elemento se ha de tener en cuenta a la hora de ponderar los intereses del responsable y los derechos de los particulares. SENTENCIA DEL TRIBUNAL DE JUSTICIA (Sala Tercera) de 24 de noviembre de 2011, asuntos C-468/10 y C-469/10, caso ASNEF, FJº 44 y 45.
Resolución disponible en:
https://curia.europa.eu/juris/document/document.jsf;jsessionid=6C74FDF9A94A64A54BAAADCBF1F6E737?text=&docid=115205&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=136248

por el mecanismo del interés legítimo[841]. La ponderación debe continuar ya que en estos casos la expectativa del titular de los datos publicados para que sean tratados de otra manera puede ser totalmente inexistente[842]. En segundo lugar, también son relevantes los mecanismos utilizados para tratar los datos. Tal y como señala el GT29, la disponibilidad de métodos alternativos para conseguir los objetivos perseguidos por el responsable del tratamiento con menos impacto negativo sobre el interesado resulta una consideración pertinente en este contexto[843]. Así, establecer sistemas de toma de decisiones automatizados más transparentes en defectos de otros más opacos puede ser un elemente a tener en cuenta que reduce el impacto sobre los particulares. En tercer lugar, otro elemento a destacar es referido a la posición del responsable del tratamiento y del interesado. Factores como el tipo de organización o el grupo de personas al que va destinada la medida que se pretende implantar también se han de tener en cuenta. Así, si dicho tratamiento va dirigido a personas vulnerables, como podría ser el colectivo de menores[844], ello puede ser motivo de una mayor protección en favor de los intereses y derechos de estos últimos. En cuarto lugar, en nuestra opinión, también será relevante la cantidad de datos que se pretendan recopilar. Este elemento es especialmente importante durante la fase de diseño de los sistemas automatizados. Como ha indicado la *Information Commissioner's Office* del Reino Unido, a través del mecanismo del interés legítimo las organizaciones pueden tener un mayor margen para experimentar y tratar con diferentes tipos de datos en las fases iniciales. Ahora bien, estas variables o datos

[841] Tal y como señala la AEPD en su resolución Nº: PS/00136/2020. El hecho de que las Administraciones Publicas o los órganos jurisdiccionales publiquen documentos con datos de carácter personal no implica en ningún caso que los datos objeto de tratamiento tengan la naturaleza de datos abiertos. (pág.11). El hecho de que un dato sea accesible por cualquiera puede ser tenido en cuenta a la hora de realizar la ponderación cuando se formule el tratamiento al amparo de una base legitimadora como el interés legítimo. Ahora bien, ello no implica necesariamente que el tratamiento vaya a ser lícito ya que se deben respetar los restantes principios del RGPD.(pág.12). Resolución disponible en: https://www.aepd.es/es/documento/ps-00136-2020.pdf

[842] Véase la Resolución de la AEPD que sancionó a Equifax por elaborar un fichero en el que se incorporaban los datos de personas que habían recibido alguna sanción administrativo por parte de la Administración. Estos datos se publicaron en los correspondientes boletines oficiales. Una vez que se recopilaban, los mismos se incorporaban a un fichero, al cual, posteriormente podían acceder terceros interesados para comprobar que esa persona tenía o no deudas con las Administraciones Públicas. Esta información se tenía en cuenta durante la concesión de créditos. Agencia Española de Protección de datos. Resolución Nº: PS/00240/2019, págs. 147 y 148. Resolución disponible en: https://www.aepd.es/es/documento/ps-00240-2019.pdf

[843] Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014, pág.47.

[844] Artículo 6.1.f) *in fine* del RGPD. También en el mismo sentido, SENTENCIA DEL TRIBUNAL DE JUSTICIA (Sala Segunda) de 4 de mayo de 2017, asunto C-13/16, caso Rīgas, FJº 33. Resolución disponible en: https://curia.europa.eu/juris/document/document.jsf?text=&docid=190322&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=138572

elegidos deberán en la medida de lo posible justificarse[845]. Por tanto, la mera posibilidad de que algunos datos puedan ser útiles para una predicción no es por sí misma suficiente para que la organización demuestre que el procesamiento de estos datos es necesario para construir el modelo[846]. Estas últimas apreciaciones resultan sobre todo relevantes cuando en la analítica de datos se pretenda utilizar técnicas de aprendizaje no supervisado.

Una vez se han sopesado los dos lados de la balanza, es posible establecer un equilibrio. Si se extrae la conclusión de que el interés legítimo del responsable prevalece sobre los derechos y los intereses de los afectados, se entiende que la base de legitimación es lícita. Cuando no se llegue a la conclusión anterior o no esté clara la preponderancia en favor del interés legítimo, el responsable deberá adoptar o prever las medidas de garantías adecuadas.

c.2) Garantías

Las garantías y medidas presentan un doble objetivo en este contexto, por un lado se reduce el impacto que puede generar el tratamiento proyectado sobre los derechos e intereses de los particulares y por otro, y como consecuencia del primero, se puede alterar el equilibrio para que prevalezca el interés legítimo del responsable del tratamiento de datos respecto de los impactos en los derechos de los titulares de los datos.

Muchas de las garantías que ahora mencionaremos ya se han ido desarrollando en las distintas páginas de este estudio. Podemos destacar tres conjuntos: el primer grupo se refiere a todo un conjunto de medidas relacionadas con el principio de minimización de datos. Aquí encontramos la aplicación de técnicas de anonimización, seudonimización, o agregación de datos. En nuestra opinión, y por lo que se refiere a la fase de diseño, será muy relevante realizar un estudio de minimización de datos que vaya justificando en cada una de las etapas que comprende esta fase las variables elegidas por parte del responsable del tratamiento. Dicho estudio se analiza en el

[845] Information Commissioner´s office. Información disponible en: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/what-do-we-need-to-do-to-ensure-lawfulness-fairness-and-transparency-in-ai-systems/

[846] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.41.

apartado referido al principio de minimización de datos[847]. Un segundo grupo de técnicas están relacionadas con las obligaciones de transparencia sobre el tratamiento proyectado por parte del responsable. En este sentido, no podemos olvidar que las organizaciones están obligadas a informar a los particulares cuando se acude al interés legítimo[848]. Finalmente y quizás el grupo más interesante correspondería con aquellas medidas que permitan a los titulares de los datos poder voluntariamente quedar excluidos de este tratamiento. Ello sobre todo puede ser adecuado cuando el equilibro y la ponderación realizada no sean del todo favorables para el responsable[849]. En este sentido, no ya como una garantía extra, sino como una medida obligatoria, el responsable siempre deberá habilitar la facultad a los interesados para poder oponerse al tratamiento de los datos personales tal y como regula el artículo 21 del RGPD.

Por último, cabe indicar que todo el proceso descrito anteriormente para justificar el mecanismo del interés legítimo como base para tratar los datos se ha de documentar debidamente y ha de quedar en disposición de la autoridad de control que lo requiera. La idea es que quede demostrado que el responsable ha realizado las diferentes pruebas para sopesar los distintos intereses en juego. Entendemos por tanto que si no existe tal documentación y justificación, el tratamiento no estará debidamente legitimado[850], aparte claro está del incumplimiento del principio de responsabilidad activa reconocido en el artículo 5.2 del RGPD.

---

[847] Véase el Capítulo IV, apartado III, punto 1 de esta tesis.
[848] Artículos 13 y 14 del RGPD.
[849] Agencia Española de Protección de Datos. Informe Jurídico. N/REF: 0017/2019, págs.21 y ss.
[850] La autoridad de protección de datos islandesa consideró que una empresa que había utilizado los datos personales de una persona con fines de marketing podía tener un interés legítimo en realizar dicho tratamiento. No obstante, dado que no se ha justificado dicho interés legítimo y no se ha realizado la prueba de sopesamiento de los intereses en juego, esto es, los del responsable y los del interesado, el principio de licitud no se cumplió. Autoridad de Protección de Datos Islandesa (Persónuvernd) Nº de resolución:2020010673. Disponible en:
https://www.personuvernd.is/urlausnir/urskurdur-um-vinnslu-personuupplysinga-af-halfu-elisu-gudrunar-ehf.-lifandi-visinda
En este mismo sentido, la Autoridad de Protección de Datos Finlandesa (Tietosuojavaltuutetun toimisto) también apreció que la falta de justificación y documentación del interés legítimo invalidaba el principio de licitud. Véase la resolución 8393/161/2019 de 26 de mayo de 2020. Texto disponible en:https://tietosuoja.fi/-/tietosuojavaltuutetun-toimiston-seuraamuskollegio-maarasi-hallinnollisen-seuraamusmaksun-useista-puutteista-henkilotietojen-kasittelyssa?_101_INSTANCE_ajcbJYZLUABn_languageId=en_US

D) Las especiales exigencias del interés legítimo en la elaboración de perfiles y la toma de decisiones automatizadas. La necesidad de legitimar determinados tratamientos de analítica de datos por vía legal

El instrumento del interés legítimo descarga en el responsable del tratamiento toda una serie de exigencias, las cuales, una vez acreditadas y justificadas permiten el tratamiento de datos pretendido. Los riesgos derivados de la automatización presentes en los tratamientos basados en la elaboración de perfiles y la toma de decisiones automatizadas son por sí mismos muy altos[851]. Los responsables que pretendan basar un tratamiento de datos de estas características a través del interés legítimo han de ser especialmente cautelosos. Las garantías y salvaguardas que garantizan una adecuada legitimación de este tipo de actividades serán superiores a otros tratamientos que presentan menos riesgos para los intereses de los particulares. En este sentido, resulta interesante traer a colación una resolución de la AEPD que examina la posible legitimación a través del interés legítimo de un tratamiento de datos basado en la elaboración de perfiles[852]. En este caso, el responsable del tratamiento pretendía legitimar la elaboración de perfiles con la finalidad de "conocer mejor al cliente y mejorar su experiencia" Pues bien, para la AEPD, el interés legítimo no puede considerarse un mecanismo de legitimación válido en este supuesto ya que el responsable no adoptó toda una serie de medidas que garantizasen un adecuado cumplimiento del principio de licitud. Entre las razones que se tuvieron en cuenta destacamos: i) El interés legítimo alegado por el responsable resultaba vago y especulativo, ii) el número de datos recopilados era excesivo y la combinación ilimitada de estos también, iii) la finalidad de obtener algoritmos y la falta de transparencia sobre la lógica del tratamiento consistente en la elaboración de perfiles puede generar problemas de discriminación. iv) la conservación de los datos para este tratamiento de dos años resulta excesiva, v) la posición dominante del responsable frente al interesado por su condición de gran empresa y una de las líderes del mercado en su sector[853].

Para la AEPD, todos estos elementos son indicadores que afectan negativamente a la consideración del interés legítimo por parte del responsable en este tipo de

[851] Agencia Española de Protección de Datos. Resolución Nº: PS/00120/2021,pág.94.
[852] La entidad bancaria BBVA utilizó como mecanismos de legitimación el interés legítimo para autorizar la elaboración de perfiles con el objetivo de desarrollar modelos algorítmicos. Agencia Española de Protección de Datos. Resolución Nº: PS/00070/2019**.** Págs. 104 y ss. Resolución disponible en: https://www.aepd.es/es/documento/ps-00070-2019.pdf
[853] Vid, págs. 128,129 y 130.

tratamientos de datos. Esta resolución marca una hoja de ruta a los futuros responsables que pretendan legitimar mediante el interés legítimo la elaboración de perfiles y la toma de decisiones automatizadas[854]. Esta resolución no veta la vía del interés legítimo como mecanismo que autoriza los tratamientos de datos algorítmicos previamente comentados pero sí deja clara la postura de esta autoridad en relación con las exigencias y garantías mínimas que han de desplegar los responsables del tratamiento si pretende acudir a dicha base de legitimación.

En nuestra opinión, para evitar entornos de inseguridad jurídica que pueden llevar a la paralización de sectores importantes que apuestan por el desarrollo de la inteligencia artificial. Sería recomendable que en estos contextos el legislador acabara legitimando expresamente tratamientos relacionados con la analítica masiva de datos donde los riesgos para los particulares sean menores a los beneficios generales que dichos tratamientos pueden generar a los responsables del tratamiento tanto en el sector público como el privado. Esta propuesta no es nueva, así, la LOPD de 2018 ha reconocido la licitud de toda una serie de tratamientos de datos con diversas finalidades[855]. Antes de la entrada en vigor de dicho texto normativo, los responsables que pretendieran llevar a cabo los mencionados tratamientos con esas finalidades debían alegar frecuentemente intereses legítimos y acudir a la prueba de ponderación que exige la normativa de protección de datos. Ello ya no es necesario ya que existe una presunción de licitud de tales tratamientos. Algo parecido creemos que ha de reconocerse para los tratamientos de analítica masiva de datos de bajo riesgo donde existe una fuerte incertidumbre sobre las pautas y garantías que los responsables han de implantar para entender respetada la normativa de protección de datos. El reconocimiento legal de este tipo de actividades facilitaría el despliegue de estos

[854] El responsable del tratamiento, en este caso la entidad bancaria BBVA, ya ha indicado que recurrirá esta decisión ante los tribunales de justicia. Si finalmente lo hace, este asunto se analizará en sede judicial. Fuente de la noticia: CANO,F ; ALBA,C: "BBVA dice que no ha vulnerado la ley de protección de datos y que recurrirá la multa de cinco millones". *El Español*. 15/12/2020. Información disponible en: https://www.elespanol.com/invertia/empresas/banca/20201215/bbva-no-vulnerado-ley-proteccion-recurrira-millones/543695987_0.html

[855] Algunos de estos tratamientos ya estaban regulados por la antigua normativa en materia de protección de datos. Sin embargo, la LOPD de 2018 ha aclarado la licitud de estos tratamientos. Estos son: tratamiento de datos de contacto, de empresarios individuales y de profesionales liberales, sistemas de información crediticia, tratamientos con fines de videovigilancia, tratamientos relacionados con la realización de determinadas operaciones mercantiles, sistemas de exclusión publicitaria, sistemas de información de denuncias internas, etc**.** Artículos 19 a 24 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

tratamientos de datos dentro de un marco de garantías adecuadas que la propia norma debería prever. Dichas garantías deberían seguir en parte el modelo ya comentado y que la propia LOPD de 2018 ha reconocido para los tratamientos de datos con fines de salud e investigación biomédica[856].

Mientras llega o no llega esa hipotética reforma legal, los responsables deberían tomar como referencia el listón de garantías previamente indicadas por la AEPD.

Recordemos por lo demás que a través del interés legítimo no es posible basar los tratamientos de datos definidos en el artículo 22 del RGPD. Estos son: la toma de decisiones plenamente automatizadas relevantes con o sin elaboración de perfiles[857].

7. La legitimación de los tratamientos de datos de categoría especial en los sistemas de toma de decisiones automatizadas

El RGPD prevé un régimen más protector para los interesados cuando el responsable del tratamiento pretenda tratar datos de categorías especial. Así, el artículo 9 establece una prohibición general para tratar estos datos por parte de los responsables. Ahora bien, este mismo precepto en su apartado segundo contiene unan serie de supuestos en los que se levanta la prohibición previamente mencionada, permitiendo así el tratamiento de estos datos para esos concretos casos. De esta manera, el legislador europeo ha considerado que los tratamientos de estos datos no pueden ser legitimados a través de cualquier mecanismo de legitimación sino sólo mediante aquellos cuya habilitación venga establecida por algunas de las excepciones previstas en ese precepto. En este sentido, tal y como señala el Tribunal Constitucional, el artículo 9 del RGPD reúne un conjunto de supuestos que engloba a: i) tratamientos que tienen un ámbito de aplicación muy acotado como el laboral, social, asociativo, sanitario, judicial, etc. ii) Tratamientos que responden a una finalidad determinada, por lo que, en sí mismas, delimitan los tratamientos específicos que autorizan como excepción a la regla general. iii) Tratamientos cuya eficacia habilitante queda en manos del derecho de Derecho de la Unión o el de los Estados miembros para que estos los prevean y regulen expresamente en sus respectivos ámbitos de sus competencias[858]. Para todos estos supuestos, el

[856] En coherencia con la ya comentado véase el Capítulo III, apartado IX, punto 3 de esta tesis.
A su vez, véase también la D.A 17ª de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.
[857] Para más información véase el Capítulo V, apartado III, punto 2 de esta tesis, así como el artículo 22 del RGPD.
[858] STC Sentencia 76/2019, de 22 de mayo de 2019. FJº 4. Texto disponible en: https://hj.tribunalconstitucional.es/HJ/docs/BOE/BOE-A-2019-9548.pdf

responsable deberá basar su tratamiento en uno de los mecanismos descritos en el artículo 6 del RGPD ya comentados y además en una de las excepciones contempladas en el artículo 9.2 de este mismo texto[859]. El tratamiento de datos de categoría especial resulta frecuentemente utilizado por parte de los responsables del tratamiento tanto en la fase de desarrollo como durante la toma de decisiones de ahí que sea muy relevante valorar estas excepciones para entender cumplido el principio de licitud.

A) Consentimiento explícito

La primera circunstancia que habilita al responsable para tratar los datos personales especialmente protegidos es el consentimiento explícito del interesado. Nos remitimos a lo indicado en el apartado referido a esta base legal[860]. Tal y como señala el artículo 9.a), el consentimiento explícito en estos contextos puede ser prohibido si así lo establece una norma de derecho de la UE o de los Estados Miembros. En este sentido, el artículo 9.1 de la LOPD de 2018 establece que el solo consentimiento del afectado no bastará para levantar la prohibición del tratamiento de datos cuya finalidad principal sea identificar su ideología, afiliación sindical, religión, orientación sexual, creencias u origen racial o étnico. Es decir, España ha recogido el guante del legislador europeo y ha considerado que el consentimiento por sí solo no puede resultar una base legítima adecuada cuando se pretendan tratar datos personales con las finalidades mencionadas en dicho precepto. Ahora bien, como se puede apreciar, las finalidades especiales que se prohíben por el artículo 9.1 de la LOPD de 2018 son menores que las contempladas por el artículo 9.1 del RGPD. Además, téngase en cuenta que el precepto de la LOPD de 2018 habla de finalidades especiales y no de datos de categoría especial.

Finalidades de los datos de categoría especial contempladas en:	
Artículo 9.1. RGPD	**Artículo 9.1. LOPD de 2018**
Datos que revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos, datos	Finalidad principal que sea identificar la ideología, afiliación sindical, religión, orientación sexual, creencias u origen racial o étnico.

[859] Comité Europeo de Protección de Datos. *Directrices 03/2020 sobre el tratamiento de datos relativos a la salud con fines de investigación científica en el contexto del brote de COVID-19* Adoptadas el 21 de abril de 2020, apartado 15, pág.6.
[860] Capítulo IV, apartado I, punto 1, epígrafe C).

biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o las orientación sexuales de una persona física.	

Es por ello que, en nuestra opinión, no todos los datos de categoría especial quedan afectados por esta prohibición general del legislador español, ni tampoco todas las finalidades consideradas especialmente sensibles por el RGPD. Así, destacamos dos supuestos donde esta prohibición del uso del consentimiento no se aplica:

i) el tratamiento de datos convencionales o de categorías especiales que no respondan a las finalidades indicadas por el artículo 9.1. LOPD de 2018, y ello, a pesar de que la finalidad pretendida sí que esté prevista en el RGPD como especialmente sensible. Por ejemplo, utilizar la forma de andar de una persona o el tipo de pisada con el fin de identificarla unívocamente a una persona a la hora de acceder a determinados edificios[861]. En este caso, el uso del dato convencional para una finalidad especial. Otro ejemplo, el uso del historial clínico para el pago de la factura de la luz. En este segundo caso, el uso del dato de categoría especial para una finalidad no sensible o especial. O por ejemplo y más explícito, el uso de ese mismo historial clínico para evaluar el riesgo de contraer o no una enfermedad. En este último caso, uso de dato de categoría especial para una finalidad tildada de sensible por el RGPD pero no prohibida por la LOPD de 2018.

ii) Tampoco quedaría afectados por esta prohibición los tratamientos de datos convencionales que resulten ser *proxies* de categorías especiales si tampoco se pretende una de las finalidades indicadas por la norma.

En ambos supuestos, tanto los datos convencionales como los estrictamente especiales son considerados especiales a los efectos del RGPD[862]. Es por ello que para el tratamiento de los mismos se seguirá requiriendo del uso de alguna de las excepciones previstas en el artículo 9.2 del RGPD y una de las bases de legitimación contempladas en el artículo 6 de ese mismo texto.

[861] Recordemos que el artículo 9.1 del RGPD establece como dato de categoría especial aquellos datos biométricos dirigidos a identificar de manera unívoca a una persona física.

[862] Recordemos que un dato convencional utilizado inicialmente puede convertirse en especial y por tanto considerarse como tal desde el inicio cuando: i) el procesamiento de ese dato inicial por el algoritmo de lugar a un dato inferido de categoría especial. ii) el procesamiento de dicho dato inicial se pretenda con alguna de las finalidades consideradas especiales o sensibles.

B) El tratamiento es necesario para el cumplimiento de obligaciones y el ejercicio de derechos específicos del responsable del tratamiento o del interesado en el ámbito del Derecho laboral y de la seguridad

Se autoriza a tratar datos de categoría especial en estos contextos, si bien, se requiere de una norma del Derecho de la Unión de los Estados miembros o en su caso un convenio colectivo para que autorice tales tratamientos. Además, se han de prever suficientes medidas de garantía. En este sentido, el convenio colectivo del sector de la banca ha regulado los derechos digitales que asisten a los trabajadores en este ámbito. Concretamente se han establecido los derechos que tienen los asalariados ante el uso de la inteligencia artificial por parte de las entidades bancarias[863]. Entre otros, este convenio reconoce: i) el derecho a no ser objeto de decisiones basadas única y exclusivamente en variables automatizadas, salvo en aquellos supuestos previstos por la Ley y también, ii) el derecho a la no discriminación en relación con las decisiones y procesos cuando ambos estén basados únicamente en algoritmos. En estos supuestos, el concurso e intervención de las personas designadas a tal efecto por la empresa será requerida en caso de discrepancia. En parte, este convenio amplía los derechos que se derivan de la normativa de protección de datos ya que las facultades previstas en el mismo no sólo afectan a las decisiones plenamente automatizadas relevantes sino también a las no relevantes.

C) El tratamiento es necesario para proteger intereses vitales del interesado

Así, la LOPD de 2018 habilita a las autoridades sanitarias e instituciones públicas con competencias en vigilancia de la salud pública a realizar estudios científicos sin el consentimiento de los afectados en situaciones de excepcional

[863] Artículo 80.5 de la Resolución de 17 de marzo de 2021, de la Dirección General de Trabajo, por la que se registra y publica el XXIV Convenio colectivo del sector de la banca. Véase este texto en: https://www.boe.es/boe/dias/2021/03/30/pdfs/BOE-A-2021-5003.pdf
Sobre la importancia de los convenios colectivos en esta materia puede verse: TODOLÍ SIGNES,A: "La gobernanza colectiva de la protección de datos en las relaciones laborales: big data, creación de perfiles, decisiones empresariales automatizadas y los derechos colectivos". *Revista de derecho social*, Núm. 84, 2018, págs. 81 y ss. En el mismo sentido véase: MERCADER UGUINA, J: "Algoritmos y derecho del trabajo". *Actualidad Jurídica Uría Menéndez*, 52, 2019, págs. 69 y 70.

relevancia y gravedad para la salud pública[864]. Estos estudios podrán realizarse con técnicas de aprendizaje automático.

D) El tratamiento es efectuado por una fundación, una asociación o cualquier otro organismo sin ánimo de lucro

Para esto supuestos se exige que dicho tratamiento: i) se realice en el ámbito de sus actividades legítimas y con las debidas garantías, ii) la finalidad sea política, filosófica, religiosa o sindical, iii) se refiera exclusivamente a los miembros actuales o antiguos de tales organismos o a personas que mantengan contactos regulares con ellos en relación con sus fines y, iv) siempre que los datos personales no se comuniquen fuera de ellos sin el consentimiento de los interesados.

E) El tratamiento se refiere a datos personales que el interesado ha hecho manifiestamente públicos

El legislador europeo considera que cuando los propios interesados hayan hecho públicos sus datos de categoría especial, la protección reforzada prevista en el artículo 9 queda eliminada. Es decir, se parte de que el propio interesado ha puesto a disposición del público esos datos especiales habilitando con ello a que terceros puedan tratar los mismos. Ahora bien, a pesar de que el particular ha publicado dichos datos de forma voluntaria y se expone al riesgo de que los mismos puedan ser utilizados por terceros, ello no puede significar una total desprotección. De ahí que este tipo de tratamientos también requieran de una serie de requisitos:

Así, en primer lugar, como norma general, al particular se la ha de informar que se está llevando a cabo ese tratamiento de datos. En este sentido, la AEPD ha indicado que las empresas tienen el deber de informar a los potenciales candidatos de un puesto de trabajo cuando estas organizaciones llevan a cabo tratamientos de datos con fines profesionales a la hora de indagar en los perfiles de redes sociales[865]. La idea es que el particular sea consciente que dichos datos que él ha publicado conscientemente están

[864] Véase la D.A.17ª.2.b) de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Véase también el informe nº 2018-0121de la AEPD, pág.6. Texto disponible en: https://www.aepd.es/es/documento/2018-0121.pdf

[865] Agencia Española de Protección de Datos. *La protección de datos en las relaciones laborales*. 2021, pág.22

siendo tratados para una determinada finalidad por parte del responsable del tratamiento.

En segundo lugar, el responsable ha de acreditar que efectivamente el particular ha hecho manifiestamente públicos esos datos. Así, el CEPD, con relación al uso de datos de categoría especial en redes sociales ha indicado que la palabra "manifiestamente" exige un umbral elevado a los responsables para acogerse a esta exención. Es decir, el responsable ha de justificar a través de distintos elementos que el sujeto de los datos ha manifestado claramente la intención de hacerlos públicos[866]. Entre estos elementos se ha de valorar: i) si el interesado realizó una acción específica para cambiar la configuración de la plataforma que por defecto era pública o privada, ii) la naturaleza de la plataforma de redes sociales, es decir, si esta plataforma está intrínsecamente vinculada con la idea de conectar con conocidos cercanos del interesado o crear relaciones íntimas o si está destinada a proporcionar un ámbito más amplio de relaciones interpersonales, como las relaciones profesionales. iii) La accesibilidad de la página donde se publican los datos sensibles, esto es, si la información es de acceso público o si, por ejemplo, es necesario crear una cuenta antes de acceder a la información, iv) si el propio interesado ha publicado los datos sensibles, o si por el contrario, los datos han sido publicados por un tercero o inferidos por un sistema[867].

Como se puede comprobar, será necesario analizar caso por caso las circunstancias para valorar si el particular ha hecho manifiestamente públicos estos datos. Teniendo en cuenta que un número importante de sistemas algorítmicos son entrenados con datos que provienen de las redes sociales, los responsables del tratamiento deben establecer especiales cautelas en asegurarse que dichos datos se han hecho manifiestamente públicos por parte de los particulares. De esta manera el responsable del tratamiento debe demostrar que los datos que obtuvo a través de estas páginas o plataformas realmente fueron manifiestamente publicados. No habrá dudas sobre los datos publicados en blogs, páginas personales, etc. Sin embargo, como hemos comprobado, cuando dicho datos provengan de plataformas o redes sociales, el análisis se vuelve más complejo. A modo de ejemplo, en Francia, una ley autoriza a las

[866] Comité Europeo de Protección de Datos. *Guidelines 8/2020 on the targeting of social media users*. Versión 1.0. Directrices adoptadas el 2 de septiembre de 2020. Apartado 120, págs.32 y 33.

[867] Por ejemplo, el consentimiento otorgado por el usuario que publica una imagen no debe confundirse con la necesidad de contar con una base legítima para el tratamiento de los datos personales de otras personas que pudieran aparecer en la imagen. En: Grupo del Artículo 29. *Dictamen 02/2012 sobre reconocimiento facial en los servicios en línea y móviles*. Adoptado el 22 de marzo de 2012. Pág.7.

autoridades públicas tributarias a recopilar datos manifestados públicamente por parte de los ciudadanos en redes sociales con el fin de elaborar un sistema de inteligencia artificial para detectar el fraude fiscal[868]. El Consejo Constitucional Francés, a la hora de interpretar el concepto de "manifiestamente público" ha indicado que este contenido ha de cumplir dos condiciones acumulativas. Por un lado debe ser contenido de libre acceso en un servicio de comunicación en línea al público de una red social o plataforma. Se excluye por tanto el contenido accesible obtenido una vez que se ingresa una contraseña o después del registro en el sitio en cuestión. Por otro lado, este contenido debe ser claramente hecho público por los usuarios de estos sitios[869]. Como resultado, solo se puede recopilar y utilizar el contenido relacionado con la persona que lo divulgó deliberadamente. Por su parte, la Ley 22/2018 de la Comunidad Valenciana que crea el sistema automatizado de alertas tempranas contra la corrupción prevé la recopilación de datos que los particulares hayan hecho manifiestamente públicos de manera voluntaria. Para ello, se prevén varias condiciones: i) que los espacios donde se publiquen los datos tengan la condición de abiertos y no estén protegidos por el derecho a la intimidad o el derecho al secreto de las comunicaciones y, ii) para el caso de las redes sociales, se trate de espacios abiertos a todos los miembros de la red, redes sociales abiertas o redes cuya naturaleza profesional, empresarial o similar permitan excluir toda expectativa de privacidad[870]. En un tono más restrictivo, la AEPD sólo considera fuente de acceso público *aquella cuya consulta puede ser realizada por cualquier persona*. Quedando por tanto excluidas otro tipo de fuentes en las que el acceso está restringido a un círculo determinado de personas[871].

En definitiva, en nuestra opinión, los responsables pueden tratar datos de categoría especial públicamente manifestados por el interesado, sin embargo, es

[868] Artículo 154 de la Loi de finances pour 2020, sous le n° 2019-796 DC, le 20 décembre 2019 que legitima el análisis masivo de datos personales de los ciudadanos disponibles públicamente en las plataformas virtuales como método para luchar contra el fraude fiscal. Texto disponible en: https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000039683923

[869] Decisión del Consejo Constitucional Francés. Décision n° 2019-796 DC du 27 décembre 2019, apartado 87. Disponible en: https://www.conseil-constitutionnel.fr/decision/2019/2019796DC.htm

[870] Artículo 17.3 de la Ley 22/2018, de 6 de noviembre, de Inspección General de Servicios y del sistema de alertas para la prevención de malas prácticas en la Administración de la Generalitat y su sector público instrumental.

[871] Artículo 5.3 de la Circular 1/2019, de 7 de marzo, de la Agencia Española de Protección de Datos, sobre el tratamiento de datos personales relativos a opiniones políticas y envío de propaganda electoral por medios electrónicos o sistemas de mensajería por parte de partidos políticos, federaciones, coaliciones y agrupaciones de electores al amparo del artículo 58 bis de la Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General.

necesario que justifiquen que realmente el particular intencionadamente divulgó esos datos de forma pública y voluntaria. Para ello, dicha información ha de estar accesible para cualquier persona, sea o no usuaria de la plataforma o red social. La justificación ha de quedar documentada y en su caso se ha de actualizar cuando puedan existir alteraciones en las condiciones de uso sobre cómo los datos son publicados en las plataformas, redes sociales o páginas web que pueden estar nutriendo de datos al sistema algorítmico sobre el cual se está diseñando el modelo o se están adoptando decisiones automatizadas.

F) El tratamiento es necesario para la formulación, el ejercicio o la defensa de reclamaciones o cuando los tribunales actúen en ejercicio de su función judicial

A través de esta excepción se podría justificar por ejemplo la recopilación de datos personales especiales para desarrollar sistemas automatizados de *legal tech* para predecir las resoluciones judiciales que adoptarán los jueces y magistrados[872]. En Francia, la recopilación de datos de los magistrados, ya sean o no de categoría especial, con el objeto de evaluar, analizar, comparar o predecir sus prácticas profesionales reales o supuestas están prohibidos[873]. Por otro lado, la AEPD en una resolución ha señalado que a pesar de que la ejecución de una sentencia judicial autorice a llevar a cabo un concreto tratamiento, en este caso, el reconocimiento facial automatizado de personas, la base de legitimación descrita en el artículo 9.2.f) del RGPD no quedará fundamentada si dicho tratamiento de datos afecta a más personas de las que en su caso autorizaba dicha resolución judicial[874]. Es decir, a través de una resolución judicial se puede

[872] Otra opción sería el uso de este tipo de sistemas por parte del poder judicial. Algunos ejemplos del uso de esta tecnología en este contexto los encontramos en: CERRILLO I MARTÍNEZ, A; VELASCO RICO, C: "Jurisdicción, algoritmos e inteligencia artificial". En: LÓPEZ RAMÓN,F; VALERO TORRIJOS,J (coord.): *20 años de la Ley de lo Contencioso-administrativo: Actas del XIV Congreso de la Asociación Española de Profesores de Derecho Administrativo : Murcia, 8-9 de febrero de 2019*. *INAP*, 2019, pág.296 y ss. Texto disponible en: http://www.aepda.es/AEPDAEntrada-2517-Actas-del-XIV-Congreso-de-la-AEPDA-20-anos-de-la-Ley-de-lo-Contencioso-Administrativo.aspx

[873] Artículo 33 de la LOI n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice. Texto disponible en: https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000038261761?r=IPa1529p3I

[874] La empresa Mercadona implantó un sistema de reconocimiento facial en sus establecimientos con el objetivo de identificar a aquellos individuos que por resolución firme se les prohibía el acceso a dichos establecimientos. Esta empresa considera que dicha resolución judicial es el mecanismo adecuado para legitimar el tratamiento de datos de categoría especial biométricos. Para la AEPD, con relación a la resolución judicial que permite de forma genérica la implantación de la medida de seguridad, *la cadena de supermercados interpreta de forma unilateral el alcance de la resolución judicial y la utiliza a los efectos de justificar que ostenta legitimación en el sentido del artículo 9.2.f) del RGPD no sólo para el condenado, sino también para el resto de los ciudadanos afectados por el sistema cuando acceden a los supermercados - que la mercantil engloba bajo el nombre de "no condenados".* La AEPD llega a la

legitimar el uso de sistemas automatizados siempre que dicha resolución judicial así lo indique y además dichas decisiones se adopten respeto de las personas o grupo de personas a los que se refiera expresamente dicha resolución. No olvidemos que además es necesario que el responsable también obtenga una de las bases previstas en el artículo 6 del RGPD.

G) El tratamiento es necesario por razones de un interés público esencial

Se establecen varios requisitos que se han de acreditar para poder acudir a esta excepción, estos son: i) dicho interés público esencial ha de reconocerse en una norma del Derecho de la Unión o de los Estados miembros, ii) esa norma debe ser proporcional al objetivo perseguido y respetar en lo esencial el derecho a la protección de datos y, iii) se han de establecer medidas adecuadas y específicas para proteger los intereses y derechos fundamentales del interesado. Es turno de analizar estos elementos.

En primer lugar, por lo que se refiere al interés público esencial, el Tribunal Constitucional exige que este se especifique. Es decir, la norma que reconozca ese interés esencial lo ha de indicar ya que si no, no será posible valorar o enjuiciar el carácter constitucionalmente legítimo de esa finalidad, ni en su caso, la proporcionalidad de la medida prevista de acuerdo con los principios de idoneidad, necesidad y proporcionalidad en sentido estricto[875]. De esta manera, el RGPD ha reconocido a los Estados miembros y al derecho europeo un margen de maniobra a la hora de especificar sus normas. Este margen de configuración legislativa se extiende a la causas habilitantes para el tratamiento de datos personales especialmente protegidos, es decir, a la identificación de los fines de interés público esencial[876]. Como es lógico, cualquier interés público no puede considerarse esencial sino sólo aquellos que revistan de una importancia relevante para los Estados Miembros o la Unión Europea. En este sentido, la PRAI, basada en el interés público esencial por reducir o evitar la adopción de decisiones discriminatoria en los sistemas de IA de alto riesgo, permite a los proveedores/desarrolladores de estos sistemas tratar datos de categoría especial a efectos

conclusión que dicha resolución judicial no puede ser una base de legitimación suficiente. Agencia Española de Protección de Datos. Resolución Nº: PS/00120/2021, págs.33, 62,66, 69 y 70.

[875] STC Sentencia 76/2019, de 22 de mayo de 2019. FJº 7, aparatado a. Texto disponible en: https://hj.tribunalconstitucional.es/HJ/docs/BOE/BOE-A-2019-9548.pdf

[876] Fundamento jurídico 4 de la misma sentencia.

de garantizar la supervisión, la detección y la corrección de sesgos presente en dichos modelos algorítmicos[877].

En segundo lugar, se ha de valorar la proporcionalidad de la medida que limita el derecho fundamental a la protección de datos. Aquí nos remitimos a la evaluación realizada durante el capítulo III de esta tesis sobre la necesidad y proporcionalidad de llevar a cabo los tratamientos de datos presentes en los sistemas de toma de decisiones automatizadas[878].

Finalmente, en tercer lugar, la norma que habilita el tratamiento ha de contener las suficientes medidas de garantías y salvaguardas. Precisamente, el Tribunal Constitucional declaró inconstitucional un precepto de la LOPD de 2018 que habilitaba al tratamiento de datos por parte de partidos políticos para fines electorales por no haber contemplado la norma garantías a la hora de legitimar este tratamiento[879]. En este mismo sentido se ha pronunciado el TJUE con relación a las normas que habilitan tratamientos de datos de categoría especial[880].

Por tanto, cualquier legislación que legitime a través de esta excepción el uso de datos personales para diseñar sistemas automatizados o legalizar su uso para la toma de decisiones automatizadas deberá: i) reconocer el interés público esencial que reside en la incorporación ese sistema en ese entorno, ii) analizar la proporcionalidad del despliegue del mismo y, iii) prever suficientes medidas de garantía que salvaguarden los derechos y libertades de los particulares que se someterán a esas decisiones[881]. Estas garantías se analizan a lo largo de esta tesis[882].

De forma provisional y un tanto forzado, a día de hoy, y a falta de una regulación como la PRAI que reconozca expresamente el tratamiento de datos de categoría especial con fines de detección y corrección de sesgos. El propio RGPD

---

[877] Véase el considerando 44 *in fine* y el artículo 10.5 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

[878] Capítulo III, apartado II, punto 5.

[879] STC Sentencia 76/2019, de 22 de mayo de 2019. FJº 6. Texto disponible en: https://hj.tribunalconstitucional.es/HJ/docs/BOE/BOE-A-2019-9548.pdf

[880] SENTENCIA DEL TRIBUNAL DE JUSTICIA (Gran Sala) de 6 de octubre de 2020, asunto C-623/17, caso Privacy International. FJº 68. Texto disponible en: https://curia.europa.eu/juris/document/document.jsf?text=&docid=232083&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=2114210

[881] Véase en este sentido: Agencia Española de Protección de Datos. Resolución Nº: PS/00120/2021, pág.48.

[882] Véase entre otras las previstas en el Capítulo V, apartado III, punto 5 de esta tesis.

podría ser un mecanismo habilitante para legitimar este tratamiento fruto de las obligaciones que tiene el responsable a la hora de evitar que los sistemas automatizados que utiliza no generen decisiones sesgadas[883].

H) El tratamiento es necesario para fines relacionados con la salud

En este apartado hemos aglutinado toda una serie de finalidades a través de las cuales los responsables pueden habilitar el tratamiento de datos de categoría especial, sobre todo, los relativos a la salud de las personas.

Así, el apartado 9.2.h) del RGPD autoriza los tratamientos que tenga como finalidad la medicina preventiva o laboral, evaluación de la capacidad laboral del trabajador, diagnóstico médico, prestación de asistencia o tratamiento de tipo sanitario o social, o gestión de los sistemas y servicios de asistencia sanitaria y social. En todos estos supuesto, el tratamiento ha de estar previsto o bien en una norma del derecho de la UE o de los Estados miembros o bien en virtud de un contrato con un profesional sanitario.

Por otro lado, el artículo 9.2.i) del RGPD permite que una norma estatal o del derecho de la UE reconozca tratamientos por razones de interés público en el ámbito de la salud pública como la protección frente a amenazas transfronterizas graves para la salud o para garantizar elevados niveles de calidad, así como tratamientos con fines para la seguridad de la asistencia sanitaria y de los medicamentos o productos sanitarios. Estos tratamientos han de legitimarse a través de una norma del derecho de la UE o de los Estados miembros. Así, la LOPD de 2018 permite a las autoridades sanitarias con competencias en vigilancia de la salud pública llevar a cabo estudios científicos sin el consentimiento de los afectados en situaciones de excepcional relevancia y gravedad para la salud pública[884]. Es decir, en situaciones como las vividas a raíz de la pandemia originada por el Covid, estas autoridades y con esa finalidad podrían desarrollar sistemas de inteligencia artificial y tratar datos de categoría especial. Como es lógico, esta situación excepcional ha de quedar acreditada y limitada en el tiempo.

[883] Véase el Capítulo IV, apartado II (principio de exactitud); Capítulo IV, apartado VII, punto 2 (principio de lealtad); Capítulo V, apartado III, punto 5 de esta tesis (Garantías).
[884] Disposición Adicional 17.2.b) de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

I) El tratamiento es necesario con fines de archivo en interés público, fines de investigación científica o histórica o fines estadísticos

Por último, también estará legitimado el tratamiento de los datos de categoría especial cuando el mismo sea necesario para los fines contemplados en el artículo 89.1 del RGPD. En estos supuestos, también se requerirá de una norma europea o estatal para legitimar el tratamiento y además se habrán de prever las suficientes medidas de garantía. En este sentido, nuevamente, el legislador español ha recogido el guante y ha reconocido como lícito el uso de datos personales seudonimizados con fines de investigación en salud y, en particular, biomédica[885]. Ello permite que los responsables puedan analizar masivamente datos seudonimizados con dicha finalidad. El análisis de estos tratamientos para las finalidades indicadas se estudia más profusamente en otros apartados de esta tesis[886].

II. EL PRINCIPIO DE LIMITACIÓN DE LA FINALIDAD

Tal y como establece el artículo 5.1.b) del RGPD, los datos personales serán:

> *recogidos con fines determinados, explícitos y legítimos, y no serán tratados ulteriormente de manera incompatible con dichos fines; de acuerdo con el artículo 89, apartado 1, el tratamiento ulterior de los datos personales con fines de archivo en interés público, fines de investigación científica e histórica o fines estadísticos no se considerará incompatible con los fines iniciales.*

Este principio se asienta en dos pilares claramente definidos. Por un lado, los responsables del tratamiento sólo han de recopilar aquellos datos para fines específicos, explícitos y legítimos. Por otro lado, una vez que los datos son recopilados para una finalidad concreta, estos datos únicamente podrían tratarse para una finalidad distinta cuando dicha finalidad posterior no sea incompatible con la finalidad inicial para la que se recopilaron los datos. En este sentido, cuando ese tratamiento ulterior de los datos tenga como fin la investigación científica e histórica, fines estadísticos o fines de

[885] Disposición Adicional 17.2.d) de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Véase entre otros: MÉNDEZ GARCÍA, M y ALFONSO FARNÓS, I: "La legitimación para el tratamiento de categorías especiales de datos con finalidades de investigación en el marco del Reglamento (UE) 2016/679 y de la Ley Orgánica 3/2018", en *Revista de derecho y genoma humano: genética, biotecnología y medicina avanzada* Nº Extra 1, 2019, págs. 205 a 231. Texto disponible en: https://www.dykinson.com/revistas/revista-de-derecho-y-genoma-humano-genetica-biotecnologia-y-medicina-avanzada/1159/

[886] Capítulo IV, apartado II, punto 2, epígrafe A). Véase también en su caso el Capítulo V, apartado III, punto 5.

archivo en interés público, tales operaciones se consideran compatibles. En el resto de supuestos se deberá valorar la compatibilidad entre el fin posterior proyectado y el inicial.

Es turno de analizar estas dos apreciaciones y su incidencia en los tratamientos de datos presentes durante el ciclo de vida de los sistemas de toma de decisiones automatizadas.

1. El responsable sólo debe recopilar datos para fines específicos, explícitos y legítimos

Esta primera vertiente del principio de limitación de la finalidad exige del responsable la obligación por la cual se han de indicar claramente las finalidades que hay detrás del tratamiento de datos que pretende llevar a cabo. Esa finalidad ha de establecerse de forma explícita y mostrarse suficientemente inequívoca[887]. Además, dicha finalidad ha de ser legítima. Es decir, la finalidad pretendida con el tratamiento debe ser conforme a la ley. El concepto de ley o norma en este ámbito se ha de interpretar de forma amplia. Por tanto, en esta definición queda abarcada no sólo la normativa de protección de datos sino cualquier otro cuerpo legal[888]. La finalidad del tratamiento se ha de fijar antes, o más tardar, durante la recogida de los datos personales. A su vez, concretar la finalidad también ayudará a respetar otros principios del tratamiento como puede ser el de minimización de datos, la limitación de los plazos de conservación o el de licitud a la hora de legitimar determinados tratamientos de datos. Así, por ejemplo, imaginemos que una organización pretende utilizar un sistema de reconocimiento facial con el objetivo de analizar el comportamiento del aspirante en una entrevista de trabajo. La organización, a priori, estaría tratando datos biométricos, por tanto, datos de categoría especial, los cuales requieren de reglas específicas a la hora de legitimar el tratamiento de dichos datos. Pues bien, dado que la finalidad de ese tratamiento no es la identificación unívoca de esa persona, dicho dato personal no se

[887] Es necesario identificar con la suficiente precisión la finalidad del tratamiento de datos. De otra manera, puede resultar imposible valorar si la misma persigue un carácter constitucionalmente legítimo. Sentencia del Tribunal Constitucional 76/2019, de 22 de mayo de 2019. FJ7º apartado a). Resolución disponible en: https://www.boe.es/buscar/doc.php?id=BOE-A-2019-9548

[888] Se ha indicado que también entrarían en este concepto todas las formas de derecho escrito y común, la legislación primaria y secundaria, los decretos municipales, los precedentes judiciales, los principios constitucionales, los derechos fundamentales, otros principios legales, así como la jurisprudencia, ya que dicha "ley" sería interpretada y tenida en cuenta por los tribunales competentes. Grupo del artículo 29. *Opinion 03/2013 on purpose limitation*. Texto adoptado el 2 de abril de 2013, págs. 12 y 20.

considera biométrico en los términos del artículo 9 del RPGD y por tanto[889], el responsable no queda afectado por las especiales limitaciones que presentan el tratamiento de estos datos específicos. Aquí como se puede apreciar, la finalidad del tratamiento juega un papel esencial a la hora de valorar el tipo de dato que se trata.

Tal y como hemos hecho en otros momentos de este trabajo, nuevamente clasificaremos la incidencia de este principio en las dos fases principales que comprende el ciclo de vida de los sistemas automatizados, esto es, la fase de diseño y la fase de despliegue.

A) La fase de diseño

Aunque para un número importante de tratamientos resulte sencillo establecer la finalidad que se pretende con los mismos antes de llevar a cabo el tratamiento. Esta claridad para fijar dicha finalidad queda en parte difuminada cuando se pretende desarrollar proyectos de minería o análisis masivo de datos. Ello es así porque, como indicamos en el capítulo inicial de este trabajo, los diseñadores de estos modelos algorítmicos en muchos supuestos desconocen el objetivo que se pretende con el diseño de los mismos.

Así, dentro de la técnica del *machine learning* encontramos dos formas o métodos de aprendizaje, el supervisado y el no supervisado. Por lo que se refiere al supervisado, normalmente, y a pesar de que se pueden descubrir patrones desconocidos, lo cierto es que la finalidad inicial puede quedar marcada desde el principio ya que comúnmente se conocerá el objetivo que se pretende con la analítica de datos[890]. Más difícil resultará cuando se pretenda llevar a cabo la analítica de datos basados en las técnicas de aprendizaje no supervisado, en estos casos, dado que no existe ninguna pauta de antemano, sino que más bien se deja a los datos que traten de buscar correlaciones ocultas, marcar la finalidad inicial resulta más compleja. Es decir, sobre todo, en esta última técnica, el responsable del tratamiento encuentra importantes

[889] Tal y como establece el artículo 9 del RGPD, Quedan prohibidos el tratamiento de datos personales que revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos**,** datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o las orientación sexuales de una persona física.

[890] Autoritat Catalana de Protecció de dades. *Intel·ligència Artificial. Decisions Automatitzades a Catalunya, 2020,* págs. 127 y 128.

escollos para fijar desde el momento inicial una finalidad específica y explícita. La solución a este problema es compleja ya que, en este contexto, un importante relajamiento de este principio supondría una afectación grave al derecho fundamental a la protección de datos pero a su vez, una interpretación estricta de este principio limitaría el uso de esta técnica y por ello, cortocircuitaría los posibles beneficios que se pueden generar de la analítica de datos basada en aprendizaje no supervisado. Son dos los enfoques que se pueden considerar para afrontar este conflicto en este contexto, por un lado, adoptar una actitud sumamente proteccionista del derecho a la protección de datos, de manera que sólo es posible el tratamiento de datos cuando la finalidad sea claramente definida desde el inicio del tratamiento. Como hemos indicado, las consecuencias de esta interpretación pueden limitar gran parte de los proyectos actuales de analítica masiva de datos, sobre todo aquellos que descansan en técnicas de aprendizaje no supervisado. Un segundo enfoque, más aperturista, permitiría llevar a cabo el análisis de datos masivos basados en finalidades más amplias y genéricas que abarquen los posibles objetivos que se pretenden dentro del proyecto de minería de datos. Es decir, dado que resulta imposible fijar una finalidad inicial meridianamente clara en muchos de los proyectos basados en aprendizaje no supervisado, se podría permitir una finalidad más amplia que abarque las hipotéticas finalidades que en su caso pueden estar presentes en el proyecto que se pretende llevar a cabo. Ahora bien, para compensar esta restricción del principio de limitación de la finalidad se han de prever suficientes medidas que mantengan un nivel adecuado que logre proteger a los titulares de los datos personales que se ven sometidos a estos tratamientos. Este segundo enfoque, ya ha sido en parte reconocido legalmente por parte de la actual LOPD de 2018. Así, la Disposición Adicional 17ª de esta norma permite que los titulares de los datos que se verán sometidos a la analítica de datos puedan prestar un consentimiento amplio para tratamiento de datos con fines de investigación en salud, los cuales, puedan englobar varias materias o áreas relacionadas[891]. Este enfoque será estudiado en un momento posterior[892].

[891] Sobre el consentimiento amplio en los tratamientos de datos con fines de investigación científica véase: ÁLVAREZ RIGAUDIAS,C: "Tratamiento de datos con fines de investigación científica y/o médica". En RALLO LOMBARTE, A (dir): Tratado de Protección de Datos. Ed. Tirant lo Blanch, Valencia, 2019, págs. 723 y ss. En el mismo sentido: SERRANO PÉREZ,M,M: "El marco jurídico de los datos relativos a la salud en el ámbito de la salud y de la investigación en salud tras la entrada en vigor del Reglamento General de Protección de Datos y de la Ley de Protección de Datos Personales y garantía de los derechos". *Estudios de Deusto: revista de la Universidad de Deusto*, ISSN 0423-4847, Vol. 68, Nº. 2, 2020, pág.276.

[892] Concretamente en el Capítulo IV, apartado II, punto 3 de esta tesis.

B) Fase de aplicación o despliegue

Llegada la fase de toma de decisiones automatizadas, el responsable ya cuenta con el conocimiento suficiente para conocer el objetivo que se pretende con dicho modelo algorítmico. En este caso, y a diferencia de la fase de diseño donde el responsable podía tener dificultades para concretar la finalidad, el responsable está obligado a indicar clara y expresamente la finalidad o finalidades específicas que se pretenden con el despliegue del sistema de toma de decisiones automatizadas. Ello es así porque dicho responsable ya cuenta con toda la información para establecer dicha finalidad. Esto es, conoce los datos que alimentarán el sistema, las principales correlaciones que se realizarán, las variables más relevantes y por supuesto el objetivo y el entorno donde ingresará el algoritmo. Resulta por tanto fundamental fijar la finalidad pretendida para que los titulares de los datos puedan conocer para qué se tratarán sus datos, o mejor dicho, qué se pretende hacer con el sistema que adoptará decisiones o elaborará perfiles sobre los mismos[893]. En este sentido, el responsable no puede camuflar las posibles pretensiones del uso de los sistemas automatizados a través de un lenguaje poco claro o ambiguo para maquillar los usos pretendidos. Así, si a través de un sistema automatizado se pretende perfilar a una persona para una finalidad específica, el titular de los datos ha de conocerlo y ser consciente de qué supone ese tratamiento para él.

Como hemos señalado previamente, el principio de limitación de la finalidad en relación con la legitimidad de las finalidades exige que las mismas sean de acuerdo a las normas. Pues bien, un sistema automatizado que se pretenda para finalidades contrarias al ordenamiento jurídico atentará no sólo contra esa norma sino también contra el mencionado principio. Este principio quedará afectado tanto si la finalidad inicialmente marcada como en su caso la encubierta realmente pretendida resulta contraria al ordenamiento jurídico.

---

[893] Como ha señalado la AEPD, finalidades como "conocerte mejor" o "mejorar la experiencia el cliente" supone el empleo de una terminología imprecisa que no supera las exigencias mínimas de transparencia. Es necesario que el propósito del tratamiento sea explícito y claro. En: Agencia Española de Protección de Datos. Resolución Nº: PS/00070/2019**.** Págs. 83 a 87 y 100. Resolución disponible en: https://www.aepd.es/es/documento/ps-00070-2019.pdf
También puede verse la resolución de la Agencia Española de Protección de Datos. Resolución Nº**:** PS/00477/2019, pág.85. Resolución disponible en: https://www.aepd.es/es/documento/ps-00477-2019.pdf
En el mismo sentido véase: ALIAGA MARTÍNEZ,L ; GUTIÉRREZ DAVID, E: "Explicando Machine Learning a través de la doctrina y práctica del Information Commissioner's Office". *LA LEY privacidad*, Nº 8, Sección Crónica de Corresponsales, Segundo trimestre de 2021, pág.9.

2. Uso posterior de los datos con una finalidad distinta a la inicial

Tal y como señala el GT29, el principio de limitación de la finalidad está diseñado para ofrecer un enfoque equilibrado donde se pretende conciliar por un lado la necesidad de previsibilidad y seguridad jurídica exigiendo que se indiquen claramente los fines de tratamiento de datos que se pretende llevar a cabo y, por otro lado, la necesidad pragmática de dotar de cierta flexibilidad el uso secundario de los datos por parte del responsable[894]. Por lo que respecta a esta segunda vertiente, el RGPD permite que los responsables del tratamiento puedan tratar los datos para una finalidad diferente a la inicialmente prevista. Para habilitar esta situación, la norma exige que la finalidad inicial y la ulterior proyectada resulten compatibles. Es por ello que el responsable deberá realizar un estudio de la compatibilidad entre ambos fines, superado ese análisis, el responsable podrá tratar los datos para un fin distinto de acuerdo al artículo 6.4 del RGPD.

La posibilidad de utilizar los datos con fines distintos a los inicialmente previstos encaja perfectamente en la operatoria en la que se basa en parte la analítica masiva de datos actual. Esto es, una organización, ya sea pública o privada cuenta con una ingente cantidad de datos que fueron recopilados para una finalidad, y ahora, siendo conscientes de las ventajas que puede generar la analítica de los mismos a través de las técnicas del *big data* y el *machine learning* deciden alterar la finalidad y desarrollar modelos que posteriormente puedan resultar útiles para las organizaciones, concretamente y de acuerdo a nuestro estudio, para la toma de decisiones automatizadas.

Pues bien, para que la operatoria descrita en el párrafo anterior encaje en la normativa de protección de datos personales los responsables del tratamiento deben conocer las vías legales que le habilitan a ello.

Estas alternativas son esencialmente tres; 1) si las operaciones del tratamiento ulterior tienen como finalidad el archivo en interés público, fines de investigación científica e histórica o fines estadísticos, dichas finalidades se consideran compatibles con la inicial, independientemente de cual haya sido ésta. Artículos 5.1.b) y 89 del RGPD. 2) Si las finalidades posteriores que se pretenden no son ninguna de las

[894] Grupo del artículo 29. *Opinion 03/2013 on purpose limitation*. Texto adoptado el 2 de abril de 2013, pág.39.

señaladas anteriormente, cabe le tratamiento ulterior si se obtiene el consentimiento del titular de los datos o una norma estatal o del derecho de la Unión Europea legitima tal tratamiento. Artículos 5.1.b), 6.4 y 23 del RGPD. 3) Si las finalidades ulteriores que se pretenden no corresponden con las señaladas anteriormente y además tampoco se obtiene el consentimiento o no existe una norma que legitime el tratamiento, aún queda la posibilidad de que el responsable pueda valorar la compatibilidad entre la finalidad inicial y la posterior, si llega la conclusión de que ambas son compatibles, también puede acudir a esta vía. Artículo 5.1.b) y 6.4) del RGPD.

Cabe señalar que en las tres vías se requiere que el responsable del tratamiento adopte medidas adecuadas de garantía.

En el trasfondo de estos preceptos se detecta que el legislador europeo consideró que el uso secundario de los datos por parte de las organizaciones puede representar intereses legítimos que también han de ser protegidos. Es decir, el derecho fundamental a la protección de datos no es un derecho absoluto y este en parte puede quedar limitado cuando se conjuga con otros derechos o intereses, siempre claro está, que dicha limitación no afecte al contenido esencial[895]. Es por ello que el RGPD permita que en su caso se puedan utilizar datos para una finalidad posterior a la inicialmente prevista, ahora bien, no vale cualquier forma sino sólo a través de los canales que prevé la normativa en materia de protección de datos. Es momento de analizarlos.

[895] Artículo 52.1 de la Carta de derechos fundamentales de la Unión Europea.

A) Tratamientos ulteriores cuya finalidad sea la investigación científica o los fines estadísticos

Cuando el tratamiento posterior tenga como finalidad el archivo en interés público, la investigación científica e histórica o los fines estadísticos dicho tratamiento no se considerará incompatible con la finalidad inicial. El legislador europeo parte por tanto de la compatibilidad de estas finalidades si bien, será necesario que el responsable del tratamiento establezca todas las garantías adecuadas ya que si estas no se establecen o no son suficientes, habría que entender que esa presunción de la compatibilidad inicial reconocida por el RGPD decae[896]. Esta compatibilidad refleja en el legislador europeo un interés relevante por los tratamientos de datos que obedezcan a estas concretas finalidades. En estos supuestos, el derecho a la protección de datos queda en mayor medida restringido en favor de otros bienes jurídicos en juego que resultan legítimos. Es decir, queda por tanto patente que el dato en estos contextos puede generar unos beneficios legítimos a las organizaciones y a la sociedad en general que está por encima de los derechos estrictos del titular de los datos personales. Es turno de analizar algunas de estas finalidades ya que las mismas pueden quedar comprendidas en las fases que forman parte del ciclo de vida de los sistemas automatizados.

En *primer lugar*, por lo que se refiere al tratamiento posterior con fines estadísticos, el RGPD los asimila a toda operación de recogida y tratamiento de datos personales necesarios para encuestas estadísticas o para la producción de resultados estadísticos. (Considerando 162). Teniendo en cuenta esta definición, es fácilmente encajable en esta finalidad el uso de técnicas de analítica masiva de datos personales con dichos fines. Así, el GT29 ha considerado que dentro de esta finalidad pueden quedar aglutinadas las llamadas herramientas analíticas de los sitios web o las aplicaciones de *big data* destinados a la investigación de mercado[897]. Ahora bien, el uso de estas técnicas ha de quedar limitado exclusivamente a este fin estadístico[898]. Es decir, los resultados estadísticos no pueden utilizarse para respaldar medidas o decisiones relativas a personas físicas concretas. (Considerando 162, in fine). De esta manera, entendemos que los resultados, hallazgos o modelos que se deriven del uso de técnicas

[896]Supervisor europeo de protección de datos. *A Preliminary Opinion on data protection and scientific research.* Texto adoptado el 6 de enero de 2020, pág 22.

[897] Grupo del artículo 29. *Opinion 03/2013 on purpose limitation*. Texto adoptado el 2 de abril de 2013,pág.29

[898] HERNÁNDEZ, J,C: "Decisiones algorítmicas de perfilado. Régimen y garantías jurídicas." *Revista Española de Derecho Administrativo*. Nº 203, 2020, págs 306 y 307.

de minería de datos o *big data* que se amparen en la finalidad estadística no podrán usarse como base para posteriormente adoptar decisiones sobre personas. Por tanto, aquellos que pretendan desarrollar modelos para su posterior despliegue o venta a terceros deben ser conscientes de esta limitación. No resulta por tanto esta vía la más idónea para los responsables del tratamiento que pretendan desarrollar modelos o sistemas algorítmicos que posteriormente se utilizarán para adoptar decisiones sobre los individuos.

En *segundo lugar*, por lo que se refiere al uso posterior para tratamientos con fines de investigación científica, las posibilidades de utilizar esta alternativa también pueden encajar en algunas de las fases que engloban el ciclo de vida de sistemas de toma de decisiones automatizadas. Así, el RGPD establece que esta finalidad ha de interpretarse de manera amplia, pudiendo comportar todos aquellos tratamientos que tienen como objetivo el desarrollo tecnológico, la demostración, la investigación fundamental, la investigación aplicada y la investigación financiada por el sector privado. (Considerando 159)[899]. Es por ello que el *machine learning* y *data mining* puedan utilizarse para esta finalidad. Como sabemos, estas técnicas facilitan la obtención de conocimiento oculto en los datos, por lo tanto, su uso puede ayudar a conseguir algunos de los objetivos que quedan integrados bajo esta finalidad. Surge la duda pues de si cualquier proyecto de analítica de datos en busca de un hipotético conocimiento oculto quedaría amparado por esta vía. La respuesta no es sencilla[900]. El CEPD señala que el régimen especial de protección de datos para la investigación científica se aplica cuando se cumplen los siguientes tres criterios: i) se tratan datos personales, ii) se aplican las normas metodológicas y éticas relacionadas con el sector de conformidad con prácticas adecuadas, iii) la investigación se lleva a cabo con el objetivo de aumentar el conocimiento y el bienestar colectivo de la sociedad, en lugar de servir principalmente a uno o varios intereses privados[901]. De esta manera, sin acotar por completo el espacio a la investigación científica con ánimo de lucro, el CEPD claramente restringe el ámbito de aplicación de esta finalidad a los requisitos arriba mencionados.

---

[899] Sobre la compatibilidad entre el fin inicial y el posterior cuando este último tiene como objetivo las investigaciones científicas y biomédicas véase: RECUERO LINARES, M: *La investigación científica con datos personales genéticos y datos relativos a la salud: perspectiva europea ante el desafío globalizado*, Premio AEPD, 2019, págs. 30 y ss. Texto disponible en: https://www.aepd.es/sites/default/files/2020-02/premio-2019-emilio-aced-accesit-mikel-recuero.pdf

[900] Autoritat Catalana de Protecció de dades. *Intel·ligència Artificial. Decisions Automatitzades a Catalunya, 2020*, pág 128. Esta autoridad plantea el problema pero tampoco arroja una solución.

[901] Supervisor europeo de protección de datos. *A Preliminary Opinion on data protection and scientific research*. Texto adoptado el 6 de enero de 2020, pág.12.

Sería recomendable por tanto que los responsables del tratamiento que pretendan acudir a esta vía justifiquen claramente las normas metodológicas en las que basan su investigación de cara a que en su caso se demuestre que realmente utilizaron adecuadamente la finalidad de la investigación científica y justifiquen que dicha investigación pretende servir al interés colectivo, además de en su caso al interés puramente privado. En muchas ocasiones la frontera entre el interés estrictamente comercial y el colectivo será difusa. Tal y como ocurre por ejemplo con el desarrollo de automóviles autónomos más seguros[902]. Las empresas automovilísticas tienen claramente un objetivo económico en dicha investigación a través de la analítica masiva de datos pero esa investigación lucrativa puede servir también al interés colectivo. A su vez, la autoridad de protección de datos de Noruega plantea otro criterio para considerar que el uso de sistemas basados en inteligencia artificial corresponde con la investigación científica. Para ello establece como criterio el tipo de modelo que se utilice. Así, si el modelo despliega sus efectos en un entorno estático y por tanto no aprende de los nuevos inputs que este va recibiendo a lo largo de su vida, hay que considerar que no está generando conocimiento. A la inversa, si un modelo se ingresa en un entorno dinámico y este continuamente se está adaptando a los nuevos cambios, el nuevo conocimiento que vaya generando en forma de resultados puede considerarse englobado en la finalidad de investigación científica[903]. En nuestra opinión, si bien el criterio de la adaptabilidad puede ser útil, sigue sin resolver todos los problemas de interpretación de este ámbito. Y es que, no todos los entornos adaptativos que vayan generando nuevos resultados y que permitirán a los sistemas algorítmicos aprender nuevas correlaciones pueden encajarse dentro de ese concepto de investigación científica. En este sentido, será fundamental que aquellos responsables que pretendan respaldar sus tratamientos a través de esta finalidad la justifiquen adecuadamente.

Independientemente del tipo de finalidad a la que acuda el responsable para legitimar el tratamiento posterior, la necesidad de establecer suficientes medidas de

[902] MESZAROS,J ; *CHIH-HSING,H:* "AI research and data protection: Can the same rules apply for commercial and academic research under the GDPR?". *Computer Law & Security Review*, Volume 41, July 2021, pág.10.
[903] Autoridad de Protección de Datos de Noruega. Datatilsynet, *Artificial intelligence and privacy* . Texto publicado en enero de 2018, págs. 16 a 18.

protección adecuadas resulta sumamente relevante. En este sentido, en los siguientes apartados se harán referencia a algunas de las medidas que deberían establecerse[904].

B) Tratamientos ulteriores amparados en el consentimiento o en una norma del estado miembro o el derecho de la Unión Europea

El artículo 6.4 del RGPD permite el tratamiento de datos con una finalidad distinta a la inicial cuando dicho tratamiento se ampare en la obtención del consentimiento otorgado por los titulares de los datos o en aquellos supuestos en los que una norma del estado miembro o del derecho de la Unión Europea lo autorice.

Por lo que se refiere al consentimiento, únicamente señalar que dicho consentimiento deberá cumplir las exigencias previstas en el RGPD, las cuales, son analizadas en el apartado de este trabajo que estudia esta base de legitimación[905].

En lo que respecta a la norma del estado miembro o del derecho de la Unión Europea que legitima el tratamiento posterior conviene realizar una serie de apreciaciones importantes.

En *primer lugar*, no toda norma que provenga del derecho estatal o de la UE autoriza al tratamiento posterior de los datos. Para que entre en juego las previsiones del artículo 6.4 del RGPD, la norma que habilita este tratamiento ulterior debe responder a algunas de las finalidades fijadas en el artículo 23 del RGPD[906]. Las finalidades previstas en dicho precepto comprenden un abanico de materias y objetivos relativamente amplio. Sin embargo, cuando no estén bajo el amparo del mencionado precepto, esa compatibilidad no se entenderá automática sino que se requerirá de un análisis de la compatibilidad, todo ello, a pesar de que haya una ley que autorice a ese

[904] Capítulo IV, apartado II, punto 2, epígrafe D) de esta tesis.
[905] Véase el Capítulo IV, apartado I, punto 1 de esta tesis.
[906] Entre otras finalidades se indica: la seguridad del Estado, la defensa, la seguridad pública, la prevención, investigación, detección o enjuiciamiento de infracciones penales o la ejecución de sanciones penales, incluida la protección frente a amenazas a la seguridad pública y su prevención; otros objetivos importantes de interés público general de la Unión o de un Estado miembro, en particular un interés económico o financiero importante de la Unión o de un Estado miembro, inclusive en los ámbitos fiscal, presupuestario y monetario, la sanidad pública y la seguridad social; la protección de la independencia judicial y de los procedimientos judiciales; la prevención, la investigación, la detección y el enjuiciamiento de infracciones de normas deontológicas en las profesiones reguladas; una función de supervisión, inspección o reglamentación vinculada, incluso ocasionalmente, con el ejercicio de la autoridad pública, la protección del interesado o de los derechos y libertades de otros; la ejecución de demandas civiles.

tratamiento o se pueda desprender tal tratamiento de la misma. Esta es la interpretación que en nuestra opinión parece desprenderse del Considerando 50 del RGPD[907].

En *segundo lugar*, además de que la norma estatal o europea se base en alguno de los objetivos establecidos en el artículo 23. La finalidad posterior autorizada por esa disposición deberá quedar explícitamente señalada. Es decir, pensamos que una norma que se ampare en la vía que ofrece el artículo 6.4 en relación con el 23 que legitime el uso secundario de los datos para una finalidad distinta no puede aludir a dicha finalidad ulterior en términos difuminados o poco claros. Así, por ejemplo, la Ley 22/2018 de la Comunidad Valenciana que crea el sistema de alertas tempranas contra la corrupción prevé el tratamiento de datos para finalidades posteriores a los usos inicialmente previstos. Aunque estas finalidades aparecen más o menos indicadas en los distintos preceptos de esta norma, no queda claro qué tipos de datos específicos se utilizarán para cada una de las finalidades que se pretende teniendo en cuenta que a través de este sistema algorítmico se pueden llevar a cabo toda una serie de operaciones[908]. Otro ejemplo donde sí que existe claridad en cuanto a la finalidad posterior lo encontramos en la PRAI, la cual habilita a los desarrolladores de los sistemas a utilizar los datos de categoría especial previstos en el artículo 9 y 10 del RGPD con el objetivo de poder detectar los posibles sesgos que pueden presentar los modelos algorítmicos[909].

La autorización de estas finalidades posteriores tanto por la vía del consentimiento como por la norma estatal o de la Unión Europea requerirá del establecimiento de suficientes medidas de garantía.

[907] Este apartado establece una compatibilidad directa cuando la norma estatal o de la UE se base en los objetivos del artículo 23 (Apartado segundo del Considerando 50). Para el resto de normas, dicha compatibilidad no se presume sino que en su caso se requiere del mencionado análisis de compatibilidad. (Apartado primero del Considerando 50).

[908] El artículo 17 de la Ley 22/2018, de 6 de noviembre, de Inspección General de Servicios y del sistema de alertas para la prevención de malas prácticas en la Administración de la Generalitat y su sector público instrumental hace referencia al conjunto de bases de datos que integrarán el sistema algorítmico. Pues bien, sería necesario que quedarán claras las finalidades exactas que se requieren para cada una de las bases de datos aportadas ya que puede que las mismas sean utilizadas para distintas finalidades. Tal y como se desprende por ejemplo del artículo 17.3 de esta misma norma.

[909] Artículo 10.5 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

C) Resto de tratamientos ulteriores, el análisis de la compatibilidad

Cuando el responsable del tratamiento no pueda acudir a ninguna de las vías previamente indicadas aún le queda la alternativa de valorar si el tratamiento posterior que pretende resulta compatible con el tratamiento inicial[910]. A la hora de valorar la compatibilidad o no de dicho tratamiento ulterior el propio RGPD establece toda una serie de criterios que se han de tener en cuenta por parte del responsable. El estudio de esos criterios es conocido como el análisis de la compatibilidad. Este estudio obliga al responsable del tratamiento a realizar un análisis particularizado de la finalidad posterior en relación con estos criterios. Este estudio será global, es decir, en determinadas ocasiones varios factores pueden indicar tal compatibilidad, pero sin embargo, puede existir uno o dos criterios que reflejan incompatibilidad y, los mismos sean de tal envergadura que no pueden ser pasados por alto por parte del responsable. Cabe indicar que los factores que ahora se analizarán son solo algunos de los elementos, entre otros, que el responsable deberá de tener en cuenta para llevar a cabo el análisis de compatibilidad.

Así, *en primer lugar* se ha de analizar la relación entre los fines para los que se hayan recogido los datos y los fines del tratamiento posterior pretendido. En este sentido cuanto más se aleje la finalidad inicial de la posterior, más probable será que se considere incompatible el tratamiento ulterior[911]. A la hora de valorar esa relación, el responsable puede tener en cuenta si la finalidad posterior pretendida se integra en algunas de las competencias, obligaciones u objetivos que dicha organización tiene asumida. Así, por ejemplo, una de las obligaciones que tienen las entidades bancarias es mantener un ecosistema bancario meridanamente solvente. Para ello, a la hora de conceder crédito han de evaluar la previsible solvencia financiera de los potenciales clientes[912]. Es entendible que una entidad bancaria pueda utilizar los datos de clientes antiguos para desarrollar modelos que tenga como objetivo evaluar la posible solvencia

[910] El SEPD ha considerado que a pesar de la presunción de compatibilidad entre la finalidad inicial y la finalidad posterior relacionada con la investigación científica, dicha prueba de compatibilidad debería seguir realizándose ya que se garantiza de mejor forma los derechos del interesado. Esta resulta más necesario cuando los datos se recogieron originalmente con fines muy diferentes o fuera del ámbito de la investigación científica. En: Supervisor europeo de protección de datos. *A Preliminary Opinion on data protection and scientific research.* Texto adoptado el 6 de enero de 2020, pág 23.

