



VNIVERSITATIS VALÈNCIA

DOCTORADO DIRECCIÓN DE EMPRESAS-3113
R.D. 99/2011

TESIS DOCTORAL

Ecosistemas de Big Data en las organizaciones y su efecto en los procesos de gestión del conocimiento y los mecanismos de coordinación y control: un estudio de casos exploratorio

PRESENTADO POR:

Erik Daniel Maldonado Ascanio

DIRIGIDO POR:

Dr. Francisco Balbastre Benavent

FECHA:

Mayo de 2022

Departament Direcció de Empreses
“Juan José Renau Piqueras”
Facultat d’Economía

Agradecimientos

A mi hijo Daniel José y mi esposa Katherine, gracias por su comprensión y afecto, son la fuerza que me mueve todos los días.

A mis padres Julio y Elsy por su apoyo incondicional.

A mi director de tesis Francisco, gracias por apoyarme en todo.

A mis amigos que siempre estuvieron ahí para apoyarme.

A las personas que participaron y apoyaron mi investigación, gracias.

TABLA DE CONTENIDO

ÍNDICE DE TABLAS	V
ÍNDICE DE FIGURAS	VI
INTRODUCCIÓN.....	1
CAPÍTULO 1: BIG DATA Y LOS SISTEMAS Y TECNOLOGÍAS DE LA INFORMACIÓN EN LAS ORGANIZACIONES.....	9
1.1 TECNOLOGÍAS Y SISTEMAS DE INFORMACIÓN EN LAS ORGANIZACIONES: CONCEPTOS Y CARACTERÍSTICAS	10
1.1.1 Concepto y Valor de la Información en las Organizaciones	10
1.1.2 Concepto e Importancia de los Sistemas y Tecnologías de la Información.....	13
1.1.3 Tipos de Sistemas de Información	19
1.2 BIG DATA Y SU RELACIÓN CON LOS SISTEMAS Y TECNOLOGÍAS DE LA INFORMACIÓN	25
1.2.1 Concepto de Big Data: las 10 V's	25
1.2.2 Concepto, Métodos, Técnicas y Herramientas de Big Data Analytics	32
1.2.3 Big Data y los Sistemas de Información para la Toma de Decisiones y el Conocimiento	42
1.2.4 Big Data y las Tecnologías Claves de la Transformación Digital	45
1.3 BIG DATA EN LAS ORGANIZACIONES	49
1.3.1 Big Data, la Ciencia de Datos y el Científico de Datos	49
1.3.2 Aplicaciones del Big Data en Diferentes Sectores.....	53
1.3.3 Niveles de Implementación de Ecosistemas de Big Data en las Organizaciones ..	57
CAPÍTULO 2: PROCESOS DE GESTIÓN DEL CONOCIMIENTO Y MECANISMOS DE COORDINACIÓN Y CONTROL EN LA ORGANIZACIÓN	65
2.1 CONCEPTO Y PROCESOS DE LA GESTIÓN DEL CONOCIMIENTO	66
2.1.1 Información, Conocimiento y Sabiduría.....	66
2.1.2 Concepto de Gestión del Conocimiento.....	73
2.1.3 Procesos de la Gestión del Conocimiento.....	78
2.2 DISEÑO ORGANIZATIVO Y LOS MECANISMOS DE COORDINACIÓN Y CONTROL	86

2.2.1	Concepto de Diseño Organizativo	86
2.2.2	Mecanismos de Coordinación y Control.....	96
CAPÍTULO 3:	BIG DATA, LOS PROCESOS DE GESTIÓN DEL CONOCIMIENTO Y LOS MECANISMOS DE COORDINACIÓN Y CONTROL.....	107
3.1	BIG DATA Y LOS PROCESOS DE GESTIÓN DEL CONOCIMIENTO	109
3.1.1	Big Data y el Proceso de Creación del Conocimiento	111
3.1.2	Big Data y el Proceso de Almacenamiento/Consulta del Conocimiento.....	114
3.1.3	Big Data y el Proceso de Transferencia del Conocimiento.....	116
3.1.4	Big Data y el Proceso de Aplicación del Conocimiento	118
3.2	BIG DATA Y LOS MECANISMOS DE COORDINACIÓN Y CONTROL.....	119
3.2.1	Big Data y los Mecanismos de Coordinación.....	120
3.2.2	Big Data y los Mecanismos de Control.....	123
3.3	BIG DATA Y SU RELACIÓN CON EL PROCESO DE GESTIÓN DEL CONOCIMIENTO Y LOS MECANISMOS DE COORDINACIÓN Y CONTROL: MODELO TEÓRICO PRELIMINAR	125
CAPÍTULO 4:	METODOLOGÍA DE INVESTIGACIÓN	129
4.1	CONSIDERACIONES Y CARÁCTERÍSTICAS GENERALES DE LA METODOLOGÍA CUANTITATIVA Y CUALITATIVA	130
4.2	EL ESTUDIO DE CASOS COMO ESTRATEGIA DE INVESTIGACIÓN CUALITATIVA	135
4.2.1	Tipos de estudio de casos	137
4.3	MUESTRA DEL ESTUDIO DE CASOS.....	141
4.3.1	Unidad de análisis	141
4.3.2	Número y criterios de selección de los datos	142
4.4	MÉTODOS Y TÉCNICAS PARA LA RECOGIDA DE DATOS	146
4.5	ANÁLISIS DE LOS DATOS.....	150
4.6	CRITERIOS DE CALIDAD	156
CAPÍTULO 5:	ANÁLISIS DE RESULTADOS	161
5.1	CASO A: CEMENTOS COLOMBIA S.A.....	163
5.1.1	Perfil de la Organización.....	163
5.1.2	Proceso de Gestión del Conocimiento en Cementos Colombia S.A.....	165

5.1.3	Mecanismos de Coordinación y Control en Cementos Colombia S.A.	167
5.1.4	Ecosistema de Big Data en Cementos Colombia S.A.	170
5.1.5	Efecto de la implementación del ecosistema de Big Data en los procesos de gestión del conocimiento y en los mecanismos de coordinación y control en Cementos Colombia S.A.	176
5.2	CASO B: FUNDACIÓN UNIVERSIDAD DEL NORTE.....	184
5.2.1	Perfil de la Organización.....	184
5.2.2	Procesos de Gestión del Conocimiento en Fundación Universidad del Norte.....	186
5.2.3	Mecanismos de Coordinación y Control en Fundación Universidad del Norte...	190
5.2.4	Ecosistema de Big Data en Fundación Universidad del Norte.	193
5.2.5	Efecto de la implementación del ecosistema de Big Data en los procesos de gestión del conocimiento y en los mecanismos de coordinación y control en Fundación Universidad del Norte.....	201
5.3	CASO C: ECO FUELS ENERGY S.A.S.....	208
5.3.1	Perfil de la Organización.....	208
5.3.2	Proceso de Gestión del Conocimiento en Eco Fuels Energy S.A.S.	209
5.3.3	Mecanismos de Coordinación y Control en Eco Fuels Energy S.A.S.	212
5.3.4	Ecosistema de Big Data en Eco Fuels Energy S.A.S.	214
5.3.5	Efecto de la implementación del ecosistema de Big Data en los procesos de gestión del conocimiento y en los mecanismos de coordinación y control en Eco Fuels Energy S.A.S.	219
5.4	ANÁLISIS CONJUNTO DE CASOS	224
5.4.1	Comparativa de los procesos de gestión del conocimiento.....	224
5.4.2	Comparativa de los mecanismos de coordinación y control	227
5.4.3	Comparativa de la implementación del ecosistema de Big Data	230
5.4.4	Comparativa del efecto de la implementación del ecosistema de Big Data en los procesos de gestión del conocimiento y en los mecanismos de coordinación y control en los casos estudiados	233
CAPÍTULO 6:	CONCLUSIONES GENERALES	241
BIBLIOGRAFÍA		251
ANEXO A: PROTOCOLO DEL ESTUDIO DE CASOS.....		271
ANEXO B: GUIÓN DE LA ENTREVISTA		279
ANEXO C: MODELO DE SOLICITUD FORMAL DE ENTREVISTA		283

ANEXO D: MODELO DE CONSENTIMIENTO INFORMADO DEL PARTICIPANTE.....284

Índice de Tablas

Tabla 1-1: Enfoques conceptuales de las TI	16
Tabla 1-2: Evolución de la terminología para el uso y análisis de datos.....	33
Tabla 1-3: Componentes adicionales del ecosistema Hadoop.....	39
Tabla 1-4: Clases de herramientas de Big Data Analytics	41
Tabla 1-5: Niveles y dimensiones de implementación de ecosistemas de Big Data en las organizaciones	60
Tabla 2-1: Perspectivas para definir el conocimiento	68
Tabla 2-2: Taxonomías del conocimiento	70
Tabla 2-3: Tipos de mecanismos de control	103
Tabla 3-1: Mecanismos de control mediante la inclusión de ecosistemas de Big Data	124
Tabla 4-1: Diferencias entre diseño metodológico cuantitativo y cualitativo en una investigación.....	132
Tabla 4-2: Condiciones para la selección de estrategias de investigación	136
Tabla 4-3: Categorías de estudio de casos en función del propósito de la investigación	138
Tabla 4-4: Planificación de las entrevistas a realizar en esta investigación	148
Tabla 4-5: Entrevistas realizadas en esta investigación.....	149
Tabla 4-6: Tácticas para cumplir los criterios de calidad de la presente investigación	159
Tabla 5-1 Presencia de Cementos Colombia S.A. en América	164
Tabla 5-2: Oferta académica de la Fundación Universidad del Norte	185
Tabla 5-3: Comparativa de los procesos de gestión del conocimiento de los casos estudiados	225
Tabla 5-4: Comparativa de los mecanismos de coordinación y control de los casos estudiados	228
Tabla 5-5: Comparativa de la implementación del ecosistema de Big Data de los casos estudiados	231
Tabla 5-6: Comparativa de las relaciones entre factores propuestos en el modelo inicial de los casos estudiados	234
Tabla 5-7: Resumen del aporte al modelo final de relaciones de la combinación de los casos estudiados.....	238

Índice de Figuras

Figura 0-1: Estructura de capítulos de la tesis doctoral.....	5
Figura 1-1: Valor de la información en la empresa	12
Figura 1-2: Actividades de un Sistema de Información	14
Figura 1-3: Dimensiones de un Sistema de Información	15
Figura 1-4: Categorización de los Sistemas de Información	20
Figura 1-5: Popularidad mundial del término Big Data de 2011 a 2018.....	26
Figura 1-6: Las 10 Vs (Atributos) del Big Data	31
Figura 1-7: Etapas del proceso de Big Data Analytics	35
Figura 1-8: Ecosistema de Big Data	58
Figura 1-9: Niveles en la implementación de ecosistemas de Big data.....	59
Figura 2-1: Visión jerárquica de los datos, información, conocimiento y sabiduría.....	67
Figura 2-2: Modos de conversión del conocimiento	71
Figura 2-3: Dimensiones de la gestión del conocimiento.....	75
Figura 2-4: Procesos de gestión del conocimiento	79
Figura 2-5: Comparación entre modelos de diseño organizativo	88
Figura 2-6: Modelo de estrella de Galbraith.....	90
Figura 2-7: Consecuencias de la desalineación del diseño organizativo.....	96
Figura 2-8: Mecanismos de coordinación en continuo.....	99
Figura 3-1: Relación entre Big Data y el proceso de creación del conocimiento.....	113
Figura 3-2: AS/RS de conocimiento.....	115
Figura 3-3: Transferencia de conocimiento en entornos con tecnologías de Big Data	117
Figura 3-4: Nuevo proceso de gestión del conocimiento tras la aplicación del Big Data en la organización.....	119
Figura 3-5: Modelo teórico de relaciones entre la implementación de ecosistemas de Big Data, procesos de gestión del conocimiento y mecanismos de coordinación y control	127
Figura 4-1: Procedimiento para estudio de casos múltiple	140
Figura 4-2: Herramienta en Microsoft Excel para la codificación de las entrevistas ...	153
Figura 4-3: Espiral de análisis del proceso de investigación cualitativa	154
Figura 4-4: Ejemplos de códigos deductivos.....	155
Figura 5-1: Modelo final de relaciones para el caso de Cementos Colombia S.A.	177