[911] Grupo del artículo 29. *Opinion 03/2013 on purpose limitation*. Texto adoptado el 2 de abril de 2013, pág.24.

[912] El artículo 29 de la Ley 2/2011, de 4 de marzo, de Economía Sostenible obliga a las entidades bancarias a llevar a cabo una adecuada evaluación de la solvencia de los clientes.

de potenciales clientes con la finalidad de prevenir un hipotético impago que puede afectar tanto a la entidad bancaria en particular, como al sistema bancario en general cuando este no esté debidamente saneado. Por el contrario, una aspiradora inteligente que toma imágenes de una casa con la finalidad de perfeccionar sus métodos de limpieza, no podría en nuestra opinión posteriormente utilizar esos datos bajo esa finalidad inicial para venderlos a terceras empresas que en su caso le envíen publicidad basada en la disposición del mobiliario de la casa[913]. En este último supuesto, la relación inicial y posterior de la finalidad están muy separadas ya que el objetivo posterior se aleja por completo de las competencias y obligaciones que se derivan de la organización que recopiló los datos inicialmente. Algo parecido ocurriría por ejemplo con los datos de telemetría que obtienen las empresas que suministran coches autónomos con el objetivo de mantener un adecuado funcionamiento de los mimos. Si esos datos iniciales obtenidos con fines de mantenimiento son revelados a compañías de seguro para que estas puedan ofrecer pólizas de seguros basadas en los comportamientos de conducción que se derivan de dichos datos, una vez más, dicha finalidad difícilmente podría considerarse compatible, requiriendo en su caso del consentimiento o una norma que legitime dicho tratamiento posterior de datos[914].

En segundo lugar, y muy relacionado con el criterio anterior, también se ha de analizar el contexto específico en el que se recogieron los datos y la relación entre el interesado y el responsable del tratamiento. En este sentido, entre otros elementos, resultará relevante valorar cómo se obtuvieron inicialmente esos datos, la disposición del interesado para facilitar esos datos, la obligación legal del interesado de proporcionar los datos, el desequilibrio de poder entre el interesado y el responsable, etc[915]. Así, por ejemplo, durante las diferentes propuestas legislativas del RGPD, la Comisión de Empleo y Asuntos Sociales del Parlamento Europeo propuso que se explicitara en el texto europeo la prohibición de tratar los datos de los trabajadores para finalidades secundarias a las inicialmente previstas por el responsable del tratamiento,

[913] Fuente de la noticia: JONES,R: "Roomba's Next Big Step Is Selling Maps of Your Home to the Highest Bidder", *Gizmodo*:24/7/2017. Disponible en: https://gizmodo.com/roombas-next-big-step-is-selling-maps-of-your-home-to-t-1797187829

[914] Grupo del Artículo 29. *Guidelines 1/2020 on processing personal data in the context of connected vehicles and mobility related applications*. Versión 1.0. Adoptadas el 28 de enero de 2020. Aparatados 50 y ss y pág.11.

[915] Grupo del artículo 29. *Opinion 03/2013 on purpose limitation*. Texto adoptado el 2 de abril de 2013, pág.,57.

esto es, el empresario[916]. Sin embargo, esta propuesta no prosperó. Relacionado con la anterior, también debe observarse la obligación que en su caso tenga el interesado en facilitar esos datos en virtud de una disposición legal. Por ejemplo, las Administraciones Públicas, con el fin de que se notifique adecuadamente a los interesados, están obligadas a publicar determinados actos administrativos que contienen datos personales en sus respectivos boletines oficiales. Esta información puede ser utilizada por terceras organizaciones para otros fines como el perfilado de esas personas[917]. Ese fin posterior puede resultar en estos casos incompatible teniendo en cuenta que el dato publicado se ha obtenido a causa de la obligación legal que se deriva de la finalidad inicial del tratamiento.

En tercer lugar, también será relevante la naturaleza de los datos que se traten para la finalidad posterior. Así, en concreto, cuando se traten categorías especiales de datos personales o datos relativos a condenas e infracciones penales, las posibilidades de que dicho tratamiento posterior sea incompatible aumentan[918]. De esta manera, si inicialmente los datos personales que se recopilaron eran convencionales y ahora esos mismos datos se utilizan para una finalidad que revele o muestre datos de categoría especial, la compatibilidad entre el fin inicial y el posterior se puede ver más comprometida. Para estos supuestos, si el responsable ya tenía previsto que en un momento posterior esos datos convencionales se iban a utilizar para finalidades sensibles, desde el momento inicial que los recopiló los debería considerar o tratar como datos de categoría especial. Así, una vez que pretenda esa finalidad secundaria sensible, el análisis de compatibilidad no resultará tan complejo o difícil de superar. En los demás supuestos, o bien el responsable obtiene una nueva base de legitimación teniendo en cuenta que además ha de contemplar alguna de las excepciones previstas por el artículo 9 del RGPD, o bien justificar adecuadamente dicho análisis de compatibilidad, el cual, como hemos indicado resultaría dificultoso.

---

[916] Informes y opiniones sobre la propuesta de Reglamento del Parlamento Europeo y del Consejo 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 (Reglamento general de protección de datos. Opinión de la Comisión de empleo y asuntos sociales. Enmienda número 19. Referida al artículo 82 del RGPD. Texto disponible en: https://www.europarl.europa.eu/doceo/document/A-7-2013-0402_ES.html?redirect#title2

[917] Las sanciones de tráfico han de ser notificadas al interesado, cuando no es posible la notificación ordinaria, la Administración Publica debe acudir a la publicación de estos actos administrativos en sus boletines oficiales. Es posible que una aseguradora esté interesada en aquellas personas que han sido sancionadas por incumplir la normativa de tráfico a la hora de realizar sus perfiles. Es cierto que normalmente los nombres de estas personas no aparecen como tal, sino su DNI o algunas cifras del mismo, aun así las posibilidades identificar a esas personas y hacer uso de esos datos persiste.

[918] Véase el artículo 9 referido a tratamientos de datos de categorías especiales y el artículo 10 sobre datos personales relativos a condenas e infracciones penales.

En *cuarto lugar*, el análisis de compatibilidad también debe comprender las posibles consecuencias que se pueden derivar del tratamiento ulterior previsto para los interesados. Se han de tener en cuenta las consecuencias tanto negativas como positivas. Como es lógico, cuanto más negativo o incierto sea el impacto del tratamiento posterior, más improbable será que se considere un uso compatible[919]. Por ejemplo, cada vez más Administraciones Públicas apuestan por utilizar los datos que los ciudadanos publican en sus redes sociales y otras web amparados en intereses públicos como el fraude fiscal[920], la corrupción o la lucha contra el terrorismo[921]. Aunque es cierto que el particular ha publicado dichos datos de forma voluntaria y se expone al riesgo de que los mismos puedan ser utilizados por terceros, ello no puede significar una total desprotección. Entre las consecuencias negativas de este uso secundario podemos encontrar la imposición de sanciones administrativas, la retirada de ayudas públicas, etc. También puede ser a la inversa, esto es, empresas privadas que obtienen datos personales de fuentes públicas procedentes de los boletines oficiales para llevar a cabo sus propios tratamientos. En estos casos, las consecuencias negativas que se pueden derivar de ese tratamiento posterior también pueden ser un elemento muy relevante para considerar la incompatibilidad de esa finalidad secundaria[922]. En general, cuanto más inesperado o sorprendente sea el uso posterior basado en ese contexto, más probable será que se considere incompatible el tratamiento posterior. En este sentido, y como ya hemos propuesto en esta tesis, una interpretación extensiva del principio de limitación de la finalidad en relación con los usos posteriores previstos podría establecer que el responsable del tratamiento, cuando tenga previsto anonimizar los datos, a la hora de

---

[919] Grupo del artículo 29. *Opinion 03/2013 on purpose limitation*. Texto adoptado el 2 de abril de 2013, págs. 25 y 28.

[920] En Francia, la Loi de finances pour 2020, sous le n° 2019-796 DC, le 20 décembre 2019 que legitima el análisis masivo de datos personales de los ciudadanos disponibles públicamente en las plataformas virtuales como método para luchar contra el fraude fiscal.

[921] Véase el 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.

[922] Recientemente la AEPD consideró que la empresa Equifax incumplió el principio de limitación de la finalidad. Esta entidad recopila datos personales de distintos individuos procedentes de los tablones oficiales de las Administraciones Públicas. Concretamente, obtiene información de aquellos actos administrativos referidos a la imposición de sanciones de diversa índole. Mientras que la finalidad inicial de esa información se refiere al deber que tienen las Administraciones de Públicas de publicar aquellos actos administrativos que no han podido notificarse (notificación infructuosa), la finalidad secundaria de que persigue Equifax es la evaluación de la solvencia financiera de esas personas. Teniendo en cuenta que las finalidades son totalmente diferentes y que además las consecuencias que se derivan de esa segunda finalidad resultan muy negativas para esas personas, la AEPD considera que la finalidad posterior no es compatible con la inicial. En: Agencia Española de Protección de Datos. Informe Nº: PS/00240/2019, págs. 136 y 137. Texto disponible en: https://files.lbr.cloud/public/2021-04/Equifax%20Spain%20fine.pdf?sVjYpVSULM6xtqnWqa4oSH06wnVm2_I8

valorar las posibles consecuencias de la anonimización tenga en cuenta los perjuicios que se pueden derivar de los potenciales usos que se pretenden con la anonimización de los datos, usos que generalmente tendrán como objetivo la elaboración de todo tipo de algoritmos y modelos para todo tipo de finalidades [923].

En definitiva, el análisis de la compatibilidad requiere de un estudio detallado de los factores previamente mencionados. Las posibles deficiencias que se puedan detectar en algunos de estos elementos podrán en su caso compensarse con el establecimiento de garantías adecuadas. El análisis de compatibilidad deberá quedar debidamente registrado ya que supone la herramienta principal que demuestra que el tratamiento de datos ulterior pretendido es conforme a la normativa de protección de datos[924]. Este será conveniente que se aporte junto a la EIPD cuando esta última sea obligatoria.

D) Las garantías necesarias

Independientemente del canal o vía que utilice un responsable para legitimar la nueva finalidad ulterior que se pretende, este último estará obligado a implementar las suficientes medidas de garantía. La idea esencial es que las restricciones que sufre el principio de limitación de la finalidad queden en parte compensadas por medidas adecuadas que reduzcan al mínimo los efectos de tal restricción. En este sentido, las garantías a implantar dependerán en gran medida del tipo de tratamiento de datos posterior que se pretende realizar. Entre otras podemos indicar; la seudonimización, la anonimización, la notificación a los interesados, la seguridad del tratamiento, la indicación de los plazos de conservación de los datos con esa nueva finalidad, etc. En este sentido, con relación al uso posterior de los datos con fines de investigación científica o estadística, si a través de la anonimización de datos se pueden alcanzar esos fines, el responsable deberá implantar esa herramienta[925]. En estos supuestos, si el responsable del tratamiento prefiere mantener los datos personales porque considera que

[923] Véase Capítulo III, apartado IX, punto 1, epígrafe E).
[924] El grupo del artículo 29 indica que sería recomendable que el responsable del tratamiento pusiera a disposición de los interesados en los avisos o declaraciones de privacidad la información sobre el análisis de compatibilidad. En: Grupo del artículo 29. *Directrices sobre la transparencia en virtud del Reglamento (UE) 2016/679.* Adoptadas el 29 de noviembre de 2017. Revisadas por última vez y adoptadas el 11 de abril de 2018. Apartado 47, pág.27**.**
[925] Artículo 89.1 in fine del RGPD.

pueden ser más relevantes para el tipo de modelo que pretende desarrollar lo lógico es que acabe por implantar la seudonimización. También se ha indicado como medida adecuada en los tratamientos que tenga como fin la investigación científica la elaboración de un plan integral de investigación, este plan debería especificar las cuestiones de la investigación y los métodos de trabajo previstos de la manera más clara posible[926], así como las distintas fases que puede comprender el proyecto.

3. Nuevos enfoques a la hora de analizar masivamente los datos bajo el paraguas del principio de limitación de la finalidad

El análisis masivo de datos por parte de las organizaciones gracias a los avances tecnológicos está permitiendo el desarrollo de nuevas aplicaciones en multitud de entornos. La normativa de protección de datos no debería convertirse en un escollo para estos avances sino más bien en la base normativa que permita el desarrollo tecnológico bajo la cual se respeten los derechos y libertades de los titulares de dichos datos. Tal y como indica el considerando 4 del RGPD, *el derecho a la protección de datos no es un derecho absoluto sino que debe considerarse en relación con su función en la sociedad y mantener el equilibrio con otros derechos fundamentales*. Es por ello que este derecho pueda limitarse y restringirse en pro de otros derechos e intereses legítimos. Esta idea es la que en parte está presente cuando el propio RGPD, con relación al principio de finalidad, autoriza a utilizar los datos con una finalidad distinta a la inicial. Sin embargo, esta misma exigencia debería implementarse cuando desde el inicio no existe una idea clara sobre la finalidad pretendida. En estos supuestos, es recomendable establecer alternativas que permitan una restricción del principio de limitación de la finalidad *compensándolo* en su caso con la implantación de toda una serie de medidas adecuadas que garanticen un tratamiento de datos respetuoso con la normativa. Este enfoque es el que se desprende de la Disposición Adicional 17 de la LOPD de 2018 referido a la investigación de datos de salud ya antes comentada[927], este precepto permite por ejemplo que los interesados otorguen el consentimiento para finalidades que

[926] Grupo del artículo 29. *Directrices 5/2020 sobre el consentimiento en el sentido del Reglamento (UE) 2016/679, versión 1.1*. Aprobadas el 4 de mayo de 2020. Apartado 161, pág.30.

[927] Disposición adicional decimoséptima. Tratamientos de datos de salud. Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Un estudio sobre esta disposición puede verse encontrarse en: RODRÍGUEZ AYUSO, J F: "Tratamiento de datos relativos a la salud del interesado en el ámbito de la sanidad pública". *Actualidad Administrativa*, Nº 10, Sección Administración del siglo XXI, Octubre 2019, págs. 7 y ss.

abarquen áreas generales vinculadas a una especialidad médica o investigadora[928]. El legislador español, al introducir este precepto, es consciente de los límites que a día de hoy encuentran los investigadores que pretenden analizar datos masivos donde inicialmente la finalidad de lo que se pretende con ese estudio no está del todo definida. De manera que se relaja el mentado principio pero a cambio se implementan otras medidas de garantía. Entre otras destacan; el papel protagonistas del tratamiento de datos seudonimizados, la necesidad de llevar a cabo una EIPD, la concesión de mayor protagonismo a los comités de ética o la limitación de los investigadores a la hora de conocer la identidad de los sujetos sometidos a la investigación. Junto a estas medidas y dependiendo de los contextos también será relevante incluir el estudio de minimización de datos que proponemos en esta tesis o el plan integral de investigación al que aludíamos previamente[929]. Por su parte, la doctrina también ha defendido la necesidad de permitir en estos entornos el tratamiento de datos con finalidades genéricas en las fases iniciales de estos proyectos. Para compensar esta medida, se apuesta por la realización de una evaluación de riesgos donde se valoren las posibles finalidades que se pretende con estos datos y todo ese procedimiento sea supervisado por parte de las autoridades de control de protección de datos[930]. De esta manera, en este modelo, lo relevante no es tanto que el individuo haya otorgado su consentimiento, el cual puede ser más genérico, sino que el origen de los datos sea legítimo, el uso secundario pretendido sea relevante y se implementen garantías suficientes que impidan que terceros no legitimados puedan acceder a esos datos[931]. Por tanto, los principios de tratamiento en materia de protección de datos deben equilibrarse con otros valores sociales e intereses legítimos donde se sopese el valor de los datos frente a los posibles

[928] En el mismo sentido, el Consejo de Europa ha señalado que: *Dado que no siempre es posible determinar de antemano los propósitos de los diferentes proyectos de investigación en el momento de la recopilación de datos, los interesados deben poder expresar su consentimiento para ciertas áreas de investigación o ciertas partes de los proyectos de investigación, en la medida en que lo permita el propósito previsto, con el debido respeto por las normas éticas reconocidas*. En: Consejo de Europa. Recomendación del Comité de Ministros a los Estados miembros sobre la protección de datos relacionados con la salud. Resolución de de 27 de marzo 2019, apartado 15.6. Texto disponible en: https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=090000168093b26e

[929] Por lo que se refiere al estudio de minimización de datos. Véase el Capítulo IV, apartado III, punto 1 de esta tesis.

[930] MANTELERO, A: "Toward a New Approach to Data Protection in the Big Data Era". En: GASSER,U; ZITTRAIN J; FARIS,R; HEACOCK JONES,R (dir.): *Internet Monitor 2014: Reflections on the Digital World* . Cambridge (MA): Berkman Center for Internet and Society at Harvard University, 2014, págs. 84 y ss.

[931] DE MONTALVO JÄÄSKELÄINEN, F: "Una reflexión desde la teoría de los derechos fundamentales sobre el uso secundario de los datos de salud en el marco del Big Data", op cit.,pág.61.

riesgos para la privacidad[932]. Como se ha indicado, cuando los posibles usos de los datos sean muy beneficiosos y los riesgos para la intimidad sean mínimos, algunas exigencias en materia de protección de datos deberían relajarse[933], compensando eso sí esos déficit con las debidas garantías. El modelo propuesto por la LOPD de 2018 para el tratamiento de datos de salud puede en nuestra opinión ser altamente recomendable y podría ser replicable en otros espacios con las salvedades y garantías que en su caso se requieran. Se evitaría así el *cortocircuito* de sectores que pueden ser altamente favorecedores sin que con ello se vulnere el derecho a la protección de datos gracias al despliegue de garantías adecuadas. Ni que decir tiene que las medidas proyectadas han de ser realmente aplicadas en la práctica y adecuarse al contexto[934]. Así, de nada servirá otorgar un mayor protagonismo a los comités de ética o a los planes de investigación integrales sí estos finalmente no tienen un papel relevante durante los tratamientos de datos que se pretende proyectar. Y es que el derecho a la protección de datos no puede ceder a cualquier precio.

Restricciones al principio de limitación de la finalidad	**Garantías para compensar las restricciones**
-Finalidades genéricas al inicio de los proyectos de analítica de datos. -Presunción de compatibilidad entre fin inicial y posterior para determinadas finalidades. (Artículo 89.1 RGPD) -Potenciar la compatibilidad entre finalidad inicial y posterior para analítica masiva de datos.	-Evaluaciones de impacto actualizadas. -Comités de ética u otros órganos. -Posibles certificaciones. -Uso de datos seudonimizados o anonimizados. -Estudio de minimización de datos. -Elaboración del plan integral del proyecto de investigación.

[932] Bienes jurídicos como la salud pública, la seguridad nacional, el cumplimiento de la ley, la protección del medio ambiente y la eficiencia económica, etc.

[933] Para los casos donde el riesgo para la privacidad resulte ínfimo se llega a decir que se debería asumir la legitimidad del tratamiento incluso si los individuos rechazan (o no se les pide) su consentimiento. En: TENE,O; POLONETSKY,J: "Big Data for All: Privacy and User Control in the Age of Analytics*"*, op cit., pág.260.

[934] A modo de ejemplo, en los proyectos de *smart city* donde esté proyectado el análisis masivo de datos personales el factor tecnológico de las herramientas que se utilicen resulta muy relevante. Pues bien, como ha señalado la doctrina, sólo podría entenderse lícito el tratamiento de la información si dicho tratamiento se somete a los estándares tecnológicos que en cada momento se encuentren aceptados. En: VALERO TORRIJOS, J: "Ciudades inteligentes y datos abiertos implicaciones jurídicas para la protección de los datos de carácter personal". *Istituzioni del federalismorivista di studi giuridici e politici*. Nº. 4, 2015, pág. 1045. Texto disponible en:
https://www.regione.emilia-romagna.it/affari_ist/Rivista_4_2015/Torrijos.pdf

III. EL PRINCIPIO DE MINIMIZACIÓN DE DATOS

El artículo 5.1.c) del RGPD establece que los datos personales *serán adecuados, pertinentes y limitados a lo necesario en relación con los fines para los que son tratados*. Este principio impone al responsable el deber de utilizar únicamente aquellos datos que sean estrictamente necesarios para llevar a cabo la finalidad que se pretende con el tratamiento. Como se ha indicado, este principio *se opone casi por definición* a la filosofía sobre la que se asientan las técnicas del *big data* y aprendizaje automático[935], esto es, el análisis masivo de datos. Y es que, la gran mayoría de los sistemas algorítmicos requieren de una gran cantidad de datos para aumentar su precisión y modelar adecuadamente el entorno sobre el que posteriormente adoptarán decisiones[936]. Es por ello que este principio resulte claramente afectado. Ahora bien, partiendo de esa premisa, en las siguientes páginas presentamos un enfoque que permite la convivencia del principio de minimización de datos con la analítica masiva de los mismos en la que, además de proteger los derechos de los particulares se logra un adecuado nivel de robustez del sistema. Para ello, se hace una división de este principio y la irradiación del mismo en las dos grandes fases que componen el ciclo de vida de los sistemas automatizados, estas son, la etapa de desarrollo y de despliegue.

1. El principio de minimización de datos durante el diseño de los algoritmos. El estudio de minimización de datos

Como se ha expuesto en otro momento de este trabajo, es habitual que durante el desarrollo de los sistemas algorítmicos las organizaciones no utilicen datos

[935] MITROU,J: "Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?", op.cit., págs 49 y50.

[936] Se ha hablado así de la paradoja de la privacidad de los datos. OCDE, *Artificial Intelligence in Society*, 2019, pág..84.

Desde el punto de vista estadístico, la muestra o conjunto de datos será normalmente más representativa a medida que aumente el número de ejemplos. De este modo, a mayor volumen de información disponible, más casos diferentes cubrirá esa base de datos y mejor se adaptará el modelo al caso de estudio real. Aunque es posible extraer conocimiento a partir de unos pocos datos, existe una alta probabilidad que el modelo generado sea demasiado específico y por tanto no sea útil para su posterior aplicación. No obstante, la respuesta no es sencilla. PANESAR, A: *Machine Learning and AI for Healthcare. Big Data for Improved Health Outcomes*, op cit., pág.149 y ss.

personales[937]. Este apartado se analizará basado en un enfoque donde los diseñadores de estos modelos utilizan datos personales. Sin embargo, como ahora se comprobará, la mayoría de los elementos que se indicarán son perfectamente atribuibles también a aquellos procesos de desarrollo de modelos donde no se utilizan datos personales. De esta manera, la aplicación de este principio se convierte en un aliado del diseño de la robustez de estos sistemas.

Pues bien, como sabemos, el principio de minimización de datos exige que la información que se pretende someter a tratamiento esté directamente vinculada a la finalidad a la que se destina dicha recopilación. Este principio debe estar presente durante todo el proceso que abarca el diseño de los sistemas, esto es, desde la recopilación de los datos, pasando por el análisis de los mismos hasta que se valida el modelo. Para ello, proponemos la realización de un estudio de minimización de datos. A través de este estudio, los responsables deberían ir justificando en cada una de las distintas fases que comprende el desarrollo de los modelos algorítmicos la pertinencia de los datos utilizados para configurar el proyecto que se pretende, esto es, la finalidad de la analítica de datos. De esta manera, ese estudio sería dinámico e iría alterándose conforme avanza el proyecto que se esté desarrollando. Es decir, a cada etapa se aplicarían distintas técnicas de minimización de datos. Así, entre otros elementos, en cada una de las fases los responsables deberían; a) explicar las razones que justifican que ese dato se mantiene en esa fase, b) indicar los datos que en su caso se han eliminado en esa fase y las razones de esa eliminación, c) indicar aquellos datos que aún no se han podido justificar plenamente pero que en fases posteriores se justificarán o en su caso se eliminarán, d) justificar que las datos elegidos no son variables *proxies* de datos de categoría especial y en el caso de que así sea, tener en cuenta que se están tratando datos de categoría especial[938]. Como es lógico, estos datos deberán tratarse en la medida de lo posible durante todas las etapas de forma seudonimizada[939].

El objetivo final que buscamos con esta propuesta es que a medida que avancen las fases, el conjunto de datos utilizados se vaya depurando hasta quedar perfilado y

[937] Como hemos dicho, es habitual que las organizaciones a la hora de diseñar modelos algorítmicos no traten datos personales ya que previamente los anonimizan. Véase Capítulo III, apartado IX, punto 1, aparatado D de esta Tesis.

[938] Respecto a esto último, la AEPD establece que los responsables han de realizar un análisis cuidadoso de las variables elegidas para establecer presunciones adecuadas sobre las posibles variables *proxy* que intervienen en el componente IA. En: Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág.21.

[939] La seudonimización durante los tratamientos de datos personales es una medida que favorece el principio de minimización de datos. Véase el Capítulo III, apartado IX, punto 2 de esta tesis.

conformado únicamente por aquellos que sean estrictamente necesarios y debidamente justificados. Es turno de analizar cómo podría integrarse este estudio de minimización de datos en la fase de desarrollo de sistemas automatizados y las técnicas que en su caso pueden aplicarse. Para ello se volverán analizar el conjunto de etapas que integran esta fase tomando como referencia la división realizada en el capítulo inicial de esta tesis[940].

Elaboración propia

A) Las técnicas de minimización en la fase inicial

Como ya se señaló a la hora de analizar el principio de limitación de la finalidad, en muchos proyectos de elaboración de modelo algorítmicos no resultaba sencillo prever de antemano y de forma específica la finalidad u objetivo pretendido. Este problema se acrecentaba cuando además se utilizan técnicas de aprendizaje no supervisado. Para superar ese escollo decíamos que era necesario que en cierta medida, el principio de limitación de la finalidad se relajara para evitar la paralización de este tipo de tratamientos pero no a cualquier precio, sino con las debidas medidas de garantía. Pues bien, una de esas medidas deviene de las exigencias establecidas por el principio de minimización de datos, el cual obliga al responsable a tratar de razonar en esta primera fase del diseño los motivos por los que se han elegido esos datos personales para el diseño del modelo algorítmico. De esta manera, el responsable en esta primera fase ha de argumentar con todo tipo de elementos las razones que le

[940] Véase el Capítulo I, apartado II, punto 2 de esta tesis.

empujan a utilizar esas bases de datos. Así, por ejemplo, una entidad bancaria que pretende utilizar técnicas de aprendizaje supervisado para analizar determinados datos de un conjunto de clientes con la finalidad de entrenar un modelo que distinga entre buenos y malos pagadores parte de una serie de premisas sobre las que decide iniciar ese proyecto de análisis. Entre esas premisas, el responsable puede basarse en hipótesis previas del uso de esos datos, aporte de otros estudios que corroboren que el uso de los mismos resulta coherente, etc. Esas hipótesis iniciales no deberían esconder ideales discriminatorios[941]. Por otro lado, un grupo de investigación podría utilizar técnicas de aprendizaje no supervisado para analizar datos genómicos de un conjunto de persona que han padecido cáncer. Aunque no se haya establecido previamente una etiqueta de salida y esencialmente se deja a los datos para que estos hablen por si solos y descubran nuevas relaciones o grupos relevantes. Incluso en estos supuestos, los responsables tienen ciertas intuiciones de la utilidad de esos datos y las mismas deben quedar reflejadas a la hora de justificar el estudio de minimización de datos en esta fase inicial. La justificación por tanto en esta fase inicial debe ser adecuada.

B) Las técnicas de minimización en la fase de pre procesamiento

Durante la fase de pre procesamiento, los desarrolladores de los modelos tratan de transformar las bases de datos de las que disponen en un formato que permita un adecuado procesamiento de las mismas por parte de los algoritmos. Así, resultará habitual que en esta etapa se lleve a cabo la eliminación de distintos datos por diversos motivos lo que favorecerá una reducción de dicha base de datos respaldando así al principio de minimización.

En *primer lugar*, resulta habitual la supresión de muestras que presentan atributos atípicos o poco representados. En estos supuestos, tanto la eliminación como en su caso el mantenimiento de este tipo de muestras deberá justificarse ya que, por un lado, la supresión puede llevar a eliminar datos de una parte minoritaria del entorno que

[941] Tal y como se ha señalado, los propios métodos de desarrollo de los sistemas de IA (por ejemplo, la programación de algoritmos) también pueden presentar sesgos injustos. Las hipótesis iniciales quedarían englobadas en este contexto. En: Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Directrices éticas para una IA fiable*.2019,pág.23.

se pretende modelar y[942], por otro, el mantenimiento de los mismos puede suponer una falta de precisión del modelo por incorporar variables no relevantes para el modelo.

En *segundo lugar*, también se deberá documentar y justificar aquellas muestras que se hayan mantenido pese a presentar valorar faltantes o aquellos supuestos en los que dichos valores faltantes hayan sido sustituidos por valores aleatorios o por otros valores resultantes de la aplicación de técnicas estadísticas como la mediana o la media.

En *tercer lugar*, también puede considerase una medida que favorece el principio de minimización de datos durante esta fase el uso de datos sintéticos en sustitución del uso de datos personales[943].

C) Las técnicas de minimización en la fase de entrenamiento/validación

Las técnicas de minimización de datos que pueden aplicarse en esta etapa son muy diversas y varían en gran medida del tipo de algoritmo o modelo que se desarrolle. Es en esta fase cuando se pueden descartar un gran número de datos ya que es en este momento del ciclo de desarrollo de los sistemas algorítmicos cuando se pueden probar y evaluar los distintos modelos y las correlaciones existentes entre las variables elegidas. Y es que, hasta la fecha, los diseñadores habían tratado de justificar los datos elegidos basados en posibles hipótesis previas o informes. Ahora es posible testar esas elecciones respecto de los modelos creados analizando las correlaciones existentes entre los distintos datos.

Así, en *primer lugar*, los diseñadores de los modelos pueden estudiar las posibles correlaciones existentes entre los datos elegidos. En este sentido, cuando dos datos presenten una correlación muy fuerte es posible que ambos estén midiendo la misma información[944]. La eliminación de alguno de ellos será necesaria para cumplir con el principio de minimización de datos.

---

[942] Hay que tener cuidado con los datos que se considera ruidosos en la medida que dichos datos pueden llegar a ser datos que si bien son distintos o diferentes a los más habituales, los mismos no tienen por qué ser incorrectos. Por ejemplo, datos de grupos minoritarios o poco representados.

[943] Consejo de Europa. Consultative Committee of the Convention for the Protection of individuals with regard to automatic processing of personal data. *Guidelines on Artificial Intelligence and Data Protection*. Resolución adoptada el 25 de enero de 2019. Pág.2. Disponible en: https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8
Para el concepto de datos sintético nos remitimos al Capítulo II, aparatado II, punto 5 de esta tesis.

[944] KUHN,M, JOHNSON,K: *Applied Predictive Modeling*. Ed. Springer Science + Business Media, New York, 2013, op.cit., págs. 43 y 44.

En *segundo lugar*, es necesario identificar y suprimir aquellas categorías de datos que no presentan ninguna influencia relevante para el desarrollo del modelo[945]. Las técnicas pueden ser muy diversas. Por ejemplo, a la hora de aplicar algoritmos de regresión[946], existen algunas técnicas que tratan de valorar si las variables elegidas presentan correlación con la variable de salida. Así, el llamado *estadístico f* se utiliza para comprobar si al menos hay una variable de entrada que presenta relación lineal con la variable de salida. Si se detecta que ninguna de las variables presenta relación lineal, hay que descartar esas variables. A su vez, el valor estadístico conocido como *P-valor de los estadísticos t* indica la relevancia de cada variable de entrada elegida respecto de la salida. Por tanto, si se muestra que existe una alta probabilidad de que una variable de entrada no presente relación lineal con la variable de salida, esta deberá eliminarse del modelo[947]. Ambos indicadores resultan fundamentales tanto para justificar la elección de una variable como en su caso para la supresión de la misma.

En *tercer lugar*, junto a las técnicas de minimización de datos ya mencionadas que coinciden en parte con las técnicas que aplican los especialistas en ciencia y analítica de datos para dotar de robustez al sistema. Existen otra serie de técnicas que se derivan estrictamente de la normativa de protección de datos, esto es, del principio de minimización de datos. Estas últimas exigen que los desarrolladores a la hora de diseñar estos sistemas valoren entre otros elementos la extensión de las categorías de datos tratadas, el grado de detalle de los mismos o el número de interesados sobre los que se tratan los datos[948]. En este sentido, los desarrolladores de estos sistemas deben justificar que han incorporado ese factor de privacidad en el desarrollo de los modelos. Por ejemplo, a la hora de elegir un modelo algorítmico que utiliza el dato de la fecha de nacimiento de las personas, se debería potenciar aquel que toma como referencia el año de nacimiento respecto de otros que identifican el año, el mes y el día concreto. Esta ponderación supone que a la hora de validar los distintos modelos que se han podido

[945] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020,pág.40.

[946] Recordemos que *la tarea de regresión* puede utilizarse para estimar la salida deseada, por ejemplo, estimar las ventas de un producto en un lapso de tiempo. Esta tarea es útil para resolver problemas que exigen como resultado el cálculo de una probabilidad. Entre los riesgos que se achacaban a estos datos encontrábamos que: *i)* los mismos llevaban consigo un mayor seguimiento y control del particular que solicita el servicio o la actividad sobre la que se pretende tomar decisiones automatizadas, ii) la fiabilidad de estos datos no está del todo constatada, iii) el proceso de recopilación de esos datos alternativos suele realizarse de forma oculta. Véase el Capítulo II, apartado II, punto 3 de esta tesis.

[947] Información obtenida en:
https://ichi.pro/es/la-prueba-f-para-el-analisis-de-regresion-33082111849049

[948] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020,pág.39.

generar no sólo se tenga como referencia el factor precisión o interpretabilidad de los mismo sino también el grado de afectación de privacidad que ese modelo puede generar fruto de los datos elegidos. La obtención de modelos algorítmicos que potencien la precisión en defecto del principio de minimización de datos en los términos descritos anteriormente deberá justificarse. Por ejemplo, puede estar justificado recopilar un volumen mayor de datos con el objetivo de que el sistema pueda generalizar de forma más adecuada la realidad donde el sistema adoptará las decisiones logrando con ello la obtención de datos de grupos minoritarios que pudieran no estar presentes en el conjunto de datos si se recopila una muestra pequeña de los mismos.

En *cuarto lugar*, las técnicas de minimización derivadas de la protección de datos también serán relevantes a la hora de justificar el uso de datos alternativos[949]. En nuestra opinión, debido a los riesgos ya señalados que presentan el uso de estos datos, consideramos que el tratamiento de los mismos requiere un mayor grado de justificación que el establecido para el resto de datos convencionales. Recordemos que el responsable del tratamiento podía tener dificultades para justificar el uso de datos alternativos en la fase inicial del desarrollo de estos sistemas ya que no disponía de elementos suficientes para compararlos con otras métricas. En esta fase ya sí que cuenta con dichos elementos y por tanto, cada uno de los datos alternativos ha de ser necesariamente justificados. Así, para justificar la pertinencia de estos nuevos datos se puede por ejemplo comparar la capacidad de predicción de un determinado modelo que usa esos datos alternativos con otros modelo predictivos que históricamente se han tenido como referencia. Por ejemplo, la calificación crediticia que establecen las agencias de calificación puede ser un buen parámetro sobre el cual se pueda comparar otro modelo algorítmico basado en datos alternativos. De esta manera, se puede utilizar la herramienta de la Curva AUC para valorar la capacidad que tiene el sistema que utiliza esos datos alternativos para detectar los verdaderos positivos[950]. Si se comprueba que al menos es tan bueno como el que establece la agencia de calificación, entonces hay que entender que está justificada esa conjunción de datos[951]. Con todo y con ello habría que valorar hasta qué punto determinadas variables alternativas que se tienen en

[949] Recordemos que el concepto de dato alternativo engloba a toda aquella información que comúnmente no ha sido utilizada por parte de una organización para un específico proceso de toma de decisiones pero ahora, y gracias al análisis masivo de datos, dicha información adquiere una fuerte relevancia. Véase el Capítulo II, apartado II, punto 4 de esta tesis.

[950] Véase el Capítulo I, apartado II, punto 2, epígrafe E) de esta tesis.

[951] BERG,T ; BURG,V ; GOMBOVI´C, A; PURI,M: "On the Rise of FinTechs: Credit Scoring Using Digital Footprints". *The Review of Financial Studies*, 2019, págs.. 17,19 y 20. Texto disponible: https://academic.oup.com/rfs/article/33/7/2845/5568311

cuenta resultan adecuadas aportarlas[952]. En otros supuestos también se podría argumentar o justificar la entrada o el uso de estos nuevos datos si por ejemplo se demuestra que el uso de los mismos permite desarrollar sistemas que amplíen el número de beneficiarios de un determinado servicio. Es decir, dado que los sistemas anteriores o los mecanismos previos se basaban en datos sesgados y no permitían el acceso a determinados grupos de población, si el uso de datos alternativos abre las fronteras de acceso a esos servicios a otros grupos históricamente perjudicados, la justificación de los mismos para desarrollar modelos también sería adecuada[953].

En *quinto lugar*, también consideramos que se debe prestar una especial atención y por tanto una debida justificación a la elección de datos difícilmente alterables por parte del particular. Así, la capacidad que tiene una persona para modificar el ADN, la raza o la edad es sumamente diferente a la capacidad para modificar los hábitos de compra o conducción de un vehículo[954]. En este sentido, ante una decisión que deniegue un determinado servicio o la adquisición de un producto, es posible que el responsable del tratamiento recomiende al particular afectado por la decisión que altere su conducta o sus características para que la decisión se vuelva más beneficiosa para él. Pues bien, aunque existirán datos que podrán ser fácilmente modificados, otros rara vez se podrán alterar y por tanto, la condena a esa personas que presentan esas características se pueden volver en muchos casos desproporcionada. Recordemos que la elaboración de perfiles grupales no distributivos se aplicaba a grupos de personas que ostentaban determinadas características similares[955]. En virtud de la pertinencia a ese perfil se asignaban determinadas consecuencias. Si dichos datos son difícilmente modificables, el particular difícilmente podrá cambiar la decisión dictada por el algoritmo.

En *sexto lugar*, el principio de minimización de datos exigirá al responsable preguntarse si, pese a que queden debidamente justificadas los datos elegidos en el modelo y estos sean por tanto pertinentes para la finalidad que se pretende con el tratamiento, los mismas resultan excesivamente incisivos con la privacidad de las

[952] Information Commisioner´s Office y The Alan Turing Institute. *Explaining decisions made with AI*. 2020. Págs.65 y 66

[953] BERG,T ; BURG,V ; GOMBOVI´C, A; PURI,M: "On the Rise of FinTechs: Credit Scoring Using Digital Footprints". *The Review of Financial Studies*, op.cit., pág.33.

[954] ZICARI,R: "Ethical Risk Assessment of Automated Decision Making Systems", *Operational Database Management Systems*, 2015. Texto disponible en: http://www.odbms.org/2015/02/ethical-risk-assessment-automated-decision-making-systems/

[955] Sobre los perfiles grupales no distributivos véase el Capítulo II, apartado I, punto 2, epígrafe B) de esta tesis.

personas. Es decir, incluso aunque se haya justificado y probado que ese dato elegido es adecuado, aún será necesario que el responsable realice un juicio de proporcionalidad para valorar las incidencias de la recopilación de ese dato en la esfera de los particulares. Así, cabe preguntarse si existen determinados datos que si bien pueden ser estadísticamente pertinentes para el modelo que se pretende desarrollar, la incidencia de los mismos en la esfera privada de las personas es tan relevante que su uso atentaría contra el derecho a la protección de datos. Por ejemplo, en el ámbito de los seguros automovilísticos, aunque pudiera ser pertinente a la hora de valorar el precio de la póliza datos tan específicos como la frecuencia de bostezos del conductor, la frecuencia de llevar o no pasajeros o el consumo de café o el alcohol que se ingiere[956]. Dichos datos afectan en tal grado al núcleo de privacidad del sujeto que, incluso resultando pertinentes, estos exceden de los estrictamente necesarios para el tratamiento. En este mismo orden de cosas se ha señalado que circunstancias tales como la actividad en redes sociales o la geolocalización de la persona no deberían considerarse adecuadas para evaluar la solvencia financiera de una persona ya que la norma que permite esa evaluación exige una valoración de la capacidad del pago del sujeto, no de su comportamiento general en la vida diaria[957]. En nuestra opinión, el uso de datos alternativos sí que puede considerarse legitimado siempre que los mismos estén debidamente justificados. Como se ha indicado anteriormente, a la hora de valorar la justificación no sólo se tendrá en cuenta los parámetros ligados a la precisión del modelo sino que los mismos se deberán de ponderar con las implicaciones que tenga la inclusión de ese valor no convencional en el derecho a la protección de datos en particular y el resto de derechos y libertades en general.

Etapas en la fase de diseño	**ESTUDIO DE MINIMIZACIÓN DE DATOS**
	Técnicas de minimización de datos en la fase de diseño
Fase inicial	-Hipótesis iniciales del proyecto debidamente justificadas. -Teorías que justifiquen los datos iniciales elegidos. -Estudios anteriores que corroboren los datos elegidos. -Tratamiento con datos seudonimizados.

[956] Comité de ética alemán. Gutachten der Datenethikkommission, 2019,pág. 106. En el mismo sentido, MUÑOZ PAREDES, M,L: "Big data y contratos de seguro: Los datos generados por los asegurados y su utilización por los aseguradores". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed. Aranzadi, Navarra, 2020, págs. 147 y 148. Esta autora indica que resulta inadecuado utilizar el dato de la solvencia financiera para valorar por ejemplo el precio de la póliza de un seguro.

[957] CASTAÑER CODINA, J: "La evaluación de la solvencia de las personas mediante el uso de algoritmos". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed. Aranzadi, Navarra, 2020, págs. 264 y 265.

Pre procesamiento	-Justificar la eliminación o mantenimiento de muestras poco representadas. -Justificar la eliminación o mantenimiento de muestras con atributos incompletos. -Uso de datos sintéticos.
Entrenamiento/ Validación	-Justificación de las correlaciones de los datos de entrada y salida. -Valorar si se ha alterado o no la finalidad inicial respecto de las datos seleccionados. -Eliminación de datos que representan la misma realidad. -Justificación adecuada de datos alternativos. -Justificación adecuada de los datos difícilmente alterables. -Análisis de la proporcionalidad de los datos elegidos. Precisión, pertinencia de los datos y grado de incisión en la privacidad.

Elaboración propia

En definitiva, el principio de minimización de datos durante la fase de diseño de los sistemas automatizados tiene como objetivo esencial la reducción al mínimo del conjunto de datos que conforman las bases de datos utilizadas para el desarrollo de los modelos algorítmicos. De esta manera, las distintas técnicas de minimización de datos que se aplican en cada una de las fases irán justificando los descartes o elecciones de los datos. La idea es que al final del proceso y una vez esté conformado el modelo, este sólo quede integrado por aquellos datos que sean pertinentes, esto es, aquellos que debidamente hayan sido justificados conforme los criterios anteriormente señalados. En este sentido, la PRAI obliga a los diseñadores de los algoritmos de alto riesgo a implementar prácticas adecuadas de gobernanza y gestión de datos sobre los datos de entrenamiento y validación que se utilicen a la hora de confeccionar estos sistemas[958]. Estas obligaciones son perfectamente compatibles con las que hemos indicado previamente y por tanto ayudarán al responsable a cumplir con la normativa de protección de datos en esta fase.

El llamado estudio de minimización de datos que planteamos puede ser visto como una especie de carta blanca que permite a las organizaciones recopilar cualquier tipo de datos aludiendo a fines analíticos. Partiendo por tanto de una base de datos, los datos que integran esta se van eliminando o justificando conforme se aplican las

[958] Artículo 10 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.
Las prácticas adecuadas de gobernanza y gestión de datos se alinearán con las normas potencialmente pertinentes como las ISO, IEEE, etc. En: Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Directrices éticas para una IA fiable*.2019, pág.36.

técnicas de minimización de datos. Como se indicó cuando se analizó el principio de limitación de la finalidad[959]. En muchos supuestos la finalidad inicial específica del tratamiento resultará muy difícil de establecer y por tanto, los datos iniciales necesarios también quedan en parte difuminados. Pues bien, para compensar ese relajamiento de ambos principios en esa fase inicial será necesario que se potencien otros principios también relevantes como la transparencia, la licitud del tratamiento y el principio de limitación de la conservación de los datos a través de compromisos claros de eliminación de los datos cuando estos no sean pertinentes para la finalidad inicial. Además, estos principios tampoco quedan eliminados del todo ya que en parte la recopilación de esos datos debe justificarse en cierta medida a través de la aportación de ciertas hipótesis u otros documentos que respalden la analítica de datos que se pretende. A partir de ahí, y tras ese primer sacrifico de los principios de minimización de datos y limitación de la finalidad –reforzados por el de licitud, transparencia y limitación del plazo de conservación- el principio de minimización de datos comienza a irradiar en todo su esplendor conforme se suceden las etapas de desarrollo de estos sistemas hasta dejar dicha base de datos en aquellos datos que sean estrictamente necesarios.

Esta propuesta no deja de ser una alternativa a las difíciles relaciones que se derivan de conjugar el principio de minimización a la analítica masiva de datos. Como se ha indicado, se intenta, por un lado, mantener un nivel alto de protección del derecho fundamental a la protección de datos y por otro, no cortocircuitar las oportunidades que brinda el desarrollo de estas tecnologías.

Es turno de analizar las incidencias del principio de minimización de datos en la fase de despliegue de los sistemas automatizados.

2. El principio de minimización en la fase de despliegue o toma de decisiones

A diferencia de la fase de diseño donde a priori podía ser complejo establecer una finalidad estricta o incluso indicándola meridianamente clara, los datos que se pretendía utilizar podían ser descartados a lo largo del diseño del sistema, es decir, no tenían por qué ser pertinentes al cien por cien en la fase inicial del desarrollo. En la fase de toma de decisiones, aquella organización que proyecta desplegar el sistema sí que es consciente ya del objetivo que se pretende con la implantación del modelo algorítmico y

[959] Capítulo IV, apartado II, punto 1, epígrafe A).

por tanto, fijada desde la fase inicial esa finalidad del tratamiento de datos, también se han de indicar aquellos datos que sean estrictamente necesarios para cumplir con el propósito marcado.

A la hora de valorar la pertinencia de los datos que se utilizarán para la toma de decisiones, el responsable del futuro tratamiento deberá analizar los tipos de datos que son requeridos para alimentar el sistema y la finalidad que se pretende para el mismo. En ese sentido, resultará sumamente útil para el responsable del tratamiento el estudio de minimización de datos que previamente se realizó durante la fase del diseño ya que en principio todas las variables que requiere ese modelo para llevar a cabo el tratamiento pretendido habrán sido debidamente justificadas durante la elaboración y conformación del mencionado modelo algorítmico. El estudio de minimización de datos debe tenerlo aquel responsable que pretenda desplegar el modelo ya que se trata de un documento que demuestra que dicho sistema cumple con el principio de minimización de datos[960]. En los casos en los que la organización que diseñó el modelo y la organización que ahora pretende implantarlo sean distintas, la entrega de ese estudio también será requerida dada la relevancia del mismo para por ejemplo ponerlo a disposición de las autoridades de control cuando estas lo requieran. No obstante, y a pesar de la importancia del mencionado estudio, el responsable que pretende implantar el sistema debe valorar específicamente el carácter necesario de los datos que se utilizarán en el contexto específico donde el modelo algorítmico adoptará las decisiones. Una recopilación excesiva de datos para un modelo que pretende adoptar decisiones sobre los individuos puede ser puesta en tela de juicio. Por tanto, incluso cuando estén debidamente justificadas dichas variables en el estudio de minimización de datos, aún se requiere ese análisis específico del contexto donde se adoptarán dichas decisiones. Este análisis sobre todo se requerirá cuando el adquirente del sistema o el que lo usa sea a una organización distinta a la que lo diseñó[961]. En este sentido, como ya se indicó en el capítulo primero, es posible que el sistema acabe implantándose en un entorno distinto para el que fue diseñado por lo que las probabilidades de que los datos requeridos no sean pertinentes para ese contexto aumentan. Así, por ejemplo, el sistema Predpol que se utiliza por parte de la policía en algunos estados de EEUU para prever ilícitos penales

[960] Se trata por tanto a su vez de una medida de responsabilidad activa. Artículos 5.2 y 24 del RGPD.

[961] Tal y como establece la ICO, una organización que adquiera un sistema deberá considerar si puede justificar la recopilación excesiva de datos que requiere el sistema que pretende utilizar. Fuente de la noticia: GALLO,V ; BINNS,R: "How using AI can require trade-offs between data protection principles, and what organizations can do to assess and balance them". 25/07/2019. *Information Commissioner's Office*. Información disponible:https://ico.org.uk/about-the-ico/news-and-events/ai-blog-trade-offs/

está integrado por un algoritmo diseñado para predecir terremotos[962]. Es posible que una vez el sistema comience a adoptar decisiones, este yerre. Ello puede llevar a la consideración de que los datos que se utilizan durante la toma de decisiones no sean considerados pertinentes para el propósito que se pretende en esa fase ya que el modelo inicial que se diseñó no respondía a esa finalidad.

IV. El PRINCIPIO DE EXACTITUD

El artículo 5.1.d) del RGPD establece que los datos personales serán *exactos y, si fuera necesario, actualizados*. Para cumplir con este principio, los responsables del tratamiento *adoptarán todas las medidas razonables para que se supriman o rectifiquen sin dilación los datos personales que sean inexactos con respecto a los fines para los que se tratan*. Este precepto obliga al responsable del tratamiento a garantizar que los datos que está tratando sean exactos y estén debidamente actualizados. Además, la obligación de garantizar la exactitud de los datos se relaciona con la finalidad del tratamiento de los mismos, de manera que, sólo los datos exactos y actualizados pueden servir para la finalidad para la cual han sido recopilados o pretenden ser tratados.

El principio de exactitud aplicado al conjunto de tratamientos de datos personales comprendidos durante el ciclo de vida de los sistemas automatizados obliga al responsable del tratamiento a utilizar durante todas las fases que comprende dicho período datos exactos y actualizados. El objetivo principal es que una correcta aplicación de este principio limite o reduzca las posibilidades de que los sistemas automatizados acaben adoptando decisiones inadecuadas o extrayendo inferencias inexactas. Así, por ejemplo, durante la fase de diseño, las bases de datos que se utilizarán para crear el modelo algorítmico han de contener datos exactos y actualizados. A su vez, en la etapa de despliegue, se han de habilitar canales adecuados para que la persona pueda rectificar las inferencias inexactas.

Ahora bien, es importante destacar que, a pesar de que el responsable utilice datos personales actualizados y exactos, aún es posible que el sistema pueda no cumplir ese objetivo de evitar decisiones inadecuadas o discriminatorias[963]. En este sentido, en

[962] DONOVAN, J, MATTHEWS, J, CAPLAN, and HANSON, L: "Algorithmic Accountability: A Primer". *DATA & SOCIETY*, 2018, págs.16 y 17. En: https://datasociety.net/output/algorithmic-accountability-a-primer/.

[963] En un mundo de grandes datos, lo que exige un escrutinio no suele ser la exactitud de los datos en bruto sino más bien la exactitud de las inferencias extraídas de los datos. Conclusiones inexactas, manipuladoras o discriminatorias pueden ser extraídos de datos perfectamente inocuos y precisos. En:

nuestra opinión, y teniendo en cuenta el considerando 71 del RGPD, junto al principio de exactitud que obliga a utilizar datos exactos y actualizados, el principio de lealtad exige del responsable en este contexto la necesidad establecer todos los medios posibles para evitar que las decisiones que adoptan estos sistemas generen efectos discriminatorios. Es decir, entendemos que las obligaciones que se derivan del principio de exactitud en este ámbito no siempre serán suficientes para lograr que un sistema adopte decisiones correctas. De ahí que el principio de lealtad complemente al de exactitud a la hora de ampliar las obligaciones que ha de asumir aquel responsable que diseña y utiliza sistemas de toma de decisiones automatizadas. Así, del principio de lealtad se desprende que los responsables, a la hora de diseñar y desplegar sistemas de toma de decisiones automatizados, han de utilizar modelos y procedimientos estadísticos adecuados y fiables con el fin de reducir al máximo los potenciales errores que pueden afectar a las decisiones que adoptan estos sistemas, decisiones que en muchos casos pueden resultar discriminatorias[964]. (Considerando 71 apartado segundo RGPD). En este sentido, el TJUE ha indicado que la fiabilidad de los modelos algorítmicos es un requisito mínimo para que un tratamiento de datos personales sea acorde a la normativa europea. Así, el este tribunal analizó el proyecto de acuerdo entre Canadá y la UE sobre transferencias de los datos del registro de nombres de los pasajeros aéreos desde la Unión a Canadá[965]. Concretamente, estos datos están principalmente destinados a ser sometidos a análisis automatizados basados en modelos algorítmicos y criterios preestablecidos.

Principio	**Obligaciones para el responsable**	**Finalidad compartida**
Exactitud	Bases de datos exactas y actualizadas	Evitar decisiones e inferencias inexactas
Lealtad	Uso de procedimientos matemáticos o estadísticos adecuados	Evitar decisiones discriminatorias

TENE, O; POLONETSKY,J: "Big Data for All: Privacy and User Control in the Age of Analytics", op.cit., apartado 87, págs.270 y 271.

[964] Sobre el principio de lealtad y la prohibición de discriminación algorítmica véase: Capítulo IV, apartado VII, punto 2 de esta tesis.

[965] El TJUE indicó que estos modelos han de ser fiables y no discriminatorios. DICTAMEN 1/15 del TJUE (Gran Sala) de 26 de julio de 2017, ECLI:EU:C:2017:592 apartados 168 a 173 y conclusión final tercera. Texto disponible en: https://curia.europa.eu/juris/document/document.jsf?text=&docid=193216&doclang=ES

Es turno de analizar las principales causas y factores que originan la adopción de decisiones incorrectas e inexactas por parte de los sistemas automatizados. Muchas de estas causas ya se han indicado a lo largo de este trabajo. Sin embargo, ahora las clasificaremos en función de la fase en las que tienen lugar y además se propondrán las medidas que ha de implantar el responsable del tratamiento para reducir los consecuencias negativas que generan estas decisiones incorrectas utilizando como base las exigencias que se derivan de los principios de exactitud y de lealtad. Artículo 5.1 letras a) y d) RGPD respectivamente.

.

1. Fase de diseño de los sistemas de toma de decisiones automatizadas

La *fase de diseño* de los sistemas automatizados está comprendida por todo un conjunto de etapas que tienen como objetivo final la creación de un modelo algorítmico. Las operaciones presentes durante este periodo tienen una fuerte repercusión una vez que el sistema se implanta en un entorno de ahí que muchas de las causas por las que un sistema genera decisiones incorrectas tengan su origen durante la elaboración del algoritmo. Así, resulta fundamental que la base de datos que se utilice para alimentar al modelo algorítmico que se pretende crear sea de calidad, entendiéndose por calidad aquellas bases de datos que reúnan las siguientes características: i) actualizadas, ii) representativas de la realidad que se pretende moldear, iii) no sesgadas.

En primer lugar, los datos que integran la base de datos han de ser lo más actuales posibles a la realidad que se pretende moldear[966]. Así, tal y como se ha indicado, las correlaciones que se derivan de una base de datos que presentan datos desactualizados e imprecisos muy posiblemente generarán modelos poco fiables[967]. Estas bases de datos, aunque pudieron albergar fuente de información que reflejaba ejemplos reales del momento en el que se recopilaron, han podido dejar de ser útiles fruto del paso del tiempo[968]. De esta manera, a medida que aumenta la antigüedad de las

[966] Tal y como recomienda el Grupo de expertos de la Comisión Europea. En: Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Directrices éticas para una IA fiable*, 2019, pág.35.

[967] Así, se ha señalado que sobre unos datos incorrectos no se pueden hacer cálculos precisos. GÓMEZ DE ÁGREDA,A: *Mundo Orwell. Manual de supervivencia para un mundo hiperconectado.* Ed. Ariel, España, 2019, pág. 93.

[968] WAYNE A, L; FERGUSON, GUTHRIE, A: "Policing Criminal Justice Data", *Minnesota Law Review* 541 2016, págs 568 y 572. Texto disponible en: https://ssrn.com/abstract=2761069

bases de datos sobre las que se pretende elaborar los algorítmicos, mayor es el riesgo de que el sistema posteriormente adopte decisiones inadecuadas, fruto de ese desfase temporal entre los datos de las personas que se incorporaron en el pasado a las bases y el despliegue del modelo algorítmico sobre el cual se adoptan decisiones que afectan a personas del presente. Para reducir estos inconvenientes derivados de ese desequilibrio temporal, corresponde al responsable valorar la antigüedad de las bases de datos y establecer periodos máximos de uso o de posibles actualizaciones de los mismos teniendo en cuenta las distintas finalidades que se pretendan con el desarrollo de los modelos. Por ejemplo, una base de datos policial de los últimos 20 años que se utiliza para crear perfiles de potenciales infractores de un determinado ilícito penal puede presentar importantes inconvenientes ya que, en ese periodo de tiempo no sólo el perfil de personas que comete ese tipo de delitos ha podido variar sino también porque el propio tipo penal ha podido ser alterado. Indicar por último que el artículo 4 de la LOPD de 2018 establece que no se imputará la inexactitud de un dato al responsable del tratamiento cuando dicho dato inexacto se haya obtenido de un registro público, siempre y cuando se hayan adoptado por este todas las medidas razonables para que se supriman o rectifiquen sin dilación indebida esos datos inexactos. Es decir, a priori, el responsable no se le puede achacar la inexactitud de esos datos cuando los mismos se obtienen de registros públicos. En estos supuestos se presume que dichos datos son exactos, por tanto, la inexactitud presente en los mismos no puede atribuirse al responsable. Ahora bien, el responsable ha de demostrar que ha hecho todo lo que está en su mano para comprobar que dichos datos son exactos, y en caso de que no lo sean, proceder a la rectificación o supresión de los mismos. Esta norma resulta interesante ya que un número importante de recopiladores de datos utilizan para elaborar perfiles los datos que se publican en los registros públicos, a priori, esos datos se presumen que son exactos. Sin embargo, a medida que transcurre el tiempo desde la publicación de los mismos, la posibilidad de que estos queden desactualizados aumenta. Corresponde a estos responsable cerciorase y asegurarse que, conforme pasa el tiempo, dichos datos siguen actualizados. Por ejemplo, el riesgo de que ese dato quede desactualizado aumenta una vez haya transcurrido el plazo máximo previsto para que dichos datos estén libremente accesibles en los registros públicos donde se recopilaron[969].

[969] Para el caso de las notificaciones administrativas, el periodo máximo que las mismas permanecerán libremente accesibles en la sede electrónica de la Agencia Estatal Boletín Oficial del Estado es de tres meses. Artículo 14 del Real Decreto 181/2008, de 8 de febrero, de ordenación del diario oficial Boletín

En segundo lugar, el conjunto de datos que conforman esas bases han de estar debidamente equilibrados y mostrar de forma adecuada la realidad sobre la que se pretende que el sistema adopte las decisiones. Es decir, el modelo debe garantizar que los datos del entrenamiento sean representativos del entorno en el que se desplegará el algoritmo entrenado[970]. Por ejemplo, el sistema VeriPol que utiliza la policía nacional para detectar posibles denuncias falsas fue entrenado con una base de datos compuesta por 1122 denuncias de robos presentadas en España, la mitad de esas denuncias eran verdaderas y la otra mitad falsas[971]. La proporción de las muestras utilizadas resulta proporcional, no sólo en cuanto a la distribución de las mismas sino también al hecho de que dichas denuncias se recopilaron de las comisarías de toda España. Por tanto, a priori, la representación de esa base de datos habrá sido adecamente generalizada por el sistema, es decir, la comisión delictiva de robos en todo el territorio nacional. Queda claro que el modelo ha de ser suficientemente capacitado con los posibles escenarios a los que se enfrentará una vez que comience a adoptar decisiones[972]. En este sentido, un problema habitual que se ha detectado en muchos sistemas de toma decisiones automatizadas es que el mismo adopta decisiones incorrectas en situaciones infrecuentes o poco usuales[973]. Ello se puede deber a distintas razones. Así, es posible que esos ejemplos o muestras se hayan eliminado intencionadamente de la base de datos por

Oficial del Estado. Este precepto ha sido recientemente reformado precisamente para reducir el riesgo de posibles usos inapropiados de los datos personales que se publican en este tablón. Véase el artículo único del Real Decreto 327/2021, de 11 de mayo por el que se modifica el Real Decreto 181/2008, de 8 de febrero, de ordenación del diario oficial "Boletín Oficial del Estado", para adaptarlo al Tablón Edictal Judicial Único.

970 MITROU,J: "Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?", op.cit., pág.53.

Si los datos son parcilamente inco

971 Exactamente, el sistema VeriPol, se capacitó con 1122 denuncias de robos presentadas en España en 2015. El corpus incluye 534 informes verdaderos y 588 informes falsos. En: QUIJANO-SÁNCHEZ,L; LIBERATORE,F; CAMACHO-COLLADOS,J, CAMACHO-COLLADOS,M: "Applying automatic text-based detection of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the police", op.cit.,pág.157.

972 La pertinencia y exactitud de los datos también se ha indicado que es sumamente relevante en cualquier modelo que adopta decisiones, independientemente de que las mismas afecten o no a personas. Por ejemplo, un sistema basado IA diseñado para detectar objetos específicos puede tener dificultades para reconocer objetos en condiciones de iluminación deficientes, por lo que los diseñadores deben incluir datos procedentes de ensayos de productos en entornos tanto típicos como mal iluminados. En: Comisión Europea. *Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica*. .2020, pág.11.

973 HOGAN-DORAN,D: "Computer says "no": automation, algorithms and artificial intelligence in Government decision-making", op.cit.,.pág.16. Este artículo analiza un sistema utilizado para recaudar deudas. El sistema tenía en cuenta a las personas que trabajan a tiempo completo o parcial pero sin prever que estas mismas pueden trabajar de forma intermitente u ocasional. Para estos últimos supuestos el sistema arrojó decisiones erróneas.

considerarse anecdóticos o residuales. A su vez, también es probable que a causa del número limitado de muestras que presentan características o variables minoritarias en la base de datos, el modelo algorítmico las haya descartado y no haya generalizado adecuadamente a las mismas. Con relación a esto último se ha indicado que una variable y/o grupo se considera modelada de forma robusta en el sistema cuando al menos la misma esté presente en el 5% del conjunto de datos o muestras que se utilizan[974]. Así, la AEPD establece la necesidad de realizar análisis que verifiquen la representatividad del conjunto de datos utilizados con relación a la población del contexto al que se orienta el componente IA y a los grupos definidos en el mismo[975]. La pregunta siguiente sería hasta qué punto se puede exigir a una organización en virtud del principio de exactitud y lealtad que el sistema sea lo suficientemente representativo del entorno al que posteriormente se enfrentará. La respuesta no es sencilla ya que pueden existir distintos factores a valorar. Por ejemplo, Australia utiliza un sistema de reconocimiento de voz para la renovación de determinados visados de residentes extranjeros. La renovación queda supeditada a que las personas superen una prueba oral de inglés, aquel que no la supera, no puede renovar ese visado. Dicha prueba la realizan las personas ante una máquina. Pues bien, se han presentado varias denuncias de personas que, aunque tienen un nivel alto de inglés, dicha máquina no reconoce adecuadamente determinados acentos, concretamente el irlandés[976]. Muy probablemente, el sistema fue entrenado con datos procedentes de ciudadanos australianos que presentan un determinado acento de la lengua inglesa específico que puede diferir del británico, el irlandés, el escocés, el norteamericano etc. Pues bien, ¿hasta qué punto se puede exigir a las autoridades australianas que en su base de datos de entrenamiento hubieran incluido todos los acentos potenciales de habla inglesa? Como hemos indicado la respuesta no es fácil. No

[974] Por debajo de este porcentaje, la variable y/o grupo puede tener una representación demasiado escasa en la base de datos, lo cual, no obstante, debe ser indicado en el estudio. Ello podría estar afectando a la precisión del modelo. En: Eticas-consulting. *Guía de Auditoría Algorítmica, 2021,*pág.47. Texto disponible en: https://www.eticasconsulting.com/wp-content/uploads/2021/01/Eticas-consulting.pdf

[975] Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág, 23.

[976] Fuente de la noticia: "Computer says no: Irish vet fails oral English test needed to stay in Australia*". The guardian.*8/8/2017. Información disponible en: https://www.theguardian.com/australia-news/2017/aug/08/computer-says-no-irish-vet-fails-oral-english-test-needed-to-stay-in-australia

Este problema es recurrente en otros servicios ofrecidos por Administraciones Públicas tal y como ocurre en EEUU. Así, a través de un estudio se ha comprobado que un sistema de reconocimiento de voz tiene una mayor tasa de errores en reconocer la voz de las personas de raza negra que blanca. Más allá del sesgo de la raza, aquí lo que se detecta es que existe un sesgo respecto de un tipo de lenguaje específico que suelen utilizar en mayor medida las personas de raza negra. Fuente de la noticia: METZ, C: "There is a Racial Divide in Speech-Recognition System". *The New York Times*. 23/03/2020. Noticia disponible en: https://www.nytimes.com/2020/03/23/technology/speech-recognition-bias-apple-amazon-google.html

obstante, una vez más consideramos que el enfoque del riesgo sobre el que se basa el RPGD para establecer unas u otras medidas de cumplimiento normativo puede tener cabida en este contexto. Así, indicábamos que los elementos que se podrían tener en cuenta a la hora de evaluar ese riesgo eran entre otros: la finalidad que se pretenda obtener con el despliegue del sistema, efectos de las decisiones en los derechos de los particulares sometidos a los tratamientos, tipo de algoritmo elegido, consecuencias de los falsos positivos y negativos, entorno en el que se despliega el sistema, presencia de muchos individuos interactuando con el sistema, etc[977]. Así, por ejemplo, el entorno puede ser un factor clave. Un sistema de reconocimiento facial utilizado para el control de fronteras en los aeropuertos deberá utilizar para su entrenamiento un conjunto de datos lo más representativo posible de la población objeto que lo utilizará, en este caso, se habrán de utilizar rostros de todo el mundo. También se habrá de establecer mayores exigencias en cuanto a la representatividad de la base de datos a los sistemas que utilicen las Administraciones Públicas. Si bien una empresa privada podría permitirse no entrenar su sistema de reconocimiento de voz con lenguas minoritarias, una Administración Pública que por ejemplo pretenda establecer un *chat bot* o asistente virtual basado en el reconocimiento de voz tendría que entrenar ese sistema en las distintas lenguas cooficiales de ese territorio[978]. No podemos olvidar que el potencial usuario de estas organizaciones puede hablar esas lenguas a la hora de solicitar un servicio público. Destacar que la PRAI obliga a los desarrolladores de sistemas catalogados como de alto riesgo a que los conjuntos de datos de entrenamiento y validación sean representativos y tengan en cuenta en la medida de lo posible las características o elementos propios del entorno específico en el que se pretende utilizar dicho sistema[979]. Por tanto, cuando coincida el ámbito de aplicación del RGPD con lo previsto por este proyecto legislativo, las exigencias previstas en este último complementarán y reforzarán a las indicadas por la normativa de protección de datos.

Finalmente, *en tercer lugar*, otra de las razones que exige que una base de datos sea adecuada es que la misma no presente datos sesgados o, en la medida de lo posible,

[977] Véase el Capítulo III, apartado I, punto 3 de esta tesis.
[978] Sobre el uso de lenguas minoritarias y el desarrollo de herramientas basadas en Inteligencia Artificial véase. TASA FUSTER,V: "Llengües minoritzades i tecnologies disruptives de comunicació", *Universitat de València*. 2020. Texto disponible en: https://www.uv.es/seminaridret/sesiones2020/vicentatasa.pdf
[979] Apartados tercero y cuarto del artículo 10 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

que estos sean reducidos. Como ya se sabe, los algoritmos tratan de generalizar los datos que se utilizan durante el entrenamiento, de manera que, una vez creado el modelo con esos datos, este mismo se implementa y adopta decisiones. Si esos datos presentan sesgos históricos previos, muy posiblemente el sistema adopte decisiones sesgadas y discriminatorias cuando despliegue sus efectos en la realidad[980]. Es por ello que los responsables deban adoptar las medidas adecuadas para reducir y controlar los potenciales sesgos que pueden estar presentes en dichas bases de datos. En este sentido, la AEPD ha indicado la necesidad de evaluar las bases de datos y comprobar que no existen datos históricos sesgados previos, o existiendo, se ha realizado una limpieza y depuración adecuada para la normalización de dicha base de datos[981]. A su vez, también se indica que resulta recomendable disponer de datos adicionales de cara a eliminar posibles sesgos. Queda por tanto clara la relación entre el principio de exactitud en relación con las bases de datos sesgadas y los riesgos que para los particulares pueden generar las mismas[982]. La discriminación algorítmica se analizará también en el apartado referido al principio de lealtad presente en este capítulo de la tesis.

En definitiva, en virtud del principio de exactitud, una base de datos será adecuada cuando dichos datos sean actualizados, representativos y en la medida de lo posible con sesgos reducidos. La calidad de los datos por tanto se muestra en muchas situaciones mucho más relevante que la cantidad de los mismos a la hora de que un sistema adopte decisiones adecuadas[983]. Los principios de exactitud y lealtad reconocidos en el RGPD obligan al responsable a velar precisamente por el cumplimiento de esos requisitos de calidad, los cuales han de estar presentes en las bases de datos que se utilizarán para entrenar a los modelos algorítmicos.

---

[980] Sobre los sesgos en la toma de decisiones existe un importante número de artículos que analizan este problema. M O,DONELL,R:"Challenging racist predictive policing algorithms under the equal protection clause", op.cit., pág. 548. También véase: BORGES BLÁZQUEZ,R: "El sesgo de la máquina en la toma de decisiones en el proceso penal", *IUS ET SCIENTIA*, Vol. 6,Nº 2, 2020 págs 69 y 70. A su vez, RICHARDSON, R; SCHULTZ, J Y CRAWFORD, K.: "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice", op.cit., pág.199. Por ultimo también se ha indicado por parte de los organismos internacionales. Australian Human Rights Commission. *Human rights and Technology. Discussion Pape*r, 2019,pág.84.

[981] Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág. 25.

[982] Agencia Española de Protección de Datos. Resolución Nº: PS/00120/2021, págs. 94 y 95.

[983] The Norwegian Data Protection Authority. *Artificial intelligence and privacy.* 2018, pág. 11.

2. Fase de despliegue de los sistemas de toma de decisiones automatizados

Una vez que el sistema comienza a adoptar decisiones y generar inferencias, el responsable del tratamiento ha de tratar de reducir al máximo aquellas que puedan ser incorrectas o inexactas. De ahí que, junto a las precauciones que haya podido tener durante la fase de diseño para prevenir tales errores, el responsable en esta fase también ha de implantar toda una serie de medidas en virtud del principio de exactitud.

Así, en *primer lugar*, una de las causas principales que generan decisiones incorrectas durante el despliegue se deriva del ingreso de modelos algorítmicos en entornos distintos a los cuales tal modelo fue entrenado. Es decir, es posible que el modelo haya utilizado datos para una finalidad inicialmente y, en un momento posterior, ese sistema se ingrese en otro entorno y para otra finalidad[984]. En estos supuestos, la posibilidad de que las decisiones sean inadecuadas aumenta. Para estos casos, resulta fundamental que el responsable realice controles de desempeño de ese sistema en ese nuevo entorno sobre el que irradia los efectos el sistema. Nos remitimos al apartado en el que tratábamos el monitoreo y la evaluación de los sistemas como medida de responsabilidad activa necesaria[985]. Aquí, una vez más, el responsable que implanta dicho sistema ha de ser consciente que existen determinados contextos donde no se recomienda su incorporación[986].

En *segundo lugar*, también hemos hecho referencia a la distinción entre sistemas que se incorporarán en entornos no adaptativos y adaptativos[987]. Para el caso de estos últimos, el responsable ha de tener en cuenta que la probabilidad de que el sistema comience a adoptar decisiones incorrectas aumenta fruto de la interrelación del algoritmo con dicho contexto. Por ello, y para evitar esta situación, dicho responsable deberá establecer las medidas oportunas para reducir ese riesgo. Nótese por ejemplo que

---

[984] Es posible que se hayan desarrollado aplicaciones para reconocer a los adultos o a las personas de cierta edad y que no reconozcan a los niños con una precisión suficientemente alta. Las tasas de error de las aplicaciones pueden aumentar drásticamente cuando se aplican a un subgrupo para el que la aplicación no fue diseñada. En: LEARNED-MILLER,E: ORDÓÑEZ,V; MORGENSTERN,J; BUOLAMWINI,J: *Facial Recognition Technologies in the Wild: A Call for a Federal Office* and the supplemental document *Facial Recognition Technologies,* Algorithmic justice league, 2020, págs 24 y 25. Texto disponible en: https://global-uploads.webflow.com/5e027ca188c99e3515b404b7/5ed1145952bc185203f3d009_FRTsFederalOfficeMay2020.pdf

[985] Capítulo III, apartado X de esta tesis.