Figura 5-2: Modelo final de relaciones para el caso de Fundación Universidad del Norte	202
Figura 5-3: Modelo final de relaciones para el caso de Eco Fuels Energy S.A.S.	219
Figura 5-4: Comparativa preliminar de factores y relaciones planteadas en el modelo inicial y de los nuevos factores y relaciones que han emergido del análisis de los casos	235
Figura 5-5: Modelo inducido final de relaciones a partir del análisis conjunto de los casos	239

Introducción

La creación de valor en una organización está siendo cada vez más ligada al procesamiento de datos y a la creación de conocimiento, fenómeno que está presente desde la década de los 90 cuando se evidenció una transición de una economía dedicada a la transformación de materias primas a una economía orientada al flujo y procesamiento de información, siendo la razón fundamental para que las nuevas formas organizacionales respondan a esta necesidad (Child, 2015). Por tal motivo, los datos y la información dejaron de ser un activo estático almacenado en bases de datos que sólo se consultan cuando la organización requiere algún tipo de análisis o para tomar alguna decisión, para pasar a convertirse en un recurso vital para la organización y cuyo valor aumenta a medida que se incrementa su recolección y procesamiento (Glazer, 1993).

La importancia adquirida por los datos y la información ha llevado a una evolución en los sistemas y tecnologías de la información (SI y TI), lo que ha promovido cambios importantes en la forma en la que se gestionan las organizaciones, a tal punto que es considerada una revolución tan grande como la primera revolución industrial (Orlikowski & Barley, 2001). Los SI y las TI son recursos omnipresentes hoy en día en el ámbito de las organizaciones (Dewett & Jones, 2001), implementados con el objetivo de mejorar la eficiencia de la compañía (Hevner *et al.*, 2004). El valor de los SI y las TI en la organización dependerá de la utilidad de la información que ofrezcan (Hilton, 1981), utilidad determinada por las características de la organización, su sistema de trabajo, la cultura, los empleados y sus capacidades en el desarrollo e implementación de dichos sistemas y tecnologías (Hevner *et al.*, 2004).

En los últimos años, la recogida y el análisis de datos ha alcanzado niveles que no son manejables a través de los SI y las TI tradicionales, lo que ha dado lugar a lo que hoy conocemos como la era del Big Data. Este término se popularizó en 2011, cuando empresas tecnológicas especializadas presentaron soluciones computacionales a los problemas expuestos por las organizaciones en el manejo de grandes cantidades de datos (Gandomi & Haider, 2015).

El Big Data alcanzó su crecimiento exponencial en 2012, cuando fue un tema central en el Foro Económico Mundial y donde Estados Unidos anunció una inversión en investigación de Big Data por 200 millones de dólares americanos (Lake & Drake, 2014). Debido a la introducción de nuevos SI y TI para abordar los problemas de Big Data, éste se convirtió en uno de los principales contribuyentes para la toma de decisiones y la creación de conocimiento, ya que los datos de alta velocidad permiten analizar los datos y la información en tiempo real para la toma de decisiones y, simultáneamente, limitar el exceso de información que no tiene valor, proceso conocido como Big Data Analytics (Intezari & Gressel, 2017).

Por lo tanto, es importante abordar el fenómeno del Big Data contemplando el conjunto de sistemas, tecnologías, métodos, técnicas, herramientas, factores humanos y cambios organizativos que conlleva su uso en las organizaciones, conformando un conjunto de factores que hemos denominado en esta investigación como el “Ecosistema de Big Data”. Los ecosistemas de Big Data están llevando a las organizaciones a construir su estrategia basada en el conocimiento, lo que implica que se requiere una mejor comprensión de los cambios organizativos que suceden en función de este nuevo entorno (Kaivo-oja *et al.*, 2015). Dadas estas razones, los ecosistemas de Big Data han adquirido gran importancia en las organizaciones actuales con el objetivo de crear conocimiento, obtener valor y establecer ventajas competitivas (Fosso-Wamba *et al.*, 2015).

Sin embargo, y a pesar de su creciente importancia para las organizaciones de hoy en día, son pocas las investigaciones en el área de la dirección de empresas que abordan los cambios organizativos que genera la implementación de los ecosistemas de Big Data. Además, cuando se ha hecho se ha abordado principalmente desde el área de la gestión del conocimiento, dado que uno de los principales objetivos del Big Data es la resolución de problemas y la toma de mejores decisiones mediante un proceso sistemático que permita obtener conocimiento valioso para la organización (Carayannis *et al.*, 2017; Rothberg & Erickson, 2017; Sumbal *et al.*, 2017). De esta manera, resulta de gran interés para esta investigación analizar de qué forma la implementación de un ecosistema de Big Data afecta a los procesos de gestión de conocimiento de las organizaciones.

Por otro lado, la llegada de los ecosistemas de Big Data a las organizaciones ha requerido de y ha impulsado una modificación en los aspectos estructurales de la compañía con el

objetivo de abordar los retos técnicos y humanos que conlleva el flujo constante de datos para la toma de decisiones (Schildt, 2017). Con la introducción de los ecosistemas de Big Data, la estructura de la organización se verá afectada debido a la redistribución de las decisiones y la creación y transformación de funciones, puestos de trabajo, procesos y estrategias para lograr gestionar el ritmo acelerado que conlleva el uso del Big Data (Galbraith, 2012; Slinger & Morrison, 2014).

En este sentido, Galbraith (2014), uno de los académicos de referencia en el campo del diseño organizativo, sugiere que las empresas deben afrontar varios cambios en sus organizaciones con el fin adoptar una estrategia enfocada en aprovechar el Big Data, y adoptar estructuras y sistemas que soporten los nuevos procesos de toma de decisiones basados en grandes cantidades de información, así como conseguir el personal adecuado con las capacidades y habilidades digitales y multidisciplinarias requeridas para llevar a cabo dicha estrategia.

A colación de lo anterior, Galbraith (2012) establece que las capacidades adicionales requeridas para procesar la información del Big Data se pueden conseguir a través de dos vías fundamentales: a) una primera consistente en aumentar la capacidad en la jerarquía para procesar datos e información a través de mecanismos de coordinación centralizados; b) y una segunda basada en la adopción de mecanismos descentralizados de coordinación lateral. Adicional a esto, estas nuevas formas de organización implicarían mantener un control de resultados de la organización centralizado en el ecosistema de Big Data, en donde se evaluaría de forma automática el desempeño de la organización, así como el control de las actividades mediante equipos multidisciplinarios capaces de responder a las necesidades de la organización en tiempo real y la toma de decisiones efectivas (Gabel & Tokarski, 2014; Galbraith, 2014; Schildt, 2017).

Asimismo, mediante estos cambios estructurales se promovería una mayor visibilidad de la información y el conocimiento que se crea a partir del Big Data, impulsando de esta manera los procesos de gestión del conocimiento en todos los miembros de la organización (Berner *et al.*, 2014). En este sentido, los equipos multidisciplinarios encargados de los procesos de coordinación y control de las actividades que surgen a raíz del Big Data son considerados como "Centros de Excelencia del Conocimiento", ya que son los encargados de promover el uso de la analítica de datos y las tecnologías asociadas

al Big Data para la creación, el almacenamiento, la transferencia y la aplicación el conocimiento en toda la organización (Gabel & Tokarski, 2014). Basándonos en lo anterior, consideramos que existe una relación lógica y teórica entre la implementación de los ecosistemas de Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación y control vinculados a la estructura organizativa. Sin embargo, el estudio de esta relación ha sido poco abordado por parte de la literatura especializada lo que nos abre una gran avenida de investigación que pretendemos empezar a recorrer a través de la presente tesis doctoral.

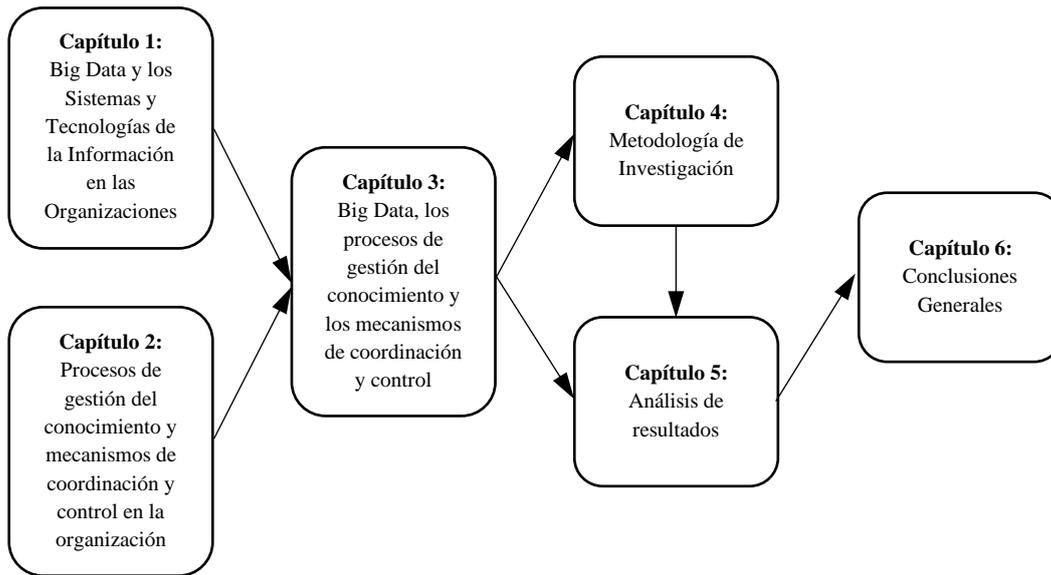
Partiendo de las consideraciones expuestas anteriormente, nos hemos planteado como objetivos de esta investigación analizar, en primer lugar, cómo las organizaciones llevan a cabo el proceso de gestión del conocimiento que surge de la recolección, el procesamiento y el análisis del Big Data y, en segundo lugar, abordar el estudio de dicho fenómeno desde la óptica del diseño organizativo, prestando una atención especial a cómo se ven afectados los mecanismos de coordinación y control del trabajo, las personas y los sistemas que surgen bajo el nuevo entorno creado por el ecosistema de Big Data. Estos objetivos quedarían expresados a través de la siguiente pregunta de investigación:

¿Cómo influye la implementación de ecosistemas de Big Data en los procesos de gestión del conocimiento y los mecanismos de coordinación y control del trabajo en las organizaciones?

Para dar respuesta a esta pregunta de investigación se ha llevado a cabo un estudio de casos exploratorio a través del cual se analizaron tres empresas de distintos sectores económicos de Colombia y que se encontraban en distintas etapas de la implementación de un ecosistema de Big Data.

Por su parte, con la intención de dar respuesta a la pregunta anterior la presente tesis doctoral ha sido estructurada en seis capítulos. La siguiente figura (ver figura 0-1) nos muestra la relación lógica entre los capítulos de esta investigación.

Figura 0-1: Estructura de capítulos de la tesis doctoral



Fuente: elaboración propia

A continuación, presentamos en detalle cómo se estructura la presente tesis doctoral. En el capítulo 1 comenzamos definiendo los conceptos clave de sistemas y tecnologías de información, su importancia en la organización y los diferentes tipos de sistemas de información existentes. Posteriormente, relacionamos dichos conceptos con el Big Data, explorando más a fondo los conceptos asociados con este fenómeno, su importancia y los diferentes métodos, técnicas, herramientas y tecnologías asociadas al análisis del Big Data. Por último, presentamos en este capítulo la importancia del Big Data en las organizaciones en distintos sectores económicos para concluir con una propuesta teórica sobre los niveles de implementación de los ecosistemas de Big Data, utilizando para ello todos los conceptos que se han revisado anteriormente a lo largo del capítulo.

El capítulo 2 ha sido dividido en dos grandes secciones para introducir los fundamentos teóricos relativos a los procesos de gestión del conocimiento y a los mecanismos de coordinación y control. En primer lugar, en este capítulo hemos presentado los conceptos de datos, información y conocimiento para dar un soporte teórico y poder definir la gestión del conocimiento y los procesos asociados a la misma. En un segundo apartado hemos presentado los conceptos clave de diseño organizativo necesarios que nos han facilitado la descripción de los mecanismos de coordinación y control del trabajo en el

ámbito organizativo, definiéndose los diferentes tipos de mecanismos tradicionales y modernos que se manejan en las organizaciones.

Con posterioridad, en el capítulo 3 relacionamos los conceptos de Big Data, procesos de gestión del conocimiento y mecanismos de coordinación y control, presentando relaciones planteadas por la literatura especializada junto a otras propias basadas en la experiencia y el conocimiento sobre el fenómeno que tiene el investigador. Así, al principio de este capítulo se plantea la relación entre el Big Data y cada uno de los cuatro procesos de gestión del conocimiento presentados en el capítulo 2, para luego realizar la comparación entre el Big Data y los mecanismos de coordinación y control. Por último, en este capítulo se plantea la relación teórica entre el Big Data (teniendo en cuenta las dimensiones para su implementación planteadas en el capítulo 1), los procesos de gestión del conocimiento y los mecanismos de coordinación y control.