[986] Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág.17.

[987] Capítulo I, apartado II, punto 3, epígrafe E) de esta tesis.

un sistema puede adaptarse a un entorno donde claramente se repiten los sesgos estructurales de la población y este acabe aprendiéndolos, ello se deberá tener en cuenta.

Finalmente, en *tercer lugar,* el responsable, siendo consciente que su sistema arrojará errores, deberá habilitar canales adecuados que permitan a los interesados modificar y alterar las inferencias y decisiones inexactas que generen estos modelos algorítmicos.

En este sentido, hemos de recordar que la elaboración de perfiles grupales en ningún momento reflejará inferencias exactas de la persona que se somete a dicho perfil. Aquellas personas que forman parte del mentado perfil se les asigna una serie de consecuencias independientemente de que dicha inferencia asignada sea o no correcta. Por tanto, siempre se ha de habilitar un canal para poder rectificar esa inferencia que le afecta y que es probable que sea errónea o inexacta. Para estos supuestos, el derecho de rectificación reconocido en el artículo 16 del RGPD se muestra como la facultad idónea que habilita a los particulares para actualizar y en su caso rectificar las inferencias y decisiones inexactas. En el apartado dedicado a esta facultad se analiza con más detalle el ejercicio de este derecho[988].

PRINCIPIO DE EXACTITUD Y LEALTAD	
FASE DISEÑO	Bases de datos actualizadas Bases de datos representativas Bases de datos en la medida de lo posible no sesgadas
FASE DESPLIEGUE	Incorporación de sistemas en entornos distintos a los de la fase de desarrollo Incorporación de sistemas en entornos adaptativos Mecanismos adecuados que faciliten la rectificación de las inferencias inexactas y discriminatorias

[988] Capítulo V, apartado II, punto 2 de esta tesis.

V. EL PRINCIPIO DE LIMITACIÓN DEL PLAZO DE CONSERVACIÓN DE LOS DATOS

De acuerdo al artículo 5.1.e) del RGPD, los datos personales *serán mantenidos de forma que se permita la identificación de los interesados durante no más tiempo del necesario para los fines del tratamiento de los datos personales.* Este periodo de conservación podrá ser superior cuando dichos datos se traten exclusivamente con *fines de archivo en interés público, fines de investigación científica o histórica o fines estadísticos*. Además, cuando se apliquen dichas políticas de conservación, se habrán de establecer las suficientes medidas técnicas y organizativas apropiadas. Este principio trata de evitar que el responsable conserve datos personales más allá del tiempo necesario que requiere la finalidad del tratamiento que se pretende ejecutar. Es por ello que se deban establecer los plazos de conservación de los datos o en su caso los criterios utilizados para determinarlos[989], obligando al responsable a establecer una política de conservación de los datos.

El principio de limitación del plazo de conservación de datos tiene su impronta durante los tratamientos de datos que están presentes en el ciclo de vida de los sistemas de toma de decisiones automatizadas. Pasamos a analizarlas.

Por lo que se refiere a la *fase de diseño* de los sistemas automatizados, el responsable del tratamiento puede establecer distintos criterios y plazos de limitación. Uno de los criterios básicos que consideramos que puede ser relevante a la hora de valorar la conservación de los datos es el referido a la antigüedad de las bases de datos sobre las que se diseñan los modelos algorítmicos. Así, a medida que esos datos son más antiguos, mayor es la posibilidad de que estos estén desactualizados y por tanto, es más probable que el modelo creado generalice una realidad desactualizada del entorno en el que posteriormente este último adoptará decisiones. Por ejemplo, imaginemos que una entidad bancaria pretende crear un modelo de evaluación de la solvencia financiera, este modelo utiliza para su entrenamiento una base de datos de los antiguos clientes de esa organización que abarca los años 1970 a 1985. A pesar de que el modelo generalice adecuadamente la realidad de esa base de datos, esta muy posiblemente diferirá del contexto en el que se pretende implantar ese sistema, esto es, el año 2021. La conservación además de esos datos respecto de esos periodos será muy posiblemente

[989] Considerando 39 y artículos 13.2.a), 14.2.a) y 15.1.d) del RGPD.

contraria a la normativa de protección de datos ya que los mismos no se pueden conservar tanto tiempo. En este sentido, la AEPD ha considerado excesivo el tratamiento de datos por un periodo superior a 2 años cuando dichos datos se almacenen con la pretensión de crear algoritmos[990]. El segundo criterio es referido a la finalidad que se pretenda con la analítica de datos y el diseño del modelo algorítmico. En este sentido, a priori, una vez que el modelo ha sido diseñado y configurado, habría que entender que el propósito de la recopilación se entiende completado. No obstante, hay que ser conscientes que este periodo se puede ampliar ya que una misma base de datos puede utilizarse para desarrollar distintos modelos fruto de las distintas técnicas y algoritmos que se pueden aplicar sobre esa base de datos. Es por ello que si bien los plazos de conservación no han de ceñirse al diseño único de un modelo algorítmico, tampoco pueden resultar totalmente inciertos.

En lo que respecta a la *fase de despliegue*, hay que partir de que la conservación de los datos cuando se está llevando a cabo la elaboración de perfiles permite ventajas significativas al responsable ya que dicho perfil puede alimentarse con más datos y por tanto el mismo se vuelve más exhaustivo y las decisiones que se toman sobre el perfilado se vuelven cada vez más precisas[991]. Pues bien, a pesar de esa ventaja, al responsable le corresponde establecer plazos máximos de conservación de esos perfiles. En este sentido, el TJUE ya ha indicado claramente que el responsable ha de fijar el plazo o criterios para determinarlo durante el cual permanecerán activas las cookies[992]. Uno de los criterios que se puede establecer es el referido a la retirada del consentimiento ejercido por parte del particular. A su vez, los perfiles y las inferencias basadas en estos sobre personas también quedarán limitados cuando finalice el contrato,

---

[990] Agencia Española de Protección de Datos. Resolución Nº: PS/00070/2019, pág.129**.** Resolución disponible en: https://www.aepd.es/es/documento/ps-00070-2019.pdf

[991] "Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679". Grupo del Artículo 29. Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.13.

[992] Así, el TJUE ha señalado que: *Si bien el tiempo durante el cual se va a proceder al tratamiento de los datos no se halla entre dichas indicaciones, de la expresión «por lo menos» que figura en el artículo 10 de la Directiva 95/46 resulta, no obstante, que esta enumeración no es exhaustiva. Pues bien, debe considerarse que la información relativa al tiempo durante el cual las cookies estarán activas responde a la exigencia, establecida en dicho artículo, de que el tratamiento de los datos sea leal, puesto que, en una situación como la controvertida en el litigio principal, un período de tiempo largo, o incluso ilimitado, implica la recogida de numerosos datos sobre los hábitos de navegación y la frecuencia de las eventuales visitas del usuario a los sitios de los socios publicitarios del organizador del juego con fines promocionales*. SENTENCIA DEL TRIBUNAL DE JUSTICIA (Gran Sala) de 1 de octubre de 2019, asunto C-673/17, caso Planet49, FJº 78. Resolución disponible en: https://curia.europa.eu/juris/document/document.jsf?text=&docid=218462&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=2440473

la obligación legal o la misión de interés público sobre la que se basa el tratamiento y dichos datos ya no sean necesarios para la finalidad inicial. En relación con las inferencias, será fundamental que queden eliminadas a la mayor brevedad posible aquellas que impliquen falsos positivos o falsos negativos. La AEPD ha considerado que un perfilado que analice masivamente las transacciones y movimientos financieros realizados por un una entidad bancaria respecto de un cliente con el fin de personalizar dichos servicios no podrá superar el plazo de un año de duración[993]. Y es que, tal y como hemos señalado, a medida que el perfil se enriquece, el conocimiento que se obtiene de la persona y las inferencias realizadas cada vez resultarán más invasivas, por ello, el riesgo para los derechos y libertades de esas personas aumentará. En nuestra opinión, ello se deberá de tener en cuenta a la hora de establecer los criterios y plazos de eliminación y conservación de los datos.

1. La conservación de los datos más allá del tiempo estrictamente necesario

Son varios los supuestos en los que el responsable puede conservar los datos personales más tiempo del estrictamente necesario una vez que la finalidad del tratamiento se ha cumplido, estos son:

En primer lugar, el responsable puede establecer plazos más amplios de tiempo cuando la finalidad del tratamiento responda a fines de archivo en interés público, fines de investigación científica o histórica o fines estadísticos. Artículo 5.1.e) RPGD. Es decir, para estos concretos tratamientos, los responsables pueden tratar los datos durante periodos más amplios. En estos casos, las organizaciones estarán obligadas a desplegar las suficientes medias técnicas y organizativas para compensar la restricción del principio de limitación de la conservación de los datos. La seudonimización de los datos será una de esas garantías básicas que deberá implantarse para que esos datos se puedan seguir tratando. A modo de ejemplo, los datos genómicos que ostente una organización en una base de datos podrán mantenerse más allá del tiempo que pueda durar el desarrollo de modelos algorítmicos siempre que dichos datos se seudonomicen. Es importante destacar que esta restricción al principio general de limitación de la conservación de los datos sólo resulta aplicable a los datos que se pretendan tratar para

[993] Agencia Española de Protección de Datos. Informe nº 0195/2017, pág.15. Texto disponible en: https://www.aepd.es/es/documento/2017-0195.pdf

las finalidades mencionadas previamente. El RGPD ha sido muy estricto en este sentido ya que habilita a esos tratamientos de datos con periodos de tiempo más amplios "exclusivamente" para esas finalidades[994]. Por tanto, si un responsable que ostenta una base de datos personales pretende conservar dichos datos por un tiempo superior al estrictamente necesario para otras finalidades diferentes a las señaladas en el artículo 5.1.e), este se verá obligado a dejar de utilizarlos, incluso, aunque los acabe seudonimizando. Para estos casos, proponemos dos alternativas: o bien el responsable se ampara en alguna norma que habilite la restricción de este principio o bien en su caso dicho responsable podría consultar a la autoridad de control para que esta última le autorice a conservar dichos datos para periodos más amplios. Esta opción aparece contemplada por el Reglamento de desarrollo de la antigua LOPD de 1999[995], este texto permite a los responsables conservar los datos por periodos más amplios siempre que lo soliciten a la autoridad de protección de datos correspondiente y esta lo autorice[996]. Curiosamente, este precepto está diseñado para que los responsables soliciten la conservación de los datos por periodos más amplios cuando los fines pretendidos sean científicos, históricos o estadísticos. Este precepto, aunque pensado para los tratamientos de esas finalidades específicas, podría servir de referencia para aquellos supuestos en los que los responsables pretendan desarrollar modelos algorítmicos y necesiten conservar los datos más allá del tiempo estrictamente necesario cuando dichas

[994] De acuerdo al artículo 5.1.e) del RGPD, los datos personales serán: *mantenidos de forma que se permita la identificación de los interesados durante no más tiempo del necesario para los fines del tratamiento de los datos personales; **los datos personales podrán conservarse durante períodos más largos siempre que se traten exclusivamente** con fines de archivo en interés público, fines de investigación científica o histórica o fines estadísticos, de conformidad con el artículo 89, apartado 1, sin perjuicio de la aplicación de las medidas técnicas y organizativas apropiadas que impone el presente Reglamento a fin de proteger los derechos y libertades del interesado («limitación del plazo de conservación»).* (La negrita y la cursiva son nuestras)

[995] Véase los artículos 157 y 158 del Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal.

[996] Este precepto forma parte de la *Sección 2.ª Procedimiento para la autorización de conservación de datos para fines históricos, estadísticos o científicos.*
Concretamente, el artículo 157 establece que:
1. El procedimiento para obtener de la Agencia Española de Protección de Datos la declaración de la concurrencia en un determinado tratamiento de datos de valores históricos, científicos o estadísticos, a los efectos previstos en la Ley Orgánica 15/1999, de 13 de diciembre, y en el presente Reglamento, se iniciará siempre a petición del responsable que pretenda obtener la declaración.
2. En el escrito de solicitud, el responsable deberá:
a) Identificar claramente el tratamiento de datos al que pretende aplicarse la excepción.
b) Motivar expresamente las causas que justificarían la declaración.
c) Exponer detalladamente las medidas que el responsable del fichero se propone implantar para garantizar el derecho de los ciudadanos.
3. La solicitud deberá acompañarse de cuantos documentos o pruebas sean necesarios para justificar la existencia de los valores históricos, científicos o estadísticos que fundamentarían la declaración de la Agencia.

finalidades no se correspondan con las indicadas previamente, esto es, archivo público, científica o estadística.

En *segundo lugar*, estos datos también podrán conservarse por periodos más amplios cuando los mismos pretendan utilizarse para una finalidad distinta a la inicial. Es decir, en estos casos la justificación que habilita un tratamiento más duradero de los datos obedece a que la finalidad inicial del tratamiento se altera por otra posterior. Para estos supuestos nos remitimos a las explicaciones ya comentadas sobre el uso de datos para fines ulteriores contenidas en el apartado referido al principio de limitación de la finalidad[997]. Recordemos nuevamente que la seudonimización de los datos será una medida básica en estos contextos.

En *tercer lugar*, esos datos también podrán conservarse cuando sobre los mismos se apliquen técnicas de anonimización y tales datos queden completamente anonimizados. Al no aplicarse sobre esa información anónima la normativa de protección de datos, el tiempo de conservación de dichos datos ya no es relevante para el derecho a la protección de datos. Únicamente el responsable deberá evitar la reidentificación de tales datos en los términos explicados en el apartado referido a la anonimización.

En *cuarto lugar*, tal y como ha indicado la APED, también deberá conservarse una muestra de datos de entrenamiento con el objeto de auditar el componente del sistema algorítmico[998].

Finalmente, *en quinto lugar*, el responsable también tendrá la obligación de conservar dichos datos mientras que no hayan prescrito las posibles responsabilidades que se puedan derivar de los tratamientos de datos personales. Por tanto, durante el periodo que va desde que el dato deja de ser necesario para la finalidad del tratamiento hasta que prescriben dichas responsabilidades el responsable tiene la obligación de conservar los datos. Durante ese tiempo estos datos quedarán bloqueados sin que se

[997] Capítulo IV, apartado II, punto 2 de esta tesis.
[998] Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág.19.

pueda proceder al borrado físico tal y como indica el artículo 32 de la LOPD 2018[999]. Así, el apartado quinto de este precepto prevé la posibilidad de que las autoridades de protección de datos establezcan excepciones a la obligación de bloqueo en los supuestos en que, atendida la naturaleza de los datos o el hecho de que se refieran a un número particularmente elevado de afectados, su mera conservación, incluso bloqueados, pudiera generar un riesgo elevado para los derechos de los afectados, así como en aquellos casos en los que la conservación de los datos bloqueados pudiera implicar un coste desproporcionado para el responsable del tratamiento. Entendemos que esta situación podría ocurrir en aquellos casos en los que un responsable alberga un gran número de datos en sus *data lakes* y este entiende que existe dicho riesgo si los mismos únicamente quedan bloqueados.

VI. EL PRINCIPIO DE TRANSPARENCIA

El principio de transparencia obliga al responsable del tratamiento a establecer todo un conjunto de medidas que garanticen un adecuado entendimiento por parte del interesado de todo lo relativo al tratamiento de sus datos personales. La transparencia es uno de los elementos que más se ha visto reforzado tras la entrada en vigor del RGPD. Así, primeramente se reconoce como un principio básico relativo al tratamiento de datos personales en el artículo 5. En segundo lugar, este principio se materializa en otra serie de preceptos. Concretamente, el artículo 12, el cual establece cómo los responsables han de facilitar y mostrar dicha información, y, los artículos 13 y 14, los cuales prevén qué información se ha de proporcionar a los interesados. En tercer lugar, junto a estas obligaciones, la transparencia también está presente en otros preceptos del RGPD como son: el derecho de acceso (artículo 15), el derecho de explicación de la decisión (considerando 71 y artículo 22) y la comunicación de las violaciones de seguridad (artículo 34). Estos preceptos se analizan en otros apartados de la tesis[1000].

[999] Tal y como ha señalado la AEPD, la supresión de los datos da lugar al bloqueo de los mismos, lo que impide el tratamiento para la finalidad que justificó su recogida, conservándose únicamente para la exigencia de posibles responsabilidades derivadas del tratamiento y solo por el plazo de prescripción de estas. Pudiendo acceder a los mismo únicamente y para ese fin los jueces y tribunales, el Ministerio Fiscal o las Administraciones Públicas competentes, en particular de las autoridades de protección de datos. Agencia Española de Protección de Datos. Consulta jurídica N/REF: 00148/2019, pág.10.

[1000] Véase el Capítulo V, aparatado II, punto 1 (Derecho de acceso); Capítulo V, apartado III, punto 3 , epígrafe A) (Derecho de explicación); Capítulo III, apartado V, punto 3 (Comunicación de las brechas de seguridad).

Los responsables del tratamiento tienen libertad para decidir cómo informar a los interesados sobre sus tratamientos de datos. A pesar de que la inmensa totalidad de responsable acude a la redacción de políticas de privacidad escritas, lo cierto es que no existe ninguna obligación en todo el RGPD para acudir a este mecanismo o modelo informativo[1001]. En este sentido, el único requisito que se exige a los responsables es que la información se ofrezca en forma concisa, transparente, inteligible y de fácil acceso, con un lenguaje claro y sencillo. (Considerando 39 y artículo 12.1.). Precisamente, con el objetivo de facilitar esa información de manera eficiente y reducir con ello la fatiga informativa, muy presente en los entornos en línea donde es frecuente la toma de decisiones automatizadas, el GT29 o la propia AEPD han recomendado el uso de declaraciones de privacidad por capas[1002]. De manera que, en un primer nivel se presentaría una información básica de forma resumida de los elementos esenciales del tratamiento de datos previstos en la normativa para después, remitir a la información adicional en un segundo nivel donde se presentarán detalladamente el resto de las informaciones en un medio más adecuado para su presentación y comprensión[1003]. La idea es clara, el responsable informa de todos los elementos que exige el RGPD con relación al principio de transparencia, lo que implica una carga de información relativamente amplia, pero, a través de un mecanismo ágil y práctico.

Es turno de analizar en primer lugar las exigencias legales que establece este principio en relación con los tratamientos de datos objetos de esta tesis tanto en la fase de diseño como en la de despliegue para después, y en segundo lugar, analizar algunas de las limitaciones principales a las que se ve sometido el principio de transparencia y diversas soluciones que se han planteado.

1. Fase de diseño de los sistemas de toma de decisiones automatizados

Son varios los elementos sobre los que se ha de prestar atención, por un lado haremos referencia a los principales deberes de información presentes en la fase de diseño y, posteriormente mencionaremos aquellos elementos que tienen que tener en

[1001] CONTISSA,G; DOCTOR,K ; LAGIOIA,F; LIPPI,M; MICKLITZ,H; PALKA,P; SARTOR,G; TORRONI,P: "Claudette meets GDPR. Automating the Evaluation of Privacy Policies using Artificial Intelligence". *Study Report,*2018, pág.8.

[1002] Grupo de Trabajo del artículo 29. *Directrices sobre la transparencia en virtud del Reglamento (UE) 2016/*679.Adoptadas el 29 de noviembre de 2017 Revisadas por última vez y adoptadas el 11 de abril de 2018, apartado 8, pág.7.

[1003] Agencia Española de Protección de Datos. Guía para el cumplimiento del deber de informar. 2017, Pág.7.

cuenta los diseñadores de los sistemas en materia de transparencia cuando pretendan desplegar estos modelos algorítmicos en la toma de decisiones automatizadas sobre particulares.

A) Deberes de información en la fase de diseño

En primer lugar, como es sabido, durante la fase de elaboración de los modelos algoritmos los ingenieros y diseñadores suelen utilizar diferentes datos con el objetivo de extraer patrones y correlaciones inesperadas, así como capacitar estos modelos para que aprendan a adoptar decisiones por sí solos. Pues bien, tanto si los datos son recopilados directamente por parte del interesado como si no lo son, el responsable estará obligado a informar a los particulares de la *finalidad* a la que se destinarán dichos datos y la base jurídica del tratamiento sobre la que se haya amparado (Artículos 13.1.c y 14.1.c del RGPD). Como norma general, en esta fase, las finalidades estarán relacionadas con el diseño o la elaboración de modelos algorítmicos. Sería recomendable que en la primera capa de información únicamente se hiciera mención a la finalidad general que se pretende con el desarrollo del modelo, para después, y ya en la segunda capa, informar de esa finalidad de forma más pormenorizada a través de ejemplos sencillos y comprensibles. Así, si una entidad bancaria pretende recopilar datos para desarrollar un sistema de *scoring*. Esta podría informar a los titulares del conjunto de datos elegidos para capacitar el modelo indicando que tales datos se utilizarán con el fin de crear un modelo algorítmico que evalúe la solvencia patrimonial de potenciales clientes de esa entidad bancaria. No se entendería en cambio superadas las exigencias de trasparencia cuando únicamente se mencione que el uso de esos datos se destina a la analítica de datos o a la elaboración de los algoritmos[1004]. Finalmente, si los datos pretenden anonimizarse, de ello también se habrá de informar (Artículos 13.3 y 14.3 RGPD). En este sentido, se habrá de informar de las potenciales finalidades que se pretende con los procesos de anonimización. Y es que la anonimización en estos contextos tendrá normalmente como objetivo el desarrollo de todo tipo de modelos algorítmicos[1005].

[1004] ALIAGA MARTÍNEZ,L ; GUTIÉRREZ DAVID, E: "Explicando Machine Learning a través de la doctrina y práctica del Information Commissioner's Office", op.cit.,pág.9.

[1005] Como ya propusimos en el apartado referido a la anonimización, a la hora de valorar la compatibilidad entre la finalidad inicial y la ulterior (anonimización de datos) se debería poner el acento no sólo en los riesgos que se pueden derivar de la potencial reidentificación de los datos que se pretende

En *segundo lugar*, por lo que se refiere a las *categorías de datos*, el responsable está obligado a informar a los interesados sobre los datos que se recopilan, sobre todo, cuando los mismos no se hayan obtenido directamente del interesado (Artículo 14.1.d). A su vez, también se le ha de informar de la posibilidad de que sea reidentificado a partir de los datos que se obtienen del modelo o, tal y como indica el artículo 11.2 del RGPD, informarle de que no es posible realizar dicha reidentificación[1006]. En este sentido, y a pesar de que durante la fase de entrenamiento normalmente no se obtienen inferencias ni se adoptan decisiones automatizadas, es posible que de las operaciones que se lleven a cabo se extraiga información relevante atribuible a una persona específica. Dado que esa información es relevante, entendemos que tomando como referencia los principios de lealtad y transparencia[1007], el responsable tiene la obligación de informar al interesado. Así, la LOPD de 2018 permite la reidentificación de las personas cuyos datos seudonimizados se estén tratando con fines de investigación en salud pública y biomédica cuando un algoritmo detecte un peligro real y concreto para la seguridad o la salud de esas personas[1008]. De esta manera, detectado por un modelo algorítmico que una persona puede sufrir una enfermedad grave o incluso que la misma puede fallecer, se autoriza a informar de esa circunstancia a la persona afectada.

---

anonimizar sino que además, dicho enfoque de compatibilidad, con la entrada en juego del *big data*, obligaría a los responsable a tener en cuenta también los usos posteriores que se pueden hacer de los modelos una vez que estos se crean basados en esos datos. Por tanto, una interpretación extensiva del principio de limitación de la finalidad en relación con los usos posteriores podría establecer que, el responsable del tratamiento, cuando tenga previsto anonimizar los datos, valore las posibles consecuencias negativas que se pueden derivar ya no sólo de la anonimización en si a efectos de reidentificación, sino también de las consecuencia negativas que se pueden derivar del uso de esos datos para genera modelos que posteriormente afecten de forma negativa a las personas. Véase Capítulo III, apartado IX, punto 1, epígrafe E) de esta tesis.

[1006] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*. Febrero de 2020, pág.33.

[1007] Capítulo IV, apartado VII, punto 1 de esta tesis.

[1008] Véase la Disposición adicional decimoséptima aparatado segundo letra d) de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Este apartado indica que: *Podrá procederse a la reidentificación de los datos en su origen, cuando con motivo de una investigación que utilice datos seudonimizados, se aprecie la existencia de un peligro real y concreto para la seguridad o salud de una persona o grupo de personas, o una amenaza grave para sus derechos o sea necesaria para garantizar una adecuada asistencia sanitaria*

Se parte de que, los datos, al estar seudonimizados están separados técnicamente entre el equipo investigador y aquellos que realizan la seudonimización. Este precepto habilita al equipo investigador a ponerse en contacto con aquel que ha seudonimizado los datos para que este último se pueda poner en contacto con la persona sobre la que el modelo algorítmico ha detectado ese peligro real.

B) La incorporación del principio de transparencia durante la fase de diseño de los sistemas automatizados como previsión a su despliegue posterior

Aquellos que diseñen los algoritmos han de tener en cuenta que si proyectan el despliegue de sus sistemas en entornos donde se tratarán datos personales, deberán adecuar estos modelos algorítmicos a las exigencias jurídicas derivadas del principio de transparencia contemplados por la normativa de protección de datos. Recordemos que es posible que la organización que diseña el sistema y la que lo adquiere o utiliza no sea la misma. Independientemente de ello, las entidades encargadas del diseño han de ser conscientes de que una incorrecta implantación de medidas de transparencia en esta fase tendrá repercusiones graves en materia de incumplimiento de la normativa lo que puede hacer que dicho sistema no sea competitivo en el mercado por falta de adaptación al contexto normativo donde este pretende desplegar sus efectos. Tanto la organización que lo diseña y lo implanta, como la tercera que lo adquiere o lo usa verán imposibilitada la incorporación de estos modelos algorítmicos a sus procesos decisorios. Ello obliga a estos responsables a valorar aquellos elementos más críticos donde el principio de transparencia pudiera quedar afectado una vez que el sistema comience a adoptar decisiones. Pasamos a analizarlos.

El principio de transparencia ha de estar presente en todas las etapas y operaciones que transcurren durante la fase de diseño de estos sistemas. Así, a la hora de recopilar los datos, resulta esencial que se fijen claramente las fuentes de dichos datos, así como la antigüedad de los mismos. Posteriormente, en la fase previa al entrenamiento, se deberán reportar y documentar la eliminación o el mantenimiento de las variables más relevantes, así como la verificación de la calidad de esos datos. Recordemos que esto último ya se deriva del estudio de minimización de datos al que hacíamos referencia cuando explicábamos este principio[1009]. Seguidamente, en la fase entrenamiento, también será elemental indicar los algoritmos elegidos y los motivos de su elección. Finalmente, y una vez que estén constituidos los distintos modelos, los diseñadores deberían tratar de apostar por aquellos que resulten más interpretables a la hora de decantarse por uno u otro modelo. Es decir, junto a los criterios de precisión y eficiencia del sistema, los cuales suelen tenerse en cuenta a la hora de optar por un modelo u otro, el principio de transparencia exige que los diseñadores también valoren

[1009] Véase el Capítulo IV, apartado III, punto 1 de esta tesis.

la interpretabilidad de dichos modelos a la hora de elegirlos[1010]. Esto último, ni quiere decir que los responsables se vean obligados a elegir modelos altamente interpretables pero totalmente imprecisos, ni tampoco a la inversa, es decir, modelos súper precisos pero opacos. Hay que lograr un justo equilibrio eligiendo aquellos modelos que en parte potencien estos criterios. Esta idea se complementa con lo ya expuesto en el apartado referido a la minimización de datos donde aludíamos a la necesidad de tener en cuenta el factor privacidad a la hora de elegir el modelo más adecuado para el entorno donde se pretende que este adopte decisiones[1011]. Por último, se han de documentar las principales correlaciones y patrones presentes en dicho modelo.

Toda la información obtenida y registrada durante esta fase resultará esencial para el responsable que pretenda desplegar estos sistemas automatizados. Ya no sólo para cumplir con el principio de transparencia, sino también con el de responsabilidad activa (Considerado 78 RGPD). Así, por un lado, al optar por modelos más interpretables, los operadores o trabajadores de esa organización que utilizarán ese sistema tendrán más facilidad para interpretar sus resultados y comunicarlos a los particulares afectados[1012]. Por otro lado, la organización que utiliza el sistema, al conocer cómo funciona, también tendrá más facilidad para describir el tratamiento que pretende implantar e informar sobre la lógica del mismo[1013], así como una explicación adecuada de la decisión. A modo de ejemplo, si un sistema se ha diseñado para detectar contenido que incita al odio, durante esta fase se han debido de incorporar al código fuente el tipo de discurso que pretende identificar y en su caso bloquear. Posteriormente, ese discurso se deberá definir en las políticas de privacidad y en los términos y condiciones de las organizaciones que pretendan utilizar estos sistemas automatizados de reconocimiento de contenidos[1014]. En este sentido, la PRAI impone

[1010] Recordemos que tras el procesamiento de los datos por diversos algoritmos era habitual que se generaran diversos modelos. Una vez que evaluaban los mismos el siguiente paso era elegir aquel modelo que más se adecuara al contexto y objetivos previamente fijados durante la planificación del proyecto. La mayor o menor opacidad del sistema se ha de tener en cuenta a efectos del cumplimiento de la normativa de protección de datos. Véase el Capítulo I, apartado II, epígrafe F) de esta tesis.

[1011] Capítulo IV, apartado III, punto 2.

[1012] Se han de prevenir los errores de interpretabilidad de los resultados algorítmicos que pueden tener las personas que operan con los sistemas. Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.35.

[1013] Debe existir suficiente documentación para que los particulares puedan comprender la lógica del componente IA. Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021,pág.15

[1014] Los instrumentos de procesamiento del lenguaje natural requieren definiciones claras y coherentes del tipo de discurso que se ha de identificar. Los debates de política sobre la moderación del contenido y la explotación de los medios de comunicación social suelen carecer de esas definiciones precisas Otros

una serie de obligaciones a los desarrolladores de sistemas de alto riesgo que favorecen la trasparencia algorítmica. Entre otros deberes destacamos: i) desarrollar modelos algorítmicos que garanticen que el funcionamiento de los mismos sea lo suficientemente transparente permitiendo a las organizaciones que posteriormente hagan uso de estos sistemas poder interpretarlos y utilizarlos adecuadamente, ii) elaborar la documentación técnica del modelo algorítmico que posteriormente se facilitará a la organización que adquiere el sistema[1015]. En aquellos supuestos en los que el ámbito de aplicación de esta propuesta coincida con el del RGPD, las obligaciones de transparencia que se derivan de este proyecto legislativo -y que se imponen a los desarrolladores del algoritmo-, facilitarán al responsable que posteriormente hará uso de estos el cumplimiento de la normativa de protección de datos en la fase de toma de decisiones.

En definitiva, la transparencia algorítmica durante el desarrollo de los sistemas automatizados resulta fundamental ya que si esta no se implementa en esta fase del ciclo de vida del modelo, difícilmente podrán posteriormente cumplirse las exigencias en materia de transparencia que se derivan de la normativa de protección datos una vez que el sistema comienza a adoptar decisiones. Es turno de analizar precisamente dichos deberes de la etapa de despliegue de los modelos algorítmicos.

2. Fase de despliegue de los sistemas de toma de decisiones automatizados

El principio de transparencia se aplica a lo largo de todo el periodo que engloba los diversos tratamientos comprendidos en la toma de decisiones automatizadas. Desde la recopilación el dato para su ingreso en el sistema, pasando por la generación de inferencias hasta llegar a la toma de decisiones automatizadas, los deberes de información han de estar presentes.

A) Deberes de información en la fase de despliegue

Corresponde al responsable informar sobre el *tratamiento de datos y la finalidad* del mismo. Para el interesado cuyos datos se tratarán por un sistema de toma de

órganos sin efectos jurídicos. En: MARKOU, CH; DEAKIN,S: "Ex Machina Lex: Exploring the Limits of Legal Computability". En: MARKOU, CH; DEAKIN,S (eds): *Is Law Computable? Critical Perspectives on Law + Artificial Intelligence* . Ed. Hart Publishing, 2020, págs 31 a 66.

[1015] Véase el artículo 15 sobre los deberes de transparencia y los artículos 11 y 18 sobre la elaboración de la documentación técnica. Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021

decisiones automatizadas debe quedar claro el tratamiento específico y el objetivo u objetivos que se pretende con el mismo. Como ha indicado el GT29, los tratamientos de datos complejos o excesivamente técnicos obligan al responsable a detallar por separado y en un lenguaje sin ambigüedades cuáles serán las consecuencias más importantes del tratamiento en cuestión[1016]. Ese plus de complejidad está presente en todos los tratamientos de datos donde entran en juego sistemas de toma de decisiones automatizadas por lo que el responsable ha de prestar especial atención a la hora de informar sobre los mismos. De esta manera, son varias las obligaciones que se derivan.

a.1) Deberes de información exigibles a cualquier tratamiento de toma de decisiones automatizadas y elaboración de perfiles

En primer lugar, el responsable ha de indicar el tratamiento, especificando concretamente el mismo. Así, si el responsable pretende llevar a cabo la elaboración de perfiles, ello ha de quedar reflejado en la información básica que se ofrece al interesado (Considerando 60 del RGPD y Artículo 11.2. LOPD 2018). La referencia al tratamiento que se pretende no debe enmascararse en términos poco comprensivos o excesivamente jurídicos.

A su vez, *en segundo lugar*, además de hacer mención expresa al tratamiento, el responsable ha de indicar la finalidad y además las consecuencias que se derivan de dicho tratamiento. Concretamente, y por lo que se refiere a la elaboración de perfiles, tanto el GT29 como la AEPD han dejado claro en diversas resoluciones que independientemente de que el perfilado se incluya en la definición prevista en el artículo 22 del RGP, el responsable del tratamiento está obligado a informar sobre las consecuencias de dicha elaboración[1017]. Ello obliga al responsable al uso de ejemplos

[1016] Grupo de Trabajo del artículo 29. *Directrices sobre la transparencia en virtud del Reglamento (UE) 2016/*679.Adoptadas el 29 de noviembre de 2017 Revisadas por última vez y adoptadas el 11 de abril de 2018, Apartado 10, pág.8.

[1017] Véase el considerando 60 del RGPD. Además, otras autoridades también se han pronunciado en el mismo sentido. Grupo de Trabajo del artículo 29. *Directrices sobre la transparencia en virtud del Reglamento (UE) 2016/*679.Adoptadas el 29 de noviembre de 2017 Revisadas por última vez y adoptadas el 11 de abril de 2018, apartado 41, pág.25. También en: Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.18. A su vez: Comité Europeo de Protección de Datos. *Guidelines 1/2020 on processing personal data in the context of connected vehicles and mobility related applications***.** Versión 1.0. Adoptadas el 28 de enero de 2020, apartado 112, pág.23. igualmente también: Agencia Española de Protección de Datos. Informe Nº: PS/00070/2019, pág.108. Resolución disponible en: https://www.aepd.es/es/documento/ps-00070-2019.pdf

donde quede claro cómo ese perfilado le afecta a ese particular en ese concreto contexto lo que le permitirá entre otros aspectos decidir si consentir o no dicho tratamiento o, en el caso de que el perfilado responda a otras bases de tratamiento que no dependen del consentimiento, comprender el tratamiento en sí y valorar las consecuencias que sobre el mismo se generan[1018]. Por ejemplo, si una plataforma digital utiliza un sistema automatizado para analizar si determinados contenidos que se suben por parte de un usuario a una determinada plataforma incitan al terrorismo, el responsable del tratamiento ha de explicitar claramente en qué consiste el perfilado y además las consecuencias hipotéticas de tal perfilado. En la práctica ello obligaría a la plataforma a describir claramente el concepto de contenido terrorista sobre el que se ha diseñado el modelo algorítmico para después señalar los efectos que se derivan del sistema cuando identifique o asimile ese contenido que el particular sube a la plataforma como terrorista[1019]. En estos casos, las consecuencias normalmente serán la retirada de la foto, la opinión o el video compartido por ese particular. Los cuales, el sistema los identifica como contenido no permitido y los bloquea. Explicitar las consecuencias del perfilado resulta por tanto sumamente necesario para los interesados ya que les permite un conocimiento real de los usos de sus datos y las repercusiones que tienen sobre estas personas dichos usos. Cuando existan colectivos vulnerables a los que potencialmente se puede dirigir ese perfilado, la información sobre las consecuencias del perfil sobre ese colectivo deberán explicitarse. Toda esta información queda en nuestra opinión justificada a causa de la excesiva complejidad que este tipo de tratamientos suele presentar para los interesados. Se consigue reducir así parte de la opacidad que

---

Véase también: Agencia Española de Protección de Datos. Resolución Nº**:** PS/00477/2019, pág.105. Resolución disponible en: https://www.aepd.es/es/documento/ps-00477-2019.pdf

[1018] Hace poco tiempo en el diario *The Guardian* se denunció que Facebook Australia permitía que terceras empresas pudieran dirigir anuncios personalizados sobre alcohol, tabaco y juegos de azar a menores que previamente habían indicado interesarse por esos temas. En otros casos, los propios algoritmos habían inferido esa información. En nuestra opinión, si esto sucediera en Europa, tanto la existencia de ese perfilado como en su caso las consecuencias que se derivan del mismo deberían mencionarse expresamente. Y además, la legalidad de la medida también quedaría en entredicho.
Fuente de la noticia: TAYLOR, J: "Facebook allows advertisers to target children interested in smoking, alcohol and weight loss", *The Guardian*, 28/04/2021. Disponible en: https://www.theguardian.com/technology/2021/apr/28/facebook-allows-advertisers-to-target-children-interested-in-smoking-alcohol-and-weight-loss

[1019] Por ejemplo, el 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 indica la necesidad de establecer una definición clara del concepto de terrorismo que ayude a las plataformas a informar adecuadamente en sus términos y condiciones de aquellos contenidos que no estarán permitidos que se compartan por identificarse con contenido terrorista. Véase los artículos 2.7 y 7.1 de la normativa mencionada. Texto disponible en: https://www.boe.es/doue/2021/172/L00079-00109.pdf

acostumbra a estar presente en este tipo de tratamientos ya que no sólo se indica que hay un sistema automatizado que está elaborando un perfil sobre el que muy posiblemente se estén adoptando decisiones automatizadas, sino que además, se explicitan algunas de las consecuencias directas que se derivan del uso de ese algoritmo. En este sentido, el Tribunal Federal de Justicia de Alemania ha establecido que Facebook está obligado a informar en sus políticas y condiciones de uso sobre la posibilidad de retirada o bloqueo del contenido de sus usuarios cuando tal contenido subido pueda ser constitutivo de delito o sea contrario a las condiciones de uso. En este supuesto, el contenido fue retirado por un sistema automatizado. Este tribunal también ha indicado que los usuarios tienen derecho a poder impugnar la decisión adoptada relativa a la retirada de contenido[1020].

En *tercer lugar,* el responsable ha de informar de los datos que se utilizarán para la elaboración de perfiles y la toma de decisiones automatizadas. Merece la pena prestar atención a los datos inferidos y los deberes de información que se pueden exigir a los responsables del tratamiento respecto de este tipo de datos personales. (Consejo de Europa indica que hay que informar de los datos inferidos. (Apartado 5.1)

https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680a46147

Hay que recordar que los datos inferidos son el producto de tratamientos analíticos probabilísticos que detectan correlaciones entre los datos personales para generar predicciones sobre el comportamiento del interesado[1021]. Por tanto, estos datos se originan una vez que los algoritmos los procesan, no antes. A la hora de analizar los deberes de información que prevé el RGPD, es necesario distinguir dos situaciones, por un lado, cuando los datos se hayan obtenido del interesado (Artículo 13) y por otro, cuando dichos datos no se hayan obtenido del interesado (Artículo 14), entendiéndose que los mismos se han obtenido de un tercero u otra fuente (Considerando 61). Pues bien, por lo que se refiere a los datos obtenidos del interesado, el GT29 ha indicado que en ellos quedan incluidos los datos que un interesado ha facilitado conscientemente a un responsable del tratamiento y además aquellos datos que se obtienen de un interesado mediante la observación, esto es, los datos observados que se derivan del uso de

[1020] Sentencia de 29 de julio de 2021. Procedimiento III ZR 179/20. Texto disponible en: https://www.bundesgerichtshof.de/SharedDocs/Pressemitteilungen/DE/2021/2021149.html

[1021] ALIAGA MARTÍNEZ, L; GUTIÉRREZ DAVID, E: "Explicando Machine Learning a través de la doctrina y práctica del Information Commissioner's Office", op.cit., pág.9. Véase también el Capítulo II, apartado II, punto 4 de esta tesis.

cámaras, sensores, equipos móviles, etc[1022]. Los datos inferidos, por su naturaleza, no forman parte de ninguna de estas categorías. Además, dado que los datos en el momento de la recopilación aún no han sido procesados por los algoritmos, el responsable no puede conocer qué datos inferidos de esa persona inferirá. Por otro lado, el artículo 14 fija los deberes de información de los datos personales cuando no se obtienen del interesado sino de otra fuente. Así, en el plazo máximo de un mes a partir de la recepción de los datos de un tercero u otra fuente, los responsables del tratamiento están obligados a revelar las categorías de datos que ostentan de los interesados. En la práctica, un responsable del tratamiento que reciba datos inferidos como puede ser un puntaje de fraude facilitado por un tercero debería proporcionarlo al particular junto al resto de datos que haya podido obtener sobre ese interesado[1023]. Informar por tanto sobre dichos datos inferidos no presenta dudas. La siguiente pregunta sería, ¿qué ocurre con los datos inferidos que el responsable va conociendo una vez que pone en marcha el algoritmo? La AEPD ha dejado claro que conforme al artículo 14 del RGPD una entidad bancaria está obligada a informar a los interesados de los datos inferidos que dicha organización obtiene a través de sus propios tratamientos[1024]. De acuerdo a esta interpretación, entendemos que los datos inferidos que origina un algoritmo una vez que comienza a funcionar representan datos que no han sido obtenidos del interesado. Por ello, hay que considerar que la fuente que proporciona los datos a la que alude la normativa de protección es el algoritmo (considerando 61 y 14.2.f del RGPD). Por tanto, hay que entender que el responsable no sólo está obligado a informar de los datos inferidos sino también de la fuente de estos, esto es, el algoritmo o sistema automatizado. Ahora bien, teniendo en cuenta la gran cantidad de datos inferidos que puede generar un sistema automatizado cuando interacciona con un particular, la obligación de informar en todo momento de las nuevas inferencias puede resultar excesiva. Por ello, consideramos que estos deberes de información quedarán superados si: i) inicialmente el responsable facilita una información de los principales datos

[1022] Grupo de Trabajo del artículo 29. *Directrices sobre la transparencia en virtud del Reglamento (UE) 2016/*679.Adoptadas el 29 de noviembre de 2017 Revisadas por última vez y adoptadas el 11 de abril de 2018. apartado 26, pág.16.

[1023] Wachter considera que el responsable debe facilitar a los interesados aquellos datos inferidos que un tercero le haya proporcionado. Sin embargo entiende que existen limitaciones prácticas ya que realmente, a lo único que estaría obligado el responsable es a facilitar las categorías de datos. WACHTER, S; MITTELSTADT, B: "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI". op.cit. págs.52 y ss.

[1024] Agencia Española de Protección de Datos. Informe Nº: PS/00070/2019, págs. 94 y 96. Resolución disponible en: https://www.aepd.es/es/documento/ps-00070-2019.pdf

inferidos que potencialmente se inferirán por el sistema y además, ii) también se deja clara la posibilidad que tiene el interesado para solicitar el acceso al resto de datos inferidos una vez que el sistema comienza a funcionar. El ejercicio de esta segunda facultad ya se correspondería con el derecho de acceso reconocido en el artículo 15 del RGPD[1025]. La importancia de informar sobre los datos que potencialmente se inferirán resulta de suma importancia ya que en muchas ocasiones el particular sólo es consciente de los datos iniciales que el responsable le solicita pero no de aquellos que posteriormente inferirá el sistema.

Para finalizar, procede indicar que sería recomendable que toda la información referida a la elaboración de perfiles y a la toma de decisiones automatizadas se trate de forma sistematizada en las políticas de privacidad. Es decir, en un apartado específico hacer referencia a los distintos perfiles, sus finalidades, en qué consisten, consecuencias de los mismos, datos inferidos principales, etc.

Información que se ha de facilitar a los interesados cuando se adoptan decisiones automatizadas o se elaboran perfiles.	
Información sobre el tratamiento	-Se ha de mencionar expresamente la elaboración de perfiles o la toma de decisiones. -Se ha de indicar la finalidad de la elaboración de perfiles y las consecuencias del perfilado con especial atención a los colectivos más vulnerables. -Tipos de perfiles que se van a realizar y toma de decisiones automatizadas.
Información sobre los datos	-Datos que se utilizarán para el perfilado o la toma de decisiones. -Datos inferidos que potencialmente se inferirán por el sistema. -Información sobre el derecho del particular a acceder a los datos inferidos una vez que el sistema algorítmico comienza a funcionar.

a.2) Deberes de información exigibles a los tratamientos basados en la toma de decisiones plenamente automatizadas relevantes

Además de la información previamente detallada aplicable a todos los tratamientos, el RGPD prevé deberes especiales de información para los tratamientos descritos en el artículo 22 del RGPD. Concretamente, los artículos 13.2.f) y 14.2.g) establecen dos obligaciones centrales para el responsable. Por un lado, se ha de comunicar la existencia de estos tratamientos y por otro lado, se ha de aportar

[1025] Capítulo V, apartado II, punto 1, de esta tesis.

información significativa sobre la lógica aplicada así como las consecuencias previstas de dicho tratamiento para el interesado.

Por lo que se refiere a la *existencia de decisiones automatizadas*, el texto europeo, al hacer hincapié en ello pretende que todo interesado que se vea sometido al tratamiento definido en el artículo 22 sea conocedor de ello[1026]. Es decir, el particular debe ser consciente de que sus datos serán tratados por una máquina y que esta adoptará decisiones totalmente automatizadas relevantes sobre el interesado. Este precepto no deja de ser una vertiente de un principio elemental en las relaciones humano-máquina por el cual, las personas debemos ser conscientes de que estamos interactuando con un aparato automatizado[1027]. Ahora bien, el RGPD no sólo exige a aquellos que tratan los datos el deber de comunicar que están llevando a cabo el tratamiento definido del artículo 22, sino que además les impone el deber de informar sobre determinados elementos concretos del mencionado tratamiento, esto es, que ofrezcan información significativa sobre la lógica aplicada de tal tratamiento, así como la importancia y las consecuencias previstas del mismo.

En lo que respecta a la *lógica del tratamiento*, la información ha de ser la suficiente para que permita al particular hacerse una idea general del funcionamiento del algoritmo que adoptará las decisiones automatizadas. Esta información debe huir de explicaciones excesivamente complejas o técnicas[1028]. Su contenido se ha de centrar en aquellos elementos que puedan ayudar al interesado a comprender ese tratamiento y cómo el mismo le puede afectar. En este sentido, tanto la doctrina como las autoridades de control ya han señalado toda una serie de elementos que han de formar parte de la información significativa.

[1026] La transparencia en estos ámbitos se muestra como una herramienta fundamental ya que en muchos casos estos tratamientos pasan desapercibidos para los particulares tal y como ocurre con los llamados precios personalizados. En: WONG,B: "Online personalised pricing as prohibited automated decision-making under Article 22 GDPR: a sceptical view", op.cit., págs.203 y ss.

[1027] Este principio se reconoce en la mayoría de textos y directrices que se han desarrollado en los últimos años sobre regulación de la inteligencia artificial. Entre otros: Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. Directrices éticas para una IA fiable.2019, págs. 22 y 35. También así lo contempla el artículo 52.1 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión. Texto aprobado por la Comisión Europea el 21 de abril de 2021. En España este principio se reconoce en el artículos 18.6.b) y 25.2.b) de la Carta de Derechos Digitales. Texto no vinculante aprobado en julio de 2021. Disponible en:
https://www.lamoncloa.gob.es/presidente/actividades/Documents/2021/140721-Carta_Derechos_Digitales_RedEs.pdf

[1028] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, págs. 27 y 28.

Tomando como referencia los señalados por la AEPD, estos elementos son: i) el detalle de los datos empleados para la toma de decisión, más allá de la categoría, ii) la importancia o el peso relativo que cada uno de ellos tiene en la toma de la decisión[1029], iii) la calidad de los datos de entrenamiento y el tipo de patrones y correlaciones utilizadas[1030], iv) los perfilados realizados y sus implicaciones, v) los valores de precisión o error según la métrica adecuada para medir la bondad de la inferencia[1031], vi) la existencia o no de supervisión humana cualificada y, vii) la referencia a auditorías, así como la certificación o certificaciones realizadas sobre el sistema de IA[1032].

Nótese que toda la información señalada anteriormente habrá sido documentada y registrada previamente por el responsable del tratamiento fruto de las distintas obligaciones de responsabilidad activa, así como las que se deriven de la PRAI previamente comentadas[1033]. Por tanto, entendemos que difícilmente podrá alegarse por parte del responsable que tal información no puede facilitarla por no disponer de ella o disponiendo de ella, mostrarla le supone esfuerzos desproporcionados. Además, en nuestra opinión, la información sobre la lógica del tratamiento comprende un plus a la hora de informar del tipo de datos que engloba el tratamiento. Así, no sólo se ha de informar de los datos personales que se utilizarán sino también de los no personales. Estos últimos suelen tener cierto peso en algunas decisiones automatizadas que adoptan los algoritmos. Piénsese por ejemplo en los datos meteorológicos o los datos de tráfico.

[1029] Por ejemplo, en los algoritmos de regresión lineal simple es fácilmente detectable el peso de las características utilizadas para predecir la variable de salida. En: MOLNAR,CH*: Interpretable Machine Learning. A Guide for Making Black Box Models Explainable*. Libro disponible on line, apartado 4.1.1. Texto disponible: https://christophm.github.io/interpretable-ml-book/limo.html#interpretation

[1030] Autoridad de protección de datos Noruega. The Norwegian Data Protection Authority. *Artificial intelligence and privacy*. 2018, pág.21. También en: Red iberoamericana de protección de datos personales. *Orientaciones Específicas para el Cumplimiento de los Principios y Derechos que Rigen la Protección de los Datos Personales en los Proyectos de Inteligencia Artificial.*2019, pág.18. A su vez, Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. Directrices éticas para una IA fiable.2019, pág.37.

[1031] Se ha se ha de informar sobre: las métricas utilizadas para evaluar la precisión, la eficiencia, la equidad y la capacidad del sistema para lograr los objetivos del gobierno, incluyendo específicamente cualquier esfuerzo realizado para probar el impacto potencial del sistema de toma de decisiones automatizado en los grupos marginados y vulnerables. En: International Human Rights Program (Faculty of Law, University of Toronto) and the Citizen Lab. *Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System*.2018, pág.63.

[1032] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.24.

[1033] Recuérdese que esta propuesta prevé que entre la documentación técnica que se ha de elaborar ha de constar entre otros elementos: la lógica general del sistema de IA y de los algoritmos, las opciones clave de diseño, incluidos los fundamentos y las suposiciones realizadas, las principales opciones de clasificación, la descripción de la arquitectura del sistema explicando cómo los componentes de software se basan o se alimentan entre sí y se integran en el procesamiento global, etc. Véase el anexo IV en relación con los artículos 11 y 18 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

De manera que, se hace necesario que se informe de estos para que exista una autentica comprensión de la lógica del tratamiento. Finalmente, el RGPD también obliga a los responsables a informar sobre las *consecuencias previstas del tratamiento*[1034]. Al igual que ocurría cuando hacíamos referencia a las consecuencias de la elaboración de perfiles, el responsable puede informar de este elemento aludiendo a los ejemplos más comunes de cómo este sistema responderá. El objetivo es que el interesado no se vea sorprendido posteriormente por la decisión que adopta el sistema ya que en parte podrá prever los posibles resultados de forma orientativa, sean o no perjudiciales.

Información que se ha de facilitar a los interesados cuando se llevan a cabo los tratamientos del artículo 22 del RGPD.	
Tratamiento	-Se ha de mencionar expresamente la toma de decisiones plenamente automatizadas relevantes.
Información significativa sobre la lógica del tratamiento	-Detalle de los datos empleados para la toma de decisión, más allá de la categoría. -La importancia o el peso relativo de los datos. -Calidad de los datos de entrenamiento y el tipo de patrones y correlaciones utilizadas. - Perfilados realizados. -Valores de precisión o error según la métrica adecuada para medir la bondad de la inferencia. - Existencia o no de supervisión humana cualificada. -Referencia a las auditorías, así como la certificación o certificaciones realizadas sobre el sistema de IA. -Datos no personales que se utilicen para el perfilado o la toma de decisiones.
Consecuencias previstas	-Ejemplos habituales del funcionamiento del sistema algorítmico en el contexto donde adopta decisiones.

B) La necesidad de complementar los deberes de información de la normativa de protección de datos con otras normas del sector privado

Uno de los objetivos que tiene la transparencia es actuar como mecanismo corrector de los desequilibrios que existen entre distintos agentes. La Unión Europea, en su afán por potenciar el mercado único digital, lleva ya varios años implementando todo

[1034] En el ámbito público la doctrina ya lleva tiempo apostando bastante tiempo por la necesidad de que las administraciones en estos contextos ofrezcan la suficiente información a los ciudadanos. En: VALERO TORRIJOS, J: "La necesaria reformulación de las garantías jurídica ante la innovación tecnológica en la Administración", *Blog de derecho, tecnología y modernización administrativa*, 01/06/2016. Disponible en: http://modernadministracion.blogspot.com/2016/06/la-necesaria-reformulacion-de-las.html

un marco legislativo que potencie la seguridad jurídica y la confianza en las actividades que se generan en los entornos en línea. Para lograr ese objetivo, es imprescindible que los usuarios de estos servicios confíen en estos. Ello en parte se logra a través de mecanismos adecuados de transparencia relativos al funcionamiento de dichos servicios, los cuales, en parte se han automatizado. Como ahora veremos, muchas de estas obligaciones reconocidas por la legislación europea se complementan o se superponen con los deberes de información previstos en la normativa de protección de datos previamente comentados. Es por ello que consideramos necesario su análisis y en su caso una propuesta de cómo encajar los distintos cuerpos legislativos para evitar un solapamiento de estas normativas que perjudica tanto a los responsables del tratamiento a la hora de cumplir con dichas obligaciones de transparencia, como a los titulares de los datos que se ven sobrecargados por la información que se les ofrece. Pues bien, esencialmente, las obligaciones de transparencia se dividen en dos grandes grupos, por un lado, aquéllas que han de fijarse en los términos y condiciones de uso de estos servicios digitales y por otro, aquellos deberes informativos que se han de facilitar durante el periodo de interacción del sistema algorítmico con los usuarios de la plataforma o el servicio.

En *primer lugar*, encontramos una serie de preceptos que obligan a determinadas organizaciones a establecer en sus términos y condiciones toda una serie de reglas de transparencia relacionadas con el uso de sistemas automatizados. Así, el Reglamento Europeo sobre la lucha contra la difusión de contenidos terroristas en línea obliga a los prestadores de servicios de alojamiento de datos a informar sobre el uso de herramientas automatizadas cuando las mismas se pretendan utilizar para la retirada de contenido considerado terrorista[1035]. A su vez, la Propuesta de Reglamento sobre un Mercado Único de Servicios Digitales impone a los prestadores de servicios intermediarios a informar en sus condiciones sobre los medios algorítmicos que utilizarán para moderar el contenido que comunica en sus plataformas. Además, este mismo texto obliga a las plataformas en línea de gran tamaño a que, cuando utilicen sistemas de recomendación de contenido, establezcan de manera clara y sencilla los parámetros principales

[1035] Artículo 7.1 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.

utilizados en sus sistemas de recomendación[1036]. En términos parecidos se ha estructurado el Reglamento contra los abusos sexuales de menores en línea, esta norma obliga a las plataformas a informar de forma clara y destacada a los usuarios de las mismas sobre el uso por parte de estas de sistemas automatizados con el fin de detectar y retirar contenido pedófilo[1037]. En todos estos supuestos es muy probable que dichos sistemas automatizados traten datos personales y por tanto, la normativa de protección de datos también resulte aplicable. Para complementar ambas legislaciones, lo ideal es que tanto las políticas de privacidad como los términos y condiciones de uso existan enlaces directos que redirijan a cada uno de los apartados referidos a la información cuando se utilizan medios automatizados. Así, por ejemplo, en los términos y condiciones de uso, cuando se haga mención a los sistemas automatizados, se puede implementar un enlace que redirija concretamente al lugar del aviso o política de privacidad donde específicamente se define ese tratamiento de datos y en el cual se hace mención explícita a la lógica del tratamiento y las consecuencias previstas del mismo. Otra opción es que se diseñe un apartado específico donde se informe sobre los distintos usos de los sistemas automatizados en el cual los derechos de información de la normativa de protección de datos y el resto de deberes de transparencia ya comentados aparezcan incorporados en un único lugar.

En *segundo lugar*, las obligaciones de transparencia sobre el uso de medios automatizados también se prevén durante la interacción del usuario con la plataforma o servicio ofrecido. Así, la Directiva relativa a la modernización de las normas de protección de los consumidores de la Unión obliga a los comerciantes que ofrezcan productos en línea con precios personalizados a informar claramente de ello[1038]. Esta norma potencia la transparencia de los tratamientos de precios personalizados basados en datos personales al obligar al responsable del tratamiento a informar claramente de

[1036] Artículo 12.1 sobre la moderación del contenido y artículo 29 sobre los sistemas de recomendación. Véase la Propuesta DEL PARLAMENTO EUROPEO Y DEL CONSEJO sobre un Mercado Único de Servicios Digitales (Ley de Servicios Digitales) y por la que se modifica la Directiva 2000/31/CE.

[1037] Artículo 3.1.g), v) del REGLAMENTO (UE) 2021/1232 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 14 de julio de 2021 por el que se establece una excepción temporal a determinadas disposiciones de la Directiva 2002/58/CE en lo que respecta al uso de tecnologías por proveedores de servicios de comunicaciones interpersonales independientes de la numeración para el tratamiento de datos personales y de otro tipo con fines de lucha contra los abusos sexuales de menores en línea.

[1038] Considerando 45 y artículo 6 de la DIRECTIVA (UE) 2019/2161 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 27 de noviembre de 2019 por la que se modifica la Directiva 93/13/CEE del Consejo y las Directivas 98/6/CE, 2005/29/CE y 2011/83/UE del Parlamento Europeo y del Consejo, en lo que atañe a la mejora de la aplicación y la modernización de las normas de protección de los consumidores de la Unión.

que dicho precio es personalizado[1039], algo que a priori la normativa de protección de datos no contempla, ya que esta última sobre todo se centra en la información previa al tratamiento de datos, es decir, antes de que se muestre el precio personalizado. Ahora bien, junto a la indicación sobre la cual se señala que dicho precio es personalizado, cuya obligación deriva de la Directiva de consumidores, sería recomendable que se facilitara un enlace que redirigiera al particular al sitio web donde se explica la lógica de dicho tratamiento y las consecuencias jurídicas del mismo, así como en su caso, una explicación personalizada de dicho precio en virtud del Considerando 71 y artículo 22 del RGPD.

Por otro lado, la Propuesta de Reglamento sobre un Mercado Único de Servicios Digitales previamente analizada obliga a las plataformas en línea a que cuando presenten publicidad personalizada en sus interfaces en línea, estas han de señalar que el contenido mostrado es un anuncio publicitario y además han de aportar información significativa acerca de los parámetros que se utilizan para determinar el destinatario al que se presenta el anuncio publicitario[1040]. Esta normativa aumenta las exigencias de transparencia con relación a la publicidad personalizada prevista en el RGPD. Así, como regla general, la publicidad personalizada no se engloba dentro de los tratamientos que aparecen definidos en el artículo 22 por no considerarse que los mismos generan efectos relevantes[1041]. Por tanto, a priori, los responsables no tendrían la obligación de informar sobre la lógica del tratamiento y las consecuencias previstas. Sin embargo, fruto de ese deber que se impone a estas plataformas para esa concreta finalidad del tratamiento en la mentada propuesta regulatoria, el responsable también estará obligado a proporcionar esa información significativa, la cual, sería recomendable que por defecto ya apareciera en las políticas de privacidad de aquellas plataformas a las que va dirigida esta obligación de transparencia.

[1039] Es lo que se conoce por la doctrina como perfiles banderas, los cuales, aparecen siempre que una decisión se basa en un perfil. En la prestación de servicios en línea, por ejemplo, cuando a alguien se le deniega una solicitud, la página web podría mostrar automáticamente un icono de perfil que podría enlazar con información sobre del perfil. En: JAAP KOOPS,B : "Some Reflections on Profiling, Power Shifts and Protection Paradigms". En: HILDEBRANDT, M & GUTWIRTH,S (eds): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. op.cit., pág.336.

[1040] Artículo 24 de la Propuesta DEL PARLAMENTO EUROPEO Y DEL CONSEJO sobre un Mercado Único de Servicios Digitales (Ley de Servicios Digitales) y por la que se modifica la Directiva 2000/31/CE,

[1041] Tal y como ha señalado el GT29, la publicidad personalizada normalmente no quedará normalmente englobada en los tratamientos contemplados en el artículo 22 del RGPD. Véase: Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, op.cit., pág.24. Sobre el carácter o no relevante de la publicidad personalizada véase el Capítulo II, apartado I, punto 1, epígrafe A, sub epígrafe a.2) de esta tesis.

3. Límites a los deberes de información en el uso de sistemas de toma de decisiones automatizadas

La transparencia como principio durante el ciclo de vida de los sistemas de toma de decisiones automatizadas se muestra como un requisito esencial y a la vez crítico. Ello es así porque existen diversas limitaciones amparadas en otros bienes e interés que restringen el desarrollo integral y pleno de este mecanismo de información. Así, el considerando 4 del RGPD establece que el derecho a la protección de datos personales no es un derecho absoluto, este, ha de lidiar con otros derechos e intereses jurídicos legítimos cuando entran en juego. Es turno de analizar algunos de esos bienes jurídicos sobre los que el responsable del tratamiento podría escudarse a la hora de conceder más o menos información.

A) Dificultad para entender el algoritmo

Un importante número de algoritmos sobre los que se sustentan los sistemas de toma de decisiones automatizados se caracterizan por ser opacos o difícilmente comprensibles tanto en su funcionamiento como a la hora de adoptar decisiones. Esa falta de interpretabilidad no sólo se manifiesta a nivel de un usuario medio sino que determinados algoritmos, sobre todo aquellos que utilizan para su entrenamiento técnicas de aprendizaje profundo[1042], son incluso inentendibles para los propios diseñadores que han participado en su elaboración fruto de la complejidad de los mismo y las interacciones que estos sufren una vez se ponen en funcionamiento[1043].

En este sentido, el artículo 14.5.b) prevé que, cuando los datos no se hayan obtenido del interesado, el responsable podrá exceptuar los deberes de información expresados en el apartado primero del artículo 14 cuando la comunicación de esa información *resulte imposible o exija un esfuerzo desproporcionado*. Es por ello que a

[1042] Sobre el aprendizaje profundo o *deep learning* véase el Capítulo I, apartado I, punto 1, epígrafe B, sub epígrafe b.3 de esta tesis.

[1043] Se habla por un lado de la opacidad analfabeta, donde un sistema solo es comprensible para aquellos con las habilidades técnicas para escribir y comprender el código. Por otro lado encontramos la opacidad intrínseca donde el complejo proceso de toma de decisiones de un sistema es difícil de explicar o incluso no es entendible para ningún humano. En: N COFONE,I: "Algorithmic Discrimination Is an Information Problem". *Hastings Law Journal*, Vol. 70:1389, 2019, págs. 1438 y 1439. Texto disponible en: https://repository.uchastings.edu/hastings_law_journal/vol70/iss6/1/ . En el mismo sentido: SELBST,A & BAROCAS, S: "The Intuitive Appeal of Explainable Machines", op.cit., pág.1094.

priori, y debido a ese déficit de comprensión del algoritmo, un responsable podría justificar la no concesión de información del tratamiento amparándose en el mencionado precepto. Pues bien, acudir a este precepto resulta en nuestra opinión cuanto menos complejo por varias razones. En *primer lugar* porque el ámbito de aplicación de la mentada excepción es muy limitado, así, este queda circunscrito a la información que se especifica en el apartado primero del artículo 14. De esta manera, las obligaciones de transparencia referidas a la información sobre la lógica y las consecuencias previstas de los tratamientos descritos en el artículo 22 del RGPD no quedan afectadas por este precepto. Pero es que además, aquella información que sí es posible que quede afectada por esa excepción a los deberes de información también resultará complicada. De esta manera, a través de esta excepción el responsable por ejemplo no tendría la obligación de informar sobre los datos que se utilizarán para la toma de decisiones o la finalidad que se pretende con la elaboración de los perfiles, así como las consecuencias que se derivan del perfilado. Sin embargo, no creemos que facilitar esa información resulte imposible o exija un esfuerzo desproporcionado cuando se utilicen algoritmos, incluso en aquellos casos en los que se acudan a técnicas de *deep learning*. Así, tanto los datos necesarios que se utilizarán como input a la hora de la toma de decisiones como el objetivo que se pretende con el perfilado son elementos que el responsable del tratamiento conoce de antemano ya que son necesarios para que el sistema pueda operar adecuadamente a la hora de adoptar decisiones en un determinado entorno. En *segundo lugar*, como ya indicábamos en otro apartado de la tesis, el responsable del tratamiento, ya sea durante la elección del modelo en la fase del diseño, o a la hora de adquirir tal modelo, ha de optar por aquel que en parte mantenga un cierto equilibrio que, por un lado sea preciso, pero a la vez, también cumpla con las exigencias en materia de protección de datos, concretamente y por lo que ahora nos interesa, las exigencias de transparencia. El factor de interpretabilidad en virtud del principio de privacidad desde el diseño ha de estar presente desde el momento que el responsable apuesta por el desarrollo del sistema o lo adquiere (Considerando 61 RGPD). Por tanto, en nuestra opinión, un responsable no podría justificar el déficit de transparencia alegando que el desarrollo de modelos altamente interpretables o explicables o el despliegue de técnicas que faciliten tal explicabilidad en sistemas automatizados opacos le supone un esfuerzo desproporcionado.

Por otro lado, el Considerando 61 del RGPD establece que *cuando el origen de los datos personales no pueda facilitarse al interesado por haberse utilizado varias fuentes, debe facilitarse información general*. Como ha señalado el GT29, la exención al requisito de facilitar a los interesados información sobre el origen de sus datos personales se aplica únicamente cuando ello no resulte posible porque los distintos elementos que componen los datos personales relativos a un mismo interesado no se pueden atribuir a un origen en particular[1044]. En este sentido, en nuestra opinión, los datos inferidos son datos personales que como mínimo devienen de dos fuentes. Por un lado, de los propios datos personales que se ingresan en el algoritmo y, por otro, del propio algoritmo que los procesa y los genera. Ahora bien, conforme es más complejo el procesado automatizado de datos, más difícil puede resultar para el responsable indicar el origen de los datos, ya sea porque se utilizan muchas fuentes de datos o porque previo a la inferencia se procesan los datos con varios algoritmos. En estos últimos supuestos, la procedencia de los datos resulta cuanto menos difícil de indicar. Pues bien, una vez más pensamos que la mera complejidad del procesamiento no puede servir de excusa para limitar los deberes de información, ello es así porque si no se primaría a aquellos responsables que acuden a vías más enmarañadas para obtener inferencias de los datos. El principio de privacidad desde el diseño exige al responsable el desarrollo de mecanismos que documenten el origen de los datos independientemente de la complejidad del sistema algorítmico.

En definitiva, en principio, el responsable está obligado a informar de los tratamientos que lleve a cabo durante el desarrollo y despliegue de los sistemas. Pese a ello, si el responsable pretende acudir a estas excepciones aduciendo que facilitar la información supondría un esfuerzo desproporcionado o imposible. Este debe sopesar por un lado el esfuerzo que implica para él facilitar la información al interesado y por otro, los efectos y las repercusiones que ello supone para el particular que no recibe esa información. El responsable del tratamiento debe documentar esta evaluación de conformidad con sus obligaciones de responsabilidad proactiva[1045]. En los supuestos en los que se limite la comunicación sobre el origen de los datos, el responsable aún estará obligado a facilitar información general sobre las fuentes de acuerdo a lo indicado en el Considerando 61 del RGPD.

[1044] Grupo de Trabajo del artículo 29. *Directrices sobre la transparencia en virtud del Reglamento (UE) 2016/*679.Adoptadas el 29 de noviembre de 2017 Revisadas por última vez y adoptadas el 11 de abril de 2018, apartado 60, pág.33.
[1045] Vid., apartado 64, pág.35.

Además de la complejidad técnica que deviene de origen, hay que indicar que ese déficit de entendimiento también puede estar presente en el personal perteneciente a las organizaciones que implanten en sus procesos internos de toma de decisiones estos sistemas automatizados. En estos supuestos, y tal y como indicábamos en el apartado referido a la formación del personal, estos trabajadores han de ostentar la suficiente capacitación para entender el sistema y por tanto facilitar los deberes de transparencia, tanto en la fase previa del tratamiento, como una vez el sistema haya adoptado la decisión[1046].

B) Secretos comerciales y propiedad intelectual

Las organizaciones utilizan diferentes medios para proteger algunos de los resultados que se derivan de sus actividades asociadas a la innovación e investigación. Para ello suelen acudir a los derechos de propiedad intelectual o a los secretos comerciales. Ambos bienes jurídicos resultan plenamente legítimos y protegidos por la legislación. A lo largo del ciclo de vida de los sistemas de toma de decisiones automatizadas existen y se crean determinados elementos que merecen precisamente protección de dichos interés jurídicos, elementos sobre los cuales la normativa de protección de datos impone deberes de transparencia respecto del responsable del tratamiento. Así, algunos de estos elementos que pueden entrar en conflicto y que pueden ser protegidos son; i) las bases de datos que se utilizan para la generación de modelos, ii) los propios modelos algorítmicos que se crean y, iii) los datos inferidos que se obtienen una vez se procesan los datos iniciales.

En *primer lugar*, por lo que se refiere a las bases de datos pueden estar protegidos por la Ley de Propiedad Intelectual[1047]. Así, esta norma define en su artículo 12.2 a las bases de datos como *las colecciones de obras, de datos, o de otros elementos independientes dispuestos de manera sistemática o metódica y accesibles individualmente por medios electrónicos o de otra forma*. Pues bien, esta legislación prevé dos mecanismos de protección para las mentadas bases de datos. Por un lado, la

[1046] Véase el Capítulo III, apartado XII de esta tesis.
[1047] 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.

protección de la estructura y disposición de la base de datos, lo que impediría la fabricación de una base de datos idéntica (artículo 12.1). Y por otro lado, la llamada protección de la base de datos por el derecho *sui generis* respecto de su contenido, impidiendo la extracción y la reutilización de ese contenido de la base de datos (Artículos 133 a 136)[1048]. De esta manera, estas previsiones indicadas en la legislación pueden ser alegadas por el responsable del tratamiento para limitar la información sobre estas bases de datos, así como limitar el acceso a las mismas por parte de los particulares.

En *segundo lugar*, el modelo o modelos algorítmicos que se utilizan para la toma de decisiones, así como los datos inferidos que se obtienen de dichos modelos también pueden recibir protección de la normativa relativa a los secretos comerciales o empresariales Concretamente, tanto la Directiva 2016/943 relativa a la protección de los secretos comerciales como la Ley 1/2019 de Secretos Empresariales protegen toda aquella información que presente los siguientes requisitos, estos son; i) que sea secreta, ii) que tenga un valor comercial o empresarial y, iii) que dicha información haya sido objeto de medidas razonables, en las circunstancias del caso, para mantenerla secreta, tomadas por la persona que legítimamente ejerza su control[1049]. Esta definición es tan genérica que en principio tanto los modelos algorítmicos como los datos inferidos resultantes de los mismos podrán perfectamente encuadrarse en la misma, por tanto, estos pueden recibir la protección que le brinda ambos textos jurídicos[1050]. La fricción por tanto entre la transparencia algorítmica y los secretos comerciales es y será cada vez

[1048] Sobre las implicaciones legales de las bases de datos véase entre otros: ORTEGA GIMÉNEZ, A: *Las aplicaciones del Big Data en el ámbito asegurador y el tratamiento legal de sus datos una perspectiva desde el derecho internacional privado*, op.cit., págs.60 y ss. También véase: PLAZA PENADÉS,J; PEDREÑO,A; MORENO,L: *Big Data e Inteligencia Artificial. Una visión económica y legal de estas herramientas disruptivas*. Ed. Fundació Parc Científic Universitat de València. Valencia, 2018, págs. 36 y 37.