En el capítulo 4 presentamos las decisiones metodológicas que se tomaron para llevar a cabo la investigación. En primer lugar, introducimos los conceptos generales de la investigación cualitativa y cuantitativa, lo que nos permite establecer y fundamentar que los objetivos de esta investigación se ajustan mejor a una metodología de corte cualitativo. En segundo lugar, seleccionamos dentro de las distintas estrategias de investigación cualitativa el estudio de casos como la más adecuada para llevar a cabo la investigación. En tercer lugar, presentamos el diseño del estudio de casos que hemos efectuado, definiendo la unidad de análisis de nuestra investigación, el número de casos a estudiar y los criterios de selección de los mismos; como resultado de lo anterior, presentamos las empresas que se han seleccionado para llevar a cabo la presente investigación. Posteriormente, en las últimas tres secciones definimos y describimos los métodos y técnicas utilizados para la recolección de los datos, cómo se llevará a cabo el análisis de dichos datos y presentamos los criterios que se han tomado en cuenta para asegurar la calidad y el rigor de la investigación realizada.

En el capítulo 5 presentamos los resultados obtenidos del análisis de la información primaria de los casos estudiados. Para ello seguimos una estructura homogénea a la hora de presentar dichos resultados. Así, en cada caso presentamos: a) el perfil de la organización; b) cómo se llevan a cabo los procesos de gestión del conocimiento; c) los mecanismos de coordinación y control utilizados; d) el nivel de implementación del

ecosistema de Big Data; e) y cómo esa aplicación del ecosistema de Big Data afecta a los procesos de gestión del conocimiento y a la forma de coordinación y control del trabajo en la organización, lo cual nos lleva a presentar un modelo inducido de relaciones entre estos grandes constructos y una serie de factores que contribuyen a explicar cómo estas relaciones tienen lugar. Luego de realizado el análisis individual de cada caso, en el último apartado de este capítulo 5 se realiza un análisis comparativo de los tres casos estudiados lo que nos lleva, en última instancia, a la elaboración de un modelo final inducido que nos permitirá dar respuesta a la pregunta de investigación de esta tesis doctoral.

Por último, en el capítulo 6 presentamos las conclusiones generales de la tesis doctoral, destacando los aportes o contribuciones que se han conseguido con la presente investigación, las implicaciones que la misma puede tener para el ámbito académico y para las organizaciones que aplican ecosistemas de Big Data, las limitaciones que presenta esta investigación y, por último, las líneas de desarrollo futuro que se pueden derivar a partir del trabajo realizado en esta tesis doctoral. Posteriormente, en otro capítulo detallamos todas las referencias bibliográficas que han dado el soporte necesario a todos los fundamentos teóricos utilizados en la presente tesis doctoral y adjuntamos los anexos complementarios de esta investigación.

Capítulo 1:

***Big Data y los Sistemas y Tecnologías de
la Información en las Organizaciones***

En el presente capítulo exploramos los conceptos, herramientas, técnicas y procesos correspondientes a los sistemas de información, las tecnologías de la información y todos los aspectos relacionados con el Big Data, Big Data Analytics y el Data Science, ejes fundamentales de esta investigación. Por lo tanto, comenzaremos describiendo los conceptos e importancia de los sistemas y tecnologías de la información; luego, estudiaremos el concepto, las técnicas, procesos y herramientas asociadas al Big Data; y para finalizar, estudiaremos la aplicación del Big Data en diferentes industrias para luego proponer una escala de madurez en la implementación de sistemas y tecnologías asociadas al Big Data.

1.1 Tecnologías y Sistemas de información en las organizaciones: conceptos y características

Los datos y la información han sido parte importante de todas las organizaciones. Es por este motivo que en esta sección comprenderemos los conceptos y valor de estos términos y qué implicaciones tienen para la organización su recolección, procesamiento y análisis. Estas definiciones introducen la necesidad de comprender los conceptos de los sistemas de información, ya que son los que permiten gestionar todas estas actividades del proceso de conversión de los datos en información y conocimiento, además de estudiar el concepto de las tecnologías de la información, y cómo estos dan soporte al diseño, construcción y administración de los sistemas de información en la organización.

1.1.1 Concepto y Valor de la Información en las Organizaciones

Los datos y la información son conceptos fundamentales que deben ser definidos claramente para poder comprender múltiples conceptos que se presentan más adelante en esta investigación, ya que son términos que en ocasiones son utilizados erróneamente como sinónimos. Autores como Cohen & Asín (2013, p. 2) establecen que esta confusión se debe a la relación directa que existe entre ambos conceptos, y diferencian a los términos definiendo los datos como “la materia prima para la producción de información” y la información como “datos que en un contexto dado tienen un significado para alguien”.

Con motivos de expandir la definición de ambos términos, determinamos que la más completa es la utilizada por Laudon & Laudon (2012, p. 15), los cuales establecen que los datos son “flujos de elementos en bruto que representan los eventos que ocurren en las organizaciones o en el entorno físico antes de ordenarlos e interpretarlos en una forma que las personas puedan comprender y usar”, y por información lo entienden como “los datos que se han modelado en una forma significativa y útil para los seres humanos”. La importancia de estas últimas definiciones de datos e información recae en la inclusión y énfasis en el procesamiento de los datos (ordenar, modelar e interpretar) para convertirse en información, proceso que convierte a la información en un recurso vital para las organizaciones debido a que la diferencia de otros recursos, ya que no es fácilmente divisible o apropiable, no es escasa e incrementa su valor a medida que aumenta su uso, y además, es un recurso renovable (Glazer, 1993).

Para comprender las razones por las que una organización procesa datos e información, tomamos el estudio de Daft & Lengel (1986), el cual establece dos principales motivos: a) el primero es reducir la incertidumbre, observación que está basada en el estudio de Galbraith (1974, p. 28), el cual afirma que “cuanto mayor sea la incertidumbre de una tarea, mayor es la cantidad de información que debe ser procesada entre los tomadores de decisiones... con el fin de alcanzar un determinado nivel de desempeño”; b) el segundo motivo es el de reducir la ambigüedad en la interpretación de la información, observación basada en el trabajo de Weick (1979). Por consiguiente, la información es uno de los componentes principales para la correcta toma de decisiones. No obstante, Child (2015, p. 29) agrega que la calidad de la toma de decisiones dependerá, en parte, de que se transmita la información necesaria a tiempo a la persona apropiada.