[1049] Artículo 2.1 de 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. Esta Directiva una serie de mecanismos que favorecen la protección de aquellos que son titulares de dichos secretos comerciales, entre ellos, se trata de impedir la obtención, la utilización o la revelación ilícita de tal información. En el mismo sentido véase el artículo 1.1 de la Ley 1/2019, de 20 de febrero, de Secretos Empresariales.

[1050] Así lo ha indicado de forma rotunda entre otros: ORTEGA GIMÉNEZ, A: *Las aplicaciones del Big Data en el ámbito asegurador y el tratamiento legal de sus datos una perspectiva desde el derecho internacional privado*, op.cit., pág. 59. También en: WACHTER, S; MITTELSTADT, B: "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI". op.cit.,págs. 116 y ss. Por último, en el mismo sentido: MCGREGOR, L; MURRAY, D.; NG, V: "International human rights law as a framework for algorithmic accountability", op.cit., págs.322 y 323.

más patente. Precisamente, hasta la fecha, las Administraciones se han excusado en la normativa de derechos de propiedad intelectual y secretos comerciales para no facilitar parte de la información solicitada a través del derecho de acceso a la información reconocido por las normas de transparencia[1051]. Es por ello que la doctrina ya ha propugnado la necesidad de que los sistemas algorítmicos que sean utilizados por las Administraciones Públicas se pongan a disposición de la ciudadanía[1052]. De manera que los organismos públicos asuman el coste de adquisición de dichos programas y lo hagan públicos. En este mismo sentido, se ha pronunciado el Consejo de Estado Italiano a través de una resolución que analizaba la legitimidad del uso de un algoritmo por parte de una Administración para contratar personal a través de esta herramienta. Así, este tribunal ha señalado la necesidad de poner en conocimiento todos los aspectos del algoritmo que utiliza un poder público, concretamente, no se podrá invocar la confidencialidad de dicho sistema automatizado relacionado con los derechos de la propiedad intelectual de los fabricantes ya que, al poner estas herramientas al servicio de la Administración, se están aceptando las consecuencias relativas a la transparencia presente en estos contextos[1053]. En España, hasta la fecha, las solicitudes de acceso a la

[1051] El derecho de acceso a la información pública está reconocido en el Artículo 105 b) de la Constitución Española y en el Artículo 12 de la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno. A su vez, existen otra serie de leyes autonómica que también reconocen este derecho.

[1052] BOIX PALOP,A :"Los algoritmos son reglamentos: La necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones". op.cit.,.pág.254 y ss. También véase: GALETTA, D,U; GUSTAVO CORVALÁN,J: "Intelligenza Artificiale per una Pubblica Amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto". *Federalismi.it*, n 3, 2019, pág.21. Texto disponible en: https://www.federalismi.it/ApplOpenFilePDF.cfm?artid=38014&dpath=document&dfile=04022019214355.pdf&content=Intelligenza%2BArtificiale%2Bper%2Buna%2BPubblica%2BAmministrazione%2B4%2E0%3F%2B%2D%2Bstato%2B%2D%2Bdottrina%2B%2D%2B

En el mismo sentido se ha pronunciado la doctrina con relación a los algoritmos que puedan utilizarse en el proceso penal a la hora de acordar unas u otras medidas cautelares. En: SIMÓN CASTELLANO, P; MAGRO SERVET,V: *Justicia cautelar e inteligencia artificial: la alternativa a los atávicos heurísticos judiciales*. Ed. Bosch, Madrid, 2021, págs. 167 a 169. También en: FENOLL NIEVA,J: *Inteligencia artificial y proceso judicial*. Ed. Marcial Pons, Madrid, 2018, págs 140 a 143.

[1053]Sentencia Consejo de Estado italiano de 13 diciembre 2019, nº. 8472. FJº 13.1. Disponible en: https://www.giustizia-amministrativa.it/provvedimenti-cds

En este mismo sentido se ha pronunciado el Tribunal Administrativo Regional de Lazio. Sentencia del Tribunal Administrativo Regional de Lazio de 22 de marzo de 2017, nº3769. FJº 2. Disponible en: https://www.giustizia-amministrativa.it/portale/pages/istituzionale/visualizza/?nodeRef=&schema=tar_rm&nrg=201611419&nomeFile=201703769_01.html&subDir=Provvedimenti

En contra, en EEUU, el Tribunal Supremo del Estado de Wisconsin consideró adecuado el uso de algoritmos de valoración del riesgo para ayudar a los jueces a imponer una u otra condena. En este caso el Tribunal no permitió a los particulares el acceso al código fuente de esa aplicación informática. Así, según este órgano judicial, el derecho de los acusados al debido proceso no fue vulnerado por el hecho de que los afectados no pudieron acceder a una explicación adecuada sobre los resultados dictados por el algoritmo. Así, la exactitud del sistema automatizado y la capacidad del juez para comprender el posible

información a través de los mecanismos que ofrecen las respectivas leyes de trasparencia han sido dispares a la hora de conceder información relacionada con el algoritmo[1054]. Así, aunque existe unanimidad en considerar un algoritmo como información pública, el Consejo de Transparencia y Buen Gobierno (en adelante CTBG) no siempre ha concedido a los ciudadanos el acceso a los mismos ya que la información que contienen está protegida por los derechos de propiedad intelectual[1055]. No obstante, recientemente, el CTBG ha obligado a la Tesorería General de la Seguridad Social a facilitar la aplicación informática que se utiliza para el cálculo de las pensiones. Para este órgano, los algoritmos están adquiriendo una relevancia decisiva en la toma de decisiones total o parcialmente automatizadas en el seno de las Administraciones Públicas. Es por ello que la sociedad cada vez con mayor frecuencia solicite conocer el funcionamiento de estos sistemas con el objetivo de fiscalizarlos y reducir los sesgos discriminatorios que pueden generar. Para el CTBG, mientras no se instauren otros mecanismos que permitan controlar el funcionamiento adecuado de estos algoritmos, el único recurso eficaz es el acceso al sistema propiamente dicho[1056]. A su vez, en el ámbito privado, recientemente se ha establecido una nueva obligación en el estatuto de los trabajadores por la cual, la empresa está obligada a informar al comité de empresa sobre los parámetros, reglas e instrucciones en los que se basan los algoritmos que afectan a la toma de decisiones que pueden incidir en las condiciones de trabajo, el

---

mal funcionamiento eran suficientes para asegurar los derechos de los acusados. Sentencia 13 de julio de 2016: State v. Loomis, 881, N.W.2d 749, 7532. Resolución disponible en: https://www.scotusblog.com/wp-content/uploads/2017/02/16-6387-op-bel-wis.pdf
Esta sentencia ha sido analizada entre otros por: DE MIGUEL BERIAIN, I.; PÉREZ ESTRADA, M.,J: "La inteligencia artificial en el proceso penal español: un análisis de su admisibilidad sobre la base de los derechos fundamentales implicados". *Revista De Derecho De La UNED (RDUNED)*, (25), 2019, págs. 531 a 561. Texto disponible en: https://doi.org/10.5944/rduned.25.2019.27013
También en: ROMEO CASABONA,C,M: "Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad". *Revista penal*, nº42, 2018, págs.48 y ss.

[1054] El artículo 14 de la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno prevé determinados límites al ejercicio del derecho de acceso. Entre los límites encontramos los secretos comerciales y a la propiedad intelectual.

[1055] El CTBG consideró que el código fuente de la aplicación algorítmica que se utiliza para conceder las ayudas referidas al bono social de la electricidad no se puede entregar ya que el mismo está protegido por los derechos de propiedad intelectual. Consejo de Transparencia y Buen Gobierno. Resolución 701/2018. Febrero de 2019. FJº5, págs. 8 y 9. Texto disponible en: https://www.consejodetransparencia.es/ct_Home/Actividad/Resoluciones/resoluciones-AGE/AGE-2019/02.html

[1056] Consejo de Transparencia y Buen Gobierno. Resolución 058/2021. Mayo de 2021.FJº 5, pág.7.Texto disponible en: https://www.consejodetransparencia.es/ct_Home/Actividad/Resoluciones/resoluciones-AGE/AGE-2021/05.html
En otro supuesto, de forma pionera en España, la Comissió de Garantia del Dret d'Accés a la Informació Pública de Cataluña concedió el acceso a un particular que solicitaba información sobre el algoritmo que se utiliza para seleccionar a los profesores que evaluaran las pruebas de acceso a la universidad. Resolución 200/2017, de 21 de junio. Texto disponible en: http://www.gaip.cat/es/detall/normativa/2017-0200

acceso y mantenimiento del empleo, incluida la elaboración de perfiles[1057]. Esta nueva incorporación[1058], obliga al responsable a informar sobre dichos tratamientos de datos personales y refuerza y complementa los deberes de transparencia previamente explicados. Volviendo al RGPD, en nuestra opinión, en ningún momento se está obligando al responsable del tratamiento a facilitar una transparencia total sobre los algoritmos, sino más bien, se impone el deber de proporcionar una información que ayude a los titulares de los datos a comprender y entender la lógica que hay detrás de las decisiones y los perfilados más relevantes que les afectan y las consecuencias previstas de las mismas. En este sentido, el considerando 35 de la Directiva de secretos comerciales indica la necesidad que tiene el poseedor de un secreto comercial de respetar la normativa de protección de datos respecto de los titulares de esos datos. Si bien no se especifica cómo se ha de conjugar ese equilibrio entre los secretos comerciales y los deberes del responsable en relación con la protección de datos[1059]. Para nosotros, cualquier limitación a los deberes de trasparencia nunca podrá suponer una restricción total sobre dicha información. No podemos olvidar que el legislador europeo a través de las disposiciones del RGPD ha aumentado las exigencias de transparencia que hasta la fecha preveía la Directiva 95/46, haciendo especial hincapié cuando se elaboran perfiles y se adoptan decisiones plenamente automatizadas relevantes. Cualquier limitación por tanto al principio de transparencia en materia de

[1057] Véase el nuevo artículo 64.4.d) del 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. Precepto incorporado por el artículo único del Real Decreto-ley 9/2021, de 11 de mayo, por el que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de plataformas digitales.

[1058] Curiosamente, la Directiva de secretos comerciales en su considerando 18 y artículo 3.1.c) habilita a los trabajadores y representantes de los trabajadores a obtener lícitamente un secreto comercial del empresario cuando los primeros ejerzan sus derechos en virtud de la normativa laboral. 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.

[1059] Tal y como establece el Considerando 35, *es importante que se respete el derecho al respeto de la vida privada y familiar y a la protección de los datos personales de toda persona cuyos datos personales puedan ser tratados por el poseedor de un secreto comercial cuando se tomen medidas para la protección del secreto comercial, o de toda persona implicada en un proceso judicial relativo a la obtención, utilización o revelación ilícitas de secretos comerciales, con arreglo a la presente Directiva, y cuyos datos personales sean objeto de tratamiento*. (…) *la presente Directiva no debe afectar a los derechos y obligaciones previstos en la Directiva 95/46/CE, en particular los derechos del interesado de acceder a aquellos de sus datos personales que sean objeto de tratamiento y de obtener la rectificación, supresión o bloqueo de los datos debido a su carácter incompleto o inexacto y, en su caso, la obligación de tratar los datos de carácter sensible de conformidad con el artículo 8, apartado 5, de esa misma Directiva*. 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.

protección de datos en relación con los secretos comerciales obliga al responsable una vez más a sopesar por una lado la afectación que tiene sobre los derechos y libertades de los titulares de los datos ese déficit de información y por otro, los efectos perjudiciales a la hora de divulgar más o menos información con relación a dichos secretos comerciales.

C) Evitar el juego del algoritmo

Los responsables del tratamiento también pueden mostrar reticencias a la hora de suministrar información sobre el algoritmo para evitar que aquellas personas sobre las que se tomarán las decisiones o se elaborarán los perfiles acaben jugando con el sistema en su beneficio. Así, el artículo 14.5.b del RGPD permite la limitación de los deberes de transparencia en aquellos supuestos en los que la información que debe proporcionar el responsable en virtud del artículo 14.1 del RGPD imposibilite u obstaculice gravemente el logro de los objetivos del tratamiento. Es decir, si a través de la información que ha de suministrar el responsable, las personas pueden trucar el algoritmo a su favor y con ello truncar los objetivos que se pretende con la implantación del mencionado sistema, el responsable puede justificar la no concesión de información prevista por la normativa de protección de datos. En este sentido, el GT29 ha indicado que para que un responsable pueda acogerse a esta excepción, los responsables del tratamiento deben demostrar que facilitar la información establecida en el artículo 14.1 por sí misma invalidaría los objetivos del tratamiento[1060].

En términos generales, en nuestra opinión, siempre que un responsable pretenda ofrecer menos información de la prevista por la normativa de protección de datos por riesgo a que los particulares acaben jugando con el algoritmo exige un proceso de análisis de la situación a través de varios pasos.

En *primer lugar*, el responsable ha de valorar si los deberes de transparencia permiten a los particulares trucar el sistema cuando vayan a interactuar con el mismo. Como regla general, los titulares afectados por los algoritmos inevitablemente tratarán de jugar con el sistema[1061], ya sea alterando su conducta de forma real o ficticia o

[1060] Grupo de Trabajo del artículo 29. *Directrices sobre la transparencia en virtud del Reglamento (UE) 2016/*679.Adoptadas el 29 de noviembre de 2017 Revisadas por última vez y adoptadas el 11 de abril de 2018, apartado 65, pág.36.

[1061] Las audiencias inevitablemente alterarán su comportamiento para alterar el algoritmo y cambiar el impacto y en su caso alterar las decisiones. En: BURK,L,D: "Algorithmic Fair Use", op.cit.,pág.295.

aportando datos inexactos o incorrectos en la búsqueda de la obtención de la decisión deseada.

En segundo lugar, el responsable debe analizar los efectos tanto negativos como positivos que puede generar el hecho de que los particulares jueguen con el algoritmo y alteren su comportamiento gracias a los deberes de información que se proporcionan. Estos efectos deben valorar las implicaciones que se generan tanto para la organización que implanta el algoritmo como para los particulares que se ven sometidos a estos, así como para la sociedad en general. De esta manera, es posible que el uso de sistemas automatizados obre en el interés legítimo tanto de la persona sobre la que se elabora un determinado perfil como de la organización que utiliza el algoritmo de perfilado[1062]. Así, en términos generales, un cliente que contrata una póliza de seguro cuyo precio depende de los hábitos de conducción que despliegue durante un determinado periodo de tiempo estará interesado en conocer en términos generales los principales parámetros que más pueden influir en los precios de esa póliza[1063]. En este caso, jugar con el algoritmo significará que el particular tratará de conducir de la forma menos temeraria posible a fin de que el algoritmo le reduzca el precio. Ello favorecerá a la propia aseguradora ya que ajustará dichos precisos y además, también se verá favorecida la sociedad en general ya que el riesgo de que esa persona sufra un accidente disminuirá. La posibilidad de alterar la conducta de los particulares que se someten a estas decisiones con el objetivo de empujar a estos a objetivos legítimos resulta también justificada cuando las Administraciones Públicas persiguen intereses públicos en virtud de las competencias que les otorga el ordenamiento jurídico. La publicación del funcionamiento de los algoritmos con el objetivo de que los ciudadanos se adecúen a sus requerimientos es por tanto lícita.

Ahora bien, la alteración de la conducta por parte de los particulares con el fin de adaptarse a las exigencias de los algoritmos puede también provocar determinados efectos negativos tanto en la esfera de los propios particulares como en las propias organizaciones que lo implantan. Por lo que se refiere a los particulares, hay que valorar el efecto que sobre los mismos puede generar el hecho de estar continuamente

[1062] Consejo de Europa. *Recomendación CM/Rec(2010)13 del Comité de Ministros a los Estados miembros sobre la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal en el contexto de la creación de perfiles*, pág.2.

[1063] Por ejemplo, si realizas conductas que disminuyan el riesgo asegurado, el precio de tu póliza bajará de manera proporcional, y a la inversa si creas conductas que aumenten tal riesgo. En: ORTEGA GIMÉNEZ, A: *Las aplicaciones del Big Data en el ámbito asegurador y el tratamiento legal de sus datos una perspectiva desde el derecho internacional privado*, op.cit., pág.47.

contentando al sistema. Esto ocurre por ejemplo en los entornos virtuales como YouTube o Instagram donde aquellos que suben contenido acaban en parte siendo esclavos de las directrices que marca el algoritmo con el fin de que su contenido aparezca entre los más reproducidos[1064]. Algo parecido puede ocurrir con los sistemas basados en el Internet de las cosas. Así, un particular en el ámbito doméstico podría auto presionarse para evitar comportamientos no habituales con el fin de impedir la detección de lo que se podría considerar anómalo. Ello resulta muy invasivo para la vida privada y la intimidad de las personas y se debe controlar muy estrechamente[1065].

Por otro lado, las organizaciones legítimamente tratarán de evitar en muchos supuestos que se juegue con el algoritmo. Ello ocurre sobre todo cuando el objetivo que se persigue con la introducción del sistema quede anulado por el exceso de información que se proporcione sobre el mismo. Piénsese por ejemplo en aquellos ámbitos donde la finalidad del sistema sea detectar posibles actuaciones fraudulentas o la prevención de determinados ilícitos penales[1066]. Por ejemplo, la información sobre el funcionamiento del sistema VeriPol que utiliza la policía nacional para detectar posibles denuncias de robos falsas interesará sobre todo a aquellas personas que habitualmente denuncian fraudulentamente el robo de sus celulares[1067]. En estos contextos, los particulares que estarán más interesados en obtener la información sobre estos sistemas serán los que precisamente más le conviene conocerlo para evitar las consecuencias que se derivan del mismo.

[1064] Muchos usuarios que se dedican a subir contenido a plataformas denominan a los algoritmos que posicionan sus contenidos como el "monstruo que hay que domar". Para ello, existe todo un mercado de compra de suscriptores, me gustas o número de clics con el fin de que el algoritmo posicione esos contenidos en los lugares más visibles. Las plataformas conocen este tipo de negocios y en muchos casos se les ha acusado de ser cómplices de esta forma de actuar y de falsear la realidad ya que a estas últimas les interesa que exista un intercambio constante de contenidos que favorece el despliegue de anuncios publicitarios. Véase los documentales cinematográficos. "*100 Million Views*" de Itamar Rose o "*#followme*" de Nicolaas Veul.

[1065] Grupo del Artículo 29. *Dictamen 8/2014 sobre la evolución reciente de la Internet de los objetos*. Adoptado el 16 de septiembre de 2014, pág.9 in fine.

[1066] Cuando utilice decisiones asistidas por IA para identificar irregularidades o conductas indebidas (por ejemplo, detección de fraude), la necesidad de limitar la información que proporcione a las personas será más fuerte, especialmente en lo que respecta a la explicación lógica. Pero aún debe proporcionar tanta información sobre el razonamiento y la lógica como pueda. En: TODOLÍ SIGNES,A: "Retos legales del uso del *big data* en la selección de sujetos a investigar por la Inspección de Trabajo y de la Seguridad Social". *Revista Galega de Administración Pública*, núm. 59, 2020, pág.324.

[1067] Véase más información sobre el sistema VeriPol en:
http://www.interior.gob.es/ca/prensa/noticias/-/asset_publisher/GHU8Ap6ztgsg/content/id/9496864
Sobre los potenciales interesados en conocer el funcionamiento del sistema véase: BAMBAUER,J ; ZARSKY,T: "The Algorithm Game", *Notre Dame Law Review***,** volume 94, Issue 1,pág.29.

Efectos de la transparencia en relación con el juego del algoritmo		
Póliza de seguros en función de los hábitos de conducción	**Beneficiados**	Cliente: Se obtiene un precio inferior si lleva a cabo una buena conducta en el volante. Aseguradora: Se personalizan los riesgos. Sociedad en general: Se reduce el número de accidentes.
	Perjudicados	Cliente: se obtiene un precio superior si lleva a cabo una mala conducción del vehículo. Aseguradora: Se intenta engañar al sistema.
Sistema para detectar actividades de fraude fiscal	**Beneficiados**	Ciudadano infractor: Puede eludir la inspección. Ciudadano no infractor: Conocer el funcionamiento del sistema.
	Perjudicados	Interés general: Se reducen las inspecciones y la supervisión de conductas ilegales.

En *tercer lugar*, una información pormenorizada del sistema puede permitir que terceros de forma intencionada traten de atacarlo tecnológicamente a efectos de que se altere su comportamiento en ese contexto específico. Nos referimos por ejemplo a los llamados ataques contradictorios o a aquellas interferencias que realizan los usuarios a los sistemas en entornos sumamente complejos y adaptativos[1068]. El exceso de información en esos ámbitos puede anular el objetivo que se persigue con la introducción del sistema. Piénsese por ejemplo en aquellos ámbitos donde la finalidad del sistema sea detectar posibles actuaciones fraudulentas o la prevención de determinados ilícitos penales[1069]. Por ejemplo, la información sobre el funcionamiento del sistema VeriPol que utiliza la policía nacional para detectar posibles denuncias de robos falsas interesará sobre todo a aquellas personas que habitualmente denuncian fraudulentamente el robo de sus celulares[1070]. En estos contextos, los particulares que estarán más interesados en obtener la información sobre estos sistemas serán los que precisamente más le conviene conocerlo para evitar las consecuencias que se derivan del mismo. El sabotaje del sistema puede legitimar la opacidad parcial del mismo.

[1068] Véase el Capítulo III, apartado V, punto 1 de esta tesis.

[1069] Cuando utilice decisiones asistidas por IA para identificar irregularidades o conductas indebidas (por ejemplo, detección de fraude), la necesidad de limitar la información que proporcione a las personas será más fuerte, especialmente en lo que respecta a la explicación lógica. Pero aún debe proporcionar tanta información sobre el razonamiento y la lógica como pueda. En: TODOLÍ SIGNES,A: "Retos legales del uso del *big data* en la selección de sujetos a investigar por la Inspección de Trabajo y de la Seguridad Social". *Revista Galega de Administración Pública*, núm. 59, 2020, pág.324.

[1070] Véase más información sobre el sistema VeriPol en: http://www.interior.gob.es/ca/prensa/noticias/-/asset_publisher/GHU8Ap6ztgsg/content/id/9496864
Sobre los potenciales interesados en conocer el funcionamiento del sistema véase: BAMBAUER,J ; ZARSKY,T: "The Algorithm Game", *Notre Dame Law Review***,** volume 94, Issue 1,pág.29.

En *cuarto lugar* conviene recordar que entre los ataques que podían afectar a la confidencialidad de los datos presentes en los algoritmos se encontraban aquellos que permitían conocer si una persona había formado parte del conjunto de datos utilizado para entrenar el modelo[1071]. En nuestra opinión, este tipo de ataques pone en entredicho la conveniencia de publicar enteramente los algoritmos, sobre todo, cuando dichos modelos sean utilizados por el sector público. La plena trasparencia de los algoritmos por la que se propugna en este ámbito puede encontrar por tanto un límite legítimo, en este caso, el derecho a la protección de datos personales de las personas que formaron parte del entrenamiento de esos modelos[1072].

En definitiva, de lo dicho en los párrafos anteriores se desprende que el responsable puede quedar exonerado de determinados deberes de transparencia que impone el RGPD. Ello queda justificado por la existencia de otros bienes e intereses legítimos que entran en juego cuando se informa de dichos tratamientos de datos. No obstante, y a pesar de ese conflicto y la aplicación de estas excepciones, el responsable nunca podrá justificar la plena opacidad de sus actividades. Y es que, aunque la transparencia no sea una herramienta de protección completa en estos contextos, siempre será un primer paso útil para el particular teniendo en cuenta el oscurantismo que caracteriza este tipo de operaciones. Gracias a esta información los interesados pueden averiguar qué se está sucediendo con sus datos y a partir de ahí y con ese conocimiento previo, poder ejercitar otras facultades y exigencias de cumplimiento normativo que se derivan del derecho a la protección de datos cuando se utilizan sistemas de toma de decisiones automatizadas. Las organizaciones deberán justificar y acreditar cómo la transparencia sobre el algoritmo en esos contextos específicos permite trucar tecnológica y fraudulentamente el sistema.

[1071] Nos referimos a los ataques por inferencia. Véase el Capítulo III, apartado V, punto 1 de esta tesis.
[1072] El artículo 15 de la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno establece la ponderación entre el derecho a la protección de datos y el derecho de acceso a la información pública que se ha de realizar cuando se solicita información que contiene datos personales. En este sentido recordar que el código fuente ha sido considerado como información pública a efectos de la normativa de transparencia. Comisión de Garantía del derecho de acceso a la información pública de Cataluña. Resolución 200/2017, de 21 de junio.
Disponible en: http://www.gaip.cat/es/detall/normativa/2017-0200

D) Garantías mínimas a desplegar como contrapeso a la restricción legítima de información

Cualquier limitación a los deberes de información previstos por la normativa de protección de datos siempre ha de ser compensada con el desarrollo y despliegue de otras medidas de garantía adecuadas. Son diversas las actuaciones que ha de llevar a cabo el responsable para amortiguar ese déficit de transparencia que exige la normativa. Una vez más, resultará ideal el enfoque del riesgo al que venimos aludiendo en toda la tesis.

En *primer lugar*, todas y cada una de estas restricciones han de quedar registradas y documentadas. En dichos documentos se ha de hacer referencia a las razones que han llevado al responsable a limitar dicha información. En concreto; los motivos o esfuerzos imperiosos para no poder facilitar esa información, la imposibilidad para conocer el origen de los datos, las causas o preceptos sobre los que se ampara los derechos de propiedad intelectual o secretos comerciales y las repercusiones negativas que genera el conocimiento del algoritmo y por tanto, la posibilidad de que los particulares jueguen con él. Así, por ejemplo, el Tribual Constitucional alemán consideró que un sistema automatizado que se encarga de reconocer matrículas de coches con fines de prevención de delitos requería del responsable del tratamiento la necesidad de documentar de manera comprensible y verificable todas las actuaciones de dicho sistema ya que la información a los particulares sobre el despliegue de ese sistema era difícilmente aplicable dado el carácter encubierto de tal medida[1073].

En *segundo lugar*, en aquellos sistemas donde la transparencia de los mismos resulte más crítica porque es posible que esta acabe frustrando los objetivos de dicho tratamiento y genere importantes consecuencias negativas, por ejemplo, en los sistemas que se encargan de detectar el fraude fiscal. Estos responsables estarán obligados a desplegar otros mecanismos de rendición de cuentas que en parte compensen ese déficit de transparencia. Para ello, deberán acudir al uso de mecanismos de certificación o de auditorías[1074], los cuales, en estos casos siempre resultarán obligatorios. De esta manera,

[1073] BVerfG, Order of the First Senate of 18 December 2018 - 1 BvR 142/15, FJ 157. Texto disponible en: http://www.bverfg.de/e/rs20181218_1bvr014215en.html

[1074] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.34. Respecto a las auditorías, véase LLANEZA GONZÁLEZ,P: *Seguridad y responsabilidad en la Internet de las cosas.* Ed. Bosch. Wolters Kluwer, España, 2018, págs. 346 y 347.

las restricciones de transparencia a las que viene obligado el responsable por el contexto en el que irradia sus efectos el sistema algorítmico se deben compensar con otras medidas que aseguran un funcionamiento adecuado del mismo. La frecuencia de las revisiones y testeos tanto externos como internos del sistema también se acentuarán. El papel de las autoridades de control para controlar y supervisar estos modelos para estos supuestos será muy relevante.

En definitiva, cualquier limitación de la transparencia al público en general deberá estar debidamente justificada. Dicha restricción nunca debería llevar al secreto total del sistema, se han de desplegar vías de transparencia alternativas que permitan el control y supervisión de estos sistemas.

VII. EL PRINCIPIO DE LEALTAD Y EL DE PROHIBICIÓN DE DISCRIMINACIÓN ALGORÍTMICA

El artículo 8.2 de la Carta de derechos fundamentales de la Unión Europea señala que los datos personales serán tratados de modo leal. En términos muy similares, el RGPD en su artículo 5.1.a) reconoce la lealtad como un principio básico del tratamiento de datos personales. Este principio elemental obliga al responsable a tratar los datos de la manera que las personas razonablemente esperarían y no usarlos de forma que tenga efectos adversos injustificados para ellos. Se exige así una actitud honesta durante todas las fases que engloban el tratamiento de datos. Visto así, el principio de lealtad es tan general y amplio que pese a que el mismo ostenta su propia entidad, normalmente aparecerá vinculado a otros principios del tratamiento ya que este consigue irradiar sus efectos sobre estos últimos. Para un análisis adecuado primeramente haremos referencia a las interrelaciones existentes entre el principio de lealtad y el resto de principios en materia de protección de datos ya analizados y, posteriormente nos centraremos en la relación especial que dicho principio ostenta con el de exactitud y la prohibición de no discriminación algorítmica.

1. El principio de lealtad y su interrelación con otros principios en materia de protección de datos

En *primer lugar*, el principio de lealtad en relación con el principio de minimización de datos obliga al responsable del tratamiento a tratar los datos de forma adecuada y justificada. En el ámbito de los sistemas de toma de decisiones automatizadas cada uno de los datos que se utilizarán han de estar debidamente justificados. Si algún dato no es idóneo para el objetivo marcado, hay que entender que el responsable está tratando esos datos de una forma deshonesta[1075]. Por tanto, han de quedar registradas y documentadas las distintas hipótesis del proyecto que se pretende desarrollar, así como los análisis posteriores con relación a la eliminación de posibles relaciones espurias entre los datos personales[1076]. Dado que los particulares pueden considerar un tratamiento inesperado del uso de sus datos alternativos, los mismos han de quedar debidamente justificados. Todas estas medidas deberán quedar incorporadas al estudio de minimización de datos previamente explicado.

En *segundo lugar*, la afectación del principio de lealtad puede devenir desde el momento inicial que se recopilan los datos, esto es, cuando la base de legitimación utilizada no es correctamente obtenida en virtud del principio de licitud. Así, si una persona ha sido engañada o influenciada de forma inadecuada para que se recopilen sus datos, dicho tratamiento podrá ser considerado no leal[1077]. Por ejemplo, la obtención del consentimiento a través de los llamados *dark patterns* a los que previamente nos referíamos es un ejemplo clarísimo donde el responsable no actúa de forma honesta y a la hora de recopilar los datos personales[1078]. Igualmente, el principio de lealtad se verá afectado cuando el responsable lleve a cabo tratamientos inesperados de datos basados en el interés legítimo y tal interés no haya superado adecuadamente la prueba de ponderación necesaria prevista en el artículo 6.1.f) del RGPD[1079].

[1075] The Norwegian Data Protection Authority. *Artificial intelligence and privacy*. 2018, pág.16.
[1076] Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág.21.
[1077] Así lo ha señalado la Information Commisioner´s Office en: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/lawfulness-fairness-and-transparency/#fairness
[1078] Véase el Capítulo IV, apartado I, punto 1, epígrafe E) de esta tesis.
[1079] Véase el Capítulo IV, apartado I, punto 6 de esta tesis.

En *tercer lugar*, tal y como señala el GT29, la transparencia se trata de una expresión del principio de lealtad en relación con el tratamiento de los datos personales[1080]. Ello significa que para que un particular no se vea sorprendido por el uso de sus datos que realiza un responsable, este último ha de informar al interesado de la existencia de las operaciones y fines que engloba dicho tratamiento entre los que se ha de incluir la elaboración de perfiles y sus consecuencias, así como la lógica del tratamiento y las consecuencias previstas de las decisiones plenamente automatizadas descritas en el artículo 22 del RGPD[1081]. Ahora bien, el principio de lealtad obliga al responsable a tratar los datos de la manera que las personas razonablemente esperarían y no usarlos de forma que tenga efectos adversos injustificados para ellos informando debidamente de los mismos. Sin embargo, este principio no queda afectado cuando la decisión que adopte el algoritmo genere un efecto negativo en la persona. Si el sistema ha sido debidamente elaborado, se han justificado las variables y además se ha informado de ello al particular, la decisión más o menos perjudicial para el particular no es relevante ya que esta estará justificada pese a que la misma sea adversa.

Incidencia del principio de lealtad en el resto de principios generales del tratamiento	
Minimización de datos	Justificación de las variables Documentación de las hipótesis del proyecto
Licitud	Recopilación de datos mediante engaño o incumplimiento de las bases de legitimación
Transparencia	Déficit de información con relación a la elaboración de perfiles y toma de decisiones automatizadas

2. El principio de lealtad y su interrelación con el principio de exactitud. La prohibición de discriminación algorítmica

Por lo que se refiere a las relaciones con el principio de exactitud, decíamos que los responsables tenían la obligación de obtener sistemas precisos, evitando con ello que los sistemas automatizados adoptarán decisiones adversas no justificadas[1082]. Para

[1080] Grupo de Trabajo del artículo 29. *Directrices sobre la transparencia en virtud del Reglamento (UE) 2016/*679.Adoptadas el 29 de noviembre de 2017 Revisadas por última vez y adoptadas el 11 de abril de 2018, apartado 2, pág.5.
[1081] Véase los considerandos 39 y 60 y los artículos 13.2 y 14.2 del RGPD.
[1082] Véase el Capítulo IV, apartado IV de esta tesis.

lograr ese objetivo, junto al deber que tiene el responsable en virtud del principio de exactitud de utilizar bases de datos actualizadas y representativas de la realidad que se pretende modelar[1083]. El principio de lealtad obliga al responsable por un lado a manejar bases de datos que presenten el menor número de sesgos y por otro[1084], a utilizar modelos y procedimientos estadísticos adecuados y fiables que limiten la posibilidad de que los sistemas arrojen decisiones discriminatorias. En este sentido, el GT29 ha señalado que la elaboración de perfiles puede ser desleal y generar discriminación por ejemplo cuando se deniega a las personas el acceso a oportunidades de empleo, crédito o seguro, o dicha elaboración se dirige a los interesados con productos financieros demasiado arriesgados o costosos[1085]. De esta manera, a pesar de que el tratamiento sea lícito, si las decisiones que arroja el mismo son injustas por ser discriminatorias[1086], se estaría afectando al principio de lealtad y al principio de prohibición de discriminación algorítmica, ambos mencionados en el Considerando 71.2 del RGPD[1087]. Así, el Consejo de Estado Italiano ha entendido que este Considerando reconoce el principio fundamental de no discriminación algorítmica[1088], esto es, prohibición de

---

[1083] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.36.

[1084] Tal y como se ha indicado, los conjuntos de datos que utilizan los sistemas de IA (tanto con fines de formación como para su funcionamiento) pueden presentar sesgos históricos inadvertidos, lagunas o modelos de gestión incorrectos. En: Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Directrices éticas para una IA fiable*.2019, pág.23.

[1085] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, págs.11 y 24.

[1086] El principal punto de fricción entre los sistemas de toma de decisiones automatizadas y el principio de lealtad es la potencial discriminación que pueden sufrir las personas como consecuencia de un algoritmo que toma decisiones sesgadas. Es necesario que estos sistemas no generen un trato desigual hacia un grupo de personas en función del género, del color de la piel, de las creencias religiosas, de la orientación sexual, etc. En: Autoritat Catalana de Protecció de Dades. *Intel·ligència Artificial. Decisions Automatitzades a Catalunya*, 2020, pág. 126.

[1087] El considerando 71 párrafo segundo establece que: ***A fin de garantizar un tratamiento leal y transparente respecto del interesado****, teniendo en cuenta las circunstancias y contexto específicos en los que se tratan los datos personales,* ***el responsable del tratamiento debe*** *utilizar procedimientos matemáticos o estadísticos adecuados para la elaboración de perfiles, aplicar medidas técnicas y organizativas apropiadas para garantizar, en particular, que se corrigen los factores que introducen inexactitudes en los datos personales y se reduce al máximo el riesgo de error,* ***asegurar los datos personales de forma que se tengan en cuenta los posibles riesgos para los intereses y derechos del interesado y se impidan,*** *entre otras cosas,* ***efectos discriminatorios en las personas físicas por motivos de raza u origen étnico, opiniones políticas, religión o creencias, afiliación sindical, condición genética o estado de salud u orientación sexual, o que den lugar a medidas que produzcan tal efecto****. Las decisiones automatizadas y la elaboración de perfiles sobre la base de categorías particulares de datos personales únicamente deben permitirse en condiciones específicas*. (la negrita es nuestra)

[1088] Consejo de Estado italiano. Sentencia 13 diciembre 2019, n.8472. Apartado 15.3. Texto disponible en: https://www.madeappalti.com/sentenze_e_normative/cons-st-sez-vi-13-dicembre-2019-n-8472/
La doctrina también ha reconocido este principio. En: MEDINA GUERRERO,M: "Categorías especiales de datos". En RALLO LOMBARTE, A (dir): Tratado de Protección de Datos. Ed. Tirant lo Blanch, Valencia, 2019, pág.258.

discriminación sobre las categorías de datos especiales del artículo 9 del RGPD[1089]. Ello obliga al responsable a evaluar la posibilidad de que haya personas o grupos que puedan verse afectados de forma desproporcionada por las implicaciones negativas del sistema[1090]. Es decir, tal y como ha señalado la AEPD, la toma de decisiones algorítmicas es un riego que se ha de tener en cuenta a la hora de llevar a cabo el análisis de riesgos que se incorpora a la EIPD[1091].

Existen esencialmente dos clases de discriminación, la directa y la indirecta.

Por lo que se refiere a la directa, la persona recibe un trato menos favorable por alguno de los motivos especialmente protegidos por la legislación en materia de no discriminación[1092]. Es decir, la presencia de esa característica o rasgo en esa persona le hace que reciba un trato discriminatorio respecto de otra persona que no la ostenta[1093]. De esta manera, la discriminación algorítmica directa puede tener lugar cuando una variable relativa a la pertinencia a un grupo protegido se asocie un valor negativo. Ese valor negativo puede conducir directamente a una consecuencia final negativa, por ejemplo, ser de procedencia hispana implica de manera automática la no concesión de un crédito. A su vez, ese valor puede contribuir a reducir la puntuación o empeorar el resultado final pero no ser determinante para su obtención[1094].

Por otro lado, con relación a la discriminación indirecta, esta se produce cuando una práctica, disposición o criterio aparentemente neutro genera un efecto negativo

---

[1089] Recordemos que las categorías especiales de datos en el contexto de la toma de decisiones automatizadas engloban: i) los datos estrictos de categoría especial del artículo 9 del RGPD, ii) los datos convencionales cuando se pretendan alguna de las finalidades especiales o sensibles del artículo 9 del RGPD, iii) los datos convencionales que resulte ser un *proxy* de otro dato de categoría especial del artículo 9 del RGPD. En: Capítulo II, apartado II, punto 2 de esta tesis.

[1090] Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Directrices éticas para una IA fiable*.2019, pág.40.

[1091] Agencia Española de Protección de Datos. *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales.* Junio 2021, pág.100. También en: Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*. Febrero de 2020, pág.30.

[1092] El artículo 14 del Convenio Europeo de Derechos Humanos establece que: *El goce de los derechos y libertades reconocidos en el presente Convenio ha de ser asegurado sin distinción alguna, especialmente por razones de sexo, raza, color, lengua, religión, opiniones políticas u otras, origen nacional o social, pertenencia a una minoría nacional, fortuna, nacimiento o cualquier otra situación*. A su vez, el artículo 21.1 de la Carta de Derechos Fundamentales de la UE señala que: *Se prohíbe toda discriminación, y en particular la ejercida por razón de sexo, raza, color, orígenes étnicos o sociales, características genéticas, lengua, religión o convicciones, opiniones políticas o de cualquier otro tipo, pertenencia a una minoría nacional, patrimonio, nacimiento, discapacidad, edad u orientación sexual*. Finalmente, el artículo 14 de la Constitución Española establece que: *Los españoles son iguales ante la ley, sin que pueda prevalecer discriminación alguna por razón de nacimiento, raza, sexo, religión, opinión o cualquier otra condición o circunstancia personal o social.*

[1093] Agencia europea de Derechos Fundamentales. *Manual de legislación europea contra la discriminación*, Edición de 2018, pág.46.

[1094] SORIANO, ARNANZ,A: "Decisiones automatizadas y discriminación: aproximación y propuestas generales".op.cit.,pág.15.

sobre uno de los grupos que presentan los rasgos o características protegidas respecto de otro grupo que no los presenta[1095]. En el caso de la discriminación algorítmica indirecta, la disposición, práctica o criterio aparentemente neutro puede ser la variable específica introducida en el sistema que, por el valor que le otorga esta, produce un resultado discriminatorio para el grupo desaventajado, o bien el algoritmo generalmente considerado [1096]. En este sentido, el Tribunal Ordinario de Bolonia consideró que las condiciones contractuales establecidas por la empresa *Deliveroo* a través de su algoritmo eran constitutivas de una discriminación indirecta[1097].Para el tribunal, la discriminación indirecta se genera desde el momento que cualquier ausencia del trabajo, independientemente de los motivos, se penaliza por el sistema algoritmo de forma similar[1098]. Como ha destacado la doctrina, se trata de un caso de discriminación a trabajadores a través de un algoritmo, no de una discriminación del algoritmo en sí[1099], si bien, tal discriminación está presente en el sistema automatizado.

Es por ello que los responsables del tratamiento, en virtud del principio de lealtad deban prever las medidas adecuadas para reducir estos posibles sesgos presentes en el algoritmo[1100]. Sesgos que pueden generar discriminación tanto directa como indirecta respecto de las categorías especialmente sensibles reconocidas en el artículo 9 del RGPD y que el considerando 71 de este mismo texto prohíbe. Recordar no obstante que tanto la discriminación directa como la indirecta pueden estar justificadas si existe una razón objetiva que habilita a dicho tratamiento menos favorable. En el derecho de la Unión Europea, los supuestos de discriminación directa únicamente pueden justificarse

---

[1095] ROIG I BATALLA,A: *Las garantías frente a las decisiones automatizadas del Reglamento general de Protección de Datos a la gobernanza algorítmica*.op.,cit, pág.167.

[1096] SORIANO, ARNANZ,A: "Decisiones automatizadas y discriminación: aproximación y propuestas generales".op.cit.,pág.19.

[1097] Sentencia del Tribunal Ordinario de Bolonia de 31 de diciembre de 2020. Resolución disponible en: https://www.algoritmolegal.com/wp-content/uploads/2021/01/Sentencia-Bologna-Italia-Deliveroo-dic-2020-Original-italiano.pdf

[1098] El tribunal aprecia discriminación indirecta ya que se puntúa negativamente de forma similar la ausencia al trabajo sin ninguna justificación con otros tipos de ausencias que si están justificados como por ejemplo la asistencia a una huelga, las ausencias por enfermedad o el cuidado de menores. Sentencia del Tribunal Ordinario de Bolonia de 31 de diciembre de 2020. FJº4, pág.9.

[1099] Se obliga a Deliveroo a eliminar los efectos de la conducta discriminatoria, lo que implica no sólo modificar las condiciones de trabajo, sino también las instrucciones del algoritmo (sus parámetros de puntuación para el ranking de riders) que ejecuta dichas condiciones de trabajo. En: CASTILLO PARRILLA, J,A: "La discriminación a través del algoritmo en una plataforma. El caso Deliveroo Bolonia y sus implicaciones para el sector público". *Observatorio de Transformación Digital del Sector Público*. Información disponible en: https://www.uv.es/catedra-pagoda/es/novedades-1286053802801/Novetat.html?id=1286182093538

[1100] Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Directrices éticas para una IA fiable*.2019, pág.13.

por razones tasadas y objetivas[1101]. Sin embargo, en el caso de la discriminación indirecta se permite la realización de un examen de proporcionalidad a través del cual se ponderen los diferentes intereses en juego[1102].

Para finalizar cabe indicar que la AEPD ha señalado que las medidas que se implanten para reducir posibles riesgos de discriminación se deberán integrar en el modelo algorítmico desde las primeras fases del diseño del mismo[1103]. Una vez puesto en marcha el sistema, es necesario desplegar toda una serie de salvaguardas que eviten o en su caso aminoren los efectos negativos que se derivan de la toma de decisiones algorítmicas discriminatorias. Dichas garantías se analizarán posteriormente[1104].

[1101] Así, se permite la discriminación directa: i) por requisitos profesionales esenciales; ii) excepciones relacionadas con las instituciones religiosas; iii) excepciones a la discriminación por motivos de edad. En: Agencia Europea de Derechos Fundamentales. *Manual de legislación europea contra la discriminación*, Edición de 2018, pág.108.

[1102] Para más información sobre ese examen de proporcionalidad véase: SORIANO, ARNANZ,A: "Decisiones automatizadas y discriminación: aproximación y propuestas generales". op.cit., pág.23.

[1103] Agencia Española de Protección de Datos. Resolución Nº: PS/00120/2021, págs. 93 y 94.

[1104] Sobre las garantías en los contexto de discriminación algorítmica véase el Capítulo V, apartado III, punto 5 de esta tesis.

CAPÍTULO V. LOS DERECHOS DE LOS INTERESADOS SOMETIDOS A SISTEMAS DE TOMA DECISIONES AUTOMATIZADAS

I. INTRODUCCIÓN

El RGPD reconoce toda una serie de facultades en favor de los interesados. Este conjunto de derechos representa uno de los pilares básicos en los que se asienta el derecho fundamental a la protección de datos. Tal y como ha señalado nuestro Tribunal Constitucional, este derecho fundamental atribuye a su titular un haz de facultades que consiste en su mayor parte en el poder jurídico de imponer a terceros la realización u omisión de determinados comportamientos con el objetivo de garantizar a la persona un poder de control sobre sus datos personales[1105]. Es por ello que ese conjunto de facultades forme parte del contenido esencial de este derecho. Por tanto, cualquier restricción a estos derechos deberá estar debidamente justificada y resultar proporcional a la finalidad perseguida[1106].

Estas facultades podrán ejercerse en cualquier fase del ciclo de vida de los sistemas de toma de decisiones automatizados siempre que haya tratamiento de datos personales[1107]. Ha de prestarse especial atención a la fase de desarrollo ya que este conjunto de derechos habrán de integrar en el diseño de los sistemas automatizados. El objetivo es que una vez el sistema comience a desplegar sus efectos, los particulares puedan ejercitar sus derechos de forma plena adaptados al contexto donde el algoritmo genere los resultados. Así, la forma de ejercer el derecho de acceso ante un sistema de recomendación de contenido diferirá de la práctica de esa misma facultad en relación con un contrato de seguro cuyo precio de la póliza se basa en la evaluación que realiza

[1105] STC 292/2000 de 30 de noviembre de 2000, FJº 6, STC 290/2000 de 30 de noviembre de 2000, FJ 7º, STC 254/1993 de 20 de julio de 1993, FJº 7.
Este conjunto de derechos son reconocidos generalmente como el *habeas data*. En: MURILLO DE LA CUEVA, P, L Y PIÑAR MAÑAS, J,L: *El derecho a la autodeterminación informativa.* Ed, Fundación Coloquio Jurídico Europeo, 2009, Madrid, pág.14.

[1106] Artículo 52 de la Carta de Derechos Fundamentales de la Unión Europea y artículo 23 del RGPD.

[1107] Hay que tener en cuenta además que la LOPD de 2018 en su TÍTULO X reconoce toda una serie de derechos digitales en favor de los interesados. Véase los artículos 79 a 97 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Un estudio de estos derechos puede consultarse en: RALLO LOMBARTE, A: "Del derecho a la protección de datos a la garantía de nuevos derechos digitales". En: GARCÍA MAHAMUT, R; TOMÁS MALLÉN, B (eds.): *El Reglamento General de Protección de Datos. Un enfoque nacional y comparado. Especial referencia a la LO 3/2018 de Protección de Datos y garantía de los derechos digitales*. Ed. Tirant lo Blanch, Valencia, 2019. Págs. 133 a 160.

un algoritmo sobre el estilo de conducción de la persona. En ambos supuestos, el derecho de acceso deberá ser plenamente ejercido por los particulares que lo soliciten, si bien, la forma variará.

La importancia de estos derechos en el contexto de la toma de decisiones automatizadas es muy relevante ya que hasta la fecha, la PRAI[1108], norma que será básica a escala de la UE en el uso de sistemas de toma de decisiones automatizadas, no contempla ningún tipo de facultad o derecho en favor de los individuos sobre los que dichos sistemas adoptarán decisiones. Es por ello que los derechos previstos en el RGPD resulten sumamente esenciales ya que complementarán el déficit mostrado por la PRAI cuando dichos sistemas traten datos personales[1109].

En este capítulo estudiaremos en primer lugar el conjunto de derechos generales que se reconocen a todos los particulares sometidos a sistemas de toma de decisiones automatizadas. En segundo lugar se analizará el derecho del interesado a no ser objeto de los tratamientos previstos en el artículo 22 y las facultades específicas que vienen aparejadas a este derecho. Por último, los apartados finales aludiremos a varias propuestas interpretativas relacionadas con el conjunto de derechos y facultades que se reconocen en el RGPD.

Sistemas de toma de decisiones automatizadas	
Derechos generales	**Derechos específicos previstos en el artículo 22 del RGPD**
Derecho de información, derecho de acceso, derecho de rectificación, derecho de supresión, derecho a la portabilidad de datos, derecho de oposición.	Derecho de información específico, prohibición general del tratamiento, derecho a solicitar intervención humana, derecho a expresar el punto de vista, derecho de impugnación de la decisión, derecho de explicación.

[1108] Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

[1109] Hasta la fecha, la PRAI focaliza su atención en los requisitos que todo sistema basado en inteligencia artificial debe contener durante su proceso de fabricación y despliegue, sin embargo, no se hace mención en ningún momento de esta propuesta legal a los derechos que tienen los particulares que se ven sometidos a las decisiones que se adopten por estos sistemas.

II. DERECHOS GENERALES EN FAVOR DE LOS PARTICULARES SOMETIDOS A SISTEMAS DE TOMA DE DECISIONES TOTAL O PARCIALMENTE AUTOMATIZADAS

El RGPD reconoce a cualquier particular que se vea sometido a la toma de decisiones total o parcialmente automatizadas o a la elaboración de perfiles una serie de facultades. Es momento de analizar la incidencia de estos derechos en el contexto del objeto de esta tesis.

1. El derecho de acceso

El derecho de acceso se encuentra reconocido en el artículo 15 del RGPD. A través del mismo el particular puede solicitar al responsable toda una serie de información relacionada con sus datos personales. Esta información es prácticamente similar a la que se deriva de los artículos 13 y 14 del RGPD. De esta manera, el legislador europeo no sólo ha considerado que el interesado debe ser notificado durante la recopilación u obtención de los datos personales sobre el tratamiento al que se van a someter dichos datos, sino que además, faculta al particular a poder conocer en cualquier momento toda aquella información que tenga relación con el tratamiento analizado una vez este ha comenzado. Hay que tener en cuenta que la naturaleza de los derechos de información y de acceso previstos por el RGPD es diferente ya que los mismos hacen referencia a momentos distintos del tratamiento[1110]. Consecuencia de ello, en el caso de que el responsable tenga más información de la que disponía en un momento inicial cuando informó conforme a los artículos 13 y 14, tal información vía derecho de acceso deberá ampliarse[1111]. Ello ocurrirá por ejemplo con los datos personales inferidos que se obtengan una vez comience a realizarse el perfilado[1112]. A diferencia de la etapa previa al procesado algorítmico de datos, el responsable ya sí

[1110] El principio de información y el derecho de acceso hacen referencia a momentos distintos: el primero a la recogida de datos, el segundo al posterior tratamiento. En: TRONCOSO, REIGADA, A: *La Protección de Datos Personales .En Busca del Equilibrio*, op.cit., pág.525.

[1111] GIL, GONZÁLEZ, E: "Aproximación al estudio de las decisiones automatizadas en el seno del Reglamento General Europeo de Protección de Datos a la luz de las tecnologías big data y de aprendizaje computacional". *Revista Española de la Transparencia*, Nº 5. Segundo Semestre 2017, pág.172.

[1112] El derecho de acceso contemplado en el artículo 15 puede ofrecer una solución cuando el interesado carece de información sobre los datos inferidos y derivados que se conservan. En: WACHTER, S; MITTELSTADT, B: "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI". op.cit, págs. 54 y 55.

conoce en esta fase del tratamiento qué datos inferidos está obteniendo del particular, por tanto, este último puede solicitar el acceso a los mismos.

Toda esta información que vamos a analizar, al igual que la que se derivaba del principio de transparencia, deberá facilitarse en forma concisa, transparente, inteligible y de fácil acceso con un lenguaje claro y sencillo. Artículo 12.1 del RGPD.

A) El Derecho de acceso en la fase de diseño

Durante la fase de desarrollo de los sistemas de toma de decisiones automatizados puede ser frecuente que se traten datos personales. En estos casos, los particulares podrán solicitar la información reconocida en el artículo 15 del RGPD.

Así, en *primer lugar*, los interesados están facultados para solicitar el acceso a los datos de entrenamiento que puedan estar utilizando los responsables para desarrollar los modelos. Siempre que dichos datos puedan atribuirse al sujeto que solicita el acceso, el responsable debería facilitar dicha información[1113].

En *segundo lugar*, decíamos que durante las etapas iniciales del diseño de los modelos algorítmicos basados en técnicas de aprendizaje no supervisado podía resultar complicado para el responsable establecer una finalidad específica. Pues bien, conforme avanza dicho proyecto, la finalidad cada vez se irá concretando. Es por ello que el responsable, ante una petición del particular solicitando conocer la finalidad para la que se están tratando los datos personales en virtud del artículo 15.1. a) del RGPD, deberá facilitar más información de la que inicialmente proporcionó ya que muy probablemente tendrá un conocimiento mayor sobre la finalidad o finalidades por las cuales está desarrollando esos algorítmicos. El estudio de minimización de datos al que hacíamos referencia al analizar este principio puede ayudar al responsable a responder a estas peticiones de información en la fase de elaboración de los sistemas automatizados.

B) El Derecho de acceso en la fase de despliegue o toma de decisiones

Una vez que el sistema se pone en marcha, este comienza a adoptar decisiones y generar inferencias. El derecho de acceso no sólo permite acceder a los datos que se

[1113] Information Commisioner´s Office. En: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/how-do-we-ensure-individual-rights-in-our-ai-systems/

facilitaron durante la fase de recopilación sino también a aquellos que en su caso se hayan generado una vez que el algoritmo los procesa. Es por ello que los particulares en muchos supuestos estén interesados en conocer aquella información que se deriva del procesamiento algorítmico de sus datos personales. En este sentido, se ha señalado que el derecho de acceso en estos contextos es claramente instrumental ya que a través del mismo se facilita el ejercicio de otras facultades[1114], tal y como puede ser la rectificación de los datos inferidos, la supresión de los perfiles, etc.

b.1) Información accesible para cualquier tratamiento de datos personales basado en la toma de decisiones automatizadas o la elaboración de perfiles

En *primer lugar*, los interesados están facultados para solicitar el acceso a los datos inferidos que esté generando el sistema. Como ya hemos señalado en otros apartados, las inferencias que realizan los algoritmos sobre los particulares son intencionadamente ocultados por las organizaciones que utilizan estos sistemas. Ello limita las posibilidades de los interesados de conocer las consecuencias reales que pueden estar generando sobre los mismos estos tratamientos. Este tipo de datos pueden ser inexactos ya que las inferencias que ha podido generar el sistema y a las que se le asigna una determinada consecuencia jurídica pueden ser erróneas. Por otro lado, la propuesta de reglamento sobre la responsabilidad civil por el funcionamiento de los sistemas de IA habilita a las personas afectadas por los daños que generen estos sistemas a usar los datos inferidos y generados por los mismos con fines probatorios o aclaratorios a la hora de redactar una hipotética demanda por responsabilidad civil[1115]. Por tanto, cuando un particular quede afectado por un daño que ha generado un sistema de inteligencia artificial, dicho interesado podrá utilizar esos datos inferidos con el fin de fundamentar su demanda.

En *segundo lugar*, tal y como ha señalado la AEPD, los interesados también pueden solicitar el acceso a los datos de entrenamiento que pudieran estar incluidos en los sistemas de IA y que puedan ser recuperados por el responsable que despliega dicho

[1114] ROIG I BATALLA,A: *Las garantías frente a las decisiones automatizadas del Reglamento general de Protección de Datos a la gobernanza algorítmica*, op.,cit, pág.53.
[1115] Artículo 10.2 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. Resolución aprobada por el Parlamento Europeo el 20 de octubre de 2020.

sistema[1116]. En estos supuestos entendemos que el acceso a esos datos quedará limitado cuando: i) el modelo algorítmico no permita técnicamente el acceso a dichos datos personales y, ii) cuando resultando posible el acceso a esos datos de formación, no se hayan utilizado datos del interesado que solicita su acceso.

En *tercer lugar*, el GT29 ha señalado que en virtud del artículo 15.3 del RGPD, ante la solicitud de acceso del particular, el responsable del tratamiento tiene el deber de facilitar el acceso a la información sobre el perfil y los detalles sobre los segmentos a los que se ha asignado al interesado[1117]. Ello permitirá al interesado verificar su perfil y las fuentes utilizadas para desarrollarlo[1118]. En este sentido, recientemente, un tribunal holandés ha reconocido el derecho de acceso a los trabajadores de una empresa de taxis a los perfiles que elabora esta última donde se evalúa el comportamiento de dichos trabajadores y se predice la probabilidad de fraude de estos últimos[1119]. En este sentido, para facilitar el derecho de acceso en los entornos en línea, resultaría recomendable el despliegue de interfaces que favorecieran el acceso rápido y sencillo de los particulares a los perfiles que se estén elaborando sobre los mismos. Así, el considerando 63 del RGPD establece que el responsable del tratamiento está facultado para facilitar acceso remoto a un sistema seguro que ofrezca al interesado un acceso directo a sus datos personales. Si dicho acceso a esos perfiles se presta de forma permanente y respecto de todos los datos hay que entender que el derecho de acceso se entiende por concedido[1120]. En este sentido, la propuesta de Reglamento de servicios digitales obliga a las grandes plataformas a implementar un repositorio para que los interesados puedan acceder al mismo cuando estas presenten publicidad. Dicho repositorio incluirá entre otra información: a) el contenido de la publicidad, b) la persona física o jurídica en cuyo

[1116] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.25.
[1117] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018. Pág.19.
[1118] Red iberoamericana de protección de datos personales. *Orientaciones Específicas para el Cumplimiento de los Principios y Derechos que Rigen la Protección de los Datos Personales en los Proyectos de Inteligencia Artificial.*2019. Págs. 18 y 37.
[1119] Tribunal de Ámsterdam. Resolución del 11-03-2021, apartados 4.44, 4.45 y 4.46. Disponible en: https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBAMS:2021:1019&showbutton=true&keyword=ola+cabs
[1120]Artículo 13.2 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

nombre se presenta el anuncio publicitario; c) el período durante el que se haya presentado la publicidad, d) grupos a los que va dirigida la publicidad, etc[1121].

b.2) Información accesible para los tratamientos de datos personales basados en la toma de decisiones automatizadas plenamente relevantes con o sin elaboración de perfiles

Además de la información descrita previamente, cuando el responsable lleve a cabo los tratamientos descritos en el artículo 22 del RGPD, el derecho de acceso también comprenderá información significativa sobre la lógica del tratamiento y las consecuencias previstas de dicho tratamiento tal y como prevé el artículo 15.1.h) del RGPD. Esta información generalmente será más amplia de aquella que en su momento se facilitó durante la recopilación de los datos de acuerdo a los deberes de información previstos en los artículos 13.2.f) y 14.2.g) del RGPD. Así, a diferencia del momento de la recopilación de datos donde el responsable podía hipotéticamente conocer las consecuencias previstas de la elaboración de perfiles o las decisiones automatizadas, ahora, en este momento, dado que ya se han procesado los datos y los algoritmos se han puesto en marcha, será más fácil proporcionar las consecuencias previstas específicas de esos tratamientos ante la solicitud de acceso del interesado.

En definitiva, el derecho de acceso faculta a los interesados a solicitar información al responsable sobre la elaboración de perfiles y la toma de decisiones automatizadas. Esta información en muchos casos será muy parecida a la que previamente y en virtud del principio de transparencia el responsable tuvo que facilitar durante la recopilación de los datos. Sin embargo, fruto de la continua alteración a la que se ven sometidos los datos y los tratamientos que sufren los mismos una vez que el sistema algorítmico los procesa, la información que se deriva del derecho de acceso y del principio de transparencia podrá diferir en mayor o menor grado.

[1121] Artículo 30 de la Propuesta de la propuesta 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.

La información ha de facilitarse de forma concisa, transparente, inteligible y de fácil acceso con un lenguaje claro y sencillo. Artículo 12 RGPD.	
Derecho de información. Artículos. 13 y 14.	**Derecho de acceso. Artículo 15.**
Finalidad general pretendida por la analítica de datos.	Finalidad específica de la analítica de datos.
Datos inferidos que potencialmente se generarán.	Datos inferidos que el sistema está generando sobre ese particular.
Información general sobre el perfilado.	Acceso al perfil que se está desarrollando o se ha elaborado sobre el interesado.
Consecuencias previstas generales del perfilado.	Consecuencias específicas del perfilado sobre el particular.
Lógica del tratamiento y consecuencias previstas generales de la toma de decisiones automatizadas	Afectación específica de la toma de decisiones automatizadas sobre el particular.

C) Límites al derecho de acceso

Cuando un particular solicita el acceso a la información previamente señalada, el responsable podría negarse a concederla por diferentes motivos. Es turno de analizarlos.

En *primer lugar*, el artículo 15.4 del RGPD establece que el derecho a obtener copia de los datos personales objeto de tratamiento no afectará negativamente a los derechos y libertades de otros. A su vez, el considerando 63 del RGPD indica que el derecho de acceso no *debe afectar negativamente a los derechos y libertades de terceros, incluidos los secretos comerciales o la propiedad intelectual y, en particular, los derechos de propiedad intelectual que protegen programas informáticos*. En este sentido, el TJUE en el *asunto C-434/16 Nowak*, ya analizado en esta tesis, reconoció el derecho que tenía el particular para acceder a las anotaciones subjetivas que realiza un examinador sobre las respuestas de su examen[1122], si bien, tal y como dijo el Tribunal, dicho acceso podía quedar limitado por el Derecho interno o en su caso el europeo[1123]. Así, la Directiva relativa a la protección de los secretos comerciales, a la que se aplica el concepto de dato inferido, prevé diferentes herramientas que protegen el acceso a dichos

[1122] Recordemos que a la hora de estudiar el concepto de dato inferido, este lo asimilábamos a las anotaciones que realizaba un examinador en el examen de un particular. Dichas anotaciones se inferían de los datos personales del interesado, en este caso, de las respuestas aportadas en el examen. Situación similar a los datos que infiere el algoritmo una vez procesa los datos iniciales del particular. Capítulo II, apartado II, punto 4.

[1123] SENTENCIA DEL TRIBUNAL DE JUSTICIA (Sala Segunda) de 20 de diciembre de 2017, asunto C-434/16, caso Nowak, FJº 59 y 61. Resolución disponible: https://curia.europa.eu/juris/document/document.jsf;jsessionid=581E9C64A0CE05E6C364706E170A864F?text=&docid=198059&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=2413779

secretos comerciales[1124]. Es por ello que una organización podría negarse a facilitar los datos inferidos de un particular ante la solicitud de acceso a los mismos por considerar que estos son secretos comerciales. A nivel nacional, las anotaciones subjetivas que realizan determinados profesionales también han encontrado respaldo legal a la hora de restringir el derecho de acceso. Así, la Ley de autonomía del paciente limita el acceso de los pacientes a su historial médico respecto de las anotaciones subjetivas realizadas por los profesionales cuando estos últimos se opongan al conocimiento de las mismas[1125]. El hecho de que el origen de los datos sometidos a tratamiento provenga no sólo del interesado sino también de terceras personas, como ocurre con los profesionales médicos a través de dichas anotaciones subjetivas respecto de ese titular, obliga a proteger la intimidad de estos últimos[1126]. En el objeto de esta tesis, ese tercero que origina el dato es el algoritmo de la organización. Esta última por tanto puede alegar la protección de los mismos. En nuestra opinión, cuando se ejercita el derecho de acceso relacionado con datos inferidos se requiere un equilibrio de los intereses del particular que realiza la solicitud y los del responsable del tratamiento. No podemos olvidar los daños que pueden generar en la esfera de los particulares las inferencias erróneas o discriminatorias sobre esa persona tanto en el presente como en el futuro[1127]. Cualquier restricción ligada al acceso de los datos inferidos deberá justificarse adecuadamente y

[1124] Ya hemos indicado que el concepto de secreto comercial es tan amplio que los datos inferidos pueden quedar incluidos en el mismo. Véase artículo 2.1 y 3 de 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 divulgada (secretos comerciales) contra su obtención, utilización y revelación ilícitas.

[1125] Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica.

[1126] En: TRONCOSO, REIGADA,A: *La Protección de Datos Personales .En Busca del Equilibrio*. op.cit, pág.553. No existe un criterio unánime actual por parte de la doctrina para considerar si los particulares tienen derecho a acceder a las anotaciones subjetivas que realizan los profesionales médicos en su documentación clínica. Véase: JOVE VILLARES, D: "El caso Nowak y sus posibles consecuencias para las anotaciones subjetivas en la historia clínica". *XVII Congreso de la Asociación de Constitucionalistas de España (ACE)* . Abril de 2019. Texto disponible en: https://www.acoes.es/congreso-xvi/wp-content/uploads/sites/2/2018/04/Jove-Villares_El-caso-Nowak-y-sus-posibles-consecuencias-para-las-anotaciones-subjetivas-en-la-historia-cl%C3%ADnica.pdf
Una versión extendida puede verse en: JOVE VILLARES, D: "Peter Nowak and Subjective Annotations in Clinical Records". *European Data Protection Law Review*. Volume 5, nº 2, 2019, págs.175 y ss. La Autoritat Catalana de Protecció de Dades considera que los facultativos pueden limitar el acceso a las anotaciones subjetivas presentes en los historiales clínicos. En Autoritat Catalana de Protecció de Dades. *Guía de protección de datos para pacientes y personas usuarias de los servicios de salud***.** Junio 2020, págs 11 y 12.

[1127] Center for Democracy and Technology. *Algorithmic Systems in Education: Incorporating Equity and Fairness When Using Student Data*, 2019, pág.18.

nunca podrá suponer la negativa total a prestar dicha información pese a la naturaleza híbrida de los mismos (Considerando 63 RGPD)[1128].

En *segundo lugar*, la restricción para permitir el acceso a los datos inferidos puede devenir de la propia tecnología. Como hemos indicado, en los entornos adaptativos, los sistemas dinámicos continuamente están alterando su comportamiento. Esto puede dar lugar a que estos algoritmos generen todo tipo de inferencias en un corto periodo de tiempo[1129]. Estos datos inferidos pueden ir variando rápidamente por lo que ante la solitud de acceso a los mismos por parte del particular, el responsable podría alegar que tiene dificultad para facilitar tal cantidad de datos. Para estos supuestos, el considerando 63 del RGPD faculta al responsable del tratamiento a que antes de facilitarse la información, requiera al interesado la especificación de la información o actividades de tratamiento a que se refiere la solicitud de acceso.

En *tercer lugar*, el artículo 13.1 de la LOPD de 2018 establece que cuando el responsable trate una gran cantidad de datos relativos al afectado y este último ejercite su derecho de acceso sin especificar a qué datos se refiere, el responsable podrá solicitar al particular que especifique los datos o actividades de tratamiento a los que se refiere la solicitud. Esta situación será muy habitual cuando el responsable utilice sistemas de toma de decisiones automatizadas ya que los mismos como sabemos recopilan una gran cantidad de datos de los particulares. Esta previsión normativa, aunque resulta lógica[1130], en el contexto de la toma de decisiones automatizadas podría desincentivar el ejercicio del derecho de acceso por parte de los particulares debido al desconocimiento que tienen estos de las tipologías de datos que puede ostentar el responsable sobre ellos. Para estos supuestos sería recomendable que en base al principio de responsabilidad activa, el responsable del tratamiento facilitara al particular las distintas categorías de datos que en su caso ostenta del particular, por ejemplo; perfiles realizados, datos

[1128] Como ya hemos indicado, los datos inferidos presentan una naturaleza híbrida ya que su origen deviene de los propios titulares de los datos y del procesamiento que haga de estos el algoritmo. Véase el Capítulo II, apartado II, punto 4, apartado B) de esta tesis.

[1129] Tal y como se ha señalado, un particular puede ser clasificado como probablemente homosexual hoy, y como probablemente heterosexual mañana en función de cómo este interactúe con el algoritmo en esos momentos. En: Privacy international. Data is power: *Towards additional guidance on profiling and automated decision-making in the GDPR*, 2017,pág.8.

[1130] El ejercicio del derecho de acceso de forma genérica de todos los datos que ostenta un responsable del tratamiento puede genera importantes distorsiones para este último fruto de la gran cantidad de datos que puede estar tratando dicho responsable sobre el particular. Resulta coherente que el particular especifique en su solicitud los datos o categorías de datos. Véase un ejemplo de ello en: Sentencia Audiencia Nacional (Sala de lo Contencioso-Administrativo, Sección1ª). Sentencia de 4 junio 2021. FJº 5º y 6º. Resolución disponible en:
https://www.poderjudicial.es/search/AN/openDocument/c919ed51764bdf0b/20210701

inferidos en el último mes, datos obtenidos de terceros, etc. Con esa información desgranada, el particular podría indicar aquellos datos específicos sobre los que pretende ejercitar el derecho de acceso.

Finalmente, en *cuarto lugar*, el artículo 12.5 del RGPD habilita al responsable del tratamiento a negarse a conceder el acceso o a cobrar un canon por dichas actuaciones cuando las solicitudes sean manifiestamente infundadas o excesivas, sobre todo, cuando las mismas sean repetitivas. En este sentido, la LOPD de 2018 indica que dicho carácter repetitivo se podrá considerar cuando el ejercicio del derecho de acceso se realice en más de una ocasión durante el plazo de seis meses salvo que exista causa legítima para ello[1131]. Esta limitación es trasladable también al contexto de los datos inferidos, si bien, consideramos necesario realizar ciertas matizaciones. Así, existirán diversas situaciones donde la solicitud de acceso a dichos datos inferidos estará más que justificada aunque la misma se realice en periodos más cortos de tiempo. Esto puede suceder cuando ese particular detecte que el algoritmo está adoptando decisiones que le afectan de modo significativo. Por ejemplo, alteración del coste de la prima de seguros, bloqueo de la cuenta que le permite ofrecer determinados servicios[1132], ofrecimiento de nuevos productos financieros por parte de la entidad bancaria, etc. En estos supuestos, el interesado legítimamente podrá conocer las inferencias que se han generado y que han derivado posiblemente en una determinada decisión que le afecta.

En definitiva, el derecho de acceso es una facultad que ostentan los interesados cuyos datos personales son tratados a lo largo del ciclo de vida de los sistemas automatizados. Teniendo en cuenta que el derecho de acceso se ejercita una vez que el algoritmo ha comenzado a procesar los datos, la información que puede obtener ese particular sobre dicho tratamiento aumenta respecto de aquella que se ofreció durante la recopilación de los datos. Así, a través del derecho de acceso se pueden conocer los datos inferidos y los perfiles, así como las consecuencias específicas de esos perfiles con relación a dicho interesado. El responsable, por su parte, podrá restringir estas facultades alegando motivos e intereses legítimos como los secretos comerciales. Sin embargo, cualquier restricción de esta facultad ha de estar justificada.

[1131] Artículo 13.3 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.
[1132] Por ejemplo, si un algoritmo cancela la cuenta de un *rider* y este último no puede ofrecer los servicios de reparto.

2. El derecho de rectificación

El artículo 16 del RGPD establece que *el interesado tendrá derecho a obtener sin dilación indebida del responsable del tratamiento la rectificación de los datos personales inexactos que le conciernan*. A su vez, y *teniendo en cuenta los fines del tratamiento, el interesado tendrá derecho a que se completen los datos personales que sean incompletos, inclusive mediante una declaración adicional*. Por tanto, de este derecho se derivan dos facultades, la rectificación en sentido estricto de los datos personales cuando estos sean inexactos y la posibilidad de completar los datos sometidos a tratamiento con información adicional[1133].

El ejercicio de este derecho puede realizarse en todas las fases comprendidas durante el ciclo de vida de los sistemas automatizados, si bien, y como ahora veremos, su principal incidencia está presente durante la fase de despliegue.