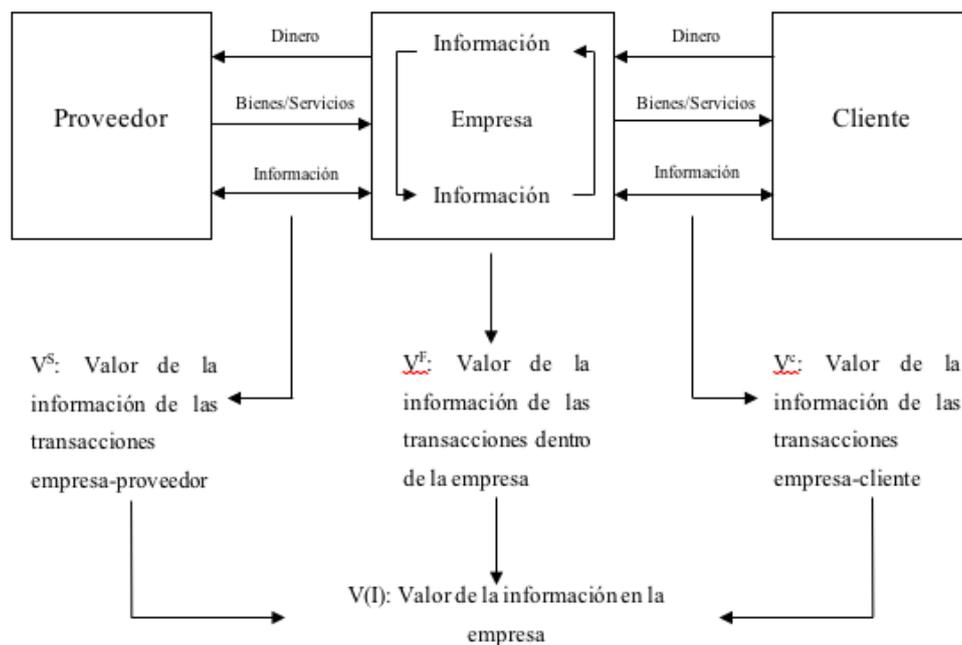
Para establecer claramente el valor que tiene la información para una organización, Glazer (1993) afirma que su valor radica en el rol que juega la información en las transacciones realizadas en el proceso de intercambio de bienes y servicios en toda la cadena de valor, incluyendo clientes, proveedores y a los procesos internos de la empresa. Teniendo en cuenta lo anterior, el autor (ibíd.) resume el valor de la información en tres medidas (V) de acuerdo al tipo de transacción:

- ✓ Vs: valor como resultado del intercambio de información en las transacciones realizadas entre la empresa y los proveedores.

- ✓ Vf: valor como resultado del intercambio de información en las transacciones realizadas en los procesos internos de la empresa.
- ✓ Vc: valor como resultado del intercambio de información en las transacciones realizadas entre la empresa y sus clientes.

La suma de estas tres medidas da como resultado el valor total de la información para una empresa $V(I)$. A continuación, se presenta gráficamente el modelo presentado por Glazer (1993):

Figura 1-1: Valor de la información en la empresa



Fuente: Glazer (1993, p. 103)

Como podemos notar en la gráfica, el valor de la información se encuentra repartido entre las diferentes transacciones o movimientos que se realizan a lo largo de la cadena de valor, y la información, al ser renovable, debe ser objeto de una constante evaluación y recirculación dentro de la organización para poder ser usada de forma correcta por los tomadores de decisiones y reducir la incertidumbre.

Por consiguiente, la creación de valor en una organización está cada vez más ligada al procesamiento de los flujos de información y del conocimiento (Child, 2015, p. 73), fenómeno que se viene presentando desde los años 90, época donde se hizo más evidente

la transición de una economía dedicada a la transformación de materias primas a una dedicada al procesamiento y flujo de información. Como notamos en la actualidad, el flujo de información no hace más que crecer día a día, y esta es la razón fundamental por la que las nuevas formas organizativas buscan responder a esta necesidad (Child, 2015, p. 62).

1.1.2 Concepto e Importancia de los Sistemas y Tecnologías de la Información

1.1.2.1 Concepto de sistemas y tecnologías de la información

De manera similar a como ocurre con los conceptos de datos e información, las tecnologías de la información y los sistemas de información son conceptos utilizados frecuentemente como sinónimos, y esto se debe a que están intrínsecamente relacionados por lo que es convencional referirse a ambos solamente como tecnologías de la información (TI) (Dewett, 2001). Sin embargo, y para efectos de esta investigación, es necesario establecer claramente la diferencia entre ambos ya que cada concepto presenta particularidades que lo diferencia del otro, aunque como veremos a continuación su relación es innegable.