A) El derecho de rectificación en la fase de diseño de los sistemas automatizados

Normalmente, las inexactitudes individuales presentes en los datos de entrenamiento tendrán un efecto mínimo o reducido en la esfera de los titulares de dichos datos. Ello es así porque si bien dichas inexactitudes pueden generar modelos imprecisos e incluso discriminatorios, los individuos no notarán sus efectos normalmente hasta que el sistema comience a adoptar decisiones, decisiones, que además no tienen por qué afectar a los individuos cuyos datos se utilizaron para el entrenamiento. Es por ello que las solicitudes para la rectificación de los datos en la fase de desarrollo sean inferiores a aquellas que se puedan presentar una vez el sistema comience a desplegar sus efectos. En la fase de diseño, como ya dijimos, la clave será que el responsable cumpla adecuadamente con los principios de tratamiento como el de exactitud o el de lealtad con el objetivo de evitar sistemas poco precisos y sesgados. Dicho lo anterior, si algún particular solicita la rectificación de sus datos, el responsable deberá proceder a la misma.

[1133] MURGA FERNÁNDEZ,J,P: "Derechos de los individuos". En: MURGA FERNÁNDEZ, J,P; FERNÁNDEZ SCAGLIUSI, M,A ; ESPEJO LERDO DE TEJADA,M: (dirs.): *Protección de datos, responsabilidad activa y técnicas de garantía*, op.cit., pág.93.

B) El derecho de rectificación en la fase de despliegue de los sistemas automatizados

En el momento que el sistema se incorpora al proceso decisorio de las organizaciones, los resultados emitidos por este pueden afectar a los particulares. Es por ello que resulte más probable que los particulares cuestionen la inexactitud de dichos resultados. Concretamente, los datos inferidos y los perfiles generados por dichos algoritmos[1134].

Así, decíamos que a través de la elaboración de perfiles se generalizaba y agrupaba a distintas personas en grupos. Cada persona presenta variables específicas propias que la pueden diferenciar del resto de miembros que forman parte del grupo al que se incorpora. Por tanto, la generalización que se pretende con estos algoritmos difícilmente será completa y podrá envolver todos los factores individuales que están presentes en esas personas[1135]. Consecuencia de ello, siempre será posible que los datos inferidos derivados de estos sistemas o la asignación de esa persona a un determinado perfil no sea la adecuada. Para corregir esas inexactitudes, el derecho de rectificación puede mostrarse como la herramienta ideal.

Correlación de características.	**Porcentaje de riesgo de impago expresado por el sistema→ Dato inferido**	**Consecuencia**
Hombres, mayores de 45 años, pertenecientes a un barrio del extrarradio.	70 % →Mal pagador	No se concede el crédito
Hombres, mayores de 25 años, pertenecientes a un barrio del casco antiguo	52 % →Mal pagador	Se concede el crédito con un interés alto.
Hombres, mayores de 45, pertenecientes a un barrio acomodado	25 % → Buen pagador.	Se concede el crédito con interés bajo

Para ejercer el derecho de rectificación adecuadamente, el artículo 14 de la LOPD de 2018 establece que el particular deberá indicar en su solicitud a qué datos se refiere y la corrección que haya de realizarse. Cuando sea preciso, ha de acompañar la solicitud con la documentación justificativa de la inexactitud o carácter incompleto de

[1134] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.27.

[1135] MORENTE,PARRA,V: "Big data o el arte de analizar datos masivos. una reflexión crítica desde los derechos fundamentales", op.cit, págs. 251 y ss.

los datos objeto de tratamiento[1136]. Para ello, será fundamental que previamente el responsable haya permitido a este particular el acceso a los datos inferidos o en su caso a los perfiles que hipotéticamente sean inexactos. Llegados a este punto, el particular deberá justificar dichas imprecisiones y en su caso el responsable deberá verificar si las inferencias que ha realizado el modelo son o no exactas en función de lo indicado por el particular. Así, por ejemplo, la catalogación de una persona como mala pagadora derivará posiblemente de un riesgo alto asignado por parte de un algoritmo. Pues bien, el interesado podrá aportar información y documentación que permita al responsable valorar si dicha inferencia realmente refleja la etiqueta asignada a dicho interesado. Por ejemplo, el particular podría aportar que pese a vivir en una determinada zona a la que el sistema asigna una alto riesgo de impago, esta persona ha pagado habitualmente los créditos hasta la fecha concedidos. Lo mismo puede ocurrir cuando un sistema asigna a una persona en un perfil que presenta una determinada afiliación política o preferencias sexuales que realmente no lo son[1137]. De esta manera, una persona puede solicitar la inclusión de información adicional en su registro que contrarreste la inferencia incorrecta. Esto ayuda a garantizar que cualquier decisión tomada sobre la base de la inferencia potencialmente incorrecta esté informada por cualquier prueba de que puede ser errónea[1138]. A través de esta interpretación no estamos defendiendo que el particular tenga derecho a rectificar el modelo o el algoritmo en sí, sino más bien los resultados o inferencias que haya establecido el modelo y que el responsable haya comprobado que son erróneos. Resultados que no dejan de ser datos personales inferidos del particular sobre los cuales pensamos que el RGPD autoriza a su rectificación.

Las ventajas que puede ofrecer el derecho de rectificación en la personalización de servicios pueden ser tanto para los particulares como para el propio responsable. Del lado de los titulares de los datos, el derecho a la rectificación permite a las personas participar y jugar con el proceso de personalización del perfil. En lugar de simplemente

[1136] Artículo 14 Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.
[1137] Tal y como ha dicho el Grupo del Artículo 29, *el derecho de rectificación se aplica tanto a los datos personales de entrada (los datos personales utilizados para crear el perfil) como a los datos de salida (el propio perfil o la puntuación asignada a la persona).* Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018. Pág.19.
[1138] Information Commisioner´s Office. Texto disponible en: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/what-do-we-need-to-do-to-ensure-lawfulness-fairness-and-transparency-in-ai-systems/#whatdoweneed

detener la personalización cuando se sienten incómodos con el perfil que se está creando de ellos[1139], los interesados pueden adoptar un papel activo en la personalización generada por los sistemas[1140]. A su vez, por el lado del responsable, la existencia de perfiles y datos inferidos exactos también le favorece ya que el proceso decisorio redundará en su beneficio.

En definitiva, la configuración del derecho de rectificación amparado en el principio de exactitud en este contexto no legitima al particular para que ponga en cuestión el modelo algorítmico. Esta facultad autoriza a que, constado por el responsable que los resultados emitidos por el sistema para ese concreto particular, ya sea en formato de inferencia o de asignación a un perfil, no resulten ser exactos, este proceda a su rectificación. Además, conforme al artículo 19 del RGPD, el responsable estará obligado a comunicar cualquier rectificación de los datos a los destinatarios a los que se hayan comunicado los datos personales, salvo que sea imposible o exija un esfuerzo desproporcionado.

C) Límites al derecho de rectificación

A diferencia de otras facultades, el RGPD no establece ninguna limitación al derecho de rectificación. Ello no quiere decir que el ejercicio del mismo no pueda presentar en determinadas situaciones alguna restricción. En este sentido, el artículo 23 del RGPD establece la posibilidad de limitar dicho derecho, por tanto, siempre será posible la misma.

Por otro lado, los órganos judiciales europeos, cuando han tenido la oportunidad de establecer restricciones a esta facultad en el contexto de los datos inferidos tampoco han sido del todo claros. Así, en el asunto *Nowak* ya comentado[1141], el TJUE consideró que las anotaciones realizadas por un examinador en un examen son datos personales del examinado ya que estas pueden tener efectos para el aspirante y además se relacionan con este[1142]. Al referirse al posible derecho de rectificación respecto de dichas anotaciones, que recordemos que es lo más parecido a los datos inferidos, el

[1139] ESKENS,S: "A right to reset your user profile and more: GDPR-rights for personalized news consumers", op.cit., pág.16.
[1140] SOENENS,ELS: "Reply: Web Usage Mining for Web Personalisation in Customer Relation Management". En: HILDEBRANDT, M & GUTWIRTH,S (eds): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008, pág.182.
[1141] Capítulo II, II, 3.
[1142] Sentencia del TJUE (Sala Segunda) de 20 de diciembre de 2017, asunto C-434/16, caso Nowak. http://curia.europa.eu/juris/document/document.jsf;jsessionid=6C361FE3B54FB5329197ADD52A787D08?text=&docid=198059&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=3496287

tribunal indicó que es posible que las correspondientes anotaciones del examinador sean inexactas cuando estas no documenten correctamente la valoración que dicho evaluador ha dado a las respuestas del aspirante de que se trate[1143]. Es decir, cuando por ejemplo la puntuación que haya otorgado a cada pregunta no se corresponda con la nota final establecida para todo el examen. Al no pronunciarse expresamente sobre el propio contenido mismo de dichas valoraciones, esto es, si lo que ha anotado el evaluador es o no inexacto, no podemos extraer ninguna conclusión sobre la posible aplicación o no del derecho de rectificación para dicho contenido. Sí que fue clara la abogada general de del asunto del TJUE previamente comentado al señalar que el derecho de rectificación que se deriva de la normativa de protección de datos no puede abarcar la justificación -sea o no adecuada- sobre la que se realiza la valoración[1144]. En palabras del Comité de Ética Alemán, a priori no existe un interés individual o general que el ordenamiento jurídico pueda respaldar cuando un tratamiento de datos genera inferencias incorrectas para mantener el tratamiento de las mismas. Por ello, los particulares, una vez han probado que esa inferencia es inexacta y la misma ha generado un daño, están facultados para solicitar la rectificación de los mismos. El responsable debería proceder a su rectificación salvo que la corrección del dato resulte desproporcionada, teniendo en cuenta la gravedad y probabilidad del daño por un lado, y el esfuerzo requerido para la corrección por otro[1145].

En definitiva, podemos extraer que, tanto la legislación como los órganos judiciales europeos no han establecido claramente hasta la fecha las posibles restricciones al derecho de rectificación respecto de los datos inferidos. En nuestra opinión, pese a la naturaleza híbrida de estos datos, con carácter general la rectificación de los mismos debería estar permitida cuando se demuestre o se compruebe que dicha inferencia no es correcta. Cualquier restricción a este derecho debería estar justificada y además, los hipotéticos daños que se pudieran generar por esas inferencias inexactas deberán ser resarcidos[1146], independientemente de si se permite la rectificación o no de los datos.

[1143] Apartado 54 de la resolución comentada

[1144] Conclusiones de la Abogada General Sra. J. Kokott, presentadas el 20 de julio de 2017. Asunto C-187/16, caso Peter Nowak contra Data Protection Commissioner. Apartado 54. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:62016CC0187

[1145] Comité de ética alemán. Gutachten der Datenethikkommission, 2019.Pág. 92.

[1146] Recordemos que existe una propuesta por parte del Parlamento Europeo sobre quién y cómo ha de responder por los daños que hipotéticamente pueden generar los sistemas basados en inteligencia artificial. En: 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.

3. El derecho de supresión

El derecho de supresión obliga al responsable del tratamiento a dejar de tratar los datos personales del interesado. El adecuado ejercicio de este derecho queda condicionado a la concurrencia de alguno de los supuestos contemplados en el artículo 17.1 del RGPD[1147], estos son: i) cuando los datos ya no sean necesarios en relación con los fines para los que fueron recogidos, ii) el interesado retire su consentimiento, iii) el interesado se oponga al tratamiento, iv) cuando los datos hayan sido tratados ilegalmente, v) cuando los datos deban suprimirse para el cumplimiento de una obligación legal, etc.

Como puede comprobarse, este derecho tiene dos vertientes, por un lado una facultad de los titulares para solicitar la supresión de sus datos. Y por otro, se obliga al responsable a dejar de tratar los datos de oficio cuando se active alguno de los supuestos contemplado previamente[1148].

Las repercusiones de este derecho tienen su incidencia en cualquier fase comprendida durante el desarrollo y despliegue de los sistemas automatizados cuando se traten datos personales. Es turno de analizarlas.

A) La supresión de los datos en la fase de diseño de los sistemas

El derecho de supresión durante esta fase tendrá relevancia en aquellas etapas de desarrollo del sistema donde se estén tratando datos personales. Así, la supresión de datos esencialmente afectará a las bases de datos o *data lakes* y a los modelos que se hayan construido y que contengan datos personales.

En *primer lugar*, el responsable tendrá la obligación de suprimir aquellos datos que ya no sean necesarios para la finalidad inicial de la recopilación. Así, el estudio de minimización de datos obliga al responsable a que conforme avanzan las fases de desarrollo del modelo, vaya eliminando aquellos datos y variables que no fueran pertinentes para la analítica de datos. A su vez, y relacionado con el principio de

[1147] MURGA FERNÁNDEZ,J,P: "Derechos de los individuos". En: MURGA FERNÁNDEZ, J,P; FERNÁNDEZ SCAGLIUSI, M,A ; ESPEJO LERDO DE TEJADA,M: (dirs): *Protección de datos, responsabilidad activa y técnicas de garantía*, op.cit., pág.95.

[1148] El Considerando 39 establece que *para garantizar que los datos personales no se conservan más tiempo del necesario, el responsable del tratamiento ha de establecer plazos para su supresión o revisión periódica. Deben tomarse todas las medidas razonables para garantizar que se rectifiquen o supriman los datos personales que sean inexactos*.

limitación de plazo de conservación, cuando se hayan previsto criterios o plazos para la supresión de los datos, el responsable, llegados los mismos, ha de proceder a su eliminación. Piénsese por ejemplo en las bases de datos desactualizadas.

En *segundo lugar*, los datos también se suprimirán cuando el interesado retire el consentimiento y el tratamiento no se base en otro fundamento jurídico. Ya hemos explicado este asunto en otro apartado de la tesis[1149]. Resumidamente cabe indicar que la retirada del consentimiento a estos efectos obligará al responsable del tratamiento a eliminar aquellos datos del interesado que estén almacenados en la base de datos, así como los datos del modelo cuando se pretenda la actualización del mismo. Sin embargo, los datos que se utilizaron para el modelo inicial y sobre los que se basó el consentimiento no se tendrán que eliminar.

En *tercer lugar*, la supresión también se llevará a efecto cuando el interesado se oponga el tratamiento de datos y no prevalezcan otros motivos legítimos o cuando explícitamente se oponga al tratamiento con fines de mercadotecnia. Nos remitimos al análisis realizado en el apartado referido al derecho de oposición[1150].

En *cuarto lugar*, los datos personales deberán suprimirse cuando los mismos se hayan tratado ilícitamente. En estos supuestos, entendemos que tanto los datos personales presentes en las bases de datos como los datos que puedan encontrarse en los modelos habrán de eliminarse. Téngase en cuenta que la supresión de los datos no tiene por qué afectar a toda la base de datos o el modelo en sí. El deber de suprimir solo alcanzará a aquellos datos de los modelos o de los *data lakes* que se hayan obtenido de forma ilegal.

En quinto lugar, los datos presentes en las bases de datos o en los modelos también se deberán suprimir cuando tal obligación se derive de una norma aplicable al responsable del tratamiento.

En *sexto lugar*, la supresión de los datos también alcanzará a los datos que estén seudonimizados, no así a los que estén anonimizados. En el caso de los seudonimizados, el responsable podrá proceder a la anonimización de los mismos a los efectos de evitar la supresión.

[1149] Véase el Capítulo IV, aparatado I, punto 1, epígrafe D) de esta tesis.
[1150] Véase el Capítulo V, apartado II, punto 5 de esta tesis.

B) La supresión de los datos en la fase de despliegue

El derecho de supresión también entra en juego cuando el sistema comienza a adoptar decisiones. Varios supuestos podemos indicar.

En *primer lugar*, el interesado puede solicitar la supresión de los datos inferidos cuando dichos datos ya no sean necesarios para cumplir con la finalidad del tratamiento de datos. Ello sucederá por ejemplo cuando finalice la relación contractual o el servicio en el que se sustentaba el procesamiento algorítmico de los datos. En este sentido, el TJUE, en el *asunto C-434/16 Nowak* ha indicado que el interesado tiene derecho a solicitar al responsable la supresión de las respuestas del examen y las correspondientes anotaciones del examinador una vez haya transcurrido un determinado período de tiempo[1151]. Estas respuestas, como ya dijimos en otro momento de la tesis[1152], son lo más parecido a lo que entendemos por datos inferidos, de manera que la posibilidad de que dichos datos inferidos se puedan suprimir de oficio o petición del interesado resulta viable.

En *segundo lugar*, algo parecido ocurrirá con los perfiles que se hayan elaborado hasta la fecha. Una vez se haya extinguido la relación de un servicio que esté basado en la personalización, tanto los datos iniciales que se recopilaron, como los perfiles creados deberán suprimirse[1153]. Cuando se ejerza este derecho y no se termine con la relación contractual o con el servicio, este derecho les da a las personas el control sobre su perfil en el que se basa la personalización, permitiéndoles por ejemplo restablecer su perfil[1154].

C) Límites al derecho de supresión. La retención ulterior de los datos

El derecho de supresión, al igual que cualquier facultad reconocida a los interesados por el RGPD no está exenta de limitaciones. Es momento de analizarlas situándolas en el contexto del objeto de esta tesis.

[1151] Sentencia del TJUE (Sala Segunda) de 20 de diciembre de 2017, asunto C-434/16, caso Nowak. FJº55. Resolución disponible en: http://curia.europa.eu/juris/document/document.jsf;jsessionid=6C361FE3B54FB5329197ADD52A787D08?text=&docid=198059&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=3496287

[1152] Véase el Capítulo II, apartado II, punto 4.

[1153] Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.26. En el mismo sentido, Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018.Pág.19. En sentido contrario véase: POLITOU, E, ALEPIS,E AND ATSAKIS,C: "Profiling tax and financial behaviour with big data under the GDPR", op.cit., pág.18.

[1154] ESKENS,S: "A right to reset your user profile and more: GDPR-rights for personalized news consumers", op.cit., pág.17.

c.1) Límites establecidos por el RGPD

El artículo 17.3 del RGPD establece explícitamente una serie de restricciones al derecho de supresión. De esta manera, la retención ulterior de los datos personales se entiende lícita en los siguientes supuestos.

En *primer lugar* cuando el tratamiento sea necesario para el ejercicio del derecho a la libertad de expresión e información o para la protección de los mismos. Así, un responsable podría negarse a suprimir los datos de un perfil de una persona que está recibiendo noticias personalizadas alegando que la supresión de dicho perfil puede afectar al funcionamiento del sistema y por tanto a la forma en que ofrece dichas noticias al resto de usuarios[1155]. Recordemos que los sistemas de recomendación de noticias se van retroalimentando con la interacción del conjunto de usuarios que visualizan el contenido ofrecido por dichos algoritmos. En nuestra opinión, si bien es cierto que dicha supresión del perfil podría afectar al sistema, la organización podría seguir ofreciendo ese servicio de personalización de forma adecuada para el resto de usuarios, de manera que, ante la solitud de borrado, el responsable debería aceptarla. Es importante destacar que aquí no se está solicitando que se elimine una determinada noticia, tal y como ocurre con el derecho al olvido, sino que se elimine el perfil que en parte influye en cómo se ofrecen las noticias.

En *segundo lugar*, el responsable se negará a suprimir los datos cuando tenga que cumplir con una obligación legal o con una misión realizada en interés público. Así, por ejemplo, la PRAI permite a los desarrolladores de los sistemas utilizar los datos de categoría especial previstos en el artículo 9 y 10 del RGPD con el objetivo de poder detectar los posibles sesgos que puede presentar el modelo algorítmico[1156]. Por tanto, el responsable del tratamiento tiene la facultad de reservarse determinados datos de categoría especial presentes en sus *data lakes* o bases de datos con esta finalidad. Dicha retención resultará muy relevante sobre todo en relación con determinados datos de colectivos que en el conjunto de datos estén menos representados y la eliminación de los mismos imposibilite evaluaciones posteriores adecuadas. Recordemos además que en virtud del principio de responsabilidad activa el responsable tenía la obligación de

[1155] Vid, pág.17.

[1156] Artículo 10.5 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

monitorear el desempeño de sus sistemas algorítmicos[1157]. Por lo tanto, la necesidad de mantenerlos para futuras evaluaciones también justifica la no supresión de los mismos en base a la normativa de protección de datos. Por otro lado, hay que tener en cuenta que la propia norma puede establecer periodos amplios de conservación de los datos, mientras perduren estos, los responsables no están obligados a eliminarlos. Esto puede ocurrir con bases de datos de las Administraciones Públicas para el desarrollo de sistemas de inteligencia artificial que pretendan luchar contra el fraude fiscal, laboral, blanqueo de capitales, etc. El interés público que reside en estos ámbitos puede justificar la ampliación de conservación de esos datos.

En *tercer lugar*, también se rechazará dicha supresión por razones de interés público en el ámbito de la salud pública o cuando el ejercicio de dicho derecho imposible u obstaculizar gravemente el logro de los objetivos de los tratamientos que tenga como finalidad el archivo en interés público, fines de investigación científica o histórica o fines estadísticos. En estos supuestos, queda patente que el RGPD considera que existen intereses y bienes jurídicos que preponderan respecto del derecho de protección de datos específico de esta persona. Es decir, la utilidad pública de ese dato es superior a los propios intereses particulares del interesado titular del mismo que solicita su supresión[1158]. En estos casos, los plazos de conservación de los datos pueden ser más amplios, para ello, dichos datos deberían estar seudonimizados[1159].

Por último, en *cuarto lugar*, los datos se mantendrán cuando el tratamiento sea necesario para la formulación, el ejercicio o la defensa de reclamaciones. Así, la propuesta de reglamento sobre la responsabilidad civil por el funcionamiento de los sistemas de IA habilita a las organizaciones a utilizar los datos generados por el sistema de IA para demostrar la negligencia concurrente de la persona afectada por un daño

[1157] Capítulo III, apartado X de esta tesis.

[1158] *En el contexto del tratamiento de datos de salud, a pesar de que el paciente sea dueño de sus datos y de que estos se ha obtenido primariamente para la asistencia sanitaria, lo cierto es que concurren intereses generales de enorme relevancia social (en materia de salud pública o de investigación científica) que deben ser atendidos y que justifican el uso de estos datos de salud siempre que, como presupuesto inexcusable, se apliquen las debidas garantía*. En: AUSÍN , T; ANDREU MARTÍNEZ, B; VALERO TORRIJOS, J ; CAYÓN DE LAS CUEVAS, J: "Diez consideraciones ético-jurídicas en relación con la reutilización y big data en el ámbito sanitario". *Bioderecho.es*, (12), 2021, pág.2, consideración cuarta. Texto disponible en: https://doi.org/10.6018/bioderecho.465761

[1159] Véase el Capítulo IV, apartado V, punto 1 de esta tesis.

generado a esta última por dicho sistema[1160].Recordemos además que la LOPD de 2018 prevé para estos supuestos el bloque de los datos[1161].

c.2) Límites derivados de la tecnología

Junto a las restricciones previstas en el RGPD, existen otra serie de limitaciones técnicas propias de los sistemas de tomas de decisiones automatizadas que condiciona el correcto ejercicio del derecho de supresión. En este sentido, a pesar de que la norma no prevea el factor técnico como límite legítimo a esta facultad[1162], consideramos necesario realizar una aproximación al problema actual y planteamos algunas soluciones que pueden resultar útiles.

En *primer lugar*, hemos señalado previamente que en determinados supuestos los responsables del tratamiento deberán suprimir los datos personales que estén presentes en los modelos algorítmicos. La eliminación de estos datos, además de ser compleja para los desarrolladores, resulta poco práctica ya que se requeriría normalmente de un nuevo reentrenamiento de datos[1163]. Para combatir este problema, se han comenzado a diseñar soluciones que permitan a los modelos olvidar los datos de entrenamiento sin necesidad de reentrenamiento. Ejemplo de ello son las funciones de influencia, estas son herramientas de estadísticas sólidas que miden el efecto de un punto de entrenamiento en los parámetros y predicciones del modelo de aprendizaje automático. Específicamente, miden el cambio en la precisión del modelo en una entrada de prueba cuando se elimina un punto del conjunto de entrenamiento. De esta manera, ante una solicitud de borrado de los datos de un particular. El responsable del tratamiento, al tener acceso completo a los parámetros del modelo, puede calcular los nuevos parámetros cuando los datos del usuario se eliminan del conjunto de entrenamiento. Estos nuevos parámetros se pueden calcular fácilmente midiendo la

[1160] Artículo 10.2 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.

[1161] Artículo 32 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Este precepto se analiza en el Capítulo IV, apartado V, punto 1 de esta tesis.

[1162] La doctrina sí que ha considerado que el elemento técnico y el coste que puede suponer le borrado son límites legítimos al derecho de supresión. SANCHO LÓPEZ,M: *Derecho al olvido y big data: Dos realidades convergentes* . Ed. Tirant lo Blanch, Valencia, 2020, págs.192 y 226.

[1163] MITROU,J: "Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?", op.cit.,pág.40.

influencia de los datos del titular de los datos y sumando las cantidades especificadas[1164]. De esta manera, se eliminan los datos y el modelo no pierde precisión.

En *segundo lugar*, el derecho de supresión también puede presentar dificultades a la hora de solicitar el borrado de datos presentes en los *data lakes* o bases de datos. En este sentido, muchas bases de datos están diseñadas para su aprovisionamiento eficiente permitiendo así una búsqueda rápida de los datos contenidos en las mismas[1165]. Dichas bases suelen presentar una serie de características específicas que chocan con la eliminación de datos puntuales. Así, una de las características es la *atomicidad*, de manera que cualquier conjunto de operaciones que se lleva a cabo ha de realizarse en su totalidad. Ejemplo: la inserción o eliminación de un registro de datos debe realizarse para todo el registro al completo. De manera que si se procede a la supresión de uno de los datos de ese registro, se ha de proceder a la eliminación de todo el registro y no sólo del dato específico que se pretende suprimir. Otro requisito es la llamada *durabilidad* de las bases de datos. De manera que los datos deben almacenarse permanentemente en la base de datos, especialmente con el objetivo de detectar errores del sistema o fallos del servidor. La necesidad de compaginar estos requisitos en las bases de datos y el derecho de supresión de los particulares es primordial para los responsables.

En *tercer lugar*, y a pesar de que sea viable la eliminación de datos de los *data lakes* o *data warehouse*, es importante valorar hasta qué punto dicha eliminación de datos puede afectar gravemente al desarrollo de los modelos de esa organización. Así, a priori, la retirada de los datos de un particular prácticamente no tendrá ningún efecto relevante en los posibles desarrollos de modelos futuros[1166]. Esa muestra puede ser sustituida por otras o incluso, la no presencia de la misma no alterará la construcción de nuevos modelos. Ahora bien, si esa persona pertenece a un conjunto de muestras minoritarias que presentan las mismas características, la eliminación de eso datos puede tener efectos muy graves a la hora de representar la realidad que se pretende generar con

---

[1164] SHINTRE,S ; ROUNDY, KA; DHALIWAL, J: "Making Machine Learning Forget". En: NALDI,M; ITALIANO, G; RANNENBERG,K; MEDINA,M; BOURKA,A (eds): *Privacy Technologies and Policy*. APF 2019. Lecture Notes in Computer Science, vol 11498. Ed. Springer, 2019, págs. 76 a 78. Texto disponible en: https://doi.org/10.1007/978-3-030-21752-5_6

[1165] FOSCH VILLARONGA,E; KIESEBERG,P; B; LI,T: "Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten". *Computer Law & Security Review*, Volume 34, Issue 2, April 2018, pág.308. Texto disponible en: ***https://www.sciencedirect.com/science/article/pii/S0267364917302091***

[1166] Según la doctrina, no están claros los efectos del borrado de datos sobre el sistema automatizado. En: ROIG I BATALLA,A: *Las garantías frente a las decisiones automatizadas del Reglamento general de Protección de Datos a la gobernanza algorítmica*.op.,cit,pág.107.

ese modelo[1167]. Aunque es posible que el responsable pueda tratar de obtener nuevas muestras que presenten esa realidad minoritarita, las dificultades aumentarán.

Todos estos supuestos podrán justificar la no supresión de los datos personales. Ello deberá quedar debidamente acreditado.

Para resolver algunos de estos problemas se han propuestos distintas soluciones. Por un lado, se ha señalado la necesidad de apostar por el principio de privacidad desde el diseño tratando de construir algoritmos que sean resistentes al borrado de determinadas entradas de datos. También será relevante que los desarrolladores almacenen muestras amplias y variadas que ayuden a mitigar el daño causado por el borrado de algunos datos fruto del ejercicio del derecho de supresión[1168]. Como dijimos anteriormente, estará justificada la conservación de determinados datos con el objetivo de detectar posibles sesgos presentes en los modelos. El objetivo es que el uso de la tecnología y el ejercicio de los derechos puedan llegar a ser compatibles.

En resumen, el derecho de supresión es una facultad del particular y a la vez un deber del responsable que obliga a este último a dejar de tratar aquellos datos en los supuestos previstos por el RGPD. La supresión podrá alcanzar tanto a los datos personales presentes durante la fase de diseño como a los utilizados en la etapa de toma de decisiones. El alcance de este derecho en relación con otros bienes e interés jurídicos en juego hace que tanto desde el punto de vista legal como tecnológico puedan reconocerse algunas restricciones al ejercicio de esta facultad.

[1167] Imaginemos que se está desarrollando un modelo para aprender a identificar caras en fotografías. Desgraciadamente no ha sido posible obtener una gran muestra de un determinado grupo minoritario pero su tamaño es lo suficientemente grande para generalizar a ese grupo y tenerlo presente a la hora de adoptar decisiones. Aun así, es necesario ajustar el algoritmo para corregir algunas imperfecciones y alimentarlo con el mismo conjunto de datos en el futuro. Sin embargo, mientras ello se realiza, un cierto número de particulares de ese grupo minoritario solicitan el borrado de sus datos y se procede a ello. Con esta diferencia de datos el algoritmo mejorado y el nuevo modelo son incapaces de identificar a las personas de ese grupo minoritario. En: SERGIO CABRAL,T: "Forgetful Al: Al and the Right to Erasure under the GDPR". *European Data Protection Law Review*, 6, no. 3, 2020, págs. 387 y 388.

[1168] Vid, págs 388 y 389.

4. Derecho a la portabilidad de datos

El artículo 20 del RGPD introduce una nueva facultad en favor de los titulares de los datos, esto es, el llamado derecho a la portabilidad de los datos. Este derecho, por un lado permite a los interesados recibir los datos personales que han proporcionado a un responsable del tratamiento en un formato estructurado, de uso común y lectura mecánica, y, por otro lado, trasmitir esos datos a otro responsable del tratamiento sin impedimentos. Son por tanto dos facultades las que están presentes en este derecho. El interesado puede por tanto obtener dichos datos personales y en su caso, decidir si transmitirlos o no a otro responsable. Para que se pueda ejercer este derecho, las operaciones de tratamiento deben basarse o bien en el consentimiento del interesado o bien en un contrato del que el interesado es parte. Artículo 20.1.a del RGPD.

De acuerdo al artículo 20, el derecho a la portabilidad engloba los datos facilitados por el interesado. Para el GT29, la categoría de dato facilitado abarca por un lado a los datos del interesado que proporciona de forma activa, y por otro, los datos observados. Sin embargo, los datos inferidos no pueden quedar englobados en este derecho ya que estos últimos son creados por el responsable del tratamiento sobre la base de los datos facilitados por el interesado y los datos observados[1169]. De esta manera, las inferencias generadas por los algoritmos sobre la base de los datos personales de los interesados no son susceptibles de esta facultad. A pesar de que esta conclusión ha sido acogida y secundada tanto por la doctrina como por las autoridades de protección de datos[1170], ya han aparecido algunas críticas a dicha restricción ya que se limita la portabilidad de un número considerable de datos[1171]. En este sentido, desde la Comisión Europea, y con el objetivo de proporcionar una mayor autonomía y control de sus datos a las personas, se ha indicado que en el futuro el derecho a la portabilidad

[1169] Grupo del Artículo 29. *Directrices sobre el derecho a la portabilidad de los datos. Adoptadas el 13 de diciembre de 2016*. Revisadas por última vez y adoptadas el 5 de abril de 2017. Págs. 11 y 12.
[1170] Agencia Española de Protección de Datos. Informe nº 0195/2017, pág.23. Resolución disponible en: https://www.aepd.es/es/documento/2017-0195.pdf
También en Information Commissioner's Office, disponible en: https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/guidance-on-ai-and-data-protection/how-do-we-ensure-individual-rights-in-our-ai-systems/
[1171] Así lo indicó el Gobierno Neerlandés durante la preparación de la posición del Consejo sobre la evaluación y revisión del Reglamento General de Protección de datos. En: General Secretariat of the Council. *Preparation of the Council position on the evaluation and review of the General Data Protection Regulation (GDPR)*. Resolución del 9 de octubre de 2019, pág.41. Documento disponible en:https://data.consilium.europa.eu/doc/document/ST-12756-2019-REV-1/en/pdf

también alcanzará a los datos inferidos de los interesados[1172]. En nuestra opinión, esta última interpretación permitiría a los interesados poder portar los perfiles e inferencias generados por los sistemas algorítmicos a otras organizaciones, sin embargo, somos conscientes que esta idea presenta serias dudas técnicas.

Vehículo inteligente/Precio póliza de seguro[1173]		**Portabilidad**
Datos facilitados de forma activa	Edad, domicilio, lugar de residencia, documentación técnica del vehículo, puntos del carnet, empleo del conductor, etc.	SI
Datos observados	Conducción del vehículo (patrón de frenado, kilometraje recorrido, aceleración rápida, etc.)	SI
Datos inferidos	Conducción temeraria, mal conductor, buen conductor	NO

Por otro lado, la portabilidad de datos sí que alcanzará a los datos que estén presentes en la base de datos y *data lakes* cuando los mismos no se hayan modificado o alterado. Aquí sobre todo, los datos que podrán ser muy interesantes serán los observados ya que muchos de estos aumentarán conforme incremente al uso de tecnologías basadas en el internet de las cosas. A priori por tanto, un interesado podrá solicitar al responsable los datos presentes en ese *data lake* y en su caso, transferirlos a un tercero. Dicha transferencia requerirá que el formato de los datos objeto de portabilidad sea interoperable. De manera que se pueda facilitar la transferencia de esos datos a ese otro responsable del tratamiento. En este sentido, puede resultar muy relevante en el futuro el papel de los de intermediarios previstos en la propuesta de reglamento sobre la gobernanza de datos cuya función es facilitar los derechos de los interesados de cara a la reutilización de los datos personales[1174]. De esta manera, a

[1172] Explorar el refuerzo del derecho a la portabilidad por parte de las personas de conformidad con el artículo 20 del Reglamento General de Protección de Datos dándoles un mayor control sobre quién puede acceder a los datos generados por las máquinas y utilizarlos es uno de los objetivos de este organismo. En: Comisión Europea. *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*. *Configurar el futuro digital de Europa*, 2020, pág.25.

[1173] Los contratos conocidos como *Pay As You Drive* basan el coste de la prima de seguros en función de cómo esa persona conduce el vehículo. Para ello, es necesario que el sistema esté continuamente recopilando datos relacionados con la conducción del usuario. (datos observados). EN: Comité Europeo de Protección de Datos. *Guidelines 1/2020 on processing personal data in the context of connected vehicles and mobility related applications*. Versión 1.0 Directrices adoptadas el 28 de enero de 2020. Apartados 75 y 103 y págs.16, 21 y 22.

[1174] Esta figura se reconoce en el Considerando 23 y artículo 9.1.b) de la Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos). Texto aprobado el 25 de noviembre de 2020.

través de estos intermediarios se pueden habilitar pasarelas que faciliten esa portabilidad de datos entre organizaciones a petición de los particulares.

5. Derecho de oposición

El derecho de oposición supone la facultad que tiene el interesado para impedir el tratamiento de sus datos personales. Al igual que ocurre con la portabilidad de datos, el derecho de oposición no está contemplado para cualquier tratamiento de datos. Así, son cuatro supuestos donde se habilita a los titulares el ejercicio de esta facultad. Estos son: i) cuando el tratamiento es necesario para el cumplimiento de una misión realizada en interés público o en el ejercicio de poderes públicos conferidos al responsable del tratamiento (artículo 6.1.e), ii) cuando el tratamiento se base en el interés legítimo (artículo 6.1.f), iii) cuando el tratamiento tenga por objeto la mercadotecnia directa y, iv) cuando los datos personales se traten con fines de investigación científica o histórica o fines estadístico.

A) El derecho de oposición a los tratamientos basados en el interés legítimo o misión realizada en interés público

Según el artículo 21.1 del RGPD, el interesado puede oponerse al tratamiento de datos, *incluida la elaboración de perfiles*, por motivos relacionados con su situación particular en todos los casos en los que el tratamiento se base en el artículo 6.1, letras e) o f). De forma expresa, el RGPD hace mención a la elaboración de perfiles como un tratamiento específico sobre el cual, el particular puede oponerse. Esta explicitud obliga al responsable a que cuando deba informar sobre los derechos del interesado, aluda específicamente al derecho de oposición que tienen los interesados con relación a ese concreto tratamiento.

Una vez que el interesado ejercita esta facultad, el responsable ha de interrumpir el tratamiento, en nuestro caso, la elaboración de perfiles o el proceso de inferencias. Ahora bien, el responsable puede negarse a dicha oposición cuando acredite motivos legítimos imperiosos para el tratamiento que prevalezcan sobre los intereses, los derechos y las libertades del interesado. De esta manera, el responsable ha de ponderar por un lado el interés que este o terceros ostentan en continuar con dicho tratamiento y por otro, los efectos que dicho tratamiento genera en los intereses y derechos de los

interesados. Nótese que el ejercicio de ponderación que realiza el responsable, aunque parecido, es distinto al que realizó cuando la elaboración de perfiles o analítica de datos se basó en el interés legítimo o el carácter necesario de la misión de interés público. El RGPD para estos casos exige que el responsable sólo pueda negarse a la oposición del interesado cuando alegue motivos imperiosos. Por tanto, no es suficiente para el responsable del tratamiento demostrar únicamente que su análisis previo sobre el interés legítimo o la misión de interés pública eran correctos. Esta prueba de ponderación exige además que el interés legítimo o la misión pública sean imperiosos, lo cual implica un límite más elevado para ignorar las objeciones[1175].

La ponderación ha de realizarse en virtud de los intereses y derechos en juego de ese concreto tratamiento en el momento que se solicita la oposición por parte del interesado. De esta manera, tanto los intereses del responsable o terceros como los impactos sobre los derechos de los interesados han podido aumentar o descender. Así, por un lado, un responsable podría alegar motivos imperiosos para negarse a paralizar la elaboración de perfiles cuando demuestre que dicho perfil de esa persona es muy relevante o tiene fuertes incidencias en el funcionamiento global del sistema que está adoptando decisiones automatizadas. Ello puede suceder cuando el perfil por ejemplo represente a un determinado colectivo poco representado y la alteración o supresión del mismo puede alterar el modelo o bajar la precisión cuando el sistema adopte decisiones automatizadas sobre personas que también presentan ese perfil. Algo parecido podría ocurrir en relación con la analítica de datos durante la fase de diseño de los modelos algorítmicos y la oposición a la misma. Por otro lado, y al igual que los intereses del responsable pueden aumentar, los impactos en la esfera de los interesados también pueden verse acrecentados. Así, en los entornos adaptativos es frecuente que el sistema cada vez perfile de forma más incisiva a la persona, si dichos perfiles cada vez son más invasivos y además se toman sobre los mismos decisiones que afectan o perjudican al particular, la oposición de este último a dichos tratamientos difícilmente podrá ser rebatida por el responsable alegando motivos imperiosos. De esta manera, la oposición se convierte en un modo de reiniciar el perfilado al que se ven sometidos los particulares. En este sentido, la propuesta de reglamento de servicios digitales obliga a las grandes plataformas que utilicen sistemas de recomendación de contenidos a

[1175] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, págs.20 y 21.

informar sobre el derecho de oposición que asiste a los particulares en materia de protección de datos sobre dichos contenidos y en su caso a facilitar una interfaz de fácil acceso que permita en su caso ejercerlo[1176].

B) El derecho de oposición para tratamientos con fines de mercadotecnia directa

Cuando el tratamiento de datos tenga como finalidad la mercadotecnia directa, el ejercicio del derecho de oposición por parte del interesado obligará al responsable a dejar de tratar los datos para ese fin. En palabras del GT29, el artículo 21.2 reconoce un derecho incondicional a los particulares para oponerse a estos concretos tratamientos. La elaboración de perfiles que tenga como finalidad la mercadotecnia directa finalizará una vez el interesado ejerza su derecho de oposición. A estos efectos, el responsable podrá conservar los datos identificativos del particular necesarios con el fin de impedir tratamientos futuros para fines de mercadotecnia directa. (Artículo 15.2 de la LOPD de 2018)

C) El derecho de oposición para tratamientos con fines de investigación científica, histórica o estadística

De forma muy excepcional, el derecho de oposición se reconoce a los interesados por motivos relacionados con su situación particular cuando los tratamientos tengan como finalidad la investigación científica, histórica o estadística. En estos supuestos, el responsable aún podría oponerse. Por ejemplo, cuando la investigación realizada a través de técnicas de aprendizaje automático esté muy avanzada. A su vez, la oposición no cabrá cuando el tratamiento sea necesario para el cumplimiento de una misión realizada por razones de interés público. En este sentido, en el ámbito del tratamiento de los datos personales con fines de investigación en salud, y en particular la biomédica, la LOPD de 2018 permite a los responsables limitar este y otros derechos cuando: i) dichos derechos se ejerzan directamente ante los investigadores o centros de investigación que empleen datos anonimizados o seudonimizados, ii) cuando el ejercicio de los derechos se refiera a los resultados de la investigación, iii) cuando la investigación tenga por objeto un interés público esencial relacionado con la seguridad

[1176] Véase el artículo 29 de la 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. Resolución adoptada el 15 de diciembre de 2020.

del Estado, la defensa, la seguridad pública u otros objetivos importantes de interés público general, siempre que en este último caso la excepción esté expresamente recogida por una norma con rango de Ley[1177]. En todos estos supuestos, el responsable deberá aludir a la causa o casusas sobre las que se ampara la restricción del derecho de oposición.

III. EL DERECHO A NO SER SOMETIDO A DECISIONES PLENAMENTE AUTOMATIZADAS RELEVANTES

El artículo 22 del RGPD reconoce el derecho a los interesados a no ser objeto de una decisión basada únicamente en el tratamiento automatizado que produzca efectos relevantes. Este derecho quedará limitado cuando el responsable utilice determinadas bases de legitimación que autoricen al tratamiento de dichas decisiones. En estos casos, el responsable deberá establecer toda una serie de garantías específicas en favor de los particulares.

Este derecho forma parte del capítulo III del RGPD referente a los derechos del interesado. Como ya hemos señalado anteriormente, el Tribunal Constitucional ha indicado que los derechos de los interesados forman parte del contenido esencial del derecho fundamental a la protección de datos[1178]. Tanto el derecho a no ser objeto de decisiones plenamente automatizadas relevantes como el conjunto de garantías que se desprenden del mismo y quedan integrado en dicho contenido esencial. Es momento de analizarlo.

1. Ámbito de aplicación y origen de este derecho

El artículo 22 del RGPD establece que *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*. Por tanto, el ámbito de aplicación de este derecho no engloba a toda decisión que adopta un sistema algorítmico sobre una persona sino solamente a algunas de ellas, estas son: i) aquellas decisiones totalmente automatizadas que generen efectos relevantes en el interesado y, ii) la elaboración de perfiles automatizados que den lugar decisiones plenamente automatizadas relevantes. De esta

[1177] Disposición Adicional 17.2.e) de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.
[1178] STC 292/2000 de 30 de noviembre de 2000, FJº 6, STC 290/2000 de 30 de noviembre de 2000, FJ 7º, STC 254/1993 de 20 de julio de 1993, FJº 7.

manera, y como ya comentamos en otro apartado de la tesis, el legislador europeo, al prever un derecho específico sobre este tipo de decisiones denota un recelo importante a la plena automatización de las decisiones que generan efectos relevantes y significativos en los particulares[1179].

Tratamientos incluidos en el artículo 22 del RGPD	**Tratamientos no incluidos en el artículo 22 del RGPD**
-Elaboración de perfiles que den lugar a decisiones relevantes totalmente automatizadas -Decisiones relevantes totalmente automatizadas	-Elaboración de perfiles sean o no relevantes no automatizados por completo -Decisiones parcialmente automatizadas sean o no relevantes

Desde un punto de vista histórico, el origen de este derecho deviene de la *Ley Francesa relativa a la tecnología de la información, los ficheros y las libertades* de 1978[1180]. Así, durante la tramitación de la Directiva 95/46 se hizo hincapié en la necesidad de afrontar los riesgos que podía comportar el uso de grandes bases de datos y de perfilado en la toma de decisiones de los particulares[1181]. Ello dio lugar a la

[1179] Véase el Capítulo II, apartado I, punto 1, epígrafe A) de esta tesis. En ese apartado de la tesis se analiza de forma pormenorizada este concreto tratamiento de datos.
Véase también. COTINO HUESO, L: "Derecho y garantías ante el uso público y privado de inteligencia artificial, robótica y big data". op.cit., pág.932.

[1180] Ley Francesa n ° 78-17 del 6 de enero de 1978 relativa al procesamiento de datos, archivos y libertades (Versión vigente el 23 de julio de 1978). El artículo 2 de esta norma establecía que: *Ninguna decisión judicial que implique una evaluación del comportamiento humano puede basarse en el procesamiento automatizado de información que dé una definición del perfil o la personalidad de la persona en cuestión. Ninguna decisión administrativa o privada que implique una evaluación del comportamiento humano puede tener como única base un tratamiento automatizado de la información que dé una definición del perfil o personalidad del interesado*.
Disponible en: https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000886460/1978-07-23/

[1181] Commission of the European Communities, COM (92) 422 final - SYN 287 Brussels, 15 October 1992. (Págs. 26 y 27). Texto disponible en:
https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:51992PC0422&from=DE
También, la Ley de protección de datos de 1992 española hacía mención a esos riesgos. Así, en el preámbulo de esta norma se venía a decir que: *Los más diversos datos -sobre la infancia, sobre la vida académica, profesional o laboral, sobre los hábitos de vida y consumo, sobre el uso del denominado <dinero plástico>, sobre las relaciones personales o, incluso, sobre las creencias religiosas e ideologías, por poner solo algunos ejemplos- relativos a las personas podrían ser, así, compilados y obtenidos sin dificultad. Ello permitiría a quien dispusiese de ellos acceder a un conocimiento cabal de actitudes, hechos o pautas de comportamiento que, sin duda, pertenecen a la esfera privada de las personas; a aquélla a la que sólo deben tener acceso el individuo y, quizás, quienes le son más próximos, o aquellos a los que él autorice. Aún más:* ***El conocimiento ordenado de esos datos puede dibujar un determinado perfil de la persona****,* ***o configurar una determinada reputación*** *o fama* ***que*** *es, en definitiva, expresión del honor; y este perfil, sin duda,* ***puede resultar luego valorado, favorable o desfavorablemente, para las más diversas actividades públicas o privadas****, como pueden ser la obtención de un empleo, la concesión de un préstamo o la admisión en determinados colectivos.* (la cursiva y negrita son nuestras).

redacción del artículo 15 de la mencionada Directiva, el cual, reconoció una serie de garantías en favor de los particulares que se veían sometidos a estas decisiones[1182]. Pues bien, la aplicación práctica de este artículo o los preceptos que han transpuesto el mismo a los ordenamientos jurídicos nacionales ha resultado hasta la fecha muy débil. Así, en pocas ocasiones los tribunales y las autoridades de protección de datos tanto europeas como nacionales han tenido la oportunidad de analizar este precepto[1183]. Ello en parte puede deberse a que hasta la fecha, pese a que la elaboración de perfiles y toma de decisiones automatizadas se viene realizando, ha sido en los últimos años cuando las organizaciones tanto públicas como privadas empiezan a generalizar estos tratamientos en sus procesos decisorios. En este sentido, y fruto de esa *tormenta perfecta* a la que aludíamos al inicio de esta tesis[1184], será cada vez más frecuente el uso de sistemas de toma de decisiones automatizadas y por consiguiente, la activación de este derecho[1185].

Reconocimiento de este derecho en otros textos europeos e internacionales.	
Norma	**Artículo**
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.	14

Apartado primero del Preámbulo de Ley Orgánica 5/1992, de 29 de octubre, de regulación del tratamiento automatizado de los datos de carácter personal.

[1182] Artículo 15 de la Directiva 95/46 del Parlamento Europeo y del Consejo, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respeta al tratamiento de datos personales y a la libre circulación de estos datos establece que:

1. *Los Estados miembros reconocerán a las personas el derecho a no verse sometidas a una decisión con efectos jurídicos sobre ellas o que les afecte de manera significativa, que se base únicamente en un tratamiento automatizado de datos destinado a evaluar determinados aspectos de su personalidad, como su rendimiento laboral, crédito, fiabilidad, conducta, etc.*

2. Los Estados miembros permitirán, sin perjuicio de lo dispuesto en los demás artículos de la presente Directiva, que una persona pueda verse sometida a una de las decisiones contempladas en el apartado 1 cuando dicha decisión:

a) se haya adoptado en el marco de la celebración o ejecución de un contrato, siempre que la petición de celebración o ejecución del contrato presentada por el interesado se haya satisfecho o que existan medidas apropiadas, como la posibilidad de defender su punto de vista, para la salvaguardia de su interés legítimo; o b) esté autorizada por una ley que establezca medidas que garanticen el interés legítimo del interesado.

[1183] En España, existe un número muy limitado de resoluciones judiciales que han aplicado el artículo 13 de la antigua Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal. (Ya derogada). A través de este precepto se traspuso a nuestro ordenamiento jurídico el artículo 15 de la Directiva 95/46. Véase: Auto AP Madrid (Sección 5ª), nº1176/2006 de 21 de marzo y Auto AP Ciudad Real (Sección1), núm.43/2005, de 7 de marzo.

[1184] Véase el Capítulo I, apartado I, punto 4 de esta tesis.

[1185] A diferencia de sus predecesores, el artículo 22 del RGPD se ha analizado ya en varias resoluciones tanto judiciales como administrativas, si bien, normalmente se considera que las decisiones automatizadas no entran dentro del ámbito de aplicación del mencionado precepto porque o bien, no son plenamente automatizadas o bien, no son relevantes. Véase por ejemplo la sentencia del Tribunal de Ámsterdam. Resolución de 11 de marzo de 2021. Apartado 4.25. Disponible en:
https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBAMS:2021:1018
En el mismo sentido véase: Agencia Española de Protección de Datos. Resolución nº PS/00037/2020, pág.117. Resolución disponible en: https://www.aepd.es/es/documento/ps-00037-2020.pdf

Reglamento (UE) 2018/1725 del parlamento europeo y del consejo de 23 de octubre de 2018 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones, órganos y organismos de la Unión.	24
Actualización del Convenio para la Protección de las Personas con respecto al tratamiento de datos personales de datos personales. Consejo de Europa. Convenio 108 +. Disponible en: https://rm.coe.int/convention-108-convention-for-the-protection-of-individuals-with-regar/16808b36f1	9.1.a)
Proyecto de ley de protección de datos de Argentina. Proyecto 147/2018. Disponible en: https://www.argentina.gob.ar/sites/default/files/mensaje_ndeg_147-2018_datos_personales.pdf	32
Ley nº 13.709, de 14 de agosto de 2018 de Brasil. Disponible en: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm	20

2. Prohibición general de este tratamiento. Las excepciones que legitiman el tratamiento

Además de un derecho, el GT29 ha indicado que este precepto también contiene un prohibición general a llevar a cabo tratamientos basados en las decisiones plenamente automatizadas relevantes[1186]. Por tanto, a priori, las organizaciones tanto públicas como privadas tienen prohibido llevar a cabo este tipo de tratamientos de datos. De esta manera, este derecho entra en juego tanto si el interesado lo ejercita como si no lo hace. A pesar de que esta prohibición general se aplica por defecto[1187], el artículo 22.2 del RGPD establece varias excepciones que habilitan a dichas decisiones automatizadas, estas son: i) cuando la decisión que se toma es necesaria para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento; ii) cuando la decisión esté autorizada por el Derecho de la Unión o de los Estados miembros, y iii) cuando la decisión se basa en el consentimiento explícito del interesado. Además, cuando las decisiones se basen en los datos de categoría especial previstos en el artículo 9 del RGPD, sólo cabra este tipo de tratamiento cuando el interesado otorgue el consentimiento explícito o exista un interés público esencial en la toma de decisiones automatizadas. Artículo 22.4 del RGPD.

Pues bien, cuando un responsable acuda a estas excepciones para legitimar la toma de decisiones plenamente automatizadas relevantes, este deberá prever suficientes medidas de garantías con el objetivo de salvaguardar los derechos, libertades y los intereses legítimos del interesado.

[1186] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679.*Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.21.

[1187] HERNÁNDEZ, J,C: "Decisiones algorítmicas de perfilado. Régimen y garantías jurídicas." op.cit., pág. 313.

Dado que el RGPD prevé diversas medidas de garantía en función del tipo de excepción sobre la cual el responsable autoriza el tratamiento de datos, en los siguientes apartados procederemos a dicha distinción.

3. Garantías en favor de los interesados cuando el tratamiento se base en el consentimiento explícito o el contrato

Cuando el responsable fundamente la toma de decisiones automatizadas en el consentimiento explícito del interesado o en la celebración o ejecución de un contrato este estará obligado a implantar toda una serie de salvaguardas relacionadas con el proceso decisorio[1188]. Como mínimo, tras la decisión algorítmica adoptada, deberá facilitar al interesado el derecho a obtener intervención humana, el derecho a que dicho particular exprese su punto de vista y además, este último, también podrá impugnar dicha decisión. (Artículo 22.3 del RGPD). A su vez, y de acuerdo al considerando 71 del RGPD, el responsable estará obligado a facilitar una explicación sobre la decisión adoptada por el sistema automatizado.

Es turno de analizar cada una de estas facultades y las implicaciones de estas en el contexto de la toma de decisiones automatizadas.

A) Derecho de explicación

a.1) El reconocimiento de este derecho

A diferencia de otras facultades, el derecho de explicación en favor del interesado sobre el que se toma una decisión algorítmica no aparece expresamente mencionado en el artículo 22.3 del RGPD. Así, esta garantía únicamente se menciona en el considerando 71. Ello ha llevado a parte de la doctrina a concluir que dado que los considerandos no son vinculantes para los destinatarios de la norma, el derecho de explicación como tal no es exigible a los responsables que utilicen sistemas algorítmicos[1189]. Estas dudas han sido disipadas por parte de diversas autoridades de

[1188] Sobre la licitud del consentimiento explícito y el carácter necesario en el uso de sistemas de toma de decisiones plenamente automatizadas relevantes véase el capítulo IV de esta tesis. Concretamente los siguientes apartados de este capítulo: consentimiento explícito (Capítulo IV, apartado I, punto 1, epígrafe C) y ejecución del contrato (Capítulo IV, apartado I, punto 2, epígrafe B).

[1189] En el lado de los que consideran que no existe un derecho de explicación: WACHTER,S; MITTELSTADT,B; FLORIDI,L: "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation". *International Data Privacy Law*, Volume 7, Issue 2, 1 May 2017, págs. 76 a 99. Texto disponible en: https://doi.org/10.1093/idpl/ipx005. Por otro lado, los que sí que apuestan por un derecho de explicación encontramos: GOODMAN,B; FLAXMAN,S: "EU Regulations on

protección de datos, las cuales, cuando han tenido la oportunidad de pronunciarse, han confirmado la existencia de este derecho. En nuestra opinión, dada la rotundidad manifestada por estos organismos, el derecho de explicación es una facultad que se reconoce a los interesados sometidos a estas decisiones y por tanto, sino se garantiza por parte del responsable, ello supondrá un incumplimiento del RGPD. Y es que, el artículo 22.2 hace referencia a las garantías mínimas que se han de exigir. Sin embargo, el responsable habrá de establecer otras tantas, entre las cuales, ha de estar presente dicho derecho de explicación.

Lo reconoce el Consejo de Europa. (apartado 5.7)

https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680a46147

Reconocimiento expreso del derecho de explicación	
Autoridad de control[1190]	**Año**
GT29 y CEPD	2018 y 2021
Autoridad de Protección de Datos Noruega	2018
Information Commissioner´s Office	2020
Agencia Española de Protección de Datos	2019 y 2021
Textos de estados europeos que lo han implantado[1191]	**Artículo**
Ley n° 78-17 del 6 de enero de 1978 relativa al procesamiento de datos, archivos y libertades. (Francia)	47.2
Ley de Autodeterminación del Derecho a la Información y Libertad de Información. (Hungría)	5.6.ba

a.2) El contenido del derecho

Algorithmic Decision-Making and a right to Explanation" . Texto disponible en: http://arxiv.org/abs/1606.08813. Véase en el mismo sentido: NÚÑEZ SEOANE,J: "El derecho de la información y acceso al funcionamiento de los algoritmos que tratan datos personales". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed. Aranzadi, Navarra, 2020, pág.308.

[1190] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.30. También en: Comité Europeo de Protección de Datos. *EDPB-EDPS. 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).* Resolución de 18 de junio de 2021. Apartado 60, pág.17. A su vez: The Norwegian Data Protection Authority. *Artificial intelligence and privacy*. 2018, pág..21. También: Information Commissioner's Office. *Explaining decisions made with AI.*,2020, pág13. Por su parte, en el mismo sentido: Agencia Española de Protección de Datos. *La protección de datos en las relaciones laborales*, 2021, pág.24. También en la resolución de este mismo órgano Nº: PS/00070/2019, pág.109. Resolución disponible en: https://www.aepd.es/es/documento/ps-00070-2019.pdf

[1191] Artículo 47.2 de la Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés
Disponible en: https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000886460/
Artículo 5.6 apartado ab) de la Törvény az információs önrendelkezési jogról és az információszabadságról. Disponible en: https://net.jogtar.hu/jogszabaly?docid=a1100112.tv

Una vez que el sistema ha adoptado la decisión automatizada, el particular tiene derecho a solicitar una explicación sobre los motivos o razones que han dado lugar a la misma. Señalamos algunos elementos que puede formar parte del contenido de la explicación:

En *primer lugar*, el responsable ha de informar al interesado sobre los datos de entrada al algoritmo que se utilizaron para la decisión, tanto los personales como no personales. Aquí no sólo se incluirán los datos directamente facilitados por el interesado, sino también los observados durante todo el proceso que abarca el tratamiento de datos y que ha podido influir en la decisión. Además, también será relevante hacer mención a los posibles datos inferidos que el sistema haya generado.

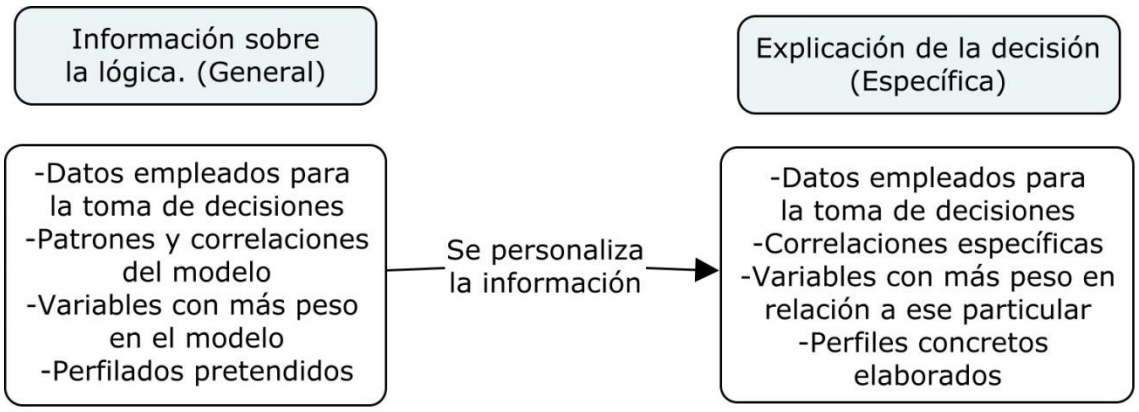
En *segundo lugar*, se ha de informar de las concretas variables o parámetros que han sido más relevantes en esa concreta decisión. Es decir, se ha de informar al particular sobre los datos que han podido tener mayor influencia en el proceso decisorio del algoritmo.

En *tercer lugar*, se ha de informar de cómo las correlaciones generales que se documentaron durante la fase de diseño del sistema afectan específicamente ahora al particular. La idea es que se pueda valorar y explicar cómo el sistema ha funcionado y operado respecto de ese interesado en concreto.

En *cuarto lugar*, se ha de informar al particular sobre la posibilidad de que la decisión pueda ser revisada por un humano en el caso de que el interesado pretenda impugnarla.

Como se puede comprobar, gran parte del contenido que abarca la explicación de la decisión algorítmica ha podido facilitarse a través de los derechos de información contemplados en los artículos 13 y 14 del RGPD. Sin embargo, esa información que se suministró inicialmente de forma general, ahora, y una vez que el sistema ha adoptado la decisión, es personalizada para indicar las causas y razones concretas del funcionamiento del sistema para ese particular. En este sentido, la propuesta de reglamento de servicios digitales obliga a los prestadores de servicios de alojamiento de datos a que, cuando utilicen sistemas automatizados para bloquear o retirar el contenido compartido por sus usuarios, indiquen que dicho contenido ha sido eliminado por un

algoritmo y además señalen los hechos y circunstancias en que se ha basado la adopción de la decisión[1192].



Elaboración propia

a.3) Elementos formales a la hora de explicar la decisión

A la hora de proceder a la explicación de la decisión, el responsable no está sujeto a ninguna obligación formal. Así, el artículo 12.1 del RGPD únicamente dispone que cualquier comunicación con arreglo a los artículos 15 a 22 relativa al tratamiento de datos se realizará de forma concisa, transparente, inteligible y de fácil acceso, con un lenguaje claro y sencillo. Trasladado a nuestro contexto, el responsable ha de explicar las razones de tal manera que una persona que no sea experta en la materia las pueda comprender[1193]. Es necesario que las organizaciones puedan traducir el lenguaje algorítmico a un formato asequible para un ciudadano medio. Y es que, en la mayoría de las ocasiones al particular afectado por la decisión le interesará conocer la motivación de la decisión y no los elementos técnicos que en su caso la han generado[1194].

Para facilitar esta comprensión tanto la doctrina como distintas autoridades administrativas han hecho referencia a distintas herramientas.

En *primer lugar*, un mecanismo que puede resultar adecuado son las llamadas explicaciones contra fácticas. Estas explicaciones consisten en indicar a los interesados los factores y variables que debería ostentar o haber aportado el particular para que la

[1192] Véase el artículo 15 de la 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. Resolución adoptada el 15 de diciembre de 2020.

[1193] Australian Human Rights Commission. *Human rights and Technology. Discussion Pape*r, 2019. pág.96.

[1194] Lo importante no es saber lo que pasa por la cabeza de un juez o de funcionario cuando adopta su decisión sino que la decisión cuente con motivación suficiente y ajustada a Derecho. COTINO HUESO, L: "Derecho y garantías ante el uso público y privado de inteligencia artificial, robótica y big data", op.cit., pág.948.

decisión hubiera sido favorable al mismo[1195]. Es decir, se informa al interesado de las circunstancias que debería alterar para que el algoritmo adopte una decisión diferente. Por ejemplo, si una entidad bancaria deniega a una persona el crédito solicitado de forma automatizada, las explicaciones contra fácticas obligarán al responsable a informar de aquellas variables específicas que tendría que alterar dicho particular para que en un futuro pueda obtener el préstamo rechazado por el algoritmo. Así, algunas organizaciones ya han implantado este mecanismo creando una especie de avatar o gemelo digital del perfil sobre la persona que se toma la decisión. Seguidamente, se van realizando pequeñas variaciones en las variables que se introducen para llegar al momento en el que la decisión que adopta el sistema favorece al interesado[1196]. Posteriormente se le informa al interesado sobre ese proceso para que en su caso pueda alterar la situación que ha derivado en la decisión algorítmica. Este mecanismo resulta muy útil para el interesado sometido a estas decisiones cuando sea fácil alterar las variables que han dado lugar a dicha decisión negativa. En este sentido, es conveniente que se generen múltiples explicaciones contra fácticas para que el sujeto sometido a la decisión tenga acceso a diversas formas viables de generar un resultado diferente[1197]. Ello es importante ya que, dado que existen datos e información que muy difícilmente podrán alterarse, la posibilidad de que estas personas queden condenadas a decisiones perjudiciales aumenta[1198]. Así, será relativamente fácil para el particular alterar datos relacionados con su conducta o comportamiento. Sin embargo, otros datos como los genéticos, los físicos o el domicilio resultarán más complejos o imposibles de alterar.

En *segundo lugar*, desde el mundo de la ciencia de datos se han ido desarrollando herramientas y estrategias que fomenten la llamada inteligencia artificial explicable. El objetivo de estos instrumentos es el de facilitar explicaciones locales relacionadas con el funcionamiento del algoritmo para el caso concreto de la persona

---

[1195] Véase por todos: WACHTER,S ; MITTELSTADT,B; RUSSELL, CH: "Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR". *Harvard Journal of Law & Technology*, Volume 31, Number 2, 2018, págs. 842 a 887.

[1196] Los llamados modelos predictivos contrafactuales. Fuente de la noticia: DARLINGTON,K: "Sistemas de IA explicables: comprender las decisiones de las máquinas". *OpenMind.BBVA*. 11/10/2017. Disponible: https://www.bbva.com/ndb/es/articulo/que-es-la-ia-explicable-xai-y-por-que-es-mas-necesaria-que-nunca/

[1197] MOLNAR,CH*: Interpretable Machine Learning. A Guide for Making Black Box Models Explainable*.. ISBN 9780244768522. Apartado 2.6.2 y apartado 6.1. Libro disponible on line. https://christophm.github.io/interpretable-ml-book/explanation.html

[1198] SELBST,A & BAROCAS,S: "The Intuitive Appeal of Explainable Machines",op.cit.,pág. 1122..

sobre la que se adopta la decisión[1199]. Por ejemplo, a través de las técnicas *shap values* es posible indicar la importancia que han tenido las distintas variables del algoritmo para un caso individual. Algo parecido ocurre con las explicaciones a través de la solución *LIME*. En este caso, a través de estas técnicas se puede llegar a conocer el funcionamiento del modelo para un concreto individuo[1200].

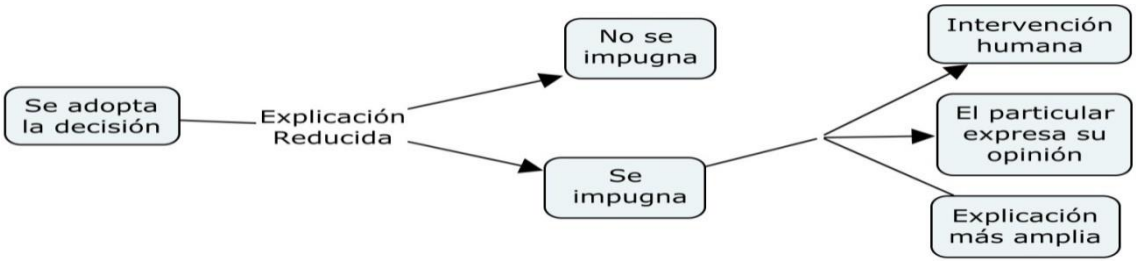
Independientemente de la técnica o herramienta que se utilice para favorecer la explicabilidad de la decisión, los responsables están obligados a desplegar modelos algorítmicos que logren satisfacer adecuadamente el derecho de explicación reconocido a los particulares[1201]. Es por ello que los responsables han de acudir a modelos interpretables o en su caso aplicar técnicas que favorezcan dicha explicabilidad. En este sentido, cabe recordar que la PRAI obliga a los desarrolladores a diseñar sistemas lo suficientemente transparentes que permitan a las organizaciones/usuarios que los despliegan interpretar los resultados del sistema y utilizarlos adecuadamente[1202]. Ello facilitará el ejercicio del derecho de explicación analizado.

Para finalizar, cabe señalar que como es lógico, y al igual que cualquier otro derecho, el derecho de explicación también puede restringirse cuando entre en conflicto con otros bienes e intereses en juego. Nos remitimos a los apartados de la tesis donde hemos analizado los límites a los deberes de información. Las conclusiones indicadas en esas líneas son perfectamente trasladables a esta facultad[1203]. De esta manera, cualquier restricción a este derecho implicará una adecuada justificación y además el despliegue de otras garantías que compensen ese déficit[1204]. Estas limitaciones en ningún caso supondrán la plena anulación del mencionado derecho.

a.4) Fase temporal en la que se ha de proceder a la explicación

[1199]Information Commissioner's Office. *Explaining decisions made with AI | Part 2: Explaining AI in practice*, pág.50.
[1200] TULIO RIBEIRO,M; SINGH,S ; GUESTRIN,C: "Why Should I Trust You? Explaining the Predictions of Any Classifier". ArXiv: 1602.04938v3, 2016. Texto disponible en: https://arxiv.org/pdf/1602.04938.pdf
[1201] Australian Human Rights Commission. *Human rights and Technology. Discussion Pape*r, 2019, pág.97.
[1202] Artículo 13.1 de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.
[1203] Véase el capítulo IV, apartado 6, punto 3, epígrafe D) de esta tesis.
[1204] Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Directrices éticas para una IA fiable*.Pág.16.

No existe una regla específica que indique en qué momento se ha de proceder a explicar la decisión al interesado. Corresponde a los responsables valorar la forma de integrar este derecho tras la adopción de la decisión. En nuestra opinión, sería conveniente que una vez el sistema ha tomado la decisión, por defecto se muestre una explicación más o menos extensa. Posteriormente, si el particular no queda satisfecho con esa primera aproximación explicativa, este pueda solicitar una explicación más pormenorizada en la cual se habiliten el resto de facultades que prevé el RGPD, esto es: el derecho a la intervención humana, la posibilidad de expresar su punto de vista y el derecho a impugnar la decisión. Aunque pensamos que la explicación mostrada de oficio por parte del responsable es conveniente pero no obligatoria , el responsable, tras la adopción de la decisión sí que deberá informar claramente sobre el derecho que asiste a los particulares para recibir dicha explicación de la decisión, así como la supervisión humana, la posibilidad de expresar su punto de vista, impugnarla, etc. A partir de ahí, el particular deberá decidir si ejercita o no esos derechos.



Elaboración propia

En conclusión, el derecho de explicación es una facultad que ostentan los particulares que se ven sometidos a la toma de decisiones automatizadas relevantes. Este derecho obliga al responsable a prever cauces específicos que faciliten la explicabilidad de la decisión. A pesar de que este derecho no esté reconocido expresamente en el texto vinculante del RGPD, una interpretación sistemática de esta norma junto al reconocimiento expreso por parte de distintas autoridades de protección de datos nos lleva a pensar que tal facultad es plenamente exigible a los responsables que implanten sistemas automatizados en sus procesos decisorios. Este derecho además, con el paso del tiempo será posiblemente *absorbido* por el resto de normas que vayan imponiendo o exigiendo la explicabilidad de las decisiones en distintos sectores.

B) La audiencia del interesado

El artículo 22.2 del RGPD reconoce expresamente un conjunto de facultades en favor del interesado cuando el responsable adopta decisiones plenamente relevantes. Estas son: el derecho a obtener intervención humana por parte del responsable, el derecho del particular a expresar su punto de vista y el derecho a impugnar la decisión. Todas estas facultades pueden englobarse en nuestra opinión en un derecho general a través del cual el responsable ha de establecer los mecanismos adecuados que faciliten un trámite de audiencia donde el interesado pueda: i) impugnar la decisión adoptada por el sistema, ii) tener contacto con alguna persona perteneciente a la organización que está utilizando el algoritmo, y, iii) poder expresar su punto de vista con relación a la decisión adoptada. En las siguientes líneas se analizarán cada una de estas facultades de forma específica con el objetivo de facilitar su comprensión. El estudio de las mismas no responde a ninguna forma secuencial o temporal predefinida de cómo estas han de presentarse por parte del responsable del tratamiento. El despliegue de estas se habrá de adaptar al contexto dónde cada organización implante el sistema de toma de decisiones. Lo que sí que está claro es que las mismas han de estar previstas[1205].

b.1) El derecho a impugnar la decisión

Una vez que el sistema ha adoptado la decisión, el particular tiene derecho a poder impugnarla. Así, se deben habilitar instrumentos adecuados a la hora de recurrir las decisiones que toma el sistema debido a que en muchos casos será la única alarma que avise o que ponga en entredicho que un sistema puede estar tomando decisiones inadecuadas. No facilitar un cauce de impugnación deja a los interesados en una situación de desprotección ante las decisiones incorrectas. En este sentido, a modo de ejemplo, en mayo de 2018 el gobierno del Reino Unido utilizó un software de reconocimiento de voz para detectar si las personas que se presentaban a dicha prueba lo eran realmente. Esta prueba la realizaban estudiantes extranjeros y lo que se evaluaba era su nivel de inglés. El sistema erró en más del 20% de los casos lo que supuso que alrededor de unas 7000 personas fueran acusadas de hacer trampa en dichos exámenes[1206]. Muchos de estos estudiantes extranjeros fueron expulsados del Reino

[1205] También así lo contempla la Carta de Derechos Digitales, el artículo 25.3 de dicho texto establece que: *las personas tienen derecho a solicitar una supervisión e intervención humana y a impugnar las decisiones automatizadas tomadas por sistemas de inteligencia artificial que produzcan efectos en su esfera personal y patrimonial.*

[1206] Fuente de la noticia: BULMAN,M: "Home Office failed to ensure innocent students were not wrongly detained in cheating scandal, report finds". *Independent*. 5/8/2020. Noticia disponible en:

Unido sin que se previera ningún cauce para impugnar la decisión. La posibilidad de impugnar los resultados algorítmicos de los sistemas se muestra fundamental.