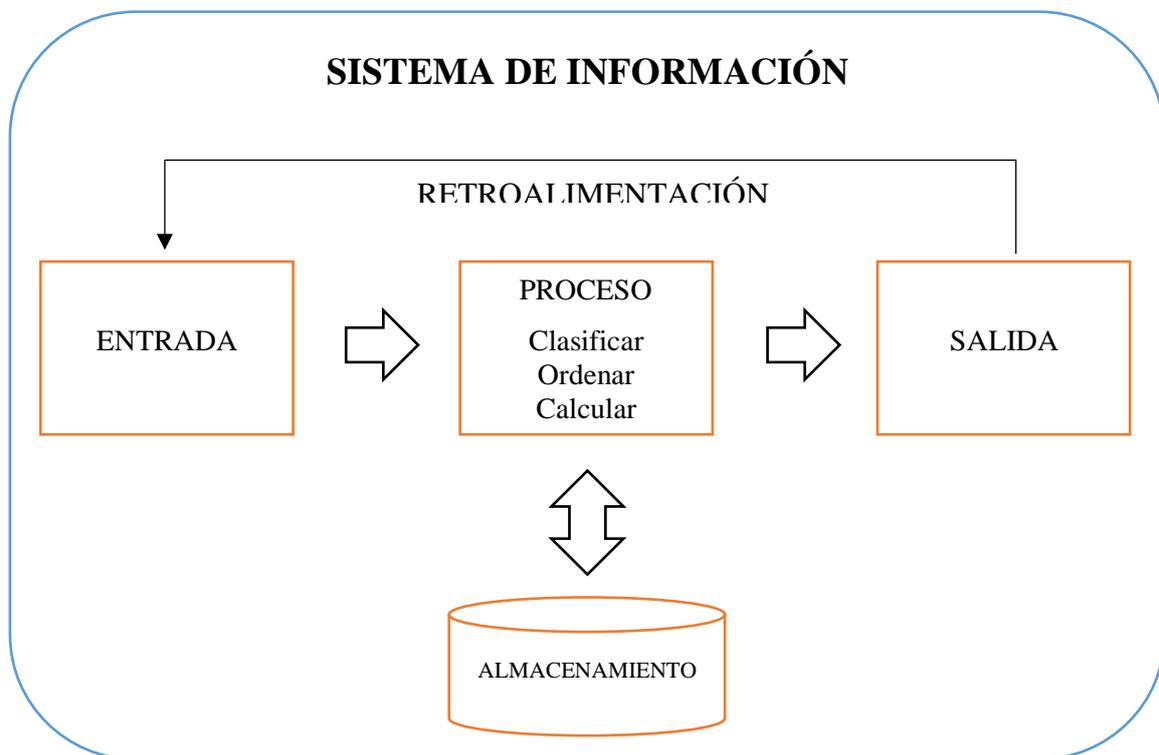
En primer lugar, estableceremos el concepto de sistema de información (SI), que Dewett & Jones (2001) corresponde a una gran variedad de plataformas de software y bases de datos que abarcan a toda la empresa, y están diseñados para administrar las funciones principales de la organización. Por otra parte, Laudon & Laudon (2012, p. 15) difieren de esta definición, ya que consideran al software como parte de las TI, y ven a los SI como “un conjunto de componentes interrelacionados que recolectan (o recuperan), procesan, almacenan y distribuyen información para apoyar los procesos de toma de decisiones y de control en una organización”. Además, mencionan que los SI sirven para que los gerentes y trabajadores del conocimiento analicen y solucionen problemas complejos, y ayude al proceso de creación de nuevos productos y servicios. Así mismo, dividen los SI en tres actividades principales para poder producir la información necesaria para ser usada por la organización (ibíd.):

- ✓ **Entrada:** captura y recolección de datos tanto del interior como del exterior de la organización.
- ✓ **Procesamiento:** convierte los datos a un formato significativo y entendible.

- ✓ **Salida:** transferencia de la información procesada a la persona o grupo que requiera de la misma.

Adicionalmente, los autores incluyen una etapa de retroalimentación, en el cual se devuelve la salida a las personas encargadas de la etapa de entrada con el objetivo de evaluar o corregir problemas que se estén presentando. Sin embargo, Cohen & Asín (2013, p. 6) agregan una etapa de almacenamiento a las actividades de un SI, actividad que permite guardar los datos de entrada y la información procesada por el SI. Tomando como base ambas aproximaciones, representamos a continuación las cinco actividades que forman parte de un SI:

Figura 1-2: Actividades de un Sistema de Información



Fuente: elaboración propia a partir de Cohen & Asín (2013) y Laudon & Laudon (2012)

En base a lo anterior, Laudon & Laudon (2012, p. 17) establecen que es necesario expandir el concepto describiendo tres dimensiones que se relacionan directamente con los SI (ver figura 1.3):

- ✓ **Organización:** corresponde a la relación de los SI con la estructura de la organización y tipo de jerarquía que maneja, sus procesos de negocio, su política y su cultura.
- ✓ **Administración:** corresponde a la relación de los SI con las personas encargadas de la toma de decisiones, utilizando el sistema para el desarrollo de nuevos productos y servicios, así como en el rediseño de la organización.
- ✓ **Tecnologías de la Información:** corresponde a la relación del SI con el hardware y software utilizado para las actividades de entrada, procesamiento y salida de información, así como todas aquellas tecnologías de comunicación entre las personas.

Figura 1-3: Dimensiones de un Sistema de Información



Fuente: Laudon & Laudon (2012, p. 18)

Por consiguiente, notamos que para los autores (ibíd.) las tecnologías de la información están directamente relacionadas con los SI, ya que proveen los medios a través de los cuales se pueden llevar a cabo todas las actividades de los SI. A continuación, ampliaremos el concepto de tecnologías de información (TI, de ahora en adelante).

Teniendo en cuenta lo anterior, las TI corresponden a todas aquellas tecnologías que constituyen la infraestructura física y de software que dan soporte al diseño, construcción y administración de los SI, y que permiten la comunicación entre el SI y las personas (D.

Cohen & Asín, 2013; Dewett, 2001; Laudon & Laudon, 2012). Para Byrd & Turner (2000), estas plataformas tecnológicas de hardware, software, redes y comunicaciones, son conocidas como “Infraestructura Técnica de TI”. No obstante, añaden otro tipo de infraestructura que corresponde a las habilidades organizacionales y humanas, experiencia, conocimiento, valores, normas y estructura organizacional, denominado “Infraestructura Humana de TI”. La relación entre ambas radica en que la efectividad en el desarrollo e implementación de la infraestructura técnica de TI dependerá de la calidad que tenga la organización en la Infraestructura Humana de TI (Byrd & Turner, 2000).

Por su parte, Orlikowski & Iacono (2001) realizan una revisión de literatura de 188 artículos de investigación sobre los diferentes enfoques y usos con los que se ha hecho referencia a las TI, agrupándolos en cinco categorías de acuerdo a con la perspectiva con la que se ve a la tecnología (ver Tabla 1-1).