Para que el proceso de impugnación sea adecuado, es recomendable que el responsable una vez que el sistema ha tomado la decisión, indique claramente al particular la posibilidad de poder impugnar la mentada decisión y además se expliciten los cauces específicos para proceder a la misma. En nuestra opinión, esta información podría incluir: i) plazos con los que cuenta el particular para impugnar la decisión[1207], ii) órgano administrativo o departamento ante el que se ha de presentar la reclamación, iii) cuando lo haya, datos de contacto del delegado de protección de datos, y a ser posible, iv) una explicación más o menos prolija de las razones o motivos principales que han dado lugar a la decisión. Así, la propuesta de reglamento europeo de servicios digitales obliga a las plataformas en línea a implantar un sistema interno de tramitación de las reclamaciones frente a la retirada de contenido por parte de los sistemas automatizados[1208]. La idea es que los usuarios de estas plataformas tengan medios adecuados para poder impugnar las decisiones automatizadas que les afectan en el seno de estas plataformas. La impugnación de la decisión, como es lógico, activará como mínimo las otras dos facultades previstas por el artículo 22.2, esto es, el derecho a la intervención humana y el derecho del interesado a expresar su punto de vista.

b.2) El derecho a obtener intervención humana del responsable y el derecho del particular a expresar su punto de vista ante la decisión tomada

El *derecho a obtener la intervención humana* del responsable del tratamiento tras la adopción de la decisión automatizada se presenta como una de las garantías más relevantes que ostenta el interesado. Así, en el origen de este derecho se vislumbra la necesidad de que ante una decisión plenamente automatizada, pueda existir una apreciación humana por parte del responsable[1209]. Visto así, este derecho no deja de ser

---

https://www.independent.co.uk/news/uk/home-news/home-office-students-english-language-tests-toeic-detained-deported-theresa-may-a8927781.html

[1207] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.37.

[1208] Artículo 17 de la 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. Texto aprobado el 15 de diciembre de 2020.

[1209] Por primera vez se hace mención a este derecho en la propuesta modificada de DIRECTIVA DEL CONSEJO relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación. Bruselas, 15 de octubre de 1992, pág.28. Texto disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:51992PC0422&from=EN

una respuesta lógica a la necesidad de que en algún momento del proceso decisorio automatizado se incorpore la presencia del humano[1210]. Concretamente, para el RGPD, esa aparición se produce una vez se ha adoptado la decisión y la misma ha generado efectos en los particulares. A través de este derecho, algún miembro de la organización que ha implantado el algoritmo puede valorar los posibles errores que dicho sistema puede generar. Aunque a priori, este derecho se activa una vez que lo ejercita el interesado afectado, el propio responsable de oficio puede establecer un procedimiento específico ante resultados discrepantes en relación con el comportamiento esperado del sistema[1211]. Tal y como ha indicado el Consejo de Estado italiano, en el proceso de toma de decisiones debe existir un aporte humano capaz de controlar, validar o negar la decisión automática[1212]. Esta presencia humana resulta muy relevante en aquellos contextos donde intencionadamente, y a pesar de los errores que puede generar el sistema, este último se introduce ya que los beneficios que reporta el mismo, pese a sus fallos, son mayores que los perjuicios generados a los particulares por dichas inexactitudes. Ello suele ocurrir en el contexto de las plataformas en línea y sus sistemas de control del contenido. Un claro ejemplo de ello quedó patente tras la eliminación de miles de videos de YouTube que documentaban atrocidades producidas en la guerra de Siria. Estos videos evidenciaban importantes violaciones de los derechos humanos. Sin embargo, el sistema automatizado de YouTube los eliminó al considerar que se estaba emitiendo contenido que incitaba al terrorismo[1213]. Pues bien, a pesar de que este cribado de contenido resulta legítimo, la necesidad de que se pueda revisar el mismo por parte de un humano resulta fundamental para enmendar los errores generados por los sistemas algorítmicos. En este sentido, tanto la Directiva de derechos de autor como la propuesta de reglamento europea de servicios digitales apuestan por la presencia del humano durante el proceso de revisión de la decisión algorítmica[1214]. Para finalizar, es

[1210] BURK,L,D: "Algorithmic Fair Use", op.cit.,pág.297.
[1211] Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág.30.
[1212] Sentencia Consejo de Estado italiano con la sentencia 13 diciembre 2019, n. 8472. Fundamento Jurídico 15.2. Disponible en: https://www.giustizia-amministrativa.it/provvedimenti-cds
[1213] MACDONALD,S; GIRO CORREIA,S; WATKIN, A-L: "Regulating terrorist content on social media: automation and the rule of law". *International Journal of Law*, 2019 , pág.190.
[1214] Artículo 17.9 de 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 y artículo 17.5 de la 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.

importante señalar que la persona de la organización que revise la decisión ha de contar con la autorización y capacidad adecuada para poder modificarla.

Junto al derecho a obtener intervención humana, el RGPD también reconoce el *derecho del particular a expresar su punto de vista*. Esta facultad está estrechamente unida a las que previamente hemos analizado. A través de la misma, el particular puede indicar al responsable las razones o motivos por los que considera que esa decisión no es adecuada. Como es lógico, el particular ha de tener la suficiente información sobre la decisión que se ha tomado para poder argumentar adecuadamente sus opiniones[1215]. De esta manera, el particular, para rebatir esa decisión, puede aportar más información y documentos que resulten relevantes para la reconsideración de la decisión. A su vez, en este momento, el particular también puede comprobar si todos los datos con los que dispone el responsable y que se han utilizado para la adopción de la decisión son exactos y actualizados. Y es que no podemos olvidar que a través de los modelos algorítmicos se tratan de generalizar distintas realidades, sin embargo, cada persona ostenta sus propias particularidades que pueden diferenciarla del resto. Esas particularidades pueden resultar esenciales para que el responsable recapacite sobre la decisión que ha adoptado el algoritmo. Por tanto, el responsable está obligado a valorar en su caso otras informaciones que exceden de las tenidas en cuenta por el algoritmo y que pueden ser relevantes para alterar la decisión[1216]. Este derecho también puede ayudar al responsable a filtrar los falsos positivos y negativos del sistema ya que si la organización acaba reconsiderando la decisión, habrá que entender que el sistema ha errado en la adopción de la misma y por tanto[1217], ha de valorar si es necesario reconfigurar o no el sistema.

Como puede comprobarse por tanto, el derecho de explicación, la impugnación de la decisión, junto con el derecho a la intervención humana y la facultad para que el interesado intervenga son garantías que por su propia naturaleza irán de la mano una vez que se adopte la decisión. Todo el proceso de audiencia del interesado pensamos que se

[1215] Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.30.

[1216] L,JANSSEN,H: "An approach for a fundamental rights impact assessment to automated decision-making", op.cit., pág.93.

[1217] ROIG, A: "Safeguards for the right not to be subject to a decision based solely on automated processing (Article 22 GDPR)". *European Journal of Law and Technology*, Vol 8, No 3, 2017, pág.6. Texto disponible en: https://core.ac.uk/download/pdf/189882259.pdf

ha de realizar en el seno de la propia organización que ha adoptado la decisión. En este sentido, conviene se señalar que la antigua Ley orgánica de Protección de Datos española de 1999 reconocía a los particulares el derecho de impugnación de los actos administrativos o decisiones privadas que implicaran una valoración de su comportamiento a través de medios exclusivamente automatizados[1218]. Este precepto llevó a parte de la doctrina española a considerar que se estaba reconociendo en favor del particular afectado una acción procesal directa para impugnar la decisión ante el orden jurisdiccional competente[1219]. Sin embargo, creemos que realmente, el artículo 22.3 del RGPD, cuando hace referencia a este conjunto de facultades y en concreto al derecho de impugnación, se está refiriendo a la posibilidad que tiene el particular afectado para recurrir la decisión ante el responsable en un contexto previo a la vía judicial en el cual se le expliquen las razones de esa decisión. A su vez, en ese proceso previo, el interesado puede expresar su punto de vista con la finalidad de revertir la decisión sin necesidad de acudir a la vía judicial. Todo ello, sin perjuicio de que posteriormente y una vez ejercidas el elenco de facultades descritas en este precepto, el afectado pueda emprender las acciones judiciales o en su caso acudir a la autoridad de control correspondiente.

C) Otras salvaguardas y medidas de garantía necesarias

Junto a las facultades descritas previamente, el responsable además habrá de desplegar toda otra serie de garantías que favorezcan o aseguren un uso adecuado de los sistemas de toma de decisiones automatizadas relevantes. El mero reconocimiento de los derechos de explicación o la audiencia del interesado sin el establecimiento de otras medidas de garantía supondrá en nuestra opinión el incumplimiento y vulneración del derecho reconocido en el artículo 22. Muchas de estas garantías ya se han ido señalando a lo largo del objeto de la tesis, ahora haremos mención a algunas de ellas, estas son:

En *primer lugar*, estos sistemas han de estar sometidos a continuos mecanismos de monitorización y supervisión. De manera que se pueda analizar en todo momento el comportamiento de los modelos algorítmicos ante la entrada de nuevos datos. Aquí, los

[1218] Artículo 13 de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal. Este precepto es la transposición a la normativa española del artículo 15 de la Directiva 95/46. El Artículo 15 de dicha Directiva como ya sabemos ha sido sustituido por el artículo 22 del RGPD.

[1219] VIZCAINO CALDERÓN, M: *Comentarios a la Ley Orgánica de Protección de datos de carácter personal.* Ed. Civitas, Madrid, 2001,pág.189.

ficheros logs y los registros de resultados resultan ideales[1220]. Esta información resultará realmente útil para las autoridades de protección de datos cuando inspecciones este tipo de tratamientos[1221].

En *segundo lugar,* y relacionado con el papel del humano una vez que el sistema adopta la decisión, se han de prever mecanismos que aseguren que las decisiones en determinados momentos puedan llegar a depender de manera exclusiva de la responsabilidad de seres humanos[1222]. Ello sobre todo podría tener su relevancia cuando el particular impugne la decisión y aporte documentación que acredite que la decisión puede ser incorrecta. El nivel de control y la implicación humana ha de estar debidamente documentada[1223].

Finalmente, en *tercer lugar*, la realización de controles externos, vía auditorías algorítmicas[1224], así como el despliegue de controles internos a través de comités éticos también ayudará a desarrollar un mejor cumplimiento de la normativa de protección de datos en la adopción de este tipo de decisiones.

D) La impugnación de la decisión ante las autoridades de protección de datos o los tribunales por falta de garantías adecuadas

Cuando un particular considere que los derechos que les reconoce el RGPD no han sido debidamente respetados o tutelados por parte del responsable del tratamiento, el interesado puede acudir a la autoridad de control competente (Artículo 77 RGPD) o a la vía judicial (Artículo 79 RGPD). Dentro de esos derechos o facultades también quedan integrados los que hemos mencionado anteriormente, estos son: el derecho de explicación, el derecho a obtener intervención humana, el derecho del particular a expresar su punto de vista y el derecho de impugnación. Es turno de analizar las posibles reclamaciones que puede plantear el particular en relación con estas facultades ante la autoridad de protección de datos o los tribunales.

[1220] *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, pág.43.

[1221] Comisión Europea. *Libro Blanco sobre la inteligencia artificial. Un enfoque europeo orientado a la excelencia y la confianza*, Bruselas, 19.2.2020, págs. 23 y 24.

[1222] Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021,pág.30.

[1223] Grupo Independiente de Expertos de Alto nivel sobre *Inteligencia Artificial creado por la Comisión Europea. Directrices éticas para una IA fiable*.2019, págs. 33 y 34.

[1224] Sobre la necesidad de llevar a cabo auditoría externas derivado de las exigencias previstas en el artículo 22.3 del RGPD véase: KAMINSKI,M; MALGIERI,G: "Algorithmic impact assessments under the GDPR: producing multi-layered explanations", op.cit.,pág.9.

En *primer lugar*, los interesados pueden acudir a la autoridad de control o a los tribunales cuando el responsable del tratamiento no haya facilitado estos derechos. Es decir, serían aquellos supuestos donde el particular ni siquiera tenga opciones para ejercitar estas facultades. Ello podría ocurrir cuando el responsable informe al particular de la decisión pero no habilite un canal para que en su caso este pueda expresar su punto de vista a través del aporte de otra información o documentos relevantes. Así, por ejemplo, el Comité de No discriminación e igualdad de Género de Finlandia, el cual ostenta competentes judiciales, consideró que una persona fue discriminada al solicitar un crédito de consumo porque no se realizó una evaluación individual de su solvencia ya que sólo se tuvo en cuenta el *scoring* que había establecido el perfil automatizado para denegar dicho crédito[1225]. Para este Tribunal, esa entidad bancaria tendría que haber valorado otros elementos específicos de la persona que pudieran alterar dicha evaluación estadística. Trasladado a nuestro contexto, si un responsable no permite al particular expresar su punto de vista y aportar documentación que acredite y demuestre que efectivamente la decisión puede ser incorrecta, el mentado derecho no sería respetado. Es decir, el responsable no está obligado a alterar la decisión en virtud de las facultades prevista en el artículo 22.3, pero si al menos, a conceder al particular la oportunidad para que pueda rebatir la decisión adoptada.

En *segundo lugar*, también es posible que el responsable haya reconocido y habilitado estas facultades pero de forma insuficiente o poco realista. Ello ocurre por ejemplo cuando el proceso de revisión humano de la decisión deje entrever que el mismo no ha sido debidamente meditado dada la celeridad en la respuesta. Así, a modo de ejemplo, está constatado que en muchos supuestos las reclamaciones frente a la retirada automática de contenido que realizan los particulares ante las plataformas sociales suelen resolverse en pocos minutos[1226]. Esto lleva a pensar que dicho proceso

[1225] Comité de No discriminación e igualdad de Género de Finlandia. Resolución de 21 de marzo de 2018, procedimiento sobre la evaluación de la solvencia. Resolución disponible en: https://www.yvtltk.fi/material/attachments/ytaltk/tapausselosteet/2SVkNzOWF/YVTltk-tapausseloste-_21.3.2018-luotto-moniperusteinen_syrjinta-S._L.pdf

[1226] La plataforma YouTube retiró un video a unos usuarios de esa red. Tras impugnar la retirada de contenido, los usuarios recibieron la respuesta a tal denuncia dos minutos después de haberla planteado. Según esta plataforma, en el momento que sucedieron los hechos dicha revisión era realizad por humanos y no por máquinas. Dado el poco tiempo de respuesta que medió entre la presentación de la denuncia y la respuesta recibida, existen dudas de que efectivamente dicha decisión fuese tomada por un humano y no por un algoritmo. Véase: Fuente de la noticia: GARCÍA,J: "YouTube elimina la parodia de Pantomima Full sobre los negacionistas de la COVID-19 por cuestionar las directrices de la OMS", *Xataka*. 2/10/2020. Noticia disponible en: https://www.xataka.com/aplicaciones/youtube-elimina-parodia-pantomima-full-negacionistas-covid-19-cuestionar-directrices-oms

está en términos generales automatizado y que la revisión del humano es, o bien ficticia, o prácticamente inexistente. Por otro lado, las deficiencias a la hora de facilitar estas facultades también pueden estar presentes cuando por ejemplo la explicación que se facilita al particular resulta muy limitada. Por ejemplo, tal y como hemos señalado anteriormente, la propuesta de directiva de servicios digitales obliga a los prestadores de servicios de alojamiento de datos a que expliciten las circunstancias y motivos en que se ha basado la adopción de la decisión[1227]. Entre estas explicaciones, si la retirada se basa en que el contenido subido es ilícito, se debe mencionar la norma legal que supuestamente se ha infringido. A su vez, si la eliminación del contenido se basa en la presunta incompatibilidad del mismo con las condiciones de uso de la plataforma, se ha de hacer referencia a la regla específica que se intenta hacer cumplir y las explicaciones de por qué la información se considera incompatible con tal fundamento[1228].

En *tercer lugar*, los particulares también podrán considerar que se han vulnerado estos derechos cuando el ejercicio de los mismos resulte imposibilitado o excesivamente restringido por parte del responsable del tratamiento. Ello puede suceder cuando el ejercicio de estos derechos quede condicionado al pago de una tasa o la asunción de determinados costes a cargo del particular. Así, por ejemplo, en Estados Unidos, el llamado "Bachillerato Internacional", un programa que otorga un prestigioso diploma a estudiantes para que accedan a universidades estadounidenses tuvo que cancelar sus exámenes a raíz de la pandemia del Covid. Esas pruebas selectivas fueron sustituidas por un algoritmo con el objetivo de que predijera las calificaciones de los estudiantes basándose en distintas variables. Los afectados que pretendieron impugnar los resultados del sistema tuvieron que pagar una tasa para apelar las decisiones adoptadas[1229]. Pues bien, de acuerdo al artículo 12.5 del RGPD. Toda comunicación, así como cualquier actuación realizada en virtud de los artículos 15 a 22 del RGPD serán a título gratuito. Por tanto, la imposición de este tipo de tarifas previo al ejercicio de estas facultades puede ser recurrida ante la autoridad de protección de datos.

[1227] Véase el artículo 15.2.d) y e) de la 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. Resolución adoptada el 15 de diciembre de 2020.

[1228] Véase en este sentido la resolución adoptada por el Consejo asesor de contenidos de Facebook. Decisión del caso 2021-002-FB-UA. 13 de abril de 2021. Apartado10.Resolución disponible en: https://www.oversightboard.com/decision/FB-S6NRTDAJ

[1229] Fuente de la noticia: SIMONITE,T: "Meet the Secret Algorithm That's Keeping Students Out of College". Wired, 07/10/2020. Noticia disponible: https://www.wired.com/story/algorithm-set-students-grades-altered-futures/

En *cuarto lugar*, el interesado también podrá recurrir la actuación llevada a cabo por el responsable cuando el mismo no haya utilizado una de las excepciones previstas que autorizan al responsable para la toma de decisiones o cuando , habiendo utilizado algunas de las excepciones previstas, no las haya justificado adecuadamente. Por ejemplo, ello podría ocurrir cuando una organización ampara su tratamiento en el carácter necesario de la decisión automatizada a la hora de ejecutar el contrato y realmente tal necesariedad no existe. En este supuesto no sólo quedaría afectado el principio de licitud[1230]. Además se estaría vulnerando el derecho del particular a no ser objeto a la toma de decisiones automatizadas relevantes ya que hay que recordar que estas están prohibidas como norma general.

En todos estos supuestos entendemos que, al quedar afectado una de las facultades que reconoce el RGPD, esto es, el derecho a no ser objeto de decisiones automatizadas relevantes, la restricción que no esté justificada supondrá la vulneración del derecho fundamental a la protección de datos[1231].

Es importante recordar que el derecho de los interesados a presentar reclamaciones ante una autoridad de control o ante los tribunales por incumplimiento de la normativa de protección de datos por parte de los responsables del tratamiento es aplicable a todas las facultades que hemos analizado en este Capítulo[1232], en este apartado hemos hecho referencia únicamente a las impugnaciones posibles que se pudieran derivar del incumplimiento del derecho a no ser objeto a decisiones automatizadas[1233].

4. Garantías en favor de los interesados cuando el tratamiento se base en una norma estatal o europea. Necesidad de establecer suficientes medidas de garantía

[1230] VALERO TORRIJOS, J: "Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración". *Revista Catalana de Dret Públic*, (58), 2019, pág.92. Texto disponible en: http://revistes.eapc.gencat.cat/index.php/rcdp/article/view/10.2436-rcdp.i58.2019.3307

[1231] Como ha dicho el Tribunal Constitucional: *se rebasa o se desconoce el contenido esencial cuando el derecho queda sometido a limitaciones que lo hacen impracticable, lo dificultan más allá de lo razonable o lo despojan de la necesaria protección.* STC Sentencia de 11/1981, de 8 de abril. FJ8. Resolución disponible en: https://hj.tribunalconstitucional.es/docs/BOE/BOE-T-1981-9433.pdf

[1232] Derecho de acceso, derecho de rectificación, derecho de supresión, derecho a la portabilidad, derecho de oposición. Además del resto de la normativa de protección de datos.

[1233] El derecho a reclamar ante una autoridad de protección de datos o ante los tribuales se analizan en el Capítulo V, apartado V, puntos 1 y 2 de esta tesis.

Junto al consentimiento y el contrato, el RGPD también permite la toma de decisiones plenamente automatizadas relevantes cuando una norma estatal o de la Unión Europea autorice tal tratamiento. El único requisito que se establece en estos casos es que dicha norma establezca medidas adecuadas para salvaguardar los derechos, libertades e intereses legítimos del interesado sobre el que se adoptan dichas decisiones. Se deja un amplio margen de actuación tanto a los Estados Miembros como a la UE para que estos regulen, con las suficientes garantías, tratamientos de datos donde la plena automatización de las decisiones esté presente. Muchas de las garantías referidas a la norma ya han sido estudiadas cuando analizamos los requisitos que ha de ostentar estas cuando autoriza el tratamiento de datos basado en el uso de sistemas de toma de decisiones parcial o totalmente automatizadas[1234], lo dicho en ese momento es perfectamente trasladable a este apartado, si bien, el carácter plenamente automatizado de estos tratamientos nos obliga a incorporar las garantías específicas que se han de prever para este tipo de tratamientos concretos.

En este sentido, los estados miembros han adoptado diversas normas en desarrollo de este precepto utilizando diversos enfoques a la hora de regular estos tratamientos[1235]. Algunos han optado por regular diferentes garantías para estos concretos tratamientos que necesariamente han de estar previstas para cualquiera de las finalidades pretendidas con la toma de decisiones[1236]. Otros Estados han optado por regular ámbitos específicos tal y como ocurre con Alemania con el sector de los seguros.[1237]. Por otro lado, y a pesar de que lo normal es que a través de este precepto se

[1234] Véase el Capítulo IV, apartado I, punto 3, epígrafe B). Esas garantías indicadas en dicho apartado ya se entienden incorporadas por defecto en la norma que autorice los tratamientos de datos a los que hace referencia el artículo 22 del RGPD.

[1235] MALGIERI,G: "Automated decision-making in the EU Member States: The right to explanation and other suitable safeguards in the national legislations". *Computer Law & Security Review***,** Volume 35, Issue 5, 2019, pág. 6.

[1236] Por ejemplo, la norma británica establece que: cuando un responsable del tratamiento toma una decisión relevante en relación con un interesado basándose únicamente en el procesamiento automatizado:(a) el controlador debe, tan pronto como sea razonablemente posible, notificar al interesado por escrito que se ha tomado una decisión basada únicamente en el procesamiento automatizado, y b) el interesado podrá, antes de que finalice el período de 1 mes, que comienza con la recepción de la notificación, solicitar al responsable del tratamiento que: (i) reconsidere la decisión, o (ii) tome una nueva decisión que no se base únicamente en el procesamiento automatizado. Artículo 14.4 de la *UK Data Protection Act* 2018. Texto disponible en:
https://www.legislation.gov.uk/ukpga/2018/12/section/14/enacted

[1237] Así, por ejemplo, la ley alemana sobre protección de datos considera las decisiones para la prestación de servicios de seguros en virtud de un contrato de seguro como un caso específico de toma de decisiones automatizada permitida. Ahora bien, estas solo se permiten cuando: i) la solicitud del interesado recibe un resultado positivo. ii) Como alternativa, si el resultado es negativo, es decir, denegación de la prestación del servicio, la decisión automatizada se permite en casos específicos. En: MALGIERI,G: "Automated

legitimen la toma de decisiones automatizadas. Algunos Estados han aprovechado este artículo para vetar determinados sectores o espacios a esa plena mecanización[1238]. Es decir, se ha reconocido una especie de esferas o ámbitos donde ha de quedar reservada la actuación del humano durante el proceso decisorio automatización[1239]. En estos supuestos, las garantías devienen de la propia prohibición que prevé la norma para este tipo de tratamientos.

Es turno de analizar las garantías que ha de tener en cuenta el legislador, tanto nacional como europeo, a la hora de acudir a esta excepción para legitimar este tipo de tratamientos. En primer lugar se analizarán las exigencias formales ligadas a la herramienta legislativa que en su caso resulta adecuada, para después y en segundo lugar, analizar el contenido material de estas normas, es decir, que garantías específicas han de preverse en dichas normas.

A) Garantías y salvaguardas. Los requisitos formales de la norma

Lo primero que tenemos que valorar es el rango legal de la norma que autorice a la toma de decisiones plenamente automatizada. En este sentido, el RGPD no establece cuál debe ser el instrumento legal adecuado. Únicamente, y para supuestos muy específicos, se hace alusión a la ley como herramienta preceptiva. Entre estos supuestos no se contemplan las normas que autorizan a los tratamientos previstos en el artículo 22.2[1240]. Por lo tanto, a priori, no existe una reserva legal a la hora de habilitar estas decisiones en el plano europeo. No obstante, como ya dijimos en otro momento de la tesis, los riesgos que pueden comportar para los particulares la elaboración de perfiles y la toma de decisiones justifican que las normas que autoricen dichos tratamientos sean como mínimo de rango legal[1241]. Las ventajas que ofrece el proceso legislativo respecto

---

decision-making in the EU Member States: The right to explanation and other suitable safeguards in the national legislations", pág 7.

[1238] Por ejemplo, la ley francesa prohíbe cualquier decisión semi automatizada o totalmente automatizada en el ámbito judicial si el tratamiento tiene como objetivo evaluar aspectos de la personalidad. Véase el artículo 47 de la Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés. Texto disponible en: https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000886460/

[1239] LE DÉAUT, Y,J:. "*Technological convergence, artificial intelligence and human rights"*. Council of Europe, Parliamentary Assembly. Committee on Culture, Science, Education and Media Session 2017 - Second part-session. 2017, págs. 15 y 16. Sobre la reserva de humanidad en: PONCE SOLÉ,J: "Inteligencia artificial, derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico" .*Revista General de Derecho Administrativo*. 50, 2019, pág.28.

[1240] Son muy pocas materias donde el RGPD hace referencia a la ley como instrumento jurídico para desarrollar los contenidos del texto europeo. Por ejemplo: artículo 8.1 con relación a la regulación de la edad para otorgar el consentimiento de los menores o el artículo 54.1 sobre las autoridades de control.

[1241] Capítulo IV, apartado I, punto 3, epígrafe B), sub epígrafe b.1)

del reglamentario favorecerán la protección de los intereses y derechos de los particulares sometidos a estas decisiones[1242]. En este sentido, la LOPD de 2018 incorpora dentro de su contenido de carácter orgánico los preceptos referidos a las decisiones plenamente automatizadas y la elaboración de perfiles[1243]. Por tanto, a priori entendemos que cualquier limitación vía texto legal que legitime los tratamientos previstos en el artículo 22.1 habrá de reconocerse por una norma con rango ley[1244]. A su vez, tal y como hemos señalado anteriormente, el derecho no ser objeto de decisiones plenamente automatizadas relevantes y las garantías que se derivan del mismo forman parte del contenido esencial del derecho fundamental a la protección de datos[1245], por ello, cualquier limitación deberá estar contemplada en la ley[1246]. Dicho lo anterior, sí que consideramos que las especificaciones técnicas del sistema algorítmico que se pretenda incorporar puedan regularse en normas o instrumentos jurídicos de rango inferior a la ley. Por ejemplo, a través de un Reglamento administrativo, convenio colectivo o incluso políticas de privacidad o términos y condiciones de uso. En este sentido, recientemente se ha autorizado por Real Decreto las actuaciones administrativas automatizadas con relación a la apertura de actas algorítmicas por parte de la inspección de trabajo[1247], pensamos que dicho tratamiento de datos personales, en el caso de que estén proyectados, debería haberse autorizado por una norma con rango de ley.

[1242] Australian Human Rights Commission. *Human rights and Technology. Discussion Pape*r, 2019, pág.91.

[1243] No todo el contenido de este nuevo texto legal tiene carácter de ley orgánica. Véase la DF1ª de la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales.

[1244] Así también lo ha entendido la Autoritat Catalana de Protecció de Dades con relación a las normas que legitimen los tratamientos del artículo 22.1. En: Autoritat Catalana de Protecció de Dades. Consulta nº: CNS 4/2021, pág.11. Texto disponible en: https://apdcat.gencat.cat/web/.content/Resolucio/Resolucions_Cercador/Dictamens/2021/Documents/ca_cns_2021_004.pdf

En sentido contrario se ha inclinado algún sector de la doctrina al considerar que estos tratamientos no requieren necesariamente del reconocimiento legal para entenderse legitimados. En: HUERGO LORA, A,J: "La evaluación de la solvencia de las personas mediante el uso de algoritmos". En: HUERGO LORA, A,J (dir): *Una aproximación a los algoritmos desde el derecho administrativo*. Ed. Aranzadi, Navarra, 2020, págs. 59 y ss.

[1245] Véase el inicio del Capítulo V, apartado III de esta tesis.

[1246] Artículo 53.1 de la Constitución Española.

[1247] Véase el artículo 1.16 del Real Decreto 688/2021, de 3 de agosto, por el que se modifica el Reglamento general sobre procedimientos para la imposición de sanciones por infracciones de orden social y para los expedientes liquidatorios de cuotas de la Seguridad Social, aprobado por el Real Decreto 928/1998, de 14 de mayo. A través de este precepto se incluye un nuevo capítulo en el Real Decreto 928/1998 denominado *procedimiento sancionador promovido por actuación administrativa automatizada en el ámbito de la Administración General del Estado*.

B) Garantías y salvaguardas. Los requisitos materiales de la norma

La incorporación de suficientes medidas de garantía en la norma que legitime la toma de decisiones plenamente automatizadas es esencial. Y es que, no podemos olvidar que tales salvaguardas son las que permiten levantar la prohibición general impuesta a los responsables del tratamiento sobre el derecho que gozan los particulares a no ser objeto de decisiones plenamente automatizadas relevantes. En este sentido, tal y como ha señalado el Tribunal Constitucional, la previsión de las garantías adecuadas no puede deferirse a un momento posterior a la regulación legal del tratamiento de datos personales de que se trate. Las garantías adecuadas deben estar incorporadas a la propia regulación legal del tratamiento, ya sea directamente o por remisión expresa y perfectamente delimitada a fuentes externas que posean el rango normativo adecuado[1248]. Es turno de analizar dichas garantías.

En *primer lugar*, estas leyes, con carácter general han de reconocer garantías similares o parecidas a las que previamente hemos analizado cuando analizábamos los tratamientos que se basen en el consentimiento o el contrato. Estas eran; el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista, impugnar la decisión (Artículo 22.3. RGPD) y en su caso a la explicación de la decisión prevista en el considerando 71 del RGPD. A pesar de que el RGPD sólo haya contemplado estas facultades para esas concretas bases de legitimación, ello no nos puede llevar a pensar que las mismas quedan vetadas para el legislador que pretenda reconocer un tratamiento de toma de decisiones automatizadas plenamente relevantes por medio de una norma. Como dijimos anteriormente, el RGPD ha permitido a los Estados un amplio margen de actuación para reconocer este tipo de tratamientos imponiéndole el deber de establecer suficientes medidas de garantías, entre estas últimas, las facultades mencionadas deben estar presentes en la medida de lo posible dada la relevancia de estas.

Por otro lado, sería recomendable que los responsables del tratamiento valoren también la incorporación de algunas de las garantías que prevé la PRAI ya que muy probablemente, cuando esta norma entre en vigor, los sistemas que adopten decisiones plenamente relevantes quedarán afectados por dicha normativa por su consideración de

[1248] Véase la STC 76/2019, de 22 de mayo de 2019. FJº 8. Texto disponible: https://hj.tribunalconstitucional.es/HJ/es/Resolucion/Show/25942

sistemas de alto riesgo[1249]. Las garantías previstas por la Carta de Derechos Digitales española también pueden servir como referencia a la hora de valorar las garantías mínimas que han de reconocerse por las leyes que autoricen los tratamientos previstos en el artículo 22 del RGPD[1250].

En *segundo lugar*, el artículo 23.2 establece toda una serie de salvaguardas que ha de tener en cuenta el legislador cuando autorice al tratamiento de datos que límite los derechos reconocidos en los artículos 15 a 22. Así, entre otros elementos se ha de indicar la finalidad, las categorías de los datos y los riesgos que ese tratamiento puede generar para los derechos y libertades de los interesados.

Por lo que se refiere a finalidad, esta debe quedar claramente reflejada en la norma. Esa finalidad ha de quedar englobada en alguno de los objetivos expresos reconocidos en el artículo 23.1[1251]. En nuestra opinión, la norma que autorice estos tratamientos expresamente ha de hacer mención a la plena automatización de la decisión o del tratamiento y, siempre que existan perfiles que finalicen en tales decisiones, dejarlo claro. Es decir, la mera alusión a la posible actuación automatizada cuando se esté llevando a cabo el perfilado de personas no resulta recomendable[1252].

También se han de establecer las categorías de datos personales que se proyectan tratar. Por tanto, si el tratamiento que se legitima pretende obtener datos inferidos, ello se ha de indicar expresamente. A modo de ejemplo, la Ley referida a los datos de registros de nombres de pasajeros establece las actuaciones que han de llevar a cabo las autoridades competentes con relación a los resultados que originan los sistemas automatizados cuando analizan los datos relativos a los datos de registros de nombres de

[1249] Esta norma prevé todo un elenco de garantías, las cuales, en gran parte ya se han analizado a lo largo de esta tesis. Véase el artículo 8 y ss. de la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

[1250] Artículos 18 y 25 de la Carta de Derechos Digitales.

[1251] a) la seguridad del Estado; b) la defensa; c) la seguridad pública; d) la prevención, investigación, detección o enjuiciamiento de infracciones penales o la ejecución de sanciones penales, incluida la protección frente a amenazas a la seguridad pública y su prevención; e) un interés económico o financiero importante de la Unión o de un Estado miembro, inclusive en los ámbitos fiscal, presupuestario y monetario, la sanidad pública y la seguridad social; f) la protección de la independencia judicial y de los procedimientos judiciales; g) la prevención, la investigación, la detección y el enjuiciamiento de infracciones de normas deontológicas en las profesiones reguladas. Artículo 23.1. del RGPD.

[1252] TODOLÍ SIGNES,A: "Retos legales del uso del big data en la selección de sujetos a investigar por la inspección de trabajo y de la seguridad social". *Revista Galega de Administración Pública* (REGAP),Núm.59, 2020, pág.321.

pasajeros[1253]. Dichos resultados son las inferencias generadas por el análisis de esos registros.

A su vez, es necesario que el legislador evalúe los potenciales riesgos que puede generar la incorporación de ese tratamiento respecto los derechos y libertades de los potenciales afectados, en este caso, las personas que se verán sometidas a la toma de decisiones automatizadas. Tal y como ha recomendado el CEPD, los riesgos potenciales que se hayan tenido en cuenta deberán incluirse en los considerandos o en la exposición de motivos de la legislación que autoriza el tratamiento de datos[1254]. Para la evaluación de estos riesgos puede ser recomendable solicitar la opinión de las autoridades de protección de datos conforme a los artículos 57.1.c) y 58.3.b) del RGPD.

En *tercer lugar*, además de las garantías previstas en el RGPD, los tribunales cada vez con más frecuencia se están pronunciando sobre las garantías que necesariamente deberían prever los textos legislativos que autorizan la toma de decisiones automatizadas. Analizamos dos supuestos que ya hemos tratado en otros momentos de la tesis.

Por un lado, Consejo Constitucional Francés ha considerado constitucional el análisis masivo de datos personales de los ciudadanos disponibles públicamente en las plataformas virtuales como método para luchar contra el fraude fiscal ya que la norma que legitima este tratamiento es proporcional y justificada ya que se prevén suficientes medidas de garantía y el fin que se persigue con la misma es legítimo[1255]. Así, por ejemplo, se establecen reglas específicas sobre el modo de recopilación de los datos de los particulares[1256]. También se asignan roles específicos para los distintos agentes que intervienen durante todo este tratamiento de datos. De esta manera, estos han de contar

[1253] Véase los artículos 12 y 16 de la Ley Orgánica 1/2020, de 16 de septiembre, sobre la utilización de los datos del Registro de Nombres de Pasajeros para la prevención, detección, investigación y enjuiciamiento de delitos de terrorismo y delitos graves.
[1254] Comité Europeo de Protección de Datos. *Guidelines 10/2020 on restrictions under Article 23 GDPR*. Versión 1.0. Texto adoptado el 15 de diciembre de 2020. Apartados 60 y 61, pág.13.
[1255] La Decisión del Consejo Constitucional Francés. Décision n° 2019-796 DC du 27 décembre 2019. Apartados 79 a 96. Disponible en: https://www.conseil-constitutionnel.fr/decision/2019/2019796DC.htm .
[1256] Los datos que se recopilen y utilicen deben cumplir dos condiciones acumulativas. Por un lado, debe ser contenido de libre acceso en un servicio de comunicación pública en línea de una de las plataformas mencionadas, excluyendo por lo tanto el contenido accesible a través de contraseña o después del registro en el sitio en cuestión. Por otro lado, este contenido debe ser claramente hecho público por los usuarios de estos sitios. Por tanto, solo se puede recopilar y utilizar el contenido relacionado con la persona que lo divulgó deliberadamente. Véase el apartado 87 de la mentada sentencia.

con el suficiente rango administrativo[1257]. Finalmente, se establecen plazos máximos de retención y almacenamiento de datos, previéndose distintos periodos en función del carácter o no sensible de los datos o la posible inutilidad de los mismos.

Por otro lado, en Países Bajos, el gobierno neerlandés implantó el ya analizado sistema SyRI. Esta aplicación algorítmica tenía el objetivo de detectar el fraude en la obtención de prestaciones sociales por parte de determinados sectores de la población. Pues bien, esta ley fue impugnada ante el Tribunal de la Haya y este consideró que la ley en la que se basaba este sistema era contrario al Convenio Europeo de Derechos Humanos. Así, y a pesar de que la norma preveía un importen elenco de garantías, el tribunal consideró que no eran suficientes[1258]. Concretamente, las deficiencias vinieron de la inexistencia de auditorías o controles por parte del personal externo del sistema y de un déficit importante de transparencia en relación con el tratamiento de datos llevado a cabo. Y es que, entre otros elementos, no se conocían los indicadores de riesgo y el modelo de riesgo ni los criterios objetivos que subyacen sobre la validez de los indicadores el modelo de riesgo[1259]. Elementos que están estrechamente relacionados con la elaboración de perfiles y los datos inferidos. Queda claro como cualquier restricción al principio de transparencia ha de quedar debidamente justificada. Esta sentencia marca un listón de garantías que en determinados sectores puede no resultar fácilmente alcanzable, sobre todo en aquellos sectores donde el conocimiento de esta información pueda dejar sin efecto los objetivos de la implantación del sistema, tal y como ocurren en aquellos ámbitos donde se pretende la detección de posibles actuaciones fraudulentas.

En *cuarto lugar*, y para finalizar, cabe indicar que el hecho de que una norma regule expresamente un determinado tratamiento y especifique una serie de garantías no desplazará la aplicación del RGPD. Las salvaguardas contempladas en la norma específica que autorice el tratamiento complementarán las reconocidas por el régimen general y ambas serán plenamente aplicables salvo que la propia norma restrinja expresamente el contenido o alcance de algunas de estas y ello esté justificado. Art 23.2.c) RGPD.

[1257] El sistema sólo puede ser utilizado por agentes de las administraciones tributaria y aduanera que tengan al menos rango de controlador y estén especialmente autorizados. Las personas que contribuyan al diseño y ejecución de las operaciones de tratamiento de que se trate están sujetas al secreto profesional. Aparatado 88 de la mentada sentencia.

[1258] Sentencia del Tribunal de la Haya de 5 de febrero de 2020. Disponible en: https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878

[1259] COTINO HUESO,L: "«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", op.cit., pág.2.

Requisitos de la norma que regula tratamientos del artículo 22.2.b) del RGPD.		
Sector	**Elementos formales**	**Elementos materiales a incluir**
Público	Norma con rango de ley.	-Designación expresa del tratamiento. -Finalidad, categoría de datos tratados, medidas de responsabilidad activa, etc. -Derechos y facultades específicas similares a las previstas en el artículo 22.3 del RGPD.
	Reglamento administrativo.	-Especificaciones técnicas del sistema algorítmico.
Privado	Norma con rango de ley.	-Designación expresa del tratamiento. -Finalidad, categoría de datos tratados, medidas de responsabilidad activa, etc. -Derechos y facultades específicas similares a las previstas en el artículo 22.3 del RGPD.
	Convenio colectivo, políticas de privacidad, términos y condiciones.	-Especificaciones técnicas del sistema algorítmico.

C) La integración de los derechos de explicación, impugnación de la decisión, supervisión humana y expresión del punto de vista en los textos legales que autoricen el tratamiento de datos

El conjunto de garantías que se desprenden del derecho a no ser objeto de decisiones plenamente automatizadas relevantes en particular y de la normativa de protección de datos en general configuran todo un elenco de salvaguardas aplicables al conjunto de etapas presentes en la fase de despliegue de estos sistemas. Garantías como la supervisión humana, la motivación o la impugnación de las decisiones representan toda una serie de medidas básicas que muy probablemente se reconocerán o están ya reconocidas en otros textos legales donde también se pretende regular el proceso decisorio automatizado.

Cono decíamos anteriormente, el RGPD ha permitido a los estados y a la UE un margen de actuación a la hora de autorizar por medio de una norma con rango de ley los tratamientos de datos basados en la toma de decisiones plenamente automatizadas relevantes contemplados en el artículo 22.1 del RGPD. Sólo se establece un requisito,

esto es, dicha norma ha de contemplar suficientes medidas de garantía[1260]. Entre esas garantías señalábamos que con carácter general deberían reconocerse de forma parecida o similar aquellas a las que hace referencia el artículo 22.3 y el considerando 71 del RGPD, estas eran: el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista, impugnar la decisión y en su caso a la explicación de la decisión.

Pues bien, cuando un tratamiento de datos personales como el descrito en el artículo 22 del RGPD queda autorizado por una norma, ya sea de derecho europeo o nacional, dicho texto legal pasa a incorporarse a un ordenamiento jurídico que puede prever o tener previsto el reconocimiento de toda una serie de garantías ligadas al proceso decisorio automatizado. Estas garantías en parte pueden ser similares a las contempladas en el artículo 22.3 del RGPD, si bien, adaptadas al contexto donde se aplicará la norma.

Así, en el sector privado, en los últimos tiempos y fruto del papel que están alcanzando las grandes plataformas tecnológicas, la Unión Europa ha decidido imponer mayores obligaciones de control a estas organizaciones para controlar los riesgos que se pueden generar en las mismas. Gran parte de estas obligaciones están centradas en el control del contenido que suben a estas plataformas los usuarios, dicho control del contenido suele ejecutarse a través de medios totalmente automatizados. Para ello, estos mismos textos dotan a los usuarios de todo tipo de mecanismos y garantías para hacer frente a las decisiones automatizadas en estos contextos[1261]. En este sentido, entre estas garantías cabe destacar el aumento de mecanismos de transparencia, la posibilidad de poder impugnar las decisiones algorítmicas que retiran el contenido o la supervisión del humano.

Por otro lado, en el sector público, por ejemplo, el ya analizado Real Decreto 688/2021 de 3 de agosto[1262], el cual autoriza la extensión de actas de infracción de

[1260] Capítulo V, apartado III, punto 4, epígrafe B).

[1261] En EEUU parece también existir la intención de imponer mayores obligaciones de control a las plataformas en línea con relación al uso de algoritmos en estos contextos. En: Véase la *Algorithmic Justice and Online Platform Transparency Act.* Proyecto de ley publicado el 28 de mayo de 2021. Texto disponible en: https://www.congress.gov/bill/117th-congress/house-bill/3611/text

[1262] Ya hemos comentado previamente que este texto normativo muy probablemente ha incumplido la normativa de protección de datos, concretamente los artículos 6 y 22 del RGPD ya que se ha reconocido la toma de decisiones plenamente automatizadas relevantes por medio de un Real Decreto y no a través de una norma con rango de Ley. Véase el Capítulo V, apartado III, punto 4, epígrafe A y el Capítulo IV, apartado I, punto 3, epígrafe B), sub epígrafe b.1) de esta tesis.

forma automatizada en el ámbito de la inspección de trabajo reconoce toda una serie de garantías que ha de desplegar la Administración cuando se adopten estas actas emitidas por el algoritmo. Estas garantías que son propias de cualquier procedimiento administrativo se asemejan en parte a las reconocidas en el artículo 22.3 del RGPD, esto es, supervisión humana, impugnación de la decisión, explicación, etc.

Texto legal	**Sistemas de toma de decisiones automatizadas**	**Derechos que se reconocen**
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.	-Bloqueo o retirada de contenido que puede incitar al terrorismo.	-Información sobre el uso de sistemas automatizados. -Derecho intervención humana. -Mecanismo de impugnación ante el bloqueo o retirada.
Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a un mercado único de servicios digitales (Ley de servicios digitales)	-Bloqueo o retirada de contenido ilícito. -Recomendación de contenido. -Personalización de anuncios.	-Información sobre la personalización y uso de sistemas automatizados. -Acceso a los perfiles publicitarios. -Alteración de los perfiles que recomiendan contenido. -Mecanismos de impugnación ante el bloqueo o retirada.
Reglamento (UE) 2021/1232 del Parlamento Europeo y del Consejo de 14 de julio de 2021 sobre la lucha contra los abusos sexuales de menores en línea.	-Bloquear contenido	-Información sobre el uso de tecnologías que bloquean el contenido. -Mecanismo de impugnación ante el bloqueo o retirada.
Real Decreto 688/2021, de 3 de agosto, que modifica el Reglamento sobre imposición de sanciones por infracciones de orden social y para los expedientes liquidatorios de cuotas de la Seguridad Social	-Emitir actas de infracción automatizadas	-Cauce de alegaciones. (impugnación, punto de vista y supervisión humana) -Se indica el órgano administrativo para realizar los actos de instrucción y ordenación. -Se indican los hechos comprobados como resultado de la actuación administrativa. (explicación)

En estos textos legales las garantías previstas se asemejan en cierto modo a las contempladas por parte del RGPD. La pregunta siguiente que habría que hacerse por

Artículo 1.16 del Real Decreto 688/2021, de 3 de agosto, por el que se modifica el Reglamento general sobre procedimientos para la imposición de sanciones por infracciones de orden social y para los expedientes liquidatorios de cuotas de la Seguridad Social, aprobado por el Real Decreto 928/1998, de 14 de mayo.

tanto en este momento es valorar cómo se habrían de integrar las garantías que propugna el RGPD teniendo en cuenta que la norma que autoriza la toma de decisiones automatizadas ya prevé unas garantías similares o parecidas adaptadas al contexto donde el sistema algorítmico irradiará sus efectos.

Para estos supuestos proponemos dos soluciones:

A) Una primera propuesta es que junto con las garantías que prevea la norma, se reconozcan a su vez las garantías ad hoc que se derivan de la normativa de protección de datos. De esta manera, ante la retirada de un contenido publicado por una particular por parte de un algoritmo automatizado en una plataforma, dicho interesado tendría dos cauces para impugnar la decisión, por un lado, a través del derecho de impugnación que reconoce la norma que legitima ese tratamiento de datos y por otro lado[1263], a través de la norma de protección de datos que reconoce el derecho de impugnación de las decisiones. A su vez, ante la emisión de un acta de infracción laboral emitida por un algoritmo, el particular podría impugnar la decisión a través del procedimiento administrativo correspondiente que autoriza ese tipo de decisiones y además por la vía que se deriva del RGPD.

B) Una segunda propuesta que proponemos es que las garantías que se derivan de la normativa de protección de datos y las presentes en la propia normativa que autoriza el tratamiento de datos queden integradas y su ejercicio se realice en el contexto jurídico donde la norma irradia sus efectos. Así, por ejemplo, la norma que autoriza las actas de infracción automatizadas de la inspección del trabajo sólo contemplaría un derecho de impugnación de acuerdo al procedimiento administrativo que prevea dicha norma sin que en su caso quepa a su vez otro derecho a poder impugnar la decisión derivado de la normativa de protección de datos.

La primera propuesta es más garantista pero en muchos casos puede resultar compleja, la segunda en cambio puede ser más práctica. En nuestra opinión, siempre que la norma que autoriza el tratamiento de datos reconozca garantías similares o

[1263] Por ejemplo el artículo 10.1 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 obliga a las plataformas a habilitar un canal que permita a los usuarios impugnar la retirada de contenido.

parecidas a las previstas por el artículo 22.3 del RGPD, la segunda propuesta resultará más idónea. A esta conclusión llegamos teniendo en cuenta que del artículo 22.2.b) del RGPD se desprende la obligación del legislador de establecer suficientes medidas de garantía, medias de garantía que como hemos dicho han de ser parecidas a las contempladas en el artículo 22.3. Reconocidas esas garantías en el contexto normativo donde irradie sus efectos el tratamiento, poco más se puede exigir en virtud del derecho a no ser objeto de decisiones plenamente automatizadas. A diferencia de los derechos de acceso, rectificación, supresión…. que tienen su propia autonomía independientemente del contexto normativo donde se implante y de la base de legitimación que se utilice, el derecho a no ser objeto de decisiones plenamente automatizadas relevantes queda en parte limitado a la base de legitimación que se utilice para autorizar al tratamiento. En el caso de que dicho tratamiento se autorice a través de una norma estatal o europea, el RGPD impone, que no es poco, la necesidad de establecer medidas de garantía, establecidas dichas medidas, este derecho al menos para esta concreta base de legitimación se entiende cumplido[1264]. Como es lógico, el resto de la normativa de protección de datos explicada a lo largo de esta tesis se seguirá aplicando para este tipo de tratamiento de datos, simplemente, el derecho a no ser objeto de decisiones plenamente autonomizadas quedará más limitado en estos contextos. Desde esta perspectiva, por ejemplo, el particular afectado por una decisión administrativa automatizada deberá acudir al cauce administrativo previsto para poder impugnar dicha resolución. A su vez, desde la esfera de protección de datos y con relación al derecho reconocido en el artículo 22 del RGPD, el particular podrá alegar ante el responsable del tratamiento o la autoridad de control que la norma que autoriza el tratamiento no contempla las suficientes medidas de garantías y por tanto, la base de legitimación que está utilizando dicho responsable no es adecuada. Para estos supuestos, el responsable habría vulnerado el principio de licitud y además el derecho del interesado a no ser sometido a decisiones plenamente automatizadas relevantes.

5. Garantías en favor de los interesados cuando el tratamiento utilice datos de categoría especial. Especial atención a los problemas de discriminación algorítmica

[1264] No así para aquellos casos en los que la base de legitimación sea el consentimiento o la ejecución/formalización del contrato, en estos supuestos explícitamente el RGPD impone el deber a los responsables del tratamiento de reconocer el derecho de intervención humana, impugnación de la decisión, expresar el punto de vista y el derecho de explicación. Artículo 22.3 y considerando 71 del RGPD.

Cuando las decisiones plenamente automatizadas relevantes se tomen sobre la base de los datos de categoría especial los responsables del tratamiento deberán prever garantías específicas aún más protectoras que las previstas para el tratamiento de datos convencionales. Son tres los factores que hacen que el riesgo presente en este tipo de decisiones aumente y por consiguiente obliguen al responsable a establecer mayores medidas de cumplimiento y garantías[1265]. Estos son: i) la plena automatización de las decisiones. ii) Los efectos relevantes que estas generan sobre el particular. iii) El uso de datos de categoría especial o sensible.

La combinación de estos tres elementos obliga a los responsables a ser especialmente cautos a la hora de llevar a cabo este tipo de tratamientos. Es turno de analizar algunas de las garantías que se han de desplegar con el objetivo de reducir o mitigar los riesgos implícitos presentes.

A) Garantías generales en los sistemas automatizados

Los responsables del tratamiento, cuando pretendan adoptar decisiones basadas en las categorías de datos del artículo 9 del RGPD, han de establecer como mínimo todas las garantías y medidas ya explicadas en relación con el uso de sistemas de toma de decisiones automatizadas[1266].

En primer lugar, entre las medidas de responsabilidad activa habrá que incluir entre otras: controles y monitorización del sistema, realización de auditorías algorítmicas tanto internas como externas, despliegue de la privacidad desde el diseño y por defectos de los sistemas que se pretende contratar y que adoptarán las decisiones automatizadas plenamente relevantes, acceso restringido del personal de la organización a los perfiles elaborados o a las bases de datos que se utilizarán para desarrollar los algoritmos[1267], utilización de técnicas de anonimización y seudonimización en el

[1265] Tal y como ha señalado el TC con relación a los datos de categoría especial. *El legislador está constitucionalmente obligado a adecuar la protección que dispensa a dichos datos personales, en su caso, imponiendo mayores exigencias a fin de que puedan ser objeto de tratamiento y previendo garantías específicas en su tratamiento, además de las que puedan ser comunes o generales.* En: STC Sentencia 76/2019, de 22 de mayo de 2019. FJ 6º.d). Texto disponible en: https://hj.tribunalconstitucional.es/HJ/es/Resolucion/Show/25942

[1266] Véase el conjunto de medidas de responsabilidad activa indicadas en el capítulo III del RGPD, así como las medidas previstas en el presente capítulo en relación con la toma de decisiones plenamente automatizadas relevantes.

[1267] Por ejemplo, en relación con historial clínico electrónico se ha indicado la necesidad de que solo tengan acceso al mismo aquellos profesionales de la salud/personal autorizado de instituciones sanitarias que participen en ese momento en el tratamiento en cuestión. En: Grupo del Artículo 29. *Documento de*

contexto de la elaboración de perfiles, posibilidad de utilizar mecanismos de certificación, etc[1268].

En segundo lugar, se deberán de garantizar el ejercicio pleno de los derechos y facultades reconocidas por el RGPD y concretamente los contemplados para las decisiones automatizadas previstas en el artículo 22.3. Estos son: el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión. A su vez, se ha de reconocer el derecho de explicación reconocido en el considerando 71 del RGPD ya comentado.

B) Garantías específicas en los sistemas automatizados

Uno de los riesgos principales que se puede generar por el uso de sistemas automatizados relevantes sobre la base de categorías especiales de datos es que los algoritmos sobre los que se sustentan las decisiones tomadas sean discriminatorios[1269]. Es por ello que resulte esencial diseñar y prever toda una serie de medidas que específicamente reduzcan los riesgos aparejados a tales tratamientos[1270].

En primer lugar, el responsable sólo puede llevar a cabo este tipo de tratamientos cuando el particular afectado por la decisión haya consentido explícitamente la misma o cuando el tratamiento sea necesario por razones de un interés público esencial sobre la base del Derecho de la Unión o de los Estados miembros. Es decir, la primera garantía deviene del propio artículo 22.4 del RGPD al reducir a dos los mecanismos de legitimación que autorizan este tipo de tratamientos de datos[1271].

En segundo lugar, por lo que se refiere a las medidas de responsabilidad activa, la AEPD en una circular expresamente se pronunció sobre las mismas en relación con la recopilación de datos con fines electorales por parte de los partidos políticos a la hora de

---

trabajo sobre el tratamiento de datos personales relativos a la salud en los historiales médicos electrónicos. Adoptado el 15 de febrero de 2007, pág.16.

[1268] "Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679". Grupo del Artículo 29. Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018, pág.36.

[1269] Sobre la discriminación y su relación con la protección de datos véase el Capítulo IV, apartado VII, punto 2 de esta tesis.

[1270] Garante per la Protezione dei Dati Personali. Resolución nº 234 de 10 de junio de 2021. Apartados 3.3.6. y 4. Resolución disponible en: **https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9675440**

[1271] Sobre el consentimiento explícito y el interés público esencial véase el capítulo IV de esta tesis. Concretamente los siguientes apartados: consentimiento explícito (Capítulo IV, apartado I, punto 1, epígrafe C) e interés público esencial (Capítulo IV, apartado I, punto 7, epígrafe G).

elaborar perfiles[1272]. Aunque dichas salvaguardas específicas no puedan trasladarse a cualquier sistema de toma de decisiones automatizadas, las mismas pueden servir como punto de partida para aquellos responsables que pretendan llevar a cabo la toma de decisiones automatizadas relevantes de categorías especiales de datos. Así, la AEPD, asumiendo que estos tratamientos de datos son de alto riesgo, obliga a los responsables a consultar a la autoridad de control antes de llevar a cabo el tratamiento o a remitir a dicha autoridad el análisis de riesgos y la EIPD realizada justificando las medidas adecuadas[1273]. En nuestra opinión, esta medida resulta muy adecuada y debería ampliarse a cualquier tratamiento de datos que implique el uso de sistemas automatizados cuando adopten decisiones plenamente automatizadas relevantes sobre datos de categoría especial. De esta manera, las autoridades, antes de que el tratamiento se lleve a cabo, analizarán los potenciales riesgos que dichos sistemas algorítmicos pueden generar y en su caso realizar recomendaciones de mejora o incluso decidir si los mismos no se pueden llevar a cabo. Insistimos, la relevancia de las decisiones, la plena automatización de las mismas y el uso de datos de categoría especiales justifican desde nuestro punto de vista esta concreta medida.

Por otro lado, junto a la obligación previamente mencionada, esta circular prevé otra serie de medidas que necesariamente habrán de implantar los responsables cuando traten datos con fines electorales. Entre otras garantías cabe destacar la designación obligatoria de un DPD, obligación de realizar una EIPD, imposición de medidas adecuadas de seguridad[1274], etc. Todas estas medidas son perfectamente trasladables al contexto que estamos ahora analizado.

[1272] Véase el artículo 7 de la Circular 1/2019, de 7 de marzo, de la Agencia Española de Protección de Datos, sobre el tratamiento de datos personales relativos a opiniones políticas y envío de propaganda electoral por medios electrónicos o sistemas de mensajería por parte de partidos políticos, federaciones, coaliciones y agrupaciones de electores al amparo del artículo 58 bis de la Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General.

[1273] El artículo 7.1.5º de la mentada circular establece: *Deberá consultarse a la AEPD antes de proceder al tratamiento conforme al artículo 36.1 del RGPD al tratarse de tratamientos que entrañan un alto riesgo, a no ser que el responsable justifique la adopción de medidas para mitigarlo. En este último caso deberá remitirse a la AEPD el análisis de riesgos y la evaluación de impacto junto a la justificación de las medidas adoptadas, al amparo de lo previsto en el artículo 58.1. a) y e) del RGPD. La solicitud de consulta a la AEPD o, en su defecto, la remisión de la documentación anteriormente indicada, deberá realizarse al menos 14 semanas antes del inicio del periodo electoral.*

[1274] Se establece por ejemplo que la supresión de los datos habrá de realizarse tras el periodo electoral a través de procedimientos específicos adecuados. Artículo 9.3 de la Circular 1/2019, de 7 de marzo. En un sentido parecido, en Francia, la ley que autoriza al tratamiento masivo de datos a través de sistemas de inteligencia artificial con el objetivo de detectar posible fraude fiscal prevé distintas reglas de supresión de los datos en función de si los mismos son o no de categoría especial. Así, cuando los datos recopilados no resulten necesarios para una investigación, si los mismos son de categoría especial, estos se suprimirán como máximo en 5 días, Para el resto de datos el plazo será de 30 días. Artículo 154 de la Loi de finances

En tercer lugar, otro elemento importante es valorar el papel de las variables *proxies* en el sistema[1275]. Resulta fundamental que las variables *proxies* sean detectadas en las fases iniciales del desarrollo de los sistemas. De esta manera, debe quedar constancia de la existencia de este tipo de *proxies* y de las categorías especiales de datos a los que se refieren[1276]. Esta información resulta útil tanto para el diseñador del sistema como para aquel que pretenda utilizarlo en la toma de decisiones. Así, por parte del diseñador esta información le ayudará a valorar si el sistema está o no sesgado. A su vez, también puede ser una documentación relevante para demostrar que el sistema cumple con las exigencias normativas. Por otro lado, en relación con la organización que despliega estos sistemas, la documentación relacionada con los *proxies* le permitirá a la misma analizar el comportamiento de dichas variables en el entorno específico donde se desplegará el algoritmo[1277]. Permitiendo así una monitorización específica sobre dichas variables con el objetivo de prever posibles decisiones discriminatorias que este vaya generando.

En cuarto lugar, junto a las variables *proxies*, el principio de no discriminación algorítmica reconocido en el considerando 71 del RGPD en relación con el principio de lealtad obliga a las organizaciones a desplegar otra serie de herramientas.

- Así, *por lo que se refiere a la fase del diseño*. Los responsables han de ser realmente cautos a la hora de elegir los datos que se utilizarán para conformar el sistema. Como ya hemos señalado, el principio de exactitud y lealtad en la fase de desarrollo obliga a los responsable a elegir bases de datos actualizadas y exactas. A su vez, el principio de minimización de datos obliga a las organizaciones a desplegar el llamado estudio de minimización de datos, el cual, impone a estas una especial diligencia a la hora de justificar las distintas variables elegidas para el

pour 2020, sous le n° 2019-796 DC, le 20 décembre 2019. Texto disponible en: https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000039683923

[1275] Recordemos que la AEPD define una variable *proxy* como aquella que se usa en lugar de la variable de interés cuando esa variable de interés no se pueda o no se quiera medir directamente Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág39.

[1276] Tal y como ha señalado la AEPD. Es necesario que se determinen las variables relevantes para el modelo, identificando las variables asociadas a categorías especiales de datos y las variables *proxy*, incluyendo la información necesaria para su interpretación. Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA*, 2021, pág.24.

[1277] La interacción con el entorno puede alterar el comportamiento de la variable *proxy*. En: Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020, págs. 40 y 41.

desarrollo del modelo algorítmico[1278]. Recordemos que no está prohibido el uso de datos de categoría especial a la hora de desarrollar sistemas de toma de decisiones automatizadas[1279], más bien, los responsables deben justificar su elección y prever las medidas que en su caso mitiguen los riesgos que pueden generar la elección de estas variables.

- Por otro lado, *en relación con la fase de toma de decisiones* consideramos que los mecanismos de control resultan esenciales en este contexto. Así, proponemos un sistema de notificación similar al previsto por el RGPD a la hora de comunicar las brechas de seguridad[1280]. De esta manera, la adopción de decisiones discriminatorias por parte del sistema sería similar al incidente de seguridad en el tratamiento de datos personales. En este sentido, en función del riesgo que pueda suponer para el interesado esa decisión algorítmica, el responsable ha de establecer unos u otros mecanismos de comunicación y en su caso desplegar unas u otras medias para mitigar los efectos generados. Por tanto, si se detecta que el sistema puede generar decisiones algorítmicas pero el mismo no ha adoptado decisiones aun sobre los particulares o las decisiones adoptadas no han generado graves efectos, el responsable únicamente deberá registrar dicho incidente y tratará de solventar ese problema aplicando las medias técnicas y organizativas para evitar que vuelva a suceder. Seguidamente, si el sistema adopta decisiones discriminatorias y los efectos que el mismo genera son graves para los particulares, el responsable estará obligado a comunicar estas incidencias tanto a los particulares afectados por las decisiones como a la autoridad de control correspondiente. En estos supuestos, a los particulares se les debería informar de las posibles actuaciones que estén en sus manos para revertir esta situación y en su caso de los mecanismos para impugnar las decisiones adoptadas. En función de la gravedad de las decisiones, las medidas a adoptar irán desde la re calibración del

[1278] Por lo que se refiere al estudio de minimización de datos, véase el Capítulo IV, apartado III, punto 1 de esta tesis.

[1279] Tal y como ha señalado la doctrina, resulta incluso recomendable el uso de datos de categoría especial ya que así es más fácil detectar posibles sesgos presentes en los sistemas de toma de decisiones automatizadas. En: SORIANO, ARNANZ, A: "Decisiones automatizadas: problemas y soluciones jurídicas. más allá de la protección de datos", op.cit., págs.37 y 38. También en: ROIG I BATALLA,A: *Las garantías frente a las decisiones automatizadas del Reglamento general de Protección de Datos a la gobernanza algorítmica*.op.,cit, págs. 177 y 178.

[1280] Véase los artículos 33 y 34 del RGPD. En parte, estos mecanismos se han analizado también en el Capítulo III, apartado V, punto 3 de esta tesis.

modelo algorítmico hasta la retirada del mismo y la necesidad de desarrollar uno nuevo.

<table>
<tr><th colspan="4">Derecho a no ser sometido a decisiones plenamente automatizadas relevantes</th></tr>
<tr><th colspan="2">Ámbito de aplicación</th><th>Prohibición general
Excepciones de legitimación</th><th>Garantías específicas</th></tr>
<tr><td rowspan="3">-Decisiones plenamente automatizadas relevantes

-Elaboración de perfiles totalmente automatizados relevantes</td><td rowspan="2">Datos no categoría especial</td><td>-Consentimiento
-Contrato</td><td>-Derecho intervención humana
-Derecho expresar punto de vista
-Derecho impugnación
-Derecho de explicación</td></tr>
<tr><td>-Normal estatal o europea</td><td>-Respeto contenido formal y material de la norma.
-Suficientes medidas de garantía.</td></tr>
<tr><td>Datos Artículo 9</td><td>-Consentimiento
-Interés público esencial</td><td>-Suficientes medidas de garantía</td></tr>
</table>

IV. PROPUESTA DE INTEGRACIÓN DE LOS DERECHOS

El derecho a no ser objeto de decisiones plenamente automatizadas relevantes reconoce una serie de facultades específicas en favor de los particulares que se ven sometidos a estos concretos tratamientos. En nuestra opinión, sería recomendable que los responsables no sólo habilitaran el ejercicio de estos derechos cuando adopten las decisiones descritas en el 22 sino también, que los mismos se ampliaran a aquellos supuestos en los que, pese a que las decisiones no fueran completamente automatizadas, las mismas generaran efectos relevantes en la esfera de los particulares. Lo que defendemos es que, la plena automatización, pudiendo ser un elemento muy relevante a tener en cuenta por parte de los responsables, no sea el definitorio para no contemplar las facultades previstas en el artículo 22. Y es que, dichas facultades son perfectamente encajables cuando se llevan a cabo tratamientos donde se toman decisiones parcialmente automatizadas relevantes. De esta manera, cuando el responsable realice el

análisis de riesgos del tratamiento en cuestión debería valorar la incorporación de estas facultades a la hora de reducir los potenciales riesgos que puede generar el sistema de toma decisiones, riesgos que en muchos casos serán similares con independencia de si la decisión es o no plenamente automatizada.

Además, pensamos que este punto de vista no sólo favorece a los particulares sino también los responsables del tratamiento. Y es que como ya dijimos, en la mayoría de los textos legales y propuestas de regulación de este fenómeno no existe una distinción entre sistemas que adoptan decisiones parcial o totalmente automatizadas sino que se pone el acento en los riesgos que puede generar esa decisión en la esfera de los particulares[1281]. Es por ello que resulte recomendable que los responsables del tratamiento desde el diseño de sus sistemas desarrollen modelos algorítmicos que habiliten al ejercicio de las mentadas facultades independientemente del tipo de decisiones que se proyecten ya que si no, estos sistemas pueden no quedar acompasados al resto de normativas donde el ámbito de aplicación coincida y el factor plenamente automatizado sea secundario[1282]. En este sentido, la PRAI, norma que en el futuro será la referencia en este contexto, prevé toda una serie de obligaciones que los responsables del tratamiento han de tener en cuenta cuando desarrollen o desplieguen un sistema de alto riesgo donde traten datos personales, ello independientemente de si los sistemas adoptan o no decisiones plenamente automatizadas[1283]. Estas garantías, en parte se complementan con algunas de las facultades descritas en el artículo 22 y el considerando 71 del RGPD tal y como ocurre con la supervisión humana o la explicación de la decisión. Es por ello que la necesidad de integrar en dichos sistemas el cumplimiento de ambas normas resultará fundamental.

Norma	**Tipo de decisiones**	**Exigencias legales y facultades**
Propuesta de Reglamento europeo sobre la regulación	Todo tipo de decisiones que generen efectos	-Obligación de que los sistemas sean transparentes

[1281] Véase el Capítulo II, apartado I, punto 3 de esta tesis.

[1282] El artículo 25.3 de la Carta de Derechos Digitales establece que: *las personas tienen derecho a solicitar una supervisión e intervención humana y a impugnar las decisiones automatizadas tomadas por sistemas de inteligencia artificial que produzcan efectos en su esfera personal y patrimonial.* Nótese que aquí no se está hablando únicamente de decisiones plenamente automatizada sino también de las parcialmente automatizadas.

[1283] Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

de los sistemas de inteligencia artificial	relevantes.	de tal manera que faciliten su comprensión. -Los sistemas deben permitir la supervisión humana de los mismos.
Reglamento General de Protección de datos	Decisiones plenamente automatizadas relevantes	-Derecho intervención humana. -Derecho expresar punto de vista. -Derecho impugnación. -Derecho de explicación.

V. LOS MECANISMOS DE TUTELA A FAVOR DEL INTERESADO ANTE EL INCUMPLIMIENTO DE LA NORMATIVA DE PROTECCIÓN DE DATOS

El RGPD contempla una serie de acciones en favor del interesado frente al incumplimiento de la normativa de protección de datos por parte de los responsables. Estos son: i) el derecho del interesado a presentar una reclamación ante la autoridad de control, ii) el derecho a recurrir por vía judicial una decisión de una autoridad de control o el incumplimiento de la normativa por parte de un responsable, iii) el derecho del interesado a solicitar la representación de una tercera organización para que actúe en el ámbito de la protección de datos.

Es turno de analizar todos estos mecanismos y su incidencia durante el ciclo de vida de los sistemas de toma de decisiones automatizadas.

1. El derecho a presentar una reclamación ante una autoridad de control

Conforme al artículo 77.1 del RGPD, todo interesado tiene derecho a presentar una reclamación ante una autoridad de control si considera que el tratamiento de datos personales que le conciernen infringe dicho texto[1284].

De esta manera, si el particular considera que el desarrollo de un sistema algorítmico o su uso en la toma de decisiones ha vulnerado su derecho fundamental a la protección de datos conforme a las prescripciones del RGPD o la LOPD de 2018, este podrá optar entre: i) dirigirse al delegado de protección de datos del responsable o, ii)

[1284] Este derecho también se reconoce en el artículo 8.3 de la Carta de Derechos Fundamentales de la Unión Europea. Así, este texto indica que el respeto de las normas relativas a la protección de datos quedarán sometidas al control de una autoridad independiente.

plantear una reclamación ante la autoridad de control correspondiente[1285]. Para este último supuesto, la autoridad de control deberá examinar y analizar la solicitud[1286] o remitir dicha reclamación al delegado de protección de datos del responsable[1287].

Normalmente, los interesados acudirán a las autoridades de control cuando los responsables no atiendan a sus solicitudes. Así, esta situación resultará más frecuentes cuando las solicitudes o peticiones entren en conflicto con los intereses o bienes jurídicos que el responsable pretenda proteger.

De esta manera, en la fase del diseño de los sistemas, el responsable puede mostrar reticencias cuando el particular solicite la supresión de los datos de los *data lakes* o retire el consentimiento para la analítica de datos. Ante esta negativa el particular puede acudir a la autoridad de control correspondiente.

Por otro lado, una vez el sistema comience a desplegar sus efectos, la mayoría de las reclamaciones ante las autoridades de control vendrán referidas a la toma de decisiones algorítmica. Entre otros supuestos; i) déficit de justificación a la hora de conceder información adecuada sobre la pertinencia de las variables elegidas, ii) límites injustificados a los derechos de acceso a los perfiles o de rectificación de los mismos aludiendo motivos de propiedad intelectual o riesgos de jugar con el sistema, iii) falta de garantías adecuadas durante la toma de decisiones automatizadas, etc.

Para finalizar cabe señalar que el responsable está obligado a informar a los interesados del derecho que estos ostentan a poder presentar reclamaciones ante las autoridades de control[1288].

2. El derecho a la tutela judicial efectiva. El recurso contra las decisiones de las autoridades de control o las actuaciones de los responsables y encargados

El RGPD faculta a los interesados a acudir ante los tribunales de justicia a través de dos mecanismos. Estos son: i) cuando pretendan impugnar una decisión jurídicamente vinculante de una autoridad de control que le concierna (Artículo 78

[1285] RECIO GAYO, M: "Los derechos a presentar reclamaciones ante la autoridad de control y a la tutela judicial efectiva". En: PIÑAR,MAÑAS, J,L (dir):*Reglamento General de Protección de Datos. Hacia un nuevo modelo de privacidad*. Ed. Reus. Madrid, 2017, pág.542.

[1286] Tal y como ha señalado el TJUE. STJUE (Gran Sala) de 6 de octubre de 2015, asunto C-362/14, caso Schrems. FJº 55,56 Y 57. Resolución disponible en: https://curia.europa.eu/juris/document/document.jsf;jsessionid=D1B7D655066A9F323C84661AB7277F8E?text=&docid=169195&pageIndex=0&doclang=ES&mode=lst&dir=&occ=first&part=1&cid=844444

[1287] Véase el artículo 37 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

[1288] Véase los artículos 12.4, 13.2.d) y 14.2.e) del RGPD.

RGPD). ii) cuando consideren que sus derechos en virtud del presente Reglamento han sido vulnerados como consecuencia de un tratamiento de sus datos personales por parte del responsable o el encargado (Artículo 79 RGPD).

Para el primero de los casos, esto es, cuando se impugne una decisión de una autoridad de control, el recurso judicial deberá ejercitarse ante los tribunales del Estado miembro en que esté establecida la autoridad de control. (Artículo 78.3 del RGPD).

En cambio, cuando el recurso se interponga frente a una conducta del responsable o el encargado, el particular podrá presentar la reclamación judicial ante los tribunales del Estado miembro en el que el responsable o encargado tenga un establecimiento o, ante los tribunales del Estado miembro en que el interesado tenga su residencia habitual. (Artículo 79.2 RGPD). Permitir al particular poder plantear el recurso en su país de residencia es muy pertinente en el contexto de la toma de decisiones automatizadas ya que es muy frecuente que dichas decisiones algorítmicas se adopten por sistemas cuyas organizaciones que los utilizan se encuentren ubicadas fuera del estado miembro al que pertenece ese particular. Ello sobre todo ocurrirá con los servicios y plataformas en línea. Esta vía puede resultar más eficaz ya que se evita acudir primeramente a la autoridad de control, la cual, y a pesar de que pueda conceder la solicitud del particular, la misma puede retrasarse si el responsable recurre tal decisión ante los tribunales. Así, a modo de ejemplo, en Holanda los particulares afectados por un sistema algorítmico que tenía como objetivo evaluar su comportamiento acudieron directamente al tribunal competente alegando el derecho de acceso a los perfiles que estaba generando este sistema ya que el responsable se negó a ello[1289], en este caso no se acudió a la autoridad de control pertinente sino que directamente se optó por la vía judicial.

3. La representación de los interesados por organizaciones o asociaciones sin ánimo de lucro

El artículo 80 del RGPD regula varios supuestos en los que un tercero puede actuar como representante del interesado con relación al cumplimiento de la normativa en materia de protección de datos. Tal y como se ha señalado, a través de este precepto se cubren varias situaciones donde estas organizaciones pueden representar a los

[1289] Tribunal de Ámsterdam. Resolución del 11-03-2021, apartados 4.44, 4.45 y 4.46. Disponible en: https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBAMS:2021:1019&showbutton=true&keyword=ola+cabs

interesados, estas son: i) cuando se pretenda reclamar ante la autoridad de control la infracción de los derechos de los interesados que representan por parte del responsable o encargado, ii) impugnar ante la justicia ordinaria las decisiones de la autoridad de control que no hayan protegido de forma adecuada esos mismos derechos, iii) demandar judicialmente al responsable o encargado del tratamiento por lesionar tales derechos , iv) solicitar una indemnización económica por la lesión sufrida[1290].

Ahora bien, no toda organización o asociación puede representar a los interesados sino que se requieren de unos requisitos, esto son: a) la organización o asociación ha de ser sin ánimo de lucro, b) debe constituirse con arreglo al derecho del estado miembro, c) deben prever en sus estatutos objetivos como la persecución de intereses públicos y en concreto los relacionados con la protección de datos. Con estas exigencias el legislador europeo pretende evitar que se creen organizaciones cuyo principal incentivo sea el interés económico a la hora de ejercer estas funciones[1291]. Se consigue además que los particulares puedan dejar a organizaciones más especializadas estos asuntos cuando los afectados no tienen tiempo o recursos para llevar a cabo estas reclamaciones[1292]. Como puede ya intuirse, la representación contemplada en este precepto puede ser realmente beneficiosa para el control de los sistemas automatizados por parte de este tipo de organizaciones ya que en muchos casos los particulares por sí solos se encuentran limitados a la hora de ejercer sus derechos o en su caso exigir el cumplimiento de la normativa de protección de datos.

Es turno de analizar los mecanismos que ha previsto el RGPD a la hora de instrumentalizar la representación de los interesados.

Por un lado, el artículo 80.1 permite que un interesado pueda dar mandato a una organización para que esta pueda ejercer las actuaciones previamente mencionadas. De esta manera, el particular ha de contactar con una de estas organizaciones para que la misma lleve a cabo esas acciones. En nuestra opinión, los interesados deben ser conscientes de la existencia de este derecho. De esta manera, cuando se adopten decisiones total o parcialmente automatizadas relevantes, al particular se le debería

[1290] RAMOS, PASCUAL,D: "Reflexiones sobre el artículo 80 del Reglamento Europeo de Protección de Datos". *LA LEY privacidad*, Nº 7, Sección El foro de la privacidad, Primer trimestre de 2021, pág.2.
[1291] ROIG I BATALLA,A: *Las garantías frente a las decisiones automatizadas del Reglamento general de Protección de Datos a la gobernanza algorítmica*.op.,cit, pág.128.
[1292] RAMOS, PASCUAL,D: "Reflexiones sobre el artículo 80 del Reglamento Europeo de Protección de Datos", op.,cit, pág.4.

informar sobre la existencia de tal facultad. La idea es que si el particular considera complejo recurrir o ejercer algunas de las facultades que se derivan de la normativa de protección de datos en relación con las decisiones algorítmicas, este pueda acudir a asociaciones especializadas que puedan representar sus intereses de forma adecuada. La información resulta por tanto relevante ya que se deja en manos de los interesados la posibilidad de que los mismo contacten con estas organizaciones. Si estos desconocen que este derecho existe, es imposible que lo ejerzan. En España, este derecho se ha regulado tímidamente y su contenido no ha supuesto un gran cambio a lo ya contemplado por el RGPD[1293].

Por otro lado, en segundo lugar, el artículo 80.2 del RGPD faculta a las organizaciones y asociaciones sin ánimo de lucro descritas previamente a que puedan ejercer tanto las reclamaciones ante las autoridades de control como en su caso en vía judicial en representación de los interesados. Así, estas organizaciones pueden ser realmente útiles a la hora de controlar y supervisar que los responsables que utilizan sistemas de toma de decisiones automatizadas cumplen con la normativa de protección de datos. El papel de las mismas puede ser realmente relevante sobre todo con relación a las medias de cumplimiento de responsabilidad activa que en su caso han de implementar los responsables cuando existan sospechas de que los sistemas automatizados no están cumpliendo con la normativa de protección de datos. De esta manera, varios afectados por un mismo sistema pueden ser representados por estas organizaciones a fin de que se respeten sus derechos. Actuaciones como el control de los sesgos algorítmicos, las peticiones de información, análisis exhaustivo de las tasas de precisión del sistema y documentación relacionada con el uso de sistemas automatizados o el ejercicio adecuado de los derechos relacionados con la elaboración de los perfiles pueden ser algunas de las acciones que estas organizaciones pueden ejercitar en representación de los interesados. Muchas de estas actuaciones resultan complejas para un particular que carece de conocimientos técnicos sobre sistemas

[1293] Así, el artículo 12.1 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales establece que *los derechos reconocidos en los artículos 15 a 22 del Reglamento (UE) 2016/679, podrán ejercerse directamente o por medio de representante legal o voluntario.* A su vez, el artículo 55.2 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 indica que: *el interesado podrá conferir su representación a una entidad, organización o asociación sin ánimo de lucro que haya sido correctamente constituida, cuyos objetivos estatutarios sean de interés público y que actúe en el ámbito de la protección de los derechos y libertades de los interesados en materia de protección de sus datos personales, para que ejerza los derechos contemplados en el apartado anterior.*

automatizados, estos intermediarios, los cuales pueden disponer de equipos o personas especializadas en la materia, pueden ayudar a dichos particulares. Pues bien, a pesar de que este precepto puede resultar muy interesante, las prescripciones contenidas en el mismo quedan condicionadas a que los Estamos Miembros así lo reconozcan en su derecho interno. En este sentido, el legislador español no ha regulado esta materia explícitamente por lo que a día de hoy esta vía queda limitada. En busca de soluciones, la doctrina ha propuesto que en aquellos supuestos en los que algún tratamiento ilícito de datos personales pudiera también considerarse una infracción que perjudicara los intereses colectivos de los consumidores, aprovechar las acciones que brinda esta legislación para en su caso poder obtener la representación de las asociaciones de consumidores[1294]. Por otro lado, si finalmente se aprueba la propuesta de reglamento europeo sobre gobernanza de datos, algunas de las restricciones del artículo 80.2 pueden solventarse. Este texto prevé la creación de los llamados proveedores de servicios de intercambio de datos que ofrecen sus servicios a los interesados en el sentido del Reglamento (UE) 2016/679. Estos proveedores tienen como objetivo principal mejorar las acciones individuales y el control de las personas sobre los datos que les conciernen. Además, les pueden asesorar y ayudar en el ejercicio de sus derechos[1295]. A pesar de que sus funciones están focalizadas en la reutilización de datos, sería recomendable que a estos mismos intermediarios se les pudiera facilitar la opción de representar a los interesados conforme a lo descrito por el artículo 80 del RGPD en dicha propuesta normativa.

Pese a esas posibles alternativas, abogamos no obstante por un reconocimiento explícito de la norma que habilite la representación de los derechos de los particulares por este tipo de asociaciones en nuestro ordenamiento jurídico interno.

VI. LA DIMENSIÓN COLECTIVA DE LOS DERECHOS

1. Límites de la protección de datos en el desarrollo y despliegue de sistemas de toma de decisiones automatizadas

[1294]Estas acciones colectivas son la de cesación y la de reparación. En: FERNÁNDEZ-SAMANIEGO Y BLAS PIÑAR GUZMÁN,J: "Las acciones colectivas en el marco del RGPD: una perspectiva desde el Derecho Civil español". *Diario La Ley*, Nº 26, Sección Ciberderecho, 11 de Febrero de 2019, pág.2.

[1295] Considerando 23 y artículo 9.1.b) de la Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos). Texto aprobado el 25 de noviembre de 2020.

A lo largo de todo este trabajo hemos analizado las implicaciones legales que presentan el uso de sistemas de toma de decisiones automatizadas en el derecho fundamental a la protección de datos. Así, podemos comprobar como el RGPD se estructura básicamente en tres pilares para regular los tratamientos de datos personales.

El primer pilar queda englobado por el conjunto de principios que han de respetar los responsables cuando traten datos personales. Así, si una organización pretende desarrollar un modelo algorítmico, ésta, por ejemplo, ha de justificar todas y cada una de las variables elegidas (minimización de datos) y además ha de suprimir los datos un vez haya pasado un determinado periodo de tiempo desde que finalizó el entrenamiento de los mismos (limitación del plazo de conservación).

El segundo pilar, está formando por el conjunto de derechos que el RGPD reconoce en favor de los particulares que se ven sometidos a estos sistemas. En este sentido, entre otras facultades, el particular tiene derecho acceder a los perfiles que elabora el responsable, a solicitar la rectificación, oponerse a los mismos, solicitar la revisión humana, el derecho de explicación, etc.

En tercer lugar, la normativa europea obliga a los responsables del tratamiento a implementar toda una serie de medidas de cumplimiento y salvaguardas adecuadas con el doble objetivo de, por un lado, prever y por consiguiente reducir los posibles riesgos en los derechos de los individuos que pueden generar los tratamientos de datos personales que llevan a cabo y por otro lado, demostrar que efectivamente se está cumpliendo con la normativa de protección de datos. Es decir, que se respetan los pilares primero y segundo antes mencionados. Así, teniendo en cuenta los riesgos que genera el uso de sistemas de toma de decisiones automatizadas y parcialmente automatizadas, los responsables deberán de desplegar medidas como la EIPD, las auditorías externas o la designación del DPD entre otras.

Pues bien, a pesar de todo este elenco de exigencias normativas, el uso de sistemas de toma de decisiones automatizadas en los entornos de *big data* genera riesgos tan dispares y difusos que en muchos casos las herramientas reguladas y reconocidas por la normativa de protección de datos escapan a la misma o en su caso resultan poco protectoras.

Así, comenzando por la fase del diseño de los sistemas, resulta habitual que los mismos se desarrollen con datos anonimizados. Es decir, las técnicas de minería de

datos y *machine learning* permiten elaborar todo tipo de perfiles grupales basados en correlaciones de datos anonimizados. Estos perfiles puedes estar representados por grupos más o menos definidos que presenten unas u otras características pero las correlaciones existentes entre los mismos no tienen por qué ser causales[1296]. La protección de datos en estas esferas, cuando se traten datos anonimizados no resulta aplicable. Es cierto que el RGPD ofrece herramientas para compensar esta situación. Así, derivado del principio de privacidad desde el diseño se desprenden diversas obligaciones, entre estas: i) los responsables están obligados a desplegar este principio desde las primeras fase del ciclo de vida siempre que los mismos pretendan desplegar ese sistema, ello, independientemente de que en la fase del diseño se traten o no datos personales. ii) A su vez, si una organización pretende adquirir un sistema de toma de decisiones automatizadas, esta está obligada a adquirir productos que respeten la normativa de protección de datos, lo que obligará a los diseñadores a elaborar sistemas que se adecuen a esta normativa. La PRAI además complementará las obligaciones del diseño de los sistemas cuando los mismos sean de alto riesgo[1297].

Por otro lado, por lo que se refiere a la fase de toma de decisiones automatizadas, el derecho a la protección de datos establece cauces para impugnar la decisión e incluso alterarla. Sin embargo, se muestra menos práctico a la hora de facilitar a los interesados herramientas adecuadas para obligar al responsable a alterar los modelos algorítmicos cuando los mismos arrojen resultados no satisfactorios a estos individuos. A modo de ejemplo, un particular podrá solicitar la revisión de una denegación de un préstamo automatizado aportando a la entidad bancaria otra información no tenida en cuenta durante el proceso algorítmico que resulte relevante a la hora de reconsiderar esa decisión. Por tanto, a pesar de que la revisión humana del modelo puede revertir ciertas situaciones[1298], el particular no podrá por ejemplo reclamar que el modelo es inadecuado o injusto por haber desarrollado esas categorías o perfiles. Es decir, el particular podría alterar la asignación individual que ha realizado el algoritmo al incluirlo a ese en ese perfil demostrando por ejemplo que esos datos son

---

[1296] KAMMOURIEH, L; BAAR,T; BERENS, J; LETOUZÉ, E; MANSKE,J; PALMER,J; SANGOKOYA, D; VINCK,P: "Group privacy in the age of big data". En: TAYLOR, L; FLORIDI, L ; VAN DER SLOOT, B (eds) Authors' final draft: *Group Privacy: new challenges of data technologies*. Ed.Springer. Dordrecht, 2017. pág.54.

[1297] Capítulo III, apartado III de la tesis.

[1298] MCGREGOR, L; MURRAY,D; NG,V: "International human rights law as a framework for algorithmic accountability". *International and Comparative Law Quarterly, 68*(2), pág.337. Texto disponible en: **https://www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6**

inexactos. Sin embargo, no puede alterar el perfil como tal que ha llegado a la conclusión que las personas que presentan esas características se les asignan unas determinadas consecuencias jurídicas. Ello es así porque esas demandas excederían de una petición individual en sentido estricto, afectando más a la esfera del colectivo representado por ese grupo del perfil.

2. La dimensión colectiva de la privacidad

De esta manera, en muchos supuestos, la elaboración de los perfiles puede afectar en mayor grado a los grupos que integran ese perfil que al propio individuo en sí, el cual, en determinados casos puede solicitar la no pertinencia al mismo o alterar la decisión que ha adoptado el algoritmo basado en dicho perfil. Sin embargo, esta misma normativa no ofrece un marco adecuado a las acciones en favor del grupo. En este sentido, la AEPD ha definido la privacidad de grupo o grupal como aquella privacidad correspondiente a grupos definidos por cualquier característica o combinación de características que se asocian a determinados individuos[1299]. La creación de estos grupos tiene su origen cuando se realiza la designación de las variables por parte de los analistas de acuerdo con algún propósito o cuando el propio modelo detecta correlaciones tras ser alimentado con los datos designados por los desarrolladores. Estos grupos puede coincidir o no con otros grupos intuitivos preexistentes[1300]. Esto explica por qué los grupos son tan dinámicos. Si se cambia el propósito, se cambia el conjunto de propiedades relevantes[1301]. Este enfoque ayuda a explicar por qué la elaboración de perfiles puede infringir la privacidad del grupo resultante desde el inicio. Es decir, si el perfilado está orientado por un objetivo que en sí mismo no pretende respetar la privacidad del grupo, esta queda afectada. También resultaría atacada la privacidad de grupo cuando los miembros del grupo no sean conscientes de ese perfilado por

[1299] Agencia Española de Protección de Datos. *Privacidad de grupo*. Información disponible en: https://www.aepd.es/es/prensa-y-comunicacion/blog/privacidad-de-grupo

[1300] FLORIDI,L: "Group Privacy: a Defence and an Interpretation". En: TAYLOR, L; FLORIDI, L ; VAN DER SLOOT, B (eds.) Authors' final draft: *Group Privacy: new challenges of data technologies*. Ed.Springer. Dordrecht, 2017, pág 107 y ss.

[1301] FLORIDI,L: "Group Privacy: a Defence and an Interpretation". En: TAYLOR, L; FLORIDI, L ; VAN DER SLOOT, B (eds.) Authors' final draft: *Group Privacy: new challenges of data technologies*, op.cit., pág.113.

pertenecer a ese grupo[1302]. De esta manera, la privacidad colectiva englobaría el derecho a limitar los daños potenciales para el propio grupo que pueden derivarse del tratamiento de datos invasivo y discriminatorio. Según esta interpretación, las dimensiones colectivas de la privacidad y la protección de datos se refieren principalmente al uso de la información. La fuente de preocupación no es la falta de secreto e intimidad que representa el objeto de la privacidad colectiva sino el uso injusto y perjudicial de los datos que se procesan mediante el uso de la analítica moderna[1303].

A) Cauces para encarar esta problemática

Para afrontar algunos de los problemas que pueden genera el uso de sistemas algorítmicos en los perfiles de grupos se han aportado distintas alternativas para complementar el marco normativo de la protección de datos en este contexto.

En primer lugar, la doctrina ha abogado por esa dimensión colectiva de la privacidad indicando que se hace necesario reconocer un enfoque diferente que no puede basarse únicamente en los derechos individuales[1304]. Estos derechos personalísimos han quedado en parte superados por la nueva escala masiva y analítica de datos. Ello supone el reconocimiento de otra capa superior de protección representada por los derechos de los grupos a la protección de su dimensión colectiva de la privacidad[1305].

En segundo lugar, se han de habilitar cauces que permitan la representación colectiva de estos grupos. En este sentido, no es la primera vez que el ordenamiento jurídico reconoce la dimensión colectiva de determinados grupos, sobre todo, cuando los individuos que pertenecen a ese grupo se encuentran en una posición de desequilibrio respecto de la otra parte. Ello ocurre por ejemplo en ámbitos jurídicos como en las relaciones entre consumidores y empresarios, trabajador y patrono o ciudadanos y grandes empresas en relación con el medio ambiente. Las actuaciones que

[1302] TAYLOR, L; FLORIDI, L ; VAN DER SLOOT, B: " Introduction: a new perspective on privacy". En: TAYLOR, L; FLORIDI, L ; VAN DER SLOOT, B (eds) Authors' final draft: *Group Privacy: new challenges of data technologies*, op.cit., pág.17.

[1303] MANTELERO, A: "Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection". *Computer law and security review*, 32 (2016), pág.245 y ss. Texto disponible en: https://www.sciencedirect.com/science/article/pii/S0267364916300280

[1304] Sobre la necesidad de trabajar con una dimensión colectiva de los derechos véase: COTINO HUESO, L: "Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales", op.cit., pág.137.

[1305] MANTELERO, A: "From group privacy to collective privacy: towards a new dimension of privacy and data protection in the big data era", En: TAYLOR, L; FLORIDI, L ; VAN DER SLOOT, B (eds) Authors' final draft: *Group Privacy: new challenges of data technologies*. op.cit., págs. 173 y ss.

llevan las organizaciones a la hora de recopilar masivamente los datos y extraer conclusiones sobre las que posteriormente se tomarán decisiones dan lugar a tratamientos de datos complejos y en muchos casos ocultos para los ciudadanos. Ello hace que el control de la normativa por parte de estos últimos los sitúe en una posición de inferioridad a la hora de reclamar sus derechos[1306]. No es por tanto descabellado que en el ámbito de la protección de datos el papel de las organizaciones y asociaciones alcance un nuevo papel. El artículo 80 del RGPD, ya comentado[1307], representa un paso adelante en esta dirección, sin embargo, se hace necesario un paso más allá. Es decir, no sólo se potencia la representación de los interesados de forma individual, sino también la representación de los colectivos pertenecientes a esos grupos. Una propuesta que podría resultar interesante es que, en la medida de lo posible, esas organizaciones que ya tienen una dilatada experiencia en representar a los individuos en sus distintos contextos vean ampliadas su capacidad para exigir el cumplimiento adecuado de la normativa de protección de datos. Estos es, asociaciones de consumidores o sindicatos pueden asumir tales tareas en sus respectivos ámbitos de actuación[1308]. Recordemos además nuevamente que el papel de los proveedores de servicios de intercambio de datos que establece la propuesta de reglamento sobre gobernanza de datos también puede tener un papel radical en este contexto[1309].

En tercer lugar, la doctrina también ha apostado por la elaboración de evaluaciones de impacto sociales y éticas respecto del contexto o sociedad en la que se pretende desplegar el sistema de toma de decisiones automatizadas. De esta manera, primeramente se deberían definir los valores relacionados con el contexto y la relación entre los intereses y derechos en conflicto con relación al uso del modelo algorítmico en ese contexto específico. Para ello, se recomienda establecer un sistema obligatorio de evaluación de riesgos múltiples teniendo en cuenta las repercusiones sociales y éticas

[1306] PLAZA PENADÉS,J; PEDREÑO,A; MORENO,L: *Big Data e Inteligencia Artificial. Una visión económica y legal de estas herramientas disruptivas*, op.cit., pág.33.
[1307] Sobre la representación de los interesados véase el Capítulo V, apartado V, punto 3 de esta tesis.
[1308] En el mundo laboral la doctrina ha apostado por la dimensión colectiva de los derechos de protección de datos por parte de los representantes de los trabajadores. EN: TODOLÍ SIGNES,A: "La gobernanza colectiva de la protección de datos en las relaciones laborales: big data, creación de perfiles, decisiones empresariales automatizadas y los derechos colectivos", op.cit., pág.84 y ss. Así, el Parlamento Europeo alienta a los interlocutores sociales a que, cuando sea necesario, actualicen los convenios colectivos de forma que las normas de protección en vigor puedan mantenerse también en el entorno laboral digital. En: Parlamento Europeo. Resolución del Parlamento Europeo, de 15 de junio de 2017, sobre una Agenda Europea para la economía colaborativa, apartado 38, pág.10.
[1309] Considerando 23 y artículo 9.1.b) de la Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos). Texto aprobado el 25 de noviembre de 2020.

del uso de estos sistemas[1310]. Tal análisis se realizaría bajo la supervisión de las autoridades de protección de datos nacionales. También deberán definir los requisitos profesionales de esos terceros. La EIPD reconocida en el RGPD, si bien focaliza su atención en la protección de datos, la misma ha de evaluar los riesgos que el tratamiento pretendido entraña para los derechos y libertades de las personas, por tanto, esa evaluación global de los riesgos podría tener encaje. En este sentido, el artículo 18.7 de la Carta de los Derechos Digitales española establece la necesidad de realizar una evaluación de impacto en los derechos digitales en el diseño de los algoritmos cuando las Administraciones Públicas desplieguen sistemas de decisiones automatizadas o semi automatizadas[1311]. A pesar de que este texto no es vinculante, existe una intención clara por el despliegue de estas herramientas.

[1310] MANTELERO, A: "From group privacy to collective privacy: towards a new dimension of privacy and data protection in the big data era", En: TAYLOR, L; FLORIDI, L ; VAN DER SLOOT, B (eds) Authors' final draft: *Group Privacy: new challenges of data technologies*. op.cit., pág.187 y ss.

[1311] Artículo 18.7 de la Carta de derechos digitales española. Texto no vinculante aprobado en julio de 2021.

CONCLUSIONES

1. La automatización del proceso decisorio y el valor económico de los datos personales son dos fenómenos con los que la normativa de protección de datos ha de lidiar en los próximos años.

En las últimas décadas la automatización de los procesos de toma de decisiones a través del despliegue de sistemas algorítmicos se ha disparado. Los factores que han potenciado este fenómeno son la disponibilidad masiva de datos, el abaratamiento de los costes de almacenamiento de esos datos y el desarrollo de tecnologías cada vez más avanzadas para el procesamiento de los mismos. Esta *tormenta perfecta* ha potenciado el desarrollo de sistemas de toma de decisiones automatizados cada vez más precisos. Este tipo de algoritmos se introducen en entornos cada vez más cambiantes con un nivel de error relativamente bajo y por tanto tolerable para una organización. Hace ya algún tiempo que los datos han dejado de ser útiles solamente para prestar un servicio público o para concertar un contrato privado. Gracias a las nuevas técnicas de análisis masivo de datos, el valor secundario de estos permite a las organizaciones tanto públicas como privadas obtener conocimiento muy relevante escondido tras los mismos. La normativa de protección de datos se muestra como una herramienta esencial de salvaguarda de los derechos y libertades de los particulares.

2. Un cumplimiento efectivo de la normativa de protección de datos requiere de conceptos claros y específicos con relación al uso de sistemas de toma de decisiones automatizadas.

2.1 El concepto de decisión definido en esta tesis y que consideramos que ha de ser tenido en cuenta a la hora de abordar el estudio del RGPD atañe a toda resolución que adopta una organización basada en el resultado que emite un sistema algorítmico una vez que este procesa los datos. Para que el resultado quede abarcado en esta definición caben dos alternativas: i) el resultado se automatiza completamente y genera efectos en el particular (decisión totalmente automatizada). ii) El resultado emitido por el sistema es de tal calibre que el mismo presenta una impronta real y material en la formación de

la decisión finalmente adoptada, todo ello a pesar de que hayan intervenido personas en dicho proceso de formación (decisión parcialmente automatizada).

2.2 Los datos inferidos se refieren a aquella información relativa a una persona que deducen los sistemas automatizados tras procesar los datos del particular. Dicha información es utilizada por las organizaciones para realizar una estimación de que una persona presenta la característica y además, sobre dicha característica estimada y atribuida se establecen una serie de consecuencias. Estos datos han de considerarse datos personales. Ello es así ya que los mismos reflejan información que, pudiendo ser o no cierta, se refieren a una persona. Fruto de esa naturaleza híbrida que caracteriza a estos datos, es decir, su origen deviene de los propios titulares de los datos y del procesamiento que haga de estos el algoritmo, algunos de los derechos que reconoce la normativa de protección de datos en favor de los interesados pueden verse restringidos.

2.3 En el contexto de la toma de decisiones automatizadas, los datos de categoría especial aglutinan no sólo los datos que inicialmente son sensibles sino también todos aquellos datos convencionales que resulten ser variables *proxy* de datos especiales o aquellos datos convencionales que se utilicen con la finalidad de inferir datos de categoría especial. Esta expansión de la categoría de dato especial se justifica ya que la analítica masiva de datos por parte de algoritmos potencia la obtención de información sensible proveniente de datos aparentemente neutros.

3. Por defecto, los responsables del tratamiento deberían reconocer los mismos derechos y prever similares garantías cuando utilicen sistemas de toma de decisiones total y parcialmente automatizadas.

3.1 El RGPD se aplica enteramente a todos los tratamientos de datos personales que se llevan a cabo durante el diseño y despliegue de sistema automatizados. La mención expresa a la elaboración de perfiles y la toma de decisiones plenamente automatizadas relevantes en este texto responde a una especial preocupación por este tipo de operaciones, sin embargo, ello no implica una desprotección para el resto de tratamientos basados en decisiones parcialmente automatizadas.

3.2 En el contexto de la toma de decisiones automatizadas, el enfoque basado en el riesgo sobre el que descansa el principio de responsabilidad activa permite a las organizaciones implementar toda una serie de herramientas de cumplimiento normativo a medida. La plena o no automatización de las decisiones pasa a un segundo plano ya que lo más relevantes son los efectos que dichas decisiones pueden causar en los particulares. De manera que en muchos escenarios las garantías de las decisiones total y parcialmente automatizadas serán similares o muy parecidas.

3.3 El derecho a no ser objeto de decisiones plenamente automatizadas relevantes reconoce una serie de facultades específicas en favor de los particulares que se ven sometidos a estos concretos tratamientos. Los responsables no sólo han de habilitar el ejercicio de estos derechos cuando adopten las decisiones descritas en el artículo 22 del RGPD sino también cuando dichas decisiones sean parcialmente automatizadas y las mismas generen efectos relevantes en la esfera de los particulares. Un análisis adecuado de los riesgos que producen estos últimos tratamientos justificaría el despliegue de dichas facultades.

4***. Las organizaciones están obligadas a integrar desde la fase inicial del diseño de los modelos algorítmicos la normativa de protección de datos. El objetivo que se pretende con ello es tratar de conseguir que una vez el sistema comience a adoptar decisiones, esta normativa sea respetada.***

4.1 La privacidad desde el diseño se convierte en un principio *bisagra* que une las fases de diseño y despliegue de los sistemas automatizados. Este principio obliga al responsable a tener en cuenta desde la fase de desarrollo del sistema algorítmico las implicaciones jurídicas que en materia de protección de datos posteriormente se exigirán una vez el sistema se despliegue en el mundo real. Es decir, la privacidad desde el diseño en este contexto no sólo logra que se cumpla la protección de datos desde las fases iniciales sino sobre todo, perfila e instrumenta las posibles implicaciones de privacidad que en un futuro el sistema algorítmico puede generar una vez que el mismo comience a adoptar decisiones.

4.2 Las organizaciones que adquieran sistemas para la toma de decisiones automatizadas están obligadas a comprar algoritmos que por diseño hayan integrado la

normativa de protección de datos. Los responsables han de actuar con la debida diligencia a la hora de elegir a encargados del tratamiento y productos tecnológicos que respeten la normativa de protección de datos.

4.3 Las organizaciones deberán integrar en el diseño de los sistemas automatizados las funcionalidades adecuadas para que los particulares puedan ejercer las facultades y derechos que les reconoce el RGPD una vez que dichos algoritmos comiencen a adoptar decisiones o elaborar perfiles. La organización que adquiera un sistema de este tipo deberá comprobar que el mismo permite el ejercicio de estas facultades, ejercicio que deberá adaptarse al contexto donde dicho sistema irradiará sus efectos.

5. Las medidas de responsabilidad activa suponen uno de los pilares básicos en los que se sustenta la normativa de protección de datos. Fruto de los riesgos que comporta el uso de sistemas automatizados y la elaboración de perfiles, los responsables del tratamiento han de desplegar todas las garantías necesarias para aminorar o reducir tales riesgos.

5.1 La realización de la Evaluación de Impacto de Protección de Datos presenta toda una serie de ventajas tanto para el responsable como para los interesados. Así, esta herramienta puede utilizarse como: i) método de evaluación o validación de un sistema algorítmico previo a su adquisición, ii) herramienta que justifique la necesidad de desplegar mayores garantías en los sistemas de toma de decisiones parcialmente automatizadas, iii) herramienta que facilita el despliegue de los derechos y principios previstos por la normativa de protección de datos gracias a toda la información que se recopila, iv) estándar sobre el cual se pueda monitorear el funcionamiento del sistema algorítmico.

5.2 A su vez, los códigos de conducta y los órganos que los supervisan pueden ser fundamentales. Los primeros permiten focalizarse en las principales fricciones que presentan el desarrollo y despliegue de sistemas automatizados en relación con las exigencias del RGPD en sectores específicos donde sea frecuente el uso de este tipo de algoritmos. Los órganos que supervisan dichos códigos de conducta además pueden facilitar listas actualizadas de nuevas técnicas, herramientas, investigaciones científicas o tecnologías que se están desarrollando por la industria o el mundo académico que

potencien la protección de la privacidad durante las fases que comprenden el ciclo de vida de estos sistemas algorítmicos.

5.3 Las certificaciones focalizadas en la normativa de protección de datos resultan una herramienta muy garantista para este derecho. Sin embargo, este mecanismo tendrá poca implementación en la práctica si dichas certificaciones únicamente se centran en el factor de privacidad. Apostamos por certificaciones que tengan en cuenta los impactos globales del uso de sistemas automatizados y parte de su contenido englobe la protección de datos.

5.4 Los responsables del tratamiento no están obligados a realizar auditorías en materia de protección de datos. No obstante, teniendo en cuenta los riesgos que conlleva el uso de sistemas automatizados, las auditorías se convierten en un mecanismo imprescindible para mitigarlos y para demostrar el cumplimiento de la normativa de protección de datos en estos contextos.

6. Los tratamientos de datos presentes durante el ciclo de vida de los sistemas automatizados son complejos y generan importantes riesgos para los interesados. Se ha de prestar atención especial a las bases de legitimación que permiten el tratamiento de dichos datos en virtud del principio de licitud.

6.1 El consentimiento como mecanismo de legitimación sólo se entenderá válidamente otorgado si el particular puede llegar a conocer realmente qué se pretende con el diseño del algoritmo, el perfilado o la toma de decisiones.

6.2 El interés legítimo es un instrumento de legitimación válido para autorizar los tratamientos basados en la elaboración de perfiles y la toma de decisiones automatizadas. No obstante, teniendo en cuenta los riesgos que dichos tratamientos comportan para los particulares, las garantías a implementar resultarán esenciales. Esas garantías deberán compensar efectivamente los impactos que dichos tratamientos pueden generan en los particulares.

6.3 La norma con rango de ley que autorice el tratamiento de datos basado en la toma de decisiones plenamente automatizadas relevantes señaladas en el artículo 22 del

RGPD deberá reconocer suficientes medidas de garantía. Al menos dichas salvaguardas deben ser similares o parecidas a las que reconoce el RGPD cuando los tratamientos de datos se basen en el consentimiento o en la ejecución/formalización del contrato. Estas garantías son: el derecho a obtener intervención humana por parte del responsable, el derecho del interesado a expresar su punto de vista, impugnar la decisión y en su caso a la explicación de la decisión.

6.4 El cumplimiento de los requisitos que impone la PRAI a los sistemas automatizados de alto riesgo no supone una autorización para llevar a cabo los tratamientos de datos que se realizan por esos sistemas. Este texto sólo acredita que dichos algoritmos cumplen unos requisitos de calidad y cumplimiento normativo exigidos por ese texto, sin embargo, sigue resultando necesaria la obtención de una base de legitimación adecuada en virtud del principio de licitud reconocido por la normativa de protección de datos.

6.5 La Administración Pública que pretenda legitimar la toma de decisiones automatizadas o la elaboración de perfiles basada en una competencia que le autoriza a llevar a cabo una misión o poder público ha de relacionar adecuadamente dichos tratamientos con esa competencia y el poder o misión pública que se alega. Para justificar la necesariedad de ese tratamiento se debe superar el juicio de proporcionalidad con relación al despliegue del sistema automatizado en el contexto donde este algoritmo irradiará sus efectos.

6.6 En el ámbito público, las normas que autoricen el tratamiento de datos personales basados en la toma de decisiones automatizadas requerirán como mínimo el rango de ley. Ese texto legal debe contener suficientes medidas de garantía e indicar claramente las finalidades que se pretende con el despliegue de dichos algoritmos. Las especificaciones técnicas del sistema pueden regularse por vía reglamentaria. A día de hoy, el artículo 41 de la Ley 40/2015 de Régimen Jurídico del Sector Público no es una base de legitimación adecuada para autorizar los tratamientos de datos mencionados.

7. El análisis masivo de datos con el objetivo de diseñar sistemas de toma de decisiones automatizadas es compatible con la normativa de protección de datos. La seudonimización de datos se muestra como el aliado perfecto en estos contextos.

7.1 La anonimización de datos personales no siempre será la estrategia adecuada a la hora de afrontar un proceso de análisis masivo de datos. Los importantes costes de una correcta anonimización, los riesgos de reidentificación y la pérdida de utilidad de los datos tras el proceso de anonimización pueden empujar a las organizaciones a la búsqueda de otras alternativas en el tratamiento masivo de datos.

7.2 La seudonimización de los datos en el contexto del diseño de sistemas automatizados beneficia tanto al responsable como al interesado. El uso de datos seudonimizados mejora la calidad de los resultados que genera la analítica masiva de datos personales sin necesidad de proceder a la anonimización de los mismos. Se conserva la utilidad potencial de los datos y además los interesados ven protegidos sus derechos ya que la normativa de protección de datos continua aplicándose a estos tratamientos.

7.3 El modelo instaurado por la LOPD de 2018 con relación al tratamiento de datos seudonimizados para la investigación en salud y biomédica facilita la analítica masiva de datos. Esta norma marca una hoja de ruta adecuada que podría replicarse en otros contextos. Para ello, se han de prever suficientes medidas de garantía. La existencia de compromisos claros sobre finalidades legítimas, el estudio de minimización de datos, el plan integral de investigación, el tratamiento de datos en entornos seguros o los comités de ética internos o externos pueden ser algunas de estas salvaguardas que favorezcan esas vías legales de análisis masivo de datos. Para evitar entornos de inseguridad jurídica que pueden llevar a la paralización de sectores importantes que apuestan por el desarrollo de la inteligencia artificial. Sería recomendable que en estos contextos el legislador acabara legitimando expresamente tratamientos relacionados con la analítica masiva de datos donde los riesgos para los particulares sean menores a los beneficios generales que dichos tratamientos pueden generar a los responsables del tratamiento tanto en el sector público como el privado. A falta de una regulación que expresamente reconozca o legitime estos tratamientos, la normativa de protección de datos actual marca las líneas rojas que no se pueden cruzar pero a la vez se convierte en la herramienta interpretativa necesaria sobre la que se han de amparar dichos análisis masivos de datos.

7.4 La figura de los proveedores de los servicios de intercambio de datos reconocida por la propuesta de Reglamento de gobernanza de datos puede adquirir un papel esencial en el contexto de la analítica de datos[1312]. Estos proveedores no sólo pueden facilitar entornos seguros para el tratamiento de dichos datos sino que ayudarán al intercambio de datos entre particulares y organizaciones. Además, debido a las implicaciones que tiene el uso de sistemas algorítmicos en los colectivos o grupos de personas, este tipo de agentes pueden representar los derechos de esos colectivos y facilitar el ejercicio de las facultades que reconoce el RGPD en favor de los interesados. Es necesario que la propuesta de Reglamento comentada otorgue legitimación a estos agentes para que puedan representar de forma colectiva a los grupos afectados por las decisiones algorítmicas, así como a los interesados afectados de forma particular por la normativa de protección de datos.

8. Los principios del tratamiento de datos reconocidos por el RGPD resultan fundamentales durante todo el ciclo de vida de los sistemas automatizados.

8.1 Los principios del tratamiento reconocidos en el artículo 5 del RGPD han de alinearse en la medida de lo posible con las exigencias técnicas de robustez propias que están presentes durante el desarrollo de los sistemas y modelos algorítmicos. Para estos supuestos, el cumplimiento de estos principios normativos convergerá con las exigencias internas mínimas que un sistema debe contener para considerase que la tecnología analizada presenta una grado de madurez suficiente para desplegar efectos en el mercado.

8.2 El principio de exactitud aplicado al conjunto de tratamientos de datos personales comprendidos durante el ciclo de vida de los sistemas automatizados obliga al responsable del tratamiento a utilizar durante todas las fases que comprende dicho período datos exactos y actualizados. El objetivo principal es que una correcta aplicación de este principio limite o reduzca las posibilidades de que los sistemas automatizados acaben adoptando decisiones inadecuadas o extrayendo inferencias inexactas. En virtud del principio de exactitud, una base de datos será adecuada cuando

[1312] Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos). Texto aprobado el 25 de noviembre de 2020.

dichos datos sean actualizados, representativos y en la medida de lo posible con sesgos reducidos o inexistentes.

8.3 El principio de transparencia obliga al responsable a informar al particular de la toma de decisiones automatizadas y de la elaboración de perfiles y sus consecuencias, así como del uso de los datos inferidos que genere el modelo. La importancia de informar sobre los datos que potencialmente se inferirán resulta de suma importancia ya que en muchas ocasiones el particular sólo es consciente de los datos iniciales que el responsable le solicita pero no de aquellos que posteriormente inferirá el sistema. Esta información ha de facilitarse con independencia de si el sistema adopta decisiones total o parcialmente automatizadas. En el caso de las decisiones totalmente automatizadas relevantes además se habrá de informar sobre la lógica del tratamiento y las consecuencias previstas del mismo.

8.4. Cualquier limitación de los deberes de información deberá estar debidamente justificada y compensarse con otras medidas de garantía. Se ha de alegar el bien o interés jurídico que se trata de proteger con esa restricción de información. Es necesario sopesar por un lado la afectación que tiene sobre los derechos y libertades de los interesados ese déficit de información y por otro, los efectos perjudiciales a la hora de divulgar más o menos información con relación a dichos bienes como pueden ser los secretos comerciales o la posibilidad de que se truque el algoritmo. Como regla general, la total opacidad no estará justificada. La transparencia, aunque no sea una herramienta de protección completa en estos contextos, siempre supondrá un primer paso en favor del particular que le permitirá ser consciente del uso de estos sistemas, elemento que resulta fundamental teniendo en cuenta el oscurantismo que está presente en estos escenarios.

8.5. El estudio de minimización de datos ha de implantarse en la fase de desarrollo de los sistemas de toma de toma de decisiones automatizadas. Los responsables deben justificar en cada una de las distintas fases que comprende el diseño de los modelos algorítmicos la pertinencia de los datos utilizados para configurar el proyecto que se pretende, esto es, la finalidad de la analítica de datos. De esta manera, conformen van avanzando las fases, el conjunto de datos utilizados se irá depurando hasta quedar perfilado y conformado únicamente por aquellos que sean estrictamente necesarios y

debidamente justificados. Se trata de una herramienta de responsabilidad activa que trata de demostrar el cumplimiento del principio de minimización de datos. Además, se consigue reducir las tensiones que a priori existen entre el principio de minimización de datos y la analítica masiva de los mismos

8.6 Por lo que se refiere el principio de limitación de la finalidad, cuando un responsable proyecte anonimizar los datos personales que se obtuvieron para una finalidad distinta de la anonimización, a la hora de realizar el juicio de compatibilidad entre ambas finalidades, ha de tener en cuenta los perjuicios que se pueden derivar de los potenciales usos que se pretenden con la anonimización de los datos, usos que generalmente tendrán como objetivo la elaboración de todo tipo de algoritmos y modelos para toda clase de finalidades. De esta manera, los propósitos posteriores que se pretende obtener con los datos anonimizados no deberían tener como objetivo crear modelos que en un futuro arrojen decisiones discriminatorias, injustas, etc.

8.7 El principio de lealtad con relación al uso de sistemas automatizados obliga al responsable a manejar bases de datos que presenten el menor número de sesgos y a utilizar modelos y procedimientos estadísticos adecuados y fiables que limiten la posibilidad de que los sistemas arrojen decisiones discriminatorias. El considerando 71.2 del RGPD reconoce el principio de prohibición de discriminación de los sistemas algorítmicos sobre la base de datos de categoría especial. Cuando las decisiones que arroje el algoritmo resulten injustas por ser discriminatorias, se estará afectando al principio de lealtad y al principio de prohibición de discriminación algorítmica.

9. ***Las facultades y derechos reconocidos por la normativa de protección de datos no suelen ejercitarse por los interesados. Es necesario que los responsables diseñen cauces y herramientas que faciliten y potencien su práctica.***

9.1 Las organizaciones han de desplegar interfaces y entornos que favorezcan y faciliten la práctica de los derechos reconocidos a los interesados. La importancia de estas facultades en estos contextos resulta esencial teniendo en cuenta que una de las normas que será la referencia en este contexto, esto es, la PRAI, no prevé hasta la fecha ningún precepto relativo a los derechos en favor de los particulares sometidos a los sistemas de inteligencia artificial.

9.2 En las políticas o avisos de privacidad, toda la información referida a la toma de decisiones automatizada y a la elaboración de perfiles debería incorporarse a un apartado específico donde se haga referencia a los distintos perfiles, sus finalidades, en qué consisten, consecuencias de los mismos, datos inferidos principales, etc. De esa manera, los usuarios podrían comprender de forma más clara y específica las consecuencias de estos tratamientos.

9.3 Dentro de las garantías que ha de adoptar el responsable frente a la toma de decisiones automatizadas discriminatorias, se propone diseñar un sistema de notificación similar al previsto en el RGPD para cuando ocurre un incidente de seguridad que afecta al tratamiento de datos personales. En este sentido, en función del riesgo que pueda suponer para el interesado esa decisión algorítmica discriminatoria, el responsable ha de establecer unos u otros mecanismos de comunicación y en su caso desplegar unas u otras medias para mitigar los efectos generados.

9.4 Las continuas interacciones a las que se ven sometidos los datos procesados por los algoritmos generan todo un conjunto de información que se puede obtener a través del derecho de acceso. Mientras que durante la recopilación de los datos el responsable sólo estaba obligado a facilitar información general sobre los perfiles, las consecuencias de estos y las decisiones automatizadas. Una vez que comienza el procesamiento de datos los particulares pueden acceder a la información específica generada por el comportamiento del sistema con relación a los datos de esa persona. Se ha de facilitar información relacionada con los datos inferidos, el perfil específico, las consecuencias concretas o la afectación de las decisiones sobre ese practicar.

9.5 La elaboración de perfiles grupales difícilmente podrá generalizar y aglutinar todas y cada una de las características específicas presentes en los individuos sobre los que se aplican dichos perfiles. La posibilidad de asignar a una persona una determinada inferencia inexacta es probable. El derecho de rectificación se muestra como el instrumento ideal para actualizar y en su caso rectificar las inferencias y decisiones inexactas generadas por los algoritmos. Dicho lo anterior, la configuración del derecho de rectificación amparado en el principio de exactitud no legitima al particular para que ponga en cuestión el modelo sobre el que se adoptan las decisiones.

9.6 El derecho a no ser objeto de decisiones plenamente automatizadas relevantes no es absoluto. El propio RGPD permite este tipo de decisiones siempre que se reconozcan todo un abanico de garantías. El objetivo de estas facultades es establecer un mecanismo adecuado que facilite un trámite de audiencia donde el interesado pueda recibir una explicación de la decisión, la pueda impugnar, pueda tener contacto con una persona y además dicho interesado pueda expresar su punto de vista con relación a la explicación. Estas salvaguardas mínimas deberán acompañarse con otras que otorguen una protección especial para este tipo de tratamientos que presentan un alto riesgo para los derechos de las personas. La virtualidad de este derecho, aunque hasta la fecha escasa, irá aumentando conforme se automaticen los procesos decisorios.

CONCLUSIONS

1. The automation of the decision-making process and the economic value of personal data are two phenomena that data protection regulation will have to deal with in the coming years.

In recent decades, the automation of decision-making processes through the deployment of algorithmic systems has surged. The factors that have driven this phenomenon are the massive availability of data, the falling cost of data storage and the development of increasingly advanced data processing technologies. This *perfect storm* has fostered the development of increasingly accurate automated decision-making systems. Such algorithms are introduced into increasingly changing environments with a relatively low level of error that is tolerable for an organisation. For some time now, data has ceased to be useful only for providing a public service or concluding a private contract. Thanks to new techniques of massive data analysis, the secondary value of data allows both public and private organisations to obtain highly relevant knowledge hidden behind the data. Data protection law is proving to be an essential tool to safeguard the rights and freedoms of individuals.

2. Effective compliance with data protection law requires clear and specific concepts regarding the use of automated decision-making systems.

2.1 The concept of decision defined in this thesis, which we consider should be taken into account when considering the GDPR, refers to any decision made by an organisation based on the result issued by an algorithmic system once it has processed the data. There are two possible scenarios in which the result issued by the system may fall under the scope of this definition: i) the result is fully automated and generates effects on the individual (fully automated decision). ii) The result issued by the system is such that it has a real and material impact on the formation of the decision finally adopted, despite the fact that people have been involved in the decision-making process (partially automated decision).

2.2 Inferred data refers to information about a person that is inferred by automated systems after processing the individual's data. Such information is used by organisations

to make an estimate that a person has a characteristic and, in addition, a number of consequences are drawn from that estimated and attributed characteristic. These data are to be considered personal data because they reflect information that, whether true or not, relates to an individual. As a result of the hybrid nature that characterises these data, that is, their origin comes from the data subjects themselves and from the processing of these data by the algorithm, some of the rights recognised by data protection regulations in favour of data subjects may be restricted.

2.3 In the context of automated decision-making, special category data encompasses not only data that are initially sensitive but also all conventional data that turn out to be *proxy* variables for special data or conventional data that are used for the purpose of inferring special category data. This expansion of the category of special data is justified since massive data analytics by algorithms makes it possible to obtain sensitive information from apparently neutral data.

3. By default, controllers should recognise the same rights and provide similar safeguards when using fully and partly automated decision-making systems.

3.1 The GDPR applies in full to all processing of personal data carried out during the design and deployment of automated systems. The express mention of profiling and relevant fully automated decision-making in this text responds to a special concern for this type of operation, but does not imply a lack of protection for other processing operations based on partially automated decisions.

3.2 In the context of automated decision-making, the risk-based approach underpinning the accountability principle allows organisations to implement a range of tailor-made compliance tools. Whether or not decisions are fully automated takes a back seat to the more relevant effects that such decisions may have on individuals. So in many scenarios the safeguards for fully and partially automated decisions will be similar or very similar.

3.3 The right not to be subject to relevant fully automated decisions recognises a number of specific safeguards in favour of individuals who are subject to these specific processing operations. Controllers must not only enable the exercise of these rights when they make the decisions described in article 22 of the GDPR, but also when such

decisions are partially automated and generate relevant effects in the sphere of individuals. A proper analysis of the risks produced by the latter processing would justify the deployment of such safeguards.

4. Organisations are mandated to integrate data protection regulations from the initial design phase of algorithmic models. The aim is to try to ensure that once the system starts to make decisions, these regulations are respected.

4.1 Privacy by design becomes a *hinge* principle linking the design and deployment phases of automated systems. This principle forces the controller to consider, from the moment the development phase of the algorithmic system starts, the legal implications in terms of data protection that will subsequently be required once the system is deployed in the real world. In other words, privacy by design in this context not only ensures that data protection is complied with from the initial stages but, above all, outlines and implements the possible privacy implications that the algorithmic system may generate in the future once it starts to make decisions.

4.2 Organisations purchasing automated decision-making systems are mandated to acquire algorithms that have integrated data protection rules by design. Controllers must exercise due diligence in choosing processor and technology products that comply with data protection rules.

4.3 Organisations should build into the design of automated systems the appropriate functionalities to enable individuals to exercise their rights and powers recognised under the GDPR once these algorithms begin to make decisions or create profiles. The organisation acquiring such a system should check that it allows the exercise of these powers, which should be adapted to the context in which the system will have its effects.

5. Accountability measures are one of the basic pillars underpinning the regulation of data protection. As a result of the risks involved in the use of automated systems and profiling, data controllers must deploy all the necessary safeguards to mitigate or reduce these risks.

5.1 Carrying out a Data Protection Impact Assessment has a number of advantages for both the controller and the data subjects. Thus, this tool can be used as: i) a method of evaluation or validation of an algorithmic system prior to its acquisition, ii) a tool that justifies the need to deploy greater safeguards in partially automated decision-making systems, iii) a tool that facilitates the deployment of the rights and principles provided for by data protection law thanks to all the information that is collected, iv) a standard against which the functioning of the algorithmic system can be monitored.

5.2 In turn, codes of conduct and the bodies that oversee them can be crucial. Codes of conduct can focus on the main frictions in the development and deployment of automated systems in relation to the requirements of the GDPR in specific sectors where such algorithms are frequently used. The bodies supervising codes of conduct can also provide updated lists of new techniques, tools, scientific research or technologies being developed by the industry or academia that enhance privacy protection during the life-cycle phases of these algorithmic systems.

5.3 Certifications focused on data protection regulations are a very protective tool for this right. However, this mechanism will have little implementation in practice if such certifications only focus on privacy aspects. We thus consider the development of certification that consider the global impacts of the use of automated systems and that include data protection as part of their content to be more useful.

5.4 Controllers are not forced to carry out data protection audits. However, given the risks associated with the use of automated systems, audits are an essential mechanism to mitigate these risks and to demonstrate compliance with data protection rules in these contexts.

6. The processing of personal data in the life cycle of automated systems is complex and generates significant risks for data subjects. Particular attention should be paid to the grounds of legitimacy allowing the processing of such data under the principle of lawfulness.

6.1 Consent is only validly given if the individual can actually know what is intended by the algorithm design, profiling or decision-making.

6.2 Legitimate interest is a valid instrument of legitimisation for authorising processing operations based on profiling and automated decision-making. However, given the risks that such processing operations entail for individuals, the safeguards to be put in place are essential. These safeguards should effectively compensate for the impacts that such processing may have on individuals.

6.3 The law authorising the processing of data based on relevant fully automated decision-making referred to in article 22 of the GDPR must recognise sufficient safeguards. These safeguards should at least be similar to those recognised by the GDPR where data processing is based on consent or on the performance/formalisation of the contract. These safeguards are: the right to obtain human intervention on the part of the controller, the right of the data subject to express his or her point of view, to challenge the decision and, where appropriate, to an explanation of the decision.

6.4 Compliance with the requirements imposed by the Artificial Intelligence Act Proposal on high-risk automated systems does not imply authorisation to carry out the data processing carried out by such systems. This text only certifies that these algorithms meet the quality and regulatory compliance requirements set out in the text, but it is still necessary to obtain an adequate basis for legitimisation by virtue of the principle of lawfulness recognised by data protection law.

6.5 A public administration seeking to legitimise automated decision-making or profiling based on a competence authorising it to carry out a public task or power must adequately relate such processing to that competence and the claimed public task or power. In order to justify the necessity of such processing, the proportionality test must be carried out in relation to the deployment of the automated system in the context where this algorithm will radiate its effects.

6.6 In the public sphere, rules authorising the processing of personal data based on automated decision-making require at least the status of law. This legal text must contain sufficient safeguards and clearly indicate the intended purposes of the

deployment of such algorithms. The technical specifications of the system can be regulated by administrative regulatory instruments. As of today, article 41 of Law 40/2015 on the Legal Regime of the Public Sector is not an adequate basis of legitimisation to authorise the aforementioned data processing.

7. Mass data analysis for the purpose of designing automated decision-making systems is compatible with data protection regulations. Pseudonymisation of data proves to be the perfect ally in these contexts.

7.1 Anonymisation of personal data will not always be the right strategy when dealing with a mass data analysis process. The significant costs of proper anonymisation, the risks of re-identification and the loss of usefulness of the data after the anonymisation process may push organisations to look for other alternatives in mass data processing.

7.2 Pseudonymisation of data in the context of the design of automated systems benefits both the controller and the data subject. The use of pseudonymised data improves the quality of the results generated by mass analysis of personal data without the need to anonymise the data. The potential usefulness of the data is preserved and data subjects' rights are protected as data protection law continues to apply to such processing.

7.3 The model established by the LOPD of 2018 in relation to the processing of pseudonymised data for health and biomedical research facilitates mass data analytics. This standard sets an appropriate roadmap that could be replicated in other contexts. To this end, sufficient safeguards must be foreseen. The existence of clear commitments on legitimate purposes, the data minimisation study, the comprehensive research plan, the processing of data in secure environments or internal or external ethics committees may be some of these safeguards that favour these legal avenues of massive data analysis. To avoid environments of legal uncertainty that could lead to the paralysis of important sectors that are committed to the development of artificial intelligence. It would be advisable that in these contexts the legislator ends up expressly legitimising treatments related to massive data analytics where the risks for individuals are lower than the general benefits that such treatments can generate for data controllers in both the public and private sectors. In the absence of a regulation that expressly recognises or legitimises such processing, current data protection regulations mark the red lines that

cannot be crossed, but at the same time become the necessary interpretative tool on which such mass data analysis must be based.

7.4 Providers of data sharing services recognised by the Data Governance Act Proposal, can play a key role in the context of data analytics[1313]. These providers can not only provide secure environments for the processing of such data but will assist in the exchange of data between individuals and organisations. In addition, due to the implications of the use of algorithmic systems on groups or groups of individuals, such actors can represent the rights of those groups and facilitate the exercise of the powers granted to data subjects under the GDPR. It is necessary that the annotated Proposal grants standing to these actors so that they can collectively represent the groups affected by algorithmic decisions, as well as the data subjects affected in particular by data protection law.

8. The principles relating to processing of personal data recognised by the GDPR are fundamental throughout the life cycle of automated systems.

8.1 The processing principles recognised in Article 5 of the GDPR must be aligned as far as possible with the technical requirements of robustness that are present during the development of systems and algorithmic models. In these cases, compliance with these regulatory principles will converge with the minimum internal requirements that a system must contain in order for the technology analysed to be considered sufficiently mature to have an impact on the market.

8.2 The principle of accuracy as applied to all processing of personal data in the life cycle of automated systems requires the controller to use accurate and up-to-date data at all stages of the life cycle. The main objective is that a correct application of this principle limits or reduces the likelihood that automated systems end up making inappropriate decisions or drawing inaccurate inferences. Under the accuracy principle, a database is adequate when the data are up-to-date, representative and as far as possible with little or no bias.

[1313] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European data governance (Data Governance Act). Resolution adopted on November 25, 2020.

8.3 The principle of transparency forces the data controller to inform the individual about automated decision-making and profiling and their consequences, as well as about the use of the inferred data generated by the model. The importance of informing about the data that will potentially be inferred is of utmost importance as the individual is often only aware of the initial data that the data controller requests but not of the data that the system will subsequently infer. This information should be provided irrespective of whether the system makes fully or partially automated decisions. In the case of relevant fully automated decisions, the logic of the processing and the intended consequences of the processing should also be disclosed.

8.4. Any limitation of information duties must be duly justified and compensated by other safeguards. The legal interest or good to be protected by the restriction of information must be invoked. The impact of the information deficit on the rights and freedoms of data subjects must be weighed against the detrimental effects of disclosing more or less information in relation to such assets, such as trade secrets or the possibility of algorithm gaming. As a general rule, total opacity will not be justified. Transparency, although not a full protection tool in these contexts, will always be a first step in favour of the individual to be aware of the use of these systems, which is essential given the opacity that is present in these scenarios.

8.5. The data minimisation study must be implemented in the development phase of automated decision-making systems. Controllers must justify in each of the different phases of the design of the algorithmic models the relevance of the data used to shape the intended project. In this way, as the phases progress, the set of data used will be refined until it is refined and made up only of those that are strictly necessary and duly justified. This is an accountability tool that seeks to demonstrate compliance with the principle of data minimisation. In addition, it manages to reduce the tensions that a priori exist between the principle of data minimisation and massive data analytics.

8.6 With regard to the purpose limitation principle, when a controller plans to anonymise personal data that were obtained for a purpose other than anonymisation, the controller must, when assessing the compatibility of the two purposes, take into account the harm that may result from the potential uses to which the anonymisation of the data is put. These uses will generally be aimed at developing all kinds of algorithms and

models for all kinds of purposes. In this way, the subsequent purposes for which anonymised data are intended should not aim at creating models that in the future will lead to discriminatory and unfair decisions.

8.7 The principle of fairness in relation to the use of automated systems forces the controller to operate databases with the least possible bias and to use appropriate and reliable statistical models and procedures that limit the possibility of discriminatory decisions being made by the systems. Recital 71(2) of the GDPR recognises the principle of prohibition of discrimination of algorithmic systems on the basis of special categories of data. Where the decisions rendered by the algorithm are unfair because they are discriminatory, the principle of fairness and the principle of prohibition of algorithmic discrimination are affected.

9. *The rights recognised by data protection legislation are not usually exercised by data subjects. It is necessary for controllers to design channels and tools that facilitate and enhance their practice.*

9.1 Controllers must deploy interfaces and environments that encourage and facilitate the exercise of the rights granted to data subjects. The importance of these powers in these contexts is essential. In this regard, the Artificial Intelligence Act Proposal does not establish any provision relating to the rights of individuals subject to artificial intelligence systems.

9.2 In privacy policies, all information relating to automated decision-making and profiling should be incorporated in a specific section referring to the different profiles, their purposes, what they consist of, their consequences, main inferred data, etc. This would allow users to understand more clearly and specifically the consequences of such processing.

9.3 As part of the safeguards to be adopted by the controller in the event of discriminatory automated decisions, the proposal is to design a system of notifications similar to that provided for by the GDPR for when a personal data breach to the data subject occurs. In this regard, depending on the risk that such a discriminatory algorithmic decision may pose for the data subject, the data controller must establish

one or other communication mechanisms and, where appropriate, deploy one or other measures to mitigate the effects generated.

9.4 The continuous interactions to which data processed by algorithms are subjected generate a whole set of information that can be obtained through the right of access whereas during data collection the controller was only mandated to provide general information on profiling, the consequences of profiling and automated decisions. Once data processing starts, data subjects can access specific information generated by the system's behaviour in relation to that person's data. Information related to the inferred data, the specific profile, the specific consequences or impact of decisions on that practice should be provided.

9.5 Group profiling is unlikely to generalise and aggregate each and every specific characteristic present in the individuals being profiled. The possibility of assigning a certain inaccurate inference to a person is likely. The right to rectification appears to be the ideal instrument for updating and, where appropriate, rectifying inaccurate inferences and decisions generated by algorithms. That said, the configuration of the right to rectification under the principle of accuracy does not legitimise the individual to challenge the model on the basis of which decisions are made.

9.6 The right not to be subject to a decision based solely on automated processing is not absolute. The GDPR itself allows such decisions provided that a whole range of safeguards are set up. The purpose of these powers is to establish an appropriate mechanism to facilitate a hearing where the data subject can receive an explanation of the decision, can challenge it, can have contact with a person and can express his or her point of view in relation to the explanation. These minimum safeguards should be accompanied by others that provide special protection for this type of treatment, which presents a high risk to the rights of individuals. The virtuality of this right, although limited to date, will increase as decision-making processes become more automatic.

BIBLIOGRAFÍA

I. ARTÍCULOS CIENTÍFICOS, CAPÍTULOS DE LIBRO Y LIBROS

ADSUARA VARELA, B: "El perfilado ideológico de los ciudadanos por los partidos políticos". *El Consultor de los Ayuntamientos*, Nº III, Sección Crónica, Julio 2019.

AGGARWAL,CH.: *Data Mining. The Textbook*. Ed.Springer, Nueva York, 2015.

AKHILA, N; BRETT, K; KAVITA,S; ROBERTO,N; JUSTIN, K: "Automated Classification of Skin Lesions: From Pixels to Practice". *Journal of Investigative Dermatology*, 138, 2018.

ALIAGA MARTÍNEZ,L ; GUTIÉRREZ DAVID, E: "Explicando Machine Learning a través de la doctrina y práctica del Information Commissioner's Office". *LA LEY privacidad*, Nº 8, Sección Crónica de Corresponsales, Segundo trimestre de 2021.

ALMADA, M,. "Human Intervention in Automated Decision-Making: Toward the Construction of Contestable Systems", *17th International Conference on Artificial Intelligence and Law* (ICAIL 2019).

ÁLVAREZ RIGAUDIAS,C: "Tratamiento de datos con fines de investigación científica y/o médica". En RALLO LOMBARTE, A (dir): Tratado de Protección de Datos. Ed. Tirant lo Blanch, Valencia, 2019.

- "El nuevo reglamento de desarrollo de la LOPD". *Actualidad Jurídica Uría y Menéndez*. Nº 21. 2008

ANISH ATHALYE,A; ENGSTROM,L; ILYAS,A; KWOK,K: "Synthesizing Robust Adversarial Examples". *Computer Vision and Pattern Recognition*, 2018.

ANRIG,B; BROWNE,W; GASSON,M: "The Role of Algorithms in Profiling". En: HILDEBRANDT, M; GUTWIRTH,S (eds.): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed. Springer, 2008.

ARELLANO TOLEDO,W: "El derecho a la transparencia algorítmica en Big Data e inteligencia artificial". *Revista General de Derecho Administrativo*. 50, 2019.

ARENAS RAMIRO,M: "Partidos políticos, opiniones políticas e internet: la lesión del derecho a la protección de datos personales". *Teoría y Realidad Constitucional*, núm. 44, 2019.

- "La sentencia del Tribunal Supremo de 19 de septiembre de 2008 protección de datos personales i apostasía". *Anuario de Derecho Eclesiástico del Estado*, vol. XXVI, 2010.

ARLOT,S Y CELISSE,A: "A survey of cross-validation procedures for model selection". *Statistics Surveys*, Vol. 4, 2010.

ATHMAJA,S; , HANUMANTHAPPA,M ; KAVITHA,V: "A survey of machine learning algorithms for big data analytics," *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, 2017.

AUSÍN, T; ANDREU MARTÍNEZ, B; VALERO TORRIJOS, J; CAYÓN DE LAS CUEVAS, J: "Diez consideraciones ético-jurídicas en relación con la reutilización y big data en el ámbito sanitario". *Bioderecho.es*, (12), 2021.

AZUAJE PIRELLA,M; FINOL GONZALEZ,D: "Transparencia algorítmica y la propiedad intelectual e industrial: tensiones y soluciones". *Revista la propiedad inmaterial*, n.º 30 - julio -diciembre de 2020.

BALCKMAN,C; FORGE,S: "Data Flows- Future Scenarios". Directorate general for internal policies policy department a: economic and scientific policy. Parlamento Europeo, 201.

BAMBAUER,J ; ZARSKY,T: "The Algorithm Game", *Notre Dame Law Review***,** volume 94, Issue 1.

BARRIOS,ANDRÉS,M: "La sentencia del TJUE YouTube y la responsabilidad de los operadores de plataformas digitales". *Diario La Ley*, Nº 53, Sección Ciberderecho, 19 de Julio de 2021.

BASTIDAS CID, Y,V: "El cumplimiento de los principios del tratamiento de datos personales establecidos en el reglamento general de protección de datos de la unión europea en proyectos de Big Data". *Informática y Derecho: Revista Iberoamericana de Derecho Informático* (segunda época), Nº. 6, 2019.

BERG, T ; BURG,V; GOMBOVIĆ, A; MANJU,P: "On the Rise of FinTechs – Credit Scoring Using Digital Footprints". *Michael J. Brennan Irish Finance Working Paper Series*, Research Paper No. 18-12, July 15, 2019.

BERMAN, E:"A government of laws and not of machines". *B.U. L. RE,* 98, 2018.

BIBAL, A; LOGNOUL, M ; DE STREEL, A:. "Legal requirements on explainability in machine learning". *Artif Intell Law*, 2020.

BOIX PALOP,A ."Los algoritmos son reglamentos: La necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por la Administración para la adopción de decisiones". *Revista de Derecho Público: Teoría y Método*, Vol. 1,Madrid, 2020.

BOOT A; HOFFMANN,P ; LAEVEN,L ; RATNOVSKI,L: "Financial Intermediation and Technology: What's Old, What's New?", *Fondo Monetario Internacional*, WP/20/161, 2020.

BORGES BLÁZQUEZ,R: "El sesgo de la máquina en la toma de decisiones en el proceso penal", *IUS ET SCIENTIA*, Vol. 6,Nº 2, 2020.

BRAMER,M: *Principles of Data Mining*. Ed.Springer, Londres, 2016.

BURK,L,D: "Algorithmic Fair Use". *University of Chicago Law Review*, 283, 2019.

CAPDEFERRO, O: "La inteligencia artificial del sector público: desarrollo y regulación de la actuación administrativa inteligente en la cuarta revolución industrial", *IDP. Revista de Internet, Derecho y Política*. N. º 30.

CASTAÑER CODINA, J: "La evaluación de la solvencia de las personas mediante el uso de algoritmos". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed. Aranzadi, Navarra, 2020, págs. 264 y 265.

CASTELLANOS CLARAMUNT, J: "La gestión de la información en el paradigma algorítmico: inteligencia artificial y protección de datos". *Métodos de Información*, 11(21), 2020.

CASTILLO PARRILLA, J,A: "Los datos personales como contraprestación en la reforma del TRLGDCU y las tensiones normativas entre la economía de los datos y la interpretación garantista del RGPD". *La Ley mercantil*, Nº 82, Sección Consumo/ Doctrina, Julio 2021.

CAVOUKIAN,A: "Privacy by Design The 7 Foundational Principles Implementation and Mapping of Fair Information Practices".

CERRILLO MARTÍNEZ,A: "¿Son fiables las decisiones de las Administraciones públicas adoptadas por algoritmos?". *European review of digital administration & law*, Vol. 1, Nº. 1-2, 2020.

- CERRILLO I MARTÍNEZ, A; VELASCO RICO, C: "Jurisdicción, algoritmos e inteligencia artificial". EN: LÓPEZ RAMÓN,F; VALERO TORRIJOS,J (coord.): *20 años de la Ley de lo Contencioso-administrativo: Actas del XIV Congreso de la Asociación Española de Profesores de Derecho Administrativo : Murcia, 8-9 de febrero de 2019*. *INAP*, 2019.

CHAPELLE,O , SCHÖLKOPF,B Y ZIEN,A: *Semi-Supervised Learning*, Ed. The MIT Press, Londres, 2006

CHAPMAN, P; CLINTON,J; KERBER,R; KHABAZA,T; REINARTZ,T; SHEARER,C ; WIRTH,R: *Step-by-step data mining guide*, 2000.

CHIAO, V: "Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice", International Journal of Law in Context, 2019, 15.

CHOULDECHOVA, A; BENAVIDES-PRADO, D; FIALKO, O Y VAITHIANATHAN, R: "A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions". *Proceedings of machine learning research*, 2018.

CLAUS, O,W: *Fundamentals of Data Visualization*. Ed. O'Reilly Media, Inc, 2019.

CONTISSA,G; DOCTOR,K ; LAGIOIA,F; LIPPI,M; MICKLITZ,H; PALKA,P; SARTOR,G; TORRONI,P: "Claudette meets GDPR. Automating the Evaluation of Privacy Policies using Artificial Intelligence". *Study Report*, 2018.

CORRALES COMPAGNUCCI, M: *Big Data, Databases and "Ownership" Rights in the Cloud*. Ed. Springer, Singapore, 2020.

COTINO HUESO,L: "El alcance e interactuación del régimen jurídico de los datos personales y *big data* relacionado con salud y la investigación biomédica". *Revista de derecho y genoma humano: genética, biotecnología y medicina avanzada,* Nº 52, 2020.

- "Exigencias constitucionales para las comunicaciones de datos entre Administraciones y la problemática de la procedencia de los datos del censo para la celebración de la consulta catalana. Posibles soluciones legales", *Boletín APEP Informa 2014,* nº 10, abril de 2014.
- "Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales". *Dilemata*, nº 24, 2017.
- "Derecho y garantías ante el uso público y privado de inteligencia artificial, robótica y big data". En: BAUZÁ REILLY,M (eds.): *El Derecho de las TIC en Iberoamérica.* Ed. La Ley Uruguay, Montevideo, 2019.
- "Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho". *Revista Catalana de Dret Públic*, núm. 58, 2019
- "Inteligencia artificial, *big data* y aplicaciones contra la COVID-19: privacidad y protección de datos". *IDP. Internet, Derecho y Política*, núm. 31.2020.
- Hacia la transparencia 4.0: el uso de la inteligencia artificial y big data para la lucha contra el fraude y la corrupción y las (muchas) exigencias constitucionales. En RAMIÓ,C (coord..): "Repensando la Administración Pública. Administración digital e innovación pública". Ed. INAP, Madrid, 2021.
- "Un análisis crítico constructivo de la Propuesta de Reglamento de la Unión Europea por el que se establecen normas armonizadas sobre la Inteligencia Artificial". *Diario La Ley*, 2 de Julio de 2021.
- "«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, Nº. 4 (Abril-junio 2020).

COVINGTON,P; ADAMS,J Y SARGIN,E: "Deep Neural Networks for YouTube Recommendations". *Google*, 2016.

DE MIGUEL BERIAIN, I.; PÉREZ ESTRADA, M.,J: "La inteligencia artificial en el proceso penal español: un análisis de su admisibilidad sobre la base de los derechos fundamentales implicados". *Revista De Derecho De La UNED (RDUNED)*, (25), 2020.

DE MONTALVO JÄÄSKELÄINEN, F: "Una reflexión desde la teoría de los derechos fundamentales sobre el uso secundario de los datos de salud en el marco del Big Data". *Revista De Derecho Político*, 1(106), 2019.

DE LA OLIVA SANTOS, A: "Justicia predictiva, interpretación matemática de las normas, sentencias robóticas y la vieja historia del justizklavier". *El cronista del estado social y democrático de derecho*, Nº 80, 2019.

DE LA PRADA, ESPINA, D: "Análisis y gestión de riesgos de los tratamientos de datos personales". En: MURGA FERNÁNDEZ, J,P; FERNÁNDEZ SCAGLIUSI, M,A ; ESPEJO LERDO DE TEJADA,M: (dirs): *Protección de datos, responsabilidad activa y técnicas de garantía*. *Curso de delegado de protección de datos.* Ed. Reus, Madrid, 2018.

DOMÉNECH PASCUAL, G: *Derechos fundamentales y riesgos tecnológicos*. Ed. Centro de Estudios Políticos y Constitucionales. 2006.

DONOVAN, J, MATTHEWS, J, CAPLAN, and HANSON, L: "Algorithmic Accountability: A Primer". *DATA & SOCIETY*, 2018

DUARTE,N Y LLANSÓ,E: "Mixed Messages? The Limits of Automated Social Media Content Analysis", *Center for Democracy and technology*, 2017.

DURÁN CARDO, B: *La figura del responsable en el derecho a la protección de datos*. Ed. Wolters Kluwer, Madrid, 2016.

ELKAN, C: "Evaluating classifiers". *University of San Diego, California*, 2012.

EL EMAM,K; ÁLVAREZ RIGAUDIAS,C: "A critical appraisal of the Article 29 Working Party Opinion 05/2014 on data anonymization techniques", *International Data Privacy Law*, Volume 5, Issue 1, February 2015.

ESKENS,S: "A right to reset your user profile and more: GDPR-rights for personalized news consumers". *International Data Privacy Law*, Volume 9, Issue 3, August 2019.

FAWCET,T: "An introduction to ROC analysis". *Pattern Recognition Letters*, 27, 2006.

FERNÁNDEZ-SAMANIEGO, J ; FERNÁNDEZ-LONGORIA,P: "El interés legítimo como principio para legitimar el tratamiento de datos". En RALLO LOMBARTE,A; GARCÍA MAHAMUT,R (eds.): *Hacia un nuevo derecho europeo de protección de datos*. Ed. Tirant lo Blanch, Valencia, 2015.

- En: FERNÁNDEZ-SAMANIEGO Y BLAS PIÑAR GUZMÁN,J: "Las acciones colectivas en el marco del RGPD: una perspectiva desde el Derecho Civil español". *Diario La Ley*, Nº 26, Sección Ciberderecho, 11 de Febrero de 2019.

FENOLL NIEVA,J: *Inteligencia artificial y proceso judicial*. Ed. Marcial Pons, Madrid, 2018.

FERNÁNDEZ SALMERÓN, M: *La protección de los datos personales en las Administraciones Públicas*. Ed. Thomson, Civitas. 2003.

FINLAY,S: "Predictive Analytics, Data Mining and Big Data. Myths, Methods and Misconceptions". *Palgrave Macmillan*. 2014.

FLORIDI,L: "Group Privacy: a Defence and an Interpretation". En: TAYLOR, L; FLORIDI, L ; VAN DER SLOOT, B (eds) Authors' final draft: *Group Privacy: new challenges of data technologies*. Ed.Springer. Dordrecht, 2017

FOSCH VILLARONGA,E; KIESEBERG,P; B; LI,T: "Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten". *Computer Law & Security Review*, Volume 34, Issue 2, April 2018.

GALETTA, D,U; GUSTAVO CORVALÁN,J: "Intelligenza Artificiale per una Pubblica Amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto". *Federalismi.it*, n 3, 2019

GARCÍA ALSINA, M: *Big data: gestión y explotación de grandes volúmenes de datos.* Barcelona: Ed. UOC, Barcelona, 2017.

GARCÍA PÉREZ, RM: "Bases jurídicas relevantes del tratamiento de datos personales en la contratación de contenidos y servicios digitales". *Cuadernos de Derecho Transnacional*, marzo 2020, Vol. 12, Nº 1.

GARCÍA-RIPOLL MONTIJANO,M: "El consentimiento al tratamiento de datos personales". En: GONZÁLEZ PACANOWSKA,I (Coord.): *Protección de datos personales*. Ed. Tirant lo Blanch, Valencia, 2020.

GARRIGA DOMÍNGUEZ, A: "La elaboración de perfiles y su impacto en los derechos fundamentales. Una primera aproximación a su regulación en el reglamento general de protección de datos de la Unión Europea". Derechos y libertades, Número 38, Época II, enero 2018.

- *Nuevos retos para la protección de datos personales: en la Era del Big Data y de la computación ubicua*. Ed. Dykinson, Madrid, 2016.

GIL GONZÁLEZ,E : *Big data, privacidad y protección de datos*, Ed. BOE, Madrid, 2016.

- "Understanding the legal provisions that allow processing and profiling of personal data—an analysis of GDPR provisions and principles". *ERA Forum* 19**,** 597–621, 2019.
- "Aproximación al estudio de las decisiones automatizadas en el seno del Reglamento General Europeo de Protección de Datos a la luz de las tecnologías big data y de aprendizaje computacional". *Revista Española de la Transparencia*, Nº 5. Segundo Semestre 2017.

GILLINGHAM, P: "Can Predictive Algorithms Assist Decision-Making in Social Work with Children and Families?" *Child Abuse Review*, 28, 2019.

GILLIS, TALIA B; SPIESS, JANN L: "Big Data and Discrimination," *University of Chicago Law Review*: Vol.86, iss.2.Article 4, 2019.

GÓMEZ DE ÁGREDA,A: *Mundo Orwell. Manual de supervivencia para un mundo hiperconectado.* Ed. Ariel, España, 2019.

GOODFELLOW, I, POUGET-ABADIE,J , MIRZA,M, XU,B , WARDE-FARLEY,D OZAIRZ,S COURVILLE,A , BENGI,Y., "Generative Adversarial Nets", 10 de junio de 2014.

GOODMAN,B; FLAXMAN,S: "EU Regulations on Algorithmic Decision-Making and a right to Explanation".

GORUNESCU,F: *Data Mining: Concepts, Models and Techniques*. Ed. Springer, 2011.

GUNNING, D., & AHA,D: "DARPA's Explainable Artificial Intelligence (XAI) Program". *AI Magazine,* Vol. 40 No. 2: Summer 2019.

HERNÁNDEZ, J,C: "Decisiones algorítmicas de perfilado. Régimen y garantías jurídicas." *Revista Española de Derecho Administrativo*. Nº 203, 2020.

HERSCHEL,R Y MIORI,V,M: "Ethics & Big Data". *Technology in Society*, Volume 49,2017.

HERTZA,V,A: "Fighting Unfair Classifications in Credit Reporting: Should the United States Adopt GDPR-Inspired Rights in Regulating Consumer Credit?". *New york university law review*, volume 93, núm. 6, 2018.

HERRERO SUÁREZ,C: "Big Data y derecho de la competencia". En PIÑAR MAÑAS,J,L Y DE LA QUADRA SALCEDEO,T (dir.): *Sociedad digital y derecho.* Ed. BOE, Madrid, 2018.

HILDEBRANDT, M: "What is profiling? Defining Profiling: A New Type of Knowledge?" En: HILDEBRANDT, M & GUTWIRTH,S (eds): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008.

HOGAN-DORAN,D: "Computer says "no": automation, algorithms and artificial intelligence in Government decision-making". *The Judicial Review*, 2017.

HUEI,KO,Y ; YU HSU,P; , MING SHIEN CHENG,M ; RUEI JHENG; CHAO LUO,Z: "Costumer Retention Prediction with CCN". En: TAN, Y SHI,Y: *Data Mining and Big Data*. 4th International Conference, DMBD, 2019.

HUERGO LORA, A,J: "Una aproximación a los algoritmos desde el derecho administrativo". En: HUERGO LORA, A,J (dir): *Una aproximación a los algoritmos desde el derecho administrativo*. Ed. Aranzadi, Navarra, 2020.

HURLEY, Y ADEBAYO, J: "Credit scoring in the era of big data*", The yale journal of law & technology*, 2017, Vol 18.

JOBIN,A ; IENCA,M Y VAYENA,E., "The global landscape of AI ethics guidelines", *Nature Machine Intelligence*, Vol 1, Septiembre 2019.

JOVE VILLARES, D: "Peter Nowak and Subjective Annotations in Clinical Records". *European Data Protection Law Review*. Volume 5, nº 2, 2019.

KAMINSKI,M; MALGIERI,G: "Algorithmic impact assessments under the GDPR: producing multi-layered explanations", *International Data Privacy Law*, 2020, ipaa020.

KAMMOURIEH, L; BAAR,T; BERENS, J; LETOUZÉ, E; MANSKE,J; PALMER,J; SANGOKOYA, D; VINCK,P: "Group privacy in the age of big data". En: TAYLOR, L; FLORIDI, L ; VAN DER SLOOT, B (eds) Authors' final draft: *Group Privacy: new challenges of data technologies*. Ed.Springer. Dordrecht, 2017.

KEAT,CITRON,D: "Technological Due Process". *Washington University Law Review*, volume 85, issue 6, 2008.

KEHL, D, GUO,P Y KESSLER,S: "Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing". *Internet & Society, Harvard Law School*, 2017.

KOSINSKI,M ;STILLWELL,D; THORE,G: "Private traits and attributes are predictable from digital records of human behavior", *PNAS*, 9/4/2013.

KOTU, V Y DESHPANDE,B:.*Predictive Analytics and Data Mining : Concepts and Practice with Rapid Miner,* Ed. Elsevier Science & Technology, 2015.

KROLL, A,J;. HUEY, J; BAROCAS, S; FELTEN, E, W; REIDENBERG, J,R;.ROBINSON, D,G;. Y YU, H:, "Accountable Algorithms". *University of Pennsylvania Law Review,* Vol. 165, 2017.

KUHN,M, JOHNSON,K: *Applied Predictive Modeling*. Ed. Springer Science + Business Media, New York, 2013.

LATZER, M; HOLLNBUCHNER, K; JUST, N. Y SAURWEIN, F: "The economics of algorithmic selection on the Internet". Working Paper – Media Change & Innovation Division. University of Zurich, Zurich, 2014.

LAZCOZ MORATINOS:G: "Análisis jurídico de la toma de decisiones algorítmicas en la asistencia sanitaria". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed.Aranzadi, Navarra, 2020.

- LAZCOZ MORATINOS, G., & CASTILLO PARRILLA, J: "Valoración algorítmica ante los derechos humanos y el Reglamento General de Protección de Datos: el caso SyRI". *Revista Chilena de Derecho y Tecnología,* 2020

LEARNED-MILLER,E; ORDÓÑEZ,V; MORGENSTERN,J Y BUOLAMWINI,J**: "**Facial recognition technologies in the wild: a call for a federal office". Algorithmic Justice League, 2020.

LE DÉAUT, Y,J:. "*Technological convergence, artificial intelligence and human rights"*. Council of Europe, Parliamentary Assembly. Committee on Culture, Science, Education and Media Session 2017 - Second part-session. 2017

LEMLEY, A,M Y CASEY,B: "Remedies for Robots." *The University of Chicago Law Review*, vol. 86, no. 5, 2019.

LESTER, T; PACHAMANOVA,D: "The Dilemma of False Positives: Making Content Id Algorithms More Conducive to Fostering Innovative Fair Use in Music Creation". *UCLA Entertainment Law Review,* Vol. 24, No. 51, 2017.

LI XAN WONG, K Y SHIELDS DOBSON,A: "We're Just Data: Exploring China's Social Credit System in Relation to Digital Platform Ratings Cultures in Westernised Democracies." *Global Media and China* 4, no. 2, June 2019.

L,JANSSEN,H:, "An approach for a fundamental rights impact assessment to automated decision-making", *International Data Privacy Law*, Volume 10, Issue 1, February 2020.

LLANEZA,P: *Datanomics: Todos los datos personales que das sin darte cuenta y todo lo que las empresas hacen con ellos*. Ed. Deusto, 2019

- *Seguridad y responsabilidad en la Internet de las cosas.* Ed. Bosch. Wolters Kluwer, España, 2018

LIANE,C: "*A Taxonomy and Classification of Data Mining", SMU Science and Technology Law Review*, 2013.

LORENZO CABRERA,S: "Posición jurídica de los intervinientes en el tratamiento de datos personales. Medidas de cumplimiento". En: MURGA FERNÁNDEZ, J,P; FERNÁNDEZ SCAGLIUSI, M,A ; ESPEJO LERDO DE TEJADA,M: (dirs.): *Protección de datos, responsabilidad activa y técnicas de garantía*. *Curso de delegado de protección de datos.* Ed. Reus, Madrid.

LOZA, CORERA,M,L: *De los microdatos a los datos masivos. Cuestiones legales.* Tesis doctoral. Universitat de València. 2017.

MACDONALD,S; GIRO CORREIA,S; WATKIN, A-L: "Regulating terrorist content on social media: automation and the rule of law". *International Journal of Law*, 2019

MALGIERI,G: "Automated decision-making in the EU Member States: The right to explanation and other "suitable safeguards" in the national legislations". *Computer Law & Security Review*, Volume 35, Issue 5, October 2019.

MANTELERO,A: "Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection", *Computer Law & Security Review*, Volume 32, Issue 2, 2016.

- "AI and big data: a blueprint for a human rights, social and ethical impact assessment", Computer Law & Security Review, Volume 34, Issue 4, August 2018.
- "Toward a New Approach to Data Protection in the Big Data Era". En: GASSER,U; ZITTRAIN J; FARIS,R; HEACOCK JONES,R (dir.): *Internet Monitor 2014: Reflections on the Digital World* . Cambridge (MA): Berkman Center for Internet and Society at Harvard University, 2014
- "From group privacy to collective privacy: towards a new dimension of privacy and data protection in the big data era", En: TAYLOR, L; FLORIDI, L ; VAN DER SLOOT, B (eds) Authors' final draft: *Group Privacy: new challenges of data technologies.* Ed.Springer. Dordrecht, 2017

MARKOU, CH; DEAKIN,S (eds): *Is Law Computable? Critical Perspectives on Law + Artificial Intelligence* . Ed. Hart Publishing, 2020.

MARTÍNEZ CRUZ, J: "El interés legítimo en el marco de la protección de datos personales". En: MURGA FERNÁNDEZ, J,P ; FERNÁNDEZ SCAGLIUSI, M,A; ESPEJO LERDO DE TEJADA,M (dirs.): *Cuestiones actuales sobre protección de datos en España y México.* Ed. Tirant lo Blanch, Valencia, 2021.

MARTÍNEZ MARTÍNEZ,R: "Cuestiones de ética jurídica al abordar proyectos de Big Data. El contexto del Reglamento general de protección de datos". *Dilemata,* año 9, nº 24, 2017.

- "Big data, investigación en salud y protección de datos personales ¿Un falso debate?". *Revista valenciana d'estudis autonòmics,* Nº 62, 2017.
- "De la transparencia estática a la dinámica en las aplicaciones móviles: lecciones aprendidas del caso de «La Liga de Futbol Profesional". *Diario La Ley,* Nº 9463, 2019.
- "El laberinto de la contratación pública en protección de datos". *Diario La Ley*, sección Ciberderecho, núm. 35, 2019.
- "Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo". *Revista Catalana de Dret Públic,* (58). 64-81, 2019.

MARGOT E KAMINSKI,M; MALGIERI,G: "Algorithmic impact assessments under the GDPR: producing multi-layered explanations". *International Data Privacy Law*, ipaa 020, 06 December 2020.

MARZAL RAGA, R: *El apercibimiento. Una nueva sanción en materia de protección de datos de carácter personal.* Ed. Tirant lo Blanch, Valencia, 2015.

MAS BADIA, M,D: *Sistemas privados de información crediticia. Nueva regulación entre la protección de datos y el crédito responsable*. Ed. Tirant lo Blanch, Valencia, 2021.

MAYSON, S,G: "Bias In, Bias Out". *Yale Law Journal,* vol 108, 2019.

MCCARTHY,J; MINSKY,M; ROCHESTER ,N ; SHANNON,C: "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence", 1955. Disponible actualmente en: AI Magazine Volumen 27 Number 4, 2006.

MCCULLOCH,W Y PITTS,W., "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics,* Vol.5, 1943.

MCGREGOR, L; MURRAY, D Y NG, V: "International human rights law as a framework for algorithmic accountability". *International and Comparative Law Quarterly,* 68(2), 2019.

MEDVEDEVA, M; VOLS, M. Y WIELING, M: "Using machine learning to predict decisions of the European Court of Human Rights". *Artificial Intelligence and Law, volume 28,* 2020.

MEDINA GUERRERO,M: "Categorías especiales de datos". En RALLO LOMBARTE, A (dir): Tratado de Protección de Datos. Ed. Tirant lo Blanch, Valencia, 2019.

MEESE, J. & JAGASIA, P. Y ARVANITAKIS, J: "Citizen or consumer? Contrasting Australia and Europe's data protection policies". *Internet Policy Review*, 2019, 8(2).

MEIKE K, KÖRFFER B & MEINTS,M: "Profiling of Customers and Consumers - Customer Loyalty Programmes and Scoring Practices". En: HILDEBRANDT, M & GUTWIRTH, S (eds.): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008.

MÉNDEZ GARCÍA, M y ALFONSO FARNÓS, I: "La legitimación para el tratamiento de categorías especiales de datos con finalidades de investigación en el marco del Reglamento (UE) 2016/679 y de la Ley Orgánica 3/2018", en *Revista de derecho y genoma humano: genética, biotecnología y medicina avanzada* Nº Extra 1, 2019.

MERCADER UGUINA, J: "Algoritmos y derecho del trabajo". *Actualidad Jurídica Uría Menéndez,* 52, 2019.

MESZAROS,J ; *CHIH-HSING,H:* "AI research and data protection: Can the same rules apply for commercial and academic research under the GDPR?". *Computer Law & Security Review,* Volume 41, July 2021.

MIRÓ-LLINARES,F: "Predictive Policing: Utopia or Dystopia? On attitudes towards the use of big data algorithms for law enforcement", *IDP. Revista de Internet, Derecho y Política.* No. 30, 2020.

MITROU,J: "Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?", 2018.

M O,DONELL,R:" Challenging racist predictive policing algorithms under the equal protection clause". ***New*** *York University Law Review*. Vol. 94:544, 2019.

MOLINA FÉLIX, L C: *"Data mining:* torturando a los datos hasta que confiesen", *UOC*, 2002.

MOLNAR,CH*: Interpretable Machine Learning. A Guide for Making Black Box Models Explainable*. Libro disponible on line. ISBN 9780244768522.

MONDAL, B: "Artificial Intelligence: State of the Art". En: BALAS V; KUMAR R; SRIVASTAVA, R; (eds.): *Recent Trends and Advances in Artificial Intelligence and Internet of Things*. Ed. Springer, Cham, vol 172, 2020.

MORENTE, PARRA,V: "Big data o el arte de analizar datos masivos. Una reflexión crítica desde los derechos fundamentales". *Derechos y libertades*, número 41, Época II, junio 2019.

MURGA FERNÁNDEZ,J,P: "Derechos de los individuos". En: MURGA FERNÁNDEZ, J,P; FERNÁNDEZ SCAGLIUSI, M,A ; ESPEJO LERDO DE TEJADA,M: (dirs.): *Protección de datos, responsabilidad activa y técnicas de garantía.* Ed. Reus. Madrid, 2018.

- "La protección de datos y los motores de búsqueda en Internet: cuestiones actuales y perspectivas de futuro acerca del derecho al olvido". *Revista de Derecho Civil*, Vol. 4, Nº. 4 (octubre–diciembre, 2017), 2017.

MURILLO DE LA CUEVA,P,L: "La construcción del derecho a la autodeterminación informativa". *Revista de Estudios Políticos* (Nueva Época) Núm. 104. Abril-Junio1999.

- MURILLO DE LA CUEVA, P, L Y PIÑAR MAÑAS, J,L: *El derecho a la autodeterminación informativa.* Ed, Fundación Coloquio Jurídico Europeo, 2009, Madrid.

MUÑOZ PAREDES, M,L: "Big data y contratos de seguro: Los datos generados por los asegurados y su utilización por los aseguradores". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed. Aranzadi, Navarra, 2020.

N COFONE,I: "Algorithmic Discrimination Is an Information Problem". *Hastings Law Journal,* Vol. 70:1389, 2019.

NÚÑEZ SEOANE,J: "El derecho de la información y acceso al funcionamiento de los algoritmos que tratan datos personales". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed. Aranzadi, Navarra, 2020.

OBERMEYER, Z; POWERS, B; VOGELI,CH Y MULLAINATHAN,S: "Dissecting racial bias in an algorithm used to manage the health of populations". *Science*, 2019.

ODINET, CHRISTOPHER K: "The New Data of Student Debt". *Southern California Law Review*, December 8, 2019.

OLSEN, H.P., SLOSSER, J., HILDEBRANDT, T., & WIESENER, C: "What's in the Box? The Legal Requirement of Explainability in Computationally Aided Decision-Making in Public Administration". *Political Economy: Structure & Scope of Government eJournal*. 2019.

OLSON,D Y DESHENG DASH,W: *Predictive Data Mining Models*. Ed. Springer, Singapore, 2017.

ORTEGA JIMÉNEZ, A: *Las aplicaciones del Big Data en el ámbito asegurador y el tratamiento legal de sus datos. Una perspectiva desde el derecho internacional privado*. Ed. Fundación Mapfre, Madrid, 2019.

ORTIZ UROZ,R: "El tratamiento de datos personales en los proyectos universitarios de investigación a la vista de la actual normativa de protección de datos personales". *LA LEY privacidad*, Nº 7, Sección El foro de la privacidad, Primer trimestre de 2021.

OSWALD,M; GRACE,J URWIN,S; BARNES,G: "Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality". *Information & Communications Technology Law*, Volume 27, 2018.

PANESAR, A: *Machine Learning and AI for Healthcare. Big Data for Improved Health Outcomes*. Ed. Apress, Berkeley, 2021.

PIÑAR MAÑAS,J,L: "Contratación pública y protección de datos". En: ORTEGA BURGOS,E; PASTOR RUIZ,F (dir): *Derecho administrativo 2021*. Ed. Tirant lo Blanch. Valencia, 2021.

- PIÑAR MAÑAS,J,L; RECIO GAYO, M: *El derecho a la protección datos en la jurisprudencia del Tribunal de Justicia de la Unión Europea*. Ed. La Ley, Madrid, 2018.

PLAZA PENADÉS,J; PEDREÑO,A; MORENO,L: *Big Data e Inteligencia Artificial. Una visión económica y legal de estas herramientas disruptivas*. Ed. Fundació Parc Científic Universitat de València. Valencia, 2018.

POLITOU,E; ALEPIS,E: PATSAKIS,C: "Profiling tax and financial behavior with big data under the GDPR". *Computer Law & Security Review*, Volume 35, Issue 3, 2019.

POLO ROCA,A: "Datos, Datos, Datos: El Dato Personal, El Dato No Personal, El Dato Personal Compuesto, La anonimización, La Pertenencia Del Dato Y Otras Cuestiones Sobre Datos". *Estudios De Deusto* 69 (1), 165-94, 2021.

PONCE SOLÉ,J: "Inteligencia artificial, derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico" .*Revista General de Derecho Administrativo*. 50, 2019.

PUENTE ESCOBAR,A: "Principios y licitud del tratamiento". En RALLO LOMBARTE, A (dir): Tratado de Protección de Datos. Ed. Tirant lo Blanch, Valencia, 2019.

- PUENTE ESCOBAR, A: "Legitimación para el tratamiento". En: MARTÍNEZ MARTÍNEZ,R (coord.): *Protección de datos. Comentarios al Reglamento de Desarrollo de la LOPD*. Ed. Tirant lo Blanch. Valencia, 2008.

PUYOL MONTERO, F, J: "Los principios del derecho a la protección de datos". En PIÑAR MAÑAS,J,L (dir): *Reglamento Europeo de Protección de Datos. Hacia un nuevo modelo europeo de privacidad.* Ed. Reus, Madrid, 2016.

QUELLE, C: "Enhancing Compliance under the General Data Protection Regulation: The Risky Upshot of the Accountability and Risk-based Approach". *European Journal of Risk Regulation, 9*(3), 2018.

QUIJANO-SÁNCHEZ,L; LIBERATORE,F; CAMACHO-COLLADOS,J, CAMACHO-COLLADOS,M: "Applying automatic text-based detection of deceptive language to police reports: Extracting behavioral patterns from a multi-step classification model to understand how we lie to the police". *Knowledge-Based Systems*, Volume 149, 2018.

RALLO LOMBARTE, A: "Del derecho a la protección de datos a la garantía de nuevos derechos digitales". En: GARCÍA MAHAMUT, R; TOMÁS MALLÉN, B (eds.): *El Reglamento General de Protección de Datos. Un enfoque nacional y comparado. Especial referencia a la LO 3/2018 de Protección de Datos y garantía de los derechos digitales*. Ed. Tirant lo Blanch, Valencia, 2019.

RAMOS, PASCUAL,D: "Reflexiones sobre el artículo 80 del Reglamento Europeo de Protección de Datos". *LA LEY privacidad*, Nº 7, Sección El foro de la privacidad, Primer trimestre de 2021.

REBOLLO DELGADO, L: "Big data, inteligencia artificial y derechos fundamentales: problemas jurídicos". En: *El derecho a la protección de datos personales en Europa y en América: diferentes visiones para una misma realidad*. Ed. Tirant lo Blanch. Valencia, 2021.

RECIO GAYO, M: "Los derechos a presentar reclamaciones ante la autoridad de control y a la tutela judicial efectiva". En: PIÑAR,MAÑAS, J,L (dir):*Reglamento General de Protección de Datos. Hacia un nuevo modelo de privacidad*. Ed. Reus. Madrid, 2017.

RECUERO LINARES, M: *La investigación científica con datos personales genéticos y datos relativos a la salud: perspectiva europea ante el desafío globalizado*, Premio AEPD, 2019.

RICHARDSON, R; SCHULTZ, J Y CRAWFORD, K.: "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice". *New York University Law Review*, 2019.

ROMEO CASABONA,C,M: "Riesgo, procedimientos actuariales basados en inteligencia artificial y medidas de seguridad". *Revista penal*, nº42, 2018.

RODRÍGUEZ AYUSO, J F: "Tratamiento de datos relativos a la salud del interesado en el ámbito de la sanidad pública". *Actualidad Administrativa*, Nº 10, Sección Administración del siglo XXI, Octubre 2019.

- *Garantía administrativa de los derechos del interesado en materia de protección de datos personales*. Ed. Bosch, Barcelona, 2021, pág.128.

ROIG I BATALLA,A: *Las garantías frente a las decisiones automatizadas del Reglamento general de Protección de Datos a la gobernanza algorítmica*. Ed. Bosch. Barcelona, 2020.

- "Safeguards for the right not to be subject to a decision based solely on automated processing (Article 22 GDPR)". *European Journal of Law and Technology*, Vol 8, No 3, 2017.

RUSSELL, S Y NORVING, P: *Inteligencia Artificial. Un enfoque moderno*, Ed. Pearson Educación, 2ªed., Madrid, 2004.

SAMUEL, A,L: "Some Studies in Machine Learning Using the Game of Checkers". *IBM Journal of Research and Development*, Volume: 3, Issue: 3, July 1959.

SANCHO LÓPEZ,M: *Derecho al olvido y big data: Dos realidades convergentes* . Ed. Tirant lo Blanch, Valencia, 2020.

SELBST,A & BAROCAS,S: "The Intuitive Appeal of Explainable Machines", 87 Fordham L. Rev. 1085, 2018.

SERGIO CABRAL,T: "Forgetful Al: Al and the Right to Erasure under the GDPR". *European Data Protection Law Review*, 6, no. 3 (2020).

SERRANO PÉREZ,M,M: "El marco jurídico de los datos relativos a la salud en el ámbito de la salud y de la investigación en salud tras la entrada en vigor del Reglamento General de Protección de Datos y de la Ley de Protección de Datos Personales y garantía de los derechos". *Estudios de Deusto: revista de la Universidad de Deusto*, ISSN 0423-4847, Vol. 68, Nº. 2, 2020.

SHINTRE,S ; ROUNDY, KA; DHALIWAL, J: "Making Machine Learning Forget". En: NALDI,M; ITALIANO, G; RANNENBERG,K; MEDINA,M; BOURKA,A (eds.): *Privacy Technologies and Policy*. APF 2019. Lecture Notes in Computer Science, vol 11498. Ed. Springer, 2019.

SIMÓN CASTELLANO,P ; MAGRO SERVET,V: *Justicia cautelar e inteligencia artificial: la alternativa a los atávicos heurísticos judiciales*. Ed. Bosch, Madrid, 2021.

SMART,A: *Más allá de ceros y unos: Robots, psicodelia y conciencia*. Ed. Clave intelectual. Madrid, 2018.

SOBRINO GARCÍA,I: "Desafíos y limitaciones en la contratación pública: el impacto de la protección de datos tras los últimos cambios legislativos". *Revista General de Derecho Administrativo*, número 56, enero 2021.

SOENENS,ELS: "Reply: Web Usage Mining for Web Personalization in Customer Relation Management". En: HILDEBRANDT, M & GUTWIRTH,S (eds.): *Profiling the European Citizen. Cross-Disciplinary Perspectives*. Ed.Springer, 2008.

SORIANO ARNANZ, A: *Data protection for the prevention of algorithmic discrimination.* Ed. Thomson Reuters Aranzadi. Pamplona. 2021.

- "Decisiones automatizadas y discriminación: aproximación y propuestas generales". *Revista General de Derecho Administrativo,* 56, 2021.
- "Decisiones automatizadas: problemas y soluciones jurídicas. más allá de la protección de datos". *Revista de Derecho Público: Teoría y Método.* Vol. 3 | 2021.

SPIELKAMP,M;KAYSER-BRIL,N: "Experimentation without a plan: automated decision-making in European public services journalist". *European Public Mosaic*, 2019.

TASA FUSTER,V: "Llengües minoritzades i tecnologies disruptives de comunicació", *Universitat de València*. 2020.

TENA ARREGUI, R: "¿Son justos los precios personalizados mediante algoritmos?". *El notario del siglo XXI: revista del Colegio Notarial de Madrid*, Nº. 87, 2019.

TENE, O; POLONETSKY,J: "Big Data for All: Privacy and User Control in the Age of Analytics". *Northwestern Journal of Technology and Intellectual Property*. Volume 11 | Issue 5, 2013.

TODOLÍ SIGNES,A: "La gobernanza colectiva de la protección de datos en las relaciones laborales: big data, creación de perfiles, decisiones empresariales automatizadas y los derechos colectivos". *Revista de derecho social*, Núm. 84, 2018.

- "Retos legales del uso del *big data* en la selección de sujetos a investigar por la Inspección de Trabajo y de la Seguridad Social". *Revista Galega de Administración Pública*, núm. 59, 2020.

TRONCOSO REIGADA,A: "Investigación, salud pública y asistencia sanitaria en el Reglamento General de Protección de Datos de la Unión Europea y en la Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales". *Revista de derecho y genoma humano: genética, biotecnología y medicina avanzada.*

- *La Protección de Datos Personales. En Busca del Equilibrio*. Ed. Tirant lo Blanch, Valencia, 2011.
- *Comentario al Reglamento General de Protección de Datos y a la Ley Orgánica de Protección de Datos y Garantía de los Derechos Digitales*. Ed. Civitas. Navarra, 2021.

TRUJILLO CABRERA,C: "Aproximación a la regulación del consentimiento en el Reglamento General de Protección de Datos". *Anales de la Facultad de Derecho*, 34; septiembre 2017.

TULIO RIBEIRO,M; SINGH,S ; GUESTRIN,C: "Why Should I Trust You? Explaining the Predictions of Any Classifier".

TURING,A,M: "Computing Machinery and Intelligence" , *Mind*, vol. 59, (1 October 1950).

URMAN,A ; MAKHORTYKH, M; ULLOA,R : "The Matter of Chance: Auditing Web Search Results Related to the 2020 U.S. Presidential Primary Elections Across Six Search Engines". *Social Science Computer Review.*

VALERO TORRIJOS, J: "Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración". *Revista Catalana de Dret Públic*, (58), 2019.

- "Ciudades inteligentes y datos abiertos implicaciones jurídicas para la protección de los datos de carácter personal". *Istituzioni del federalismorivista di studi giuridici e politici*. Nº. 4, 2015.
- VALERO TORRIJOS, J; CERDÁ MESEGUER, J,L: "Transparencia, acceso y reutilización de la información ante la transformación digital del sector público: enseñanzas y desafíos en tiempos del COVID-19". *EUNOMÍA. Revista En Cultura De La Legalidad*, (19), 2020.

VALLS PRIETO,J: "Combating terrorist financing with artificial intelligence systems", *Repositorio universidad de Granada*, 2020.

VEDASCHI, A: "Intelligenza artificiale e misure antiterrorismo alla prova del diritto costituzionale". En VV.AA : *Liber Amicorum per Pasquale Costanzo – Diritto costituzionale in trasformazione Vol. I – Costituzionalismo, Reti e Intelligenza artificiale. Ed.* Associazione giuridica scientifico-culturale. Génova, 2020.

- "Privacy and data protection versus national security in transnational flights: the EU–Canada PNR agreement", *International Data Privacy Law*, Volume 8, Issue 2, May 2018.

VEGA GARCÍA, P: "El algoritmo de YouTube, el artículo 17 de la Directiva 2019/790 y la protección de los derechos de autor". En: HUERGO LORA, A,J (dir): *La regulación de los algoritmos*. Ed. Aranzadi, Navarra, 2020.

VELASCO RICO, C: "Personalización, proactividad e inteligencia artificial. ¿Un nuevo paradigma para la prestación electrónica de servicios públicos?". *IDP. Revista d'Internet, Dret i Política*, Núm. 30, 2020.

- "La ciudad inteligente: entre la transparencia y el control". *Revista General de Derecho Administrativo* 50, 2019.

VILASAU SOLANA,M: "El consentimiento general y de menores". En RALLO LOMBARTE, A (dir): Tratado de Protección de Datos. Ed. Tirant lo Blanch, Valencia, 2019.

VILLALBA SÁNCHEZ,A: "El principio de transparencia en la ejecución automatizada del contrato de trabajo: una aproximación jurídica a la tecnología blockchain y a la inteligencia artificial". *Revista Española de Derecho del Trabajo*. Num.224/2019.

VIZCAINO CALDERÓN, M: *Comentarios a la Ley Orgánica de Protección de datos de carácter personal*. Ed. Civitas, Madrid, 2001.

WACHTER, S; MITTELSTADT, B: "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI". *Columbia Business Law Review*, 2019.

- "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation". *International Data Privacy Law*, Volume 7, Issue 2, 1 May 2017.
- "Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR". *Harvard Journal of Law & Technology,* Volume 31, Number 2, 2018.

WAYNE A, L; FERGUSON, GUTHRIE, A: "Policing Criminal Justice Data", *Minnesota Law Review* 541 2016.

WONG,B: "Online personalised pricing as prohibited automated decision-making under Article 22 GDPR: a sceptical view". Information & Communications Technology Law, 30:2, 193-207, 2021.

XINDONG,W ; XINGQUAN, ZHU; GONG-QING, WU, Y WEI, D: "Data mining with big data", *Transactions on knowledge and data engineering*, vol. 26, no. 1, january 2014.

YINGZHE,H ; GUOZH, M; KAI,CH;, XINGBO,H; JINWEN,H:" Towards Security Threats of Deep Learning Systems: A Survey". *IEEE Transactions on software engineering*, Octubre 2020.

YOUSRA ABDUL ALSAHIB, A; MAZLEENA, S; MOHAMMAD ABDUR, R: "A comprehensive review on privacy preserving data mining". *Springer Plus* 4**,** 694 (2015).

II. NOTICIAS DE PRENSA, BLOGS, WEBS

BOIX PALOP, A: "El control de las fake news y la calidad del debate público en las sociedades democráticas*"*. La página definitiva. 21/03/2019.

BULMAN,M: "Home Office failed to ensure innocent students were not wrongly detained in cheating scandal, report finds". *Independent*. 5/8/2020.

CANO,F ; ALBA,C: "BBVA dice que no ha vulnerado la ley de protección de datos y que recurrirá la multa de cinco millones". *El Español*. 15/12/2020.

CARRIE WONG,J: "Google reportedly targeted people with 'dark skin' to improve facial recognition". *The Guardian*. 03/10/2019.

CASTILLO PARRILLA, J,A: "La discriminación a través del algoritmo en una plataforma. El caso Deliveroo Bolonia y sus implicaciones para el sector público". *Observatorio de Transformación Digital del Sector Público*.

CHA,S: "Smile with your eyes': How to beat South Korea's AI hiring bots and land a job". *Thomson Reuter*. 13/01/2020.

CRIDDLE,C: "Instagram fixes mistake promoting harmful diet content", *BBC News*. 15/04/2021.

DARLINGTON,K: "Sistemas de IA explicables: comprender las decisiones de las máquinas". *OpenMind.BBVA*. 11/10/2017.

DAVALOS,J: "Google Blocks Search Suggestions to Stop Election Misinformation", *Bloomberg,* 10/10/2020.

DIAS OLIVE,T: "Content Moderation Technologies: Applying Human Rights Standards to Protect Freedom of Expression". *Human Rights Law Review*, 2020.

GALINDO, F, J,C: "Aprendizaje federado, la técnica de IA para proteger la privacidad". *Muy interesante.* 26/05/2020.

GALLO,V ; BINNS,R: "How using AI can require trade-offs between data protection principles, and what organizations can do to assess and balance them". 25/07/2019. *Information Commissioner's Office*.

GARCÍA,J: "YouTube elimina la parodia de Pantomima Full sobre los negacionistas de la COVID-19 por cuestionar las directrices de la OMS", *Xataka*. 2/10/2020.

GARCÍA HERRERO,J: "Fuentes accesibles al público y RGPD (OjoAlDato)". *Blog de Jorge García Herrero*. 07/04/2021.

HERNÁNDEZ, M; IGNACIO ESCRIBANO,J:: "Adversarial Machine Learning: una introducción. ¿Motivo de preocupación?", *BBVA Next Technologies,* 15/09/2020.

HUERGO LORA, A,J: "*Regular la inteligencia artificial (en Derecho administrativo"). Blog de la Revista de Derecho Público*. 08/03/2021.

JONES,R: "Roomba's Next Big Step Is Selling Maps of Your Home to the Highest Bidder", *Gizmodo*:24/7/2017.

KARP,H: "Industry Out of Harmony With YouTube on Tracking of Copyrighted Music", *The wall Street Journal*. 28/0672016.

LAPOWSKY,I: "After sending content moderators home, YouTube doubled its video removals". *Protocol*, 25/08/2020.

MATUSZEWSKA,K: "When design goes away – How dark patterns conflict with GDPR and CCPA". *Piwik*, 26/04/2021.

MALIK,K: "Technology will never replace human judgment. Look at football…" The Guardian. 16/11/2019.

MERINO,M: "Conceptos de inteligencia artificial: qué es la inteligencia artificial antagónica y cómo puede manipular a otras IAs", *Xataka*. 18/08/2019.

METZ, C: "There is a Racial Divide in Speech-Recognition System". *The New York Times*. 23/03/2020.

- "A.I Is learning form Humans. Many Humans". The New York Times. 16/08/2019.

MOZUR,P: "Los militares que usaron Facebook para incentivar un genocidio". *The New York Time*. 18/10/2018.

REILLYARCHIVE,M: "Is Facebook Targeting Ads at Sad Teens?", *MIT Technology Review*, 01/05/2017.

RUS,C: "Una base de datos filtrada desvela un esquema con cientos de miles de personas implicadas en reseñas falsas en Amazon", *Xataka*, 10/05/2021.

SIMONITE,T: "Meet the Secret Algorithm That's Keeping Students Out of College". Wired, 07/10/2020.

TAYLOR, J: "Facebook allows advertisers to target children interested in smoking, alcohol and weight loss", *The Guardian*, 28/04/2021.

VALERO TORRIJOS, J: "La necesaria reformulación de las garantías jurídica ante la innovación tecnológica en la Administración", *Blog de derecho, tecnología y modernización administrativa*, 01/06/2016.

III. DOCUMENTOS E INFORMES DE GRUPOS DE TRABAJO, ORGANISMOS INTERNACIONES, ASOCIACIONES Y FUNDACIONES

Agencia europea de Derechos Fundamentales. *Manual de legislación europea contra la discriminación*, Edición de 2018.

AI 360 | COPENHAGEN workshop, *Human Brain Project***,** 2019.

Asociación empresarial del seguro. *Guía para el tratamiento de los datos personales por las entidades aseguradoras,* febrero 2019.

Australian Human Rights Commission. *Human rights and Technology. Discussion Pape*r, 2019.

Center for Democracy and Technology. *Digital Decisions,2017.*

- *Algorithmic Systems in Education: Incorporating Equity and Fairness When Using Student Data*, 2019.

Centro Criptológico Nacional. *Guía de auditoría. Guía de Seguridad de las TIC CCN-STIC 802.*

Comité de ética alemán. Gutachten der Datenethikkommission, 2019.

Comisión Europea. *Libro Blanco sobre la inteligencia artificial. Un enfoque europeo orientado a la excelencia y la confianza*, Bruselas, 19.2.2020.

- *Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica*, 2020.
- Commission of the European Communities, COM (92) 422 final - SYN 287 Brussels, 15 October 1992.
- Comunicación de la Comisión sobre la protección de las personas en lo referente al tratamiento de datos personales en la Comunidad y a la seguridad de los sistemas de Información. COM(90) 314 final - SYN 288, 24 de septiembre de 1990.
- *Ethics and data protection*. Texto aprobado el 14 de noviembre de 2018.
- *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. Configurar el futuro digital de Europa*, 2020.
- Code of practice on disinformation. First annual reports. 2019

Consejo de Europa. *Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data*, 2017.

- *European ethical charter on the use of artificial intelligence in judicial systems and their environment*, 2018.
- *Recomendación CM/Rec(2010)13 del Comité de Ministros a los Estados miembros sobre la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal en el contexto de la creación de perfiles.*

- *Artificial Intelligence and Data Protection: Challenges and Possible Remedies*. Informe adoptado el 25 de enero de 2019.

Eticas-consulting. *Guía de Auditoría Algorítmica, 2021*.

European Parliamentary Research Service. *A governance framework for algorithmic accountability and transparency*. 2019.

- *The ethics of artificial intelligence: Issues and initiatives,*2020.
- Resolución del Parlamento Europeo, de 15 de junio de 2017, sobre una Agenda Europea para la economía colaborativa.

Experian. *The State of Alternative Credit Data. How the financial services industry is adopting and benefiting from these new data sources.* 2018.

Expert Group on Liability for New Technologies. Comisión Europea. *Liability for Artificial Intelligence and other emerging digital technologies*, 2019.

Fundación 29. Playbook Health Data.

General Secretariat of the Council. *Preparation of the Council position on the evaluation and review of the General Data Protection Regulation (GDPR)*. Resolución del 9 de octubre de 2019.

Gobierno del Reino Unido. *Guidelines for AI procurement*, junio de 2020.

Grupo Independiente de Expertos de Alto nivel sobre Inteligencia Artificial creado por la Comisión Europea. *Una definición de la inteligencia artificial: Principales capacidades y disciplinas científicas*, 2019.

- *Directrices éticas para una IA fiable, 2019.*

Informe del Relator Especial sobre la extrema pobreza y los derechos humanos, Philip Alston, presentado de conformidad con la resolución 35/19 del Consejo de Derechos Humanos. 11 de octubre de 2019

International Human Rights Program (Faculty of Law, University of Toronto) and the Citizen Lab. *Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System*.2018.

Memorándum de la Casa Blanca dirigido a las agencias federales sobre cómo regular las aplicaciones de inteligencia aparece una definición de inteligencia artificial. 17 de noviembre de 2020.

New York City Automated Decision Systems (ADS). "Automated Decision Systems Task Force Report, November",2019.

OCDE, *Artificial Intelligence in Society*, Paris, 2019.

- "Directrices sobre protección de la privacidad y flujos transfronterizos de datos personales" de 1980.

Office for artificial intelligence. Government Digital Service of UK. *A guide to using artificial intelligence in the public sector*. 2020.

Privacy international. Data is power: *Towards additional guidance on profiling and automated decision-making in the GDPR*, 2017.

Red DAIA (Derecho Administrativo e Inteligencia Artificia. *Declaración de Valencia*, octubre 2019.

Red iberoamericana de protección de datos personales. *Orientaciones Específicas para el Cumplimiento de los Principios y Derechos que Rigen la Protección de los Datos Personales en los Proyectos de Inteligencia Artificial.*2019.

Secretary's Advisory Committee on Automated Personal Data Systems. *Records, Computers and the Rights of Citizens*. July, 1973.

Select Committee on Artificial Intelligence: House of Lord. *AI in the UK: ready, willing and able?* Report of Session 2017–19, HL Paper 100.

Supervisor europeo de protección de datos. *A Preliminary Opinion on data protection and scientific research.* Texto adoptado el 6 de enero de 2020.

- EDPS Ethics Advisory Group | Report, *Towards a digital ethics*, 2018

The International Committee of the Red Cross (ICRC). *Autonomy, artificial intelligence and robotics: Technical aspects of human control*, agosto 2019.

IV RESOLUCIONES DE AUTORIDADES DE PROTECCIÓN DE DATOS

1. Grupo del Artículo 29 y Comité Europeo de Protección de Datos

A) Grupo del Artículo 29

Grupo del Artículo 29. *Documento de trabajo sobre el tratamiento de datos personales relativos a la salud en los historiales médicos electrónicos*. Adoptado el 15 de febrero de 2007.

Grupo del Artículo 29. *Dictamen 4/2007 sobre el concepto de datos personales.* Adoptado el 20 de junio de 2007.

Grupo del Artículo 29 y Grupo de Policía y Justicia. *The Future of Privacy Joint contribution to the Consultation of the European Commission on the legal framework for the fundamental right to protection of personal data*. Resolución adoptada el 1 de diciembre de 2009.

Grupo del Artículo 29. *Opinion 1/2010 on the concepts of controller and processor*. Adoptada el 16 de febrero de 2010.

Grupo del Artículo 29. *Dictamen 3/2010 sobre el principio de responsabilidad.* Adoptado el 13 de julio de 2010.

Grupo del Artículo 29. *Advice paper on special categories of data ("sensitive data"),* 2011.

Grupo del Artículo 29. *Dictamen 02/2012 sobre reconocimiento facial en los servicios en línea y móviles*. Adoptado el 22 de marzo de 2012.

Grupo del Artículo 29. *Dictamen 3/2012 sobre la evolución de las tecnologías biométricas.* Resolución adoptada el 27 de abril de 2012.

Grupo del artículo 29. *Opinion 03/2013 on purpose limitation*. Texto adoptado el 2 de abril de 2013.

Dictamen 03/2014 sobre la notificación de violación de datos personales. Adoptado el 25 de marzo de 2014.

Grupo del artículo 29. *Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE*. Adoptado el 9 de abril de 2014.

Grupo del Artículo 29. Dictamen 05/2014 sobre técnicas de anonimización. Adoptado el 10 de abril de 2014.

Grupo del Artículo 29. *Statement on the role of a risk-based approach in data protection legal frameworks.* Resolución adoptada el 30 de mayo de 2014.

Grupo del Artículo 29. *Dictamen 8/2014 sobre la evolución reciente de la Internet de los objetos*. Adoptado el 16 de septiembre de 2014.

Grupo del Artículo 29. *Directrices sobre el derecho a la portabilidad de los datos. Adoptadas el 13 de diciembre de 2016*. Revisadas por última vez y adoptadas el 5 de abril de 2017.

Grupo del artículo 29. *Directrices sobre los delegados de protección de datos (DPD). Adoptadas el 13 de diciembre de 2016.* Revisadas por última vez y adoptadas el 5 de abril de 2017.

Grupo del artículo 29. *Dictamen 2/2017 sobre el tratamiento de datos en el trabajo.* Adoptado el 8 de junio de 2017**.**

Grupo del Artículo 29. *Directrices del grupo del artículo 29 sobre la aplicación y la fijación de multas administrativas a efectos del Reglamento 2016/679*. Adoptadas el 3 de octubre de 2017.

Grupo del Artículo 29. *Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679***.** Adoptadas el 4 de abril de 2017. Revisadas por última vez y adoptadas el 4 de octubre de 2017.

Grupo Artículo 29. *Directrices sobre la notificación de las violaciones de la seguridad de los datos personales de acuerdo con el Reglamento 2016/679*. Adoptadas el 3 de octubre de 2017. Revisadas por última vez y adoptadas el 6 de febrero de 2018.

Grupo del artículo 29. Dictamen sobre algunas cuestiones fundamentales de la Directiva sobre protección de datos en el ámbito penal (UE 2016/680). Aprobado el 29 de noviembre de 2017.

Grupo del Artículo 29. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*.Adoptadas el 3 de octubre de 2017 y revisadas el 6 de febrero de 2018.

Grupo del artículo 29. *Directrices sobre la transparencia en virtud del Reglamento (UE) 2016/679.* Adoptadas el 29 de noviembre de 2017. Revisadas por última vez y adoptadas el 11 de abril de 2018.

Grupo del Artículo 29. *Guidelines 1/2020 on processing personal data in the context of connected vehicles and mobility related applications*. Versión 1.0. Adoptadas el 28 de enero de 2020.

Grupo del Artículo 29. *Guidelines 8/2020 on the targeting of social media users*. Adoptadas el 2 septiembre de 2020.

B) Comité Europeo de Protección de datos

Comité Europeo de Protección de Datos. *Opinion 8/2018 on the legislative package "A New Deal for Consumers"*. Resolución adoptada el 5 de octubre de 2018.

Comité Europeo de Protección de Datos. *Directrices 1/2018 sobre la certificación y la determinación de los criterios de certificación de conformidad con los artículos 42 y 43 del Reglamento*. Versión 3.0. Adoptadas el 4 de junio de 2019.

Comité Europeo de Protección de Datos. *Directrices 2/2019 sobre el tratamiento de datos personales en virtud del artículo 6, apartado 1, letra b), del RGPD en el contexto de la prestación de servicios en línea a los interesados*. Versión 2.0. Aprobadas el 8 de octubre de 2019.

Comité Europeo de Protección de Datos. *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default*. Directrices adoptadas el 13 de Noviembre de 2019.

Comité Europeo de Protección de Datos. *Guidelines 1/2020 on processing personal data in the context of connected vehicles and mobility related applications*. Versión 1.0 Directrices adoptadas el 28 de enero de 2020.

Comité Europeo de Protección de Datos. *Directrices 03/2020 sobre el tratamiento de datos relativos a la salud con fines de investigación científica en el contexto del brote de COVID-19* Adoptadas el 21 de abril de 2020.

Comité Europeo de Protección de Datos. *Directrices 5/2020 sobre el consentimiento en el sentido del Reglamento (UE) 2016/679*, versión 1.1. Directrices adoptadas el 4 de mayo de 2020.

Comité Europeo de Protección de Datos. *Guidelines 8/2020 on the targeting of social media users.* Versión 1.0. Directrices adoptadas el 2 de septiembre de 2020.

Comité Europeo de Protección de Datos. "Guidelines 07/2020 on the concepts of controller and processor in the GDPR". Adoptadas el 2 de septiembre de 2020.

Comité Europeo de Protección de Datos. *Guidelines 10/2020 on restrictions under Article 23 GDPR*. Versión 1.0. Texto adoptado el 15 de diciembre de 2020.

Comité Europeo de Protección de Datos. *Guidelines 01/2021 on Examples regarding Data Breach Notification*. Resolución adoptada el 14 de junio de 2021.

Comité Europeo de Protección de Datos. EDPB-EDPS. Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act). Resolución de 18 de junio de 2021.

2. Agencia Española de Protección de Datos

A) Documentos de interés

Agencia Española de Protección de Datos. *Guía sobre el uso de las cookies*.

Agencia Española de Protección de Datos. *Orientaciones y garantías en los procedimientos de anonimización de datos personales*. Adoptadas en 2016.

Agencia Española de Protección de Datos. *Orientaciones sobre protección de datos en la reutilización de la información del sector público*, 2016.

Agencia Española de Protección de Datos. Guía para el cumplimiento del deber de informar. 2017.

Agencia Española de Protección de Datos. *Guía Práctica de análisis de riesgos en los tratamientos de datos personales sujetos al RGPD*. 2018.

Agencia Española de Protección de Datos. *Guía práctica para las Evaluaciones de Impacto en la Protección de los Datos sujetas al RGPD,* 2018.

Agencia Española de Protección de Datos. *El delegado de protección de datos en las administraciones públicas*, 2018.

Agencia española de protección de datos. La k-anonimidad como medida de la privacidad Anonimización. 2019, pág.3.

Agencia Española de Protección de Datos. Estudio Fingerprinting o Huella digital del dispositivo". Adoptado en Febrero 2019.

Agencia Española de Protección de datos. *Listas de tipos de tratamientos de datos que requieren evaluación de impacto relativa a protección de datos (art 35.4).* Resolución adoptada el 4 de septiembre de 2019.

Agencia Española de Protección de Datos. *Introducción al hash como técnica de seudonimización de datos personales*. Octubre de 2019.

Agencia Española de Protección de Datos. *Guía de Privacidad desde el Diseño*. Octubre de 2019.

Agencia Española de Protección de Datos. *Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, febrero 2020.

Agencia Española de Protección de Datos. Guía sobre el uso de las cookies de julio de 2020.

Agencia Española de Protección de Datos. *La protección de datos en las relaciones laborales*. 2021.

Agencia Española de Protección de Datos. *Requisitos para Auditorías de Tratamientos que incluyan IA,* 2021.

Agencia Española de Protección de Datos. *Gestión del riesgo y evaluación de impacto en tratamientos de datos personales.* Junio 2021.

B) Resoluciones, consultas e informes

Agencia Española de Protección de Datos. Resolución: Nº AP/00075/2015.

Agencia Española de Protección de Datos. Informe jurídico 0195/2017.

Agencia Española de Protección de Datos. Informe nº 0121/2018.

Agencia Española de Protección de Datos. Informe 175/2018.

Agencia Española de Protección de Datos. Resolución Nº: PS/00326/2018.

Agencia Española de Protección de Datos. Informe AEPD Nº: 073667/2018.

Agencia Española de Protección de Datos. Informe N/REF: *0017/2019.*

Agencia Española de Protección de Datos. Informe Nº: PS/00070/2019.

Agencia Española de Protección de Datos. Consulta jurídica N/REF: 00148/2019.

Agencia Española de Protección de Datos. Resolución Nº: PS/00187/2019.

Agencia Española de Protección de Datos. Resolución Nº: PS/00240/2019.

Agencia Española de Protección de Datos. Resolución Nº**:** PS/00477/2019.

Agencia Española de Protección de Datos. Informe de 8 de mayo de 2021. Informe N/REF: 0036/2020.

Agencia Española de Protección de Datos. . Resolución Nº: PS/00037/2020.

Agencia Española de Protección de datos. Informe N/REF: 0089/2020.

Agencia Española de Protección de datos. Resolución Nº: PS/00136/2020

Agencia Española de Protección de Datos., Resolución de la AEPD Nº E/02666/2020.

Agencia Española de Protección de Datos. Informe N/REF: 0014/2021.

Agencia Española de Protección de Datos. Informe N/REF: 0047/2021.

Agencia Española de Protección de Datos. Resolución Nº: PS/00120/2021.

3. Resto de autoridades de control

A) Autoritat Catalana de Protecció de dades

Autoritat Catalana de Protecció de dades. *Guía práctica. Evaluación de impacto relativa a la protección de datos*. Enero de 2018, versión 2.0.

Autoritat Catalana de Protecció de dades. *Guia Pràctica. Avaluació d'impacte relativa a la protecció de dades*. 2019.

Autoritat Catalana de Protecció de Dades. Informe nº PD 9/2019.

Autoritat Catalana de Protecció de dades. Consulta CNS 15/2019.

Autoritat Catalana de Protecció de Dades. *Intel·ligència Artificial. Decisions Automatitzades a Catalunya*, 2020.

Autoritat Catalana de Protecció de Dades. *Guía de protección de datos para pacientes y personas usuarias de los servicios de salud***.** Junio 2020.

Autoritat Catalana de Protecció de dades. Consulta CNS 6/2020

Autoritat Catalana de Protecció de dades. Consulta nº CNS 36/2020.

Autoritat Catalana de Protecció de Dades. Consulta nº: CNS 4/2021.

B) Information Commissioner´s Office

Information Commissioner´s Office. *Anonymisation: managing data protection risk code of practice*.2012.

ICO. Guidance on AI and data protection. How do we ensure individual rights in our AI systems?

Information Commisioner´s Office . Resolución IC-45844-V2X8

Information Commisioner´s Office. *Data Protection Impact Assessments (DPIAs)*. Documento adoptado el 22 de marzo de 2018.

Information Commissioner's Office. *Explaining decision made with AI | Part 1: The basics of explaining AI*, 2019.

Information Commissioner's Office. *Explaining decisions made with AI. Part 1. The basics of explaining AI*, mayo de 2020.

C) Otras autoridades de control

Autoridad de Protección de Datos Sueca (datainspektionen). *List regarding Data Protection Impact Assessments according to article 35.4 of the Data Protection Regulation*. Documento adoptado el 16 de enero de 2019.

Autoridad de Protección de Datos Eslovena (Information Commisioner). The list of the kind of processing operations1 which are subject to the requirement for a Data Protection Impact Assessment under the Article 35(4) of the General Data Protection Regulation (EU) 2016/679 (GDPR). Resolución adoptada el 21 de diciembre de 2018.

Autoridad de Protección de datos Irlandesa establece como obligatorio la realización de la EIPD cuando se traten datos de personas vulnerables. (Data protection comission). *List of Types of Data Processing Operations which require a Data Protection Impact Assessment.*

Autoridad de protección de datos francesa (CNIL), Decisión MED-2020-015 de 15 julio de 2020.

Autoridad de Protección de Datos Islandesa (Persónuvernd) Nº de resolución: 2020010673.

Autoridad de Protección de Datos Finlandesa (Tietosuojavaltuutetun toimisto) resolución 8393/161/2019 de 26 de mayo de 2020.

Comisionado de Hamburgo para la Protección de Datos y la Libertad de Información. Resolución de 11 de mayo de 2021.

Comissão Nacional de Proteção de Dados. Deliberação 622/2021.

Garante per la Protezione dei Dati Personali. Resolución nº 234 de 10 de junio de 2021.

The Norwegian Data Protection Authority. *Artificial intelligence and privacy*. Report, 2018.

- Datatilsynet. *Software development with Data Protection by Design and by Default.*

4. Consejos de transparencia

Comisión de Garantía del derecho de acceso e información púbica de Cataluña (GAIP) Resolución 200/2017, de 21 de junio de 2017.

Consejo de Transparencia y Buen Gobierno. Resolución 701/2018. Febrero de 2019.

Consejo de Transparencia y Buen Gobierno. Resolución 058/2021. Mayo de 2021.

V. SENTENCIAS JUDICIALES

1. Tribunal de Justicia de la Unión Europea

Dictamen 1/15 del TJUE (Gran Sala) de 26 de julio de 2017 sobre el Proyecto de Acuerdo entre Canadá y la Unión Europea con relación a la transferencia de los datos del registro de nombres de los pasajeros aéreos desde la Unión a Canadá.

Sentencia del TJUE (Sala Cuarta) de 14 de septiembre de 2000, asunto C-369/98, caso Fisher.

Sentencia del TJUE, Sentencia de 20 Mayo de 2003, asuntos acumulados C-465/00, C-138/01 y C-139/01, caso Rechnungshof.

Sentencia del TJUE de 7 de mayo de 2009, asunto C-553/07, caso Rijkeboer.

Sentencia del TJUE (Gran Sala) de 16 de diciembre de 2008, asunto C-524/06, caso Heinz Huber.

Sentencia del TJUE (Sala Tercera) de 30 de mayo de 2013, asunto C-342/12, caso Worten.

Sentencia del TJUE (Sala Tercera) de 24 de noviembre de 2011. Asunto C-70/10, caso Scarlet Extended SA.

Sentencia del TJUE (Sala Tercera) de 24 de noviembre de 2011, asuntos C-468/10 y C-469/10, caso ASNEF.

Sentencia del TJUE (Sala Tercera) de 16 de febrero de 2012, Asunto C-360/10, caso SABAM.

Sentencia del TJUE (Gran Sala) de 13 de mayo de 2014, asunto C-131/12, caso Mario Costeja y Google Spain, S.L

Sentencia del TJUE (Sala Tercera) de 17 de julio de 2014, asuntos acumulados, C-141/12 y C-372/12, caso YS. y M. y S.

Sentencia del TJUE (Sala Cuarta) de 11 de diciembre de 2014, asunto C-212/13, caso Ryneš.

Sentencia del TJUE (Gran Sala) de 6 de octubre de 2015, asunto C-362/14, caso Schrems.

Sentencia del TJUE de 19 de octubre de 2016, asunto C-582/14, caso Patrick Breyer**.**

Sentencia del TJUE (Sala Segunda) de 4 de mayo de 2017, asunto C-13/16, caso Rīgas.

Sentencia del TJUE (Sala Segunda) de 27 de septiembre de 2017, asunto C-73/16, caso Peter Puškár.

Sentencia del TJUE (Sala Segunda) de 20 de diciembre de 2017, asunto C-434/16, caso Nowak.

Sentencia del TJUE (Gran Sala) de 10 de julio de 2018, asunto C–25/17, caso Jehovan todistajat.

Sentencia del TJUE (Sala Segunda) de 29 de julio de 2019, asunto C-40/17, caso Fashion ID.

Sentencia del TJUE (Gran Sala) de 1 de octubre de 2019, asunto C-673/17, caso Planet49.

Sentencia del TJUE (Gran Sala) de 6 de octubre de 2020, asunto C-623/17, caso Privacy International.

Sentencia del TJUE (Sala Segunda) de 11 de noviembre de 2020, asunto C-61/19, caso Orange Romania SA / ANSPDCP.

Sentencia del TJUE de 15 de junio de 2021, asunto C-645/19, caso Facebook Ireland y otros.

Sentencia del TJUE de 22 de junio de 2021, asuntos acumulados C-682/18 (YouTube) C-683/18 (Cyando).

2. Tribunal Europeo de Derechos Humanos

Tribunal Europeo de Derechos Humanos. Sentencia CASE OF BIG BROTHER WATCH AND OTHERS v. THE UNITED KINGDOM *(Applications nos. 58170/13, 62322/14 and 24960/15*) de 25 de mayo de 2021.

3. Tribunales españoles

A) Tribunal Constitucional

STC Sentencia 11/1981, de 8 de abril.

STC Sentencia 254/1993 de 20 de julio de 1993.

STC Sentencia 290/2000 de 30 de noviembre de 2000.

STC Sentencia 292/2000, de 30 de noviembre de 2000.

STC Sentencia 17/2013 de 31 de enero de 2013.

STC Sentencia 76/2019, de 22 de mayo de 2019.

B) Tribunal Supremo

STS (Sala de lo Contencioso-administrativo), Sentencia de 19/09/2008.

STS (Sala Tercera, de lo Contencioso-administrativo, Sección 6ª) Sentencia de 12 Marzo de 2012, Rec. 2453/2009.

STS (Sala de lo Contencioso-Administrativo, Sección 3ª) Sentencia núm. 1007/2019 de 8 julio

STS (Sala de lo Contencioso-Administrativo, Sección 3ª) Sentencia núm. 1062/2019 de 12 julio.

STS 518/2021 sala de lo social, Sentencia de 08/02/2021.

STS Sala de lo Penal, Sentencia núm. 1153/2021, Sentencia de 22/03/2021.

C) Audiencia Nacional

Sentencia de la Audiencia Nacional AN (Sala de lo Contencioso-Administrativo, Sección 1ª) Sentencia de 23 noviembre 2012.

Sentencia Audiencia Nacional (Sala de lo Contencioso-Administrativo, Sección1ª). Sentencia de 4 junio 2021.

D) Juzgados y audiencias provinciales

Auto Audiencia Provincial Ciudad Real (Sección1), núm.43/2005, de 7 de marzo

Auto Audiencia Provincial Madrid (Sección 5ª), nº1176/2006 de 21 de marzo.

Audiencia Provincial de Barcelona, Sección 9ª, Auto 72/2021 de 15 Feb. 2021, Rec. 840/2021.

4. Tribunales de países europeos

Sentencia del Tribunal Administrativo Regional de Lazio de 22 de marzo de 2017.

Comité de No discriminación e igualdad de Género de Finlandia. Resolución de 21 de marzo de 2018. (Finlandia)

Tribunal Constitucional Alemán. *Bundesverfas-sungsgericht (1 BvR 142/15), 18 de diciembre* 2018.

Consejo de Estado Francés. Resolución Nº 434376 de 6 de noviembre de 2019.

Sentencia Consejo de Estado italiano de 13 diciembre 2019, nº. 8472.

Consejo Constitucional Francés. Décision n° 2019-796 DC du 27 décembre 2019

Court of appeal (civil division) on appeal form the high court of justice queen´s bench division (administrative court). Case No: C1/2019/2670, (2020) EWCA Civ 1058. (Reino Unido)

Tribunal de la Haya. Sentencia de 5 de febrero de 2020.

Tribunal Ordinario de Bolonia de 31 de diciembre de 2020.

Tribunal de Ámsterdam. Resolución de 11 de marzo de 2021.

Tribunal Federal de Justicia de Alemania. Sentencia de 29 de julio de 2021. Procedimiento III ZR 179/20.

5. Tribunales en el resto del mundo

Naperville Smart Meter Awareness v. City of Naperville, No. 16-3766 (7th Cir. 2018). (EEUU)

Sentencia State v.Loomis, 881.N.W.2d 749 (Wisc.2016). (EEUU)

VI. TEXTOS NORMATIVOS UTILIZADOS

1. Normativa Unión Europea

DIRECTIVA 95/46/CE del PARLAMENTO EUROPEO Y DEL CONSEJO, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos.

CARTA DE DERECHOS FUNDAMENTALES DE LA UNIÓN EUROPA.

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.

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.

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.

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.

DIRECTIVA (UE) 2019/770 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 20 de mayo de 2019 relativa a determinados aspectos de los contratos de suministro de contenidos y servicios digitales.

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.

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.

DIRECTIVA (UE) 2019/1936 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 23 de octubre de 2019 por la que se modifica la directiva 2008/96/CE sobre gestión de la seguridad de las infraestructuras viarias.

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. Resolución del Parlamento Europeo, de 20 de octubre de 2020.

Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO relativo a la gobernanza europea de datos (Ley de Gobernanza de Datos), Bruselas. Resolución del 25 de noviembre de 2020.

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. Resolución de 15 de diciembre de 2020.

Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas sobre la inteligencia artificial (ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión. Resolución de la Comisión Europea de 21 de abril de 2021.

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.

REGLAMENTO (UE) 2021/694 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 29 de abril de 2021 por el que se establece el Programa Europa Digital y por el que se deroga la Decisión (UE) 2015/2240.

REGLAMENTO (UE) 2021/1232 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 14 de julio de 2021 por el que se establece una excepción temporal a determinadas disposiciones de la Directiva 2002/58/CE en lo que respecta al uso de tecnologías por proveedores de servicios de comunicaciones interpersonales independientes de la numeración para el tratamiento de datos personales y de otro tipo con fines de lucha contra los abusos sexuales de menores en línea.

2. Normativa española

Constitución Española de 1978.

Ley 7/1985, de 2 de abril, Reguladora de las Bases del Régimen Local.

Real Decreto 190/1996, de 9 de febrero, por el que se aprueba el Reglamento Penitenciario.

Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal.

Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica.

Ley 11/2007, de 22 de junio, de acceso electrónico de los ciudadanos a los Servicios Públicos.

Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público.

Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal.

Decreto Legislativo 1/2009, de 21 de julio, por el que se aprueba el Texto refundido de la Ley reguladora de los residuos. (Cataluña)

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.

Ley 2/2011, de 4 de marzo, de Economía Sostenible.

Orden EHA/2899/2011, de 28 de octubre, de transparencia y protección del cliente de servicios bancarios.

Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno.

Ley 21/2013, de 9 de diciembre, de evaluación ambiental.

Ley 20/2015, de 14 de julio, de ordenación, supervisión y solvencia de las entidades aseguradoras y reaseguradoras.

Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público.

Ley 9/2017, de 8 de noviembre, de Contratos del Sector Público, por la que se transponen al ordenamiento jurídico español las Directivas del Parlamento Europeo y del Consejo 2014/23/UE y 2014/24/UE, de 26 de febrero de 2014.

Resolución de 13 de abril de 2018, de la Secretaría de Estado de Función Pública, por la que se aprueba la Instrucción Técnica de Seguridad de Notificación de Incidentes de Seguridad.

Ley 22/2018, de 6 de noviembre, de Inspección General de Servicios y del sistema de alertas para la prevención de malas prácticas en la Administración de la Generalitat y su sector público instrumental. (Comunidad Valenciana)

Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

Ley 1/2019, de 20 de febrero, de Secretos Empresariales.

Circular 1/2019, de 7 de marzo, de la Agencia Española de Protección de Datos, sobre el tratamiento de datos personales relativos a opiniones políticas y envío de propaganda electoral por medios electrónicos o sistemas de mensajería por parte de partidos políticos, federaciones, coaliciones y agrupaciones de electores al amparo del artículo 58 bis de la Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General.

Ley Orgánica 1/2020, de 16 de septiembre, sobre la utilización de los datos del Registro de Nombres de Pasajeros para la prevención, detección, investigación y enjuiciamiento de delitos de terrorismo y delitos graves.

Resolución de 17 de marzo de 2021, de la Dirección General de Trabajo, por la que se registra y publica el XXIV Convenio colectivo del sector de la banca.

Real Decreto 203/2021, de 30 de marzo sobre el funcionamiento del sector público por medios electrónicos viene prácticamente a calcar lo ya previsto y comentado por el artículo 41 de la Ley 40/2015.

Real Decreto 327/2021, de 11 de mayo por el que se modifica el Real Decreto 181/2008, de 8 de febrero, de ordenación del diario oficial "Boletín Oficial del Estado", para adaptarlo al Tablón Edictal Judicial Único.

Real Decreto-ley 9/2021, de 11 de mayo, por el que se modifica el texto refundido de la Ley del Estatuto de los Trabajadores, aprobado por el Real Decreto Legislativo 2/2015, de 23 de octubre, para garantizar los derechos laborales de las personas dedicadas al reparto en el ámbito de plataformas digitales.

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.

Carta de Derechos Digitales española. Julio de 2021.

Real Decreto 688/2021, de 3 de agosto, por el que se modifica el Reglamento general sobre procedimientos para la imposición de sanciones por infracciones de orden social y para los expedientes liquidatorios de cuotas de la Seguridad Social, aprobado por el Real Decreto 928/1998, de 14 de mayo.

3. Normativa resto de Europa

Convenio Europeo de Derechos Humanos. Consejo de Europa.

Ley Francesa n ° 78-17 del 6 de enero de 1978 relativa al procesamiento de datos, archivos y libertades (Versión vigente el 23 de julio de 1978). (Francia)

Convenio 108 para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal. 1981. Consejo de Europa.

Ley de Autodeterminación del Derecho a la Información y Libertad de Información de 2011. (Hungría).

Besluit van 1 september 2014 tot wijziging van het Besluit SUWI in verband met regels voor fraudeaanpak door gegevensuitwisselingen en het effectief gebruik van binnen de overheid bekend zijnde gegevens met inzet van SyRI. (Países Bajos)

Loi de finances pour 2020, sous le n° 2019-796 DC, le 20 décembre 2019. (Francia)

4. Normativa de países no europeos

E-Governance Act of 2002. (EEUU)

California Consumer Privacy Act of 2018. (EEUU)

Directive on Automated Decision-Making de 1 de abril de 2019. (Canadá)

Algorithm Charter for Aotearoa New Zealand de Julio de 2020. (Australia)

Artificial Intelligence Video Interview Act, Public Act 101-0260,1/1/2020. (EEUU)

Algorithmic Justice and Online Platform Transparency Act. Proyecto de ley publicado el 28 de mayo de 2021. (EEUU)