Tabla 1-1: Enfoques conceptuales de las TI

<i>Enfoques conceptuales de las tecnologías de la información</i>	
Categoría	Enfoque
Tecnología vista como herramienta	Las TI son herramientas que buscan incrementar la productividad y el desempeño, mejoran cómo la organización procesa información y mejora las comunicaciones y relaciones sociales
Tecnología vista como indicador	El valor de las TI puede ser calculado por medio de medidas sustitutas, como percepción del usuario, penetración de las TI en la organización e inversión económica en TI
Tecnología vista como conjunto	Las TI tratan sobre las interacciones dinámicas entre las personas y la tecnología durante su construcción, implementación o uso, en las organizaciones
Tecnología vista como poder computacional	Las TI se analizan en base a su capacidad para manipular, guardar, recuperar y transmitir información
Tecnología vista como ausente	Las TI no son conceptualizadas, y es utilizado como antecedente o para referirse a otro concepto

Fuente: elaboración propia a partir de Orlikowski & Iacono (2001)

Como notamos anteriormente, las definiciones de SI y TI que se han presentado anteriormente, caen en la categoría de “Tecnología vista como herramienta”, debido a que es la más común en las investigaciones y la que muestra con mayor facilidad la

utilidad de estos sistemas. Así mismo, y como veremos en secciones posteriores, será en el enfoque que más se adapta a los objetivos de esta investigación. Dado este motivo, expandimos este enfoque en sus cuatro conceptualizaciones de la tecnología como herramienta en las organizaciones (Orlikowski & Iacono, 2001):

- ✓ **La tecnología como herramienta de sustitución del trabajo:** las nuevas tecnologías, como la automatización y la mecanización, son herramientas que llevarán a la mano de obra a ser reemplazada, conduciendo a la reducción de puestos de trabajo y a la reestructuración de la organización. Pero, a su vez, permite a las organizaciones realizar las operaciones de forma más barata y eficiente.
- ✓ **La tecnología como herramienta de productividad:** las tecnologías son herramientas productivas que permiten a las personas extender su alcance y beneficios en el desempeño de sus actividades socio-económicas, a través de la inclusión de nuevas formas de trabajar.
- ✓ **La tecnología como herramienta para procesar información:** el principal beneficio de las tecnologías es mejorar las capacidades de las personas y las organizaciones para procesar información, aumentando los flujos de información, la retroalimentación y el aprendizaje.
- ✓ **La tecnología como herramienta de relaciones sociales:** las nuevas tecnologías han llevado a un cambio en las relaciones sociales y en el comportamiento de las comunicaciones, y en una organización esto conlleva a que los roles sociales cambien, se alteren los sistemas jerárquicos, y se modifiquen los procesos de negocio.

De acuerdo con estas conceptualizaciones, podemos afirmar que las TI son herramientas que nos permitirán llevar a cabo actividades que normalmente las personas no podríamos hacer con tanta habilidad, rapidez y efectividad. Y, en conjunto con los SI, conllevan un cambio en las formas de trabajar, comunicarnos y gestionar, interna y externamente, los procesos de una organización.

Por lo tanto, podemos concluir que los SI y las TI son conceptos ligados por su propia naturaleza, debido a que las TI están integradas y forman parte importante del SI, por lo que es normal encontrarse en la literatura a ambos términos como uno solo (González,

1994). Teniendo en cuenta los conceptos, destacaremos a continuación la importancia que tienen los SI y las TI en las organizaciones y en la investigación.

1.1.2.2 Importancia de los SI y las TI en las organizaciones y la investigación

El avance de los sistemas y tecnologías de la información está promoviendo cambios importantes en la forma como nos comunicamos, trabajamos y vivimos, e inclusive es considerado que será algo tan grande como la revolución industrial (Orlikowski & Barley, 2001). En las organizaciones, los SI y las TI son recursos que hoy día están omnipresentes (Dewett, 2001), y son implementados con el objetivo de mejorar la eficiencia de la organización (Hevner et al., 2004). El valor que tengan los SI y las TI en la organización dependerá de la utilidad de la información que estos provean (Hilton, 1981), utilidad que viene determinada por las características de la organización, su sistema de trabajo, cultura, empleados y sus capacidades en el desarrollo e implementación de metodologías (Hevner et al., 2004).

Para Dewett & Jones (2001), las TI (utilizado por los autores como término conjunto entre SI y TI) promueven la eficiencia en muchos aspectos de la organización, entre los cuales destacamos:

- ✓ Habilidad para comunicar a los empleados de manera fácil sin importar la ubicación geográfica.
- ✓ Mejorar la toma de decisiones al permitir el almacenamiento, acceso y consulta de grandes volúmenes de información de forma rápida y económica.
- ✓ Reducir costos en el monitoreo del trabajo en equipo al agrupar los miembros en espacios de trabajo electrónicos.
- ✓ Reducir los costos de procesamiento de la información entre actores, permitiendo además una estructura más eficiente.

Por consiguiente, la eficiencia que promueve el uso de los SI y las TI llevará a la organización a un incremento de la productividad laboral. Adicionalmente, Dewett & Jones (2001) destacan que uno de los beneficios más importantes del uso de los SI y las TI radica en su habilidad para enlazar a los empleados y darles autoridad para la toma de decisiones, suplantando en las organizaciones muchos de los roles de control y coordinación de la jerarquía, creando oportunidades para que las nuevas formas de

organización se enfoquen en procesos en vez de funciones (Zammuto et. al., 2007). Por ende, las organizaciones que invierten en SI y TI y adoptan estas nuevas formas que promueven su uso, tienen un efecto positivo en el desempeño y las capacidades de innovación y conocimiento (Tsoukatos et. al. , 2017), a los cambios disruptivos de los mercados, a la globalización de las empresas y a los cambios demográficos que pueda sufrir un sector (Dewett, 2001).

La investigación no ha sido ajena al fenómeno de los SI y las TI. Desde el año 1958, con el libro de Joan Woodward titulado “Management and Technology”, los investigadores se han interesado en crear teorías que expliquen cómo la tecnología se relaciona con la estructura organizativa (Orlikowski & Barley, 2001). Los autores establecen que la investigación se ha centrado principalmente en resolver qué factores organizativos que afectan el uso de estas tecnologías, y cómo éstas moldean a las organizaciones. Sin embargo, existe una diferencia principal entre la investigación en temas de SI y TI y la investigación organizativa, ya que la primera tiene mayor interés en explicar, desarrollar y proponer soluciones prácticas al rol que juega estas tecnologías en la vida cotidiana, mientras que la segunda tiene una mayor aproximación hacia la ciencia tradicional, y ven las tecnologías como un tema especializado, menos central que temas como la toma de decisiones, dinámica de grupos y la estructura organizativa, donde se desarrollan y prueban explicaciones a diferentes clases de fenómenos, principalmente sobre el comportamiento humano a nivel individual y grupal en las organizaciones.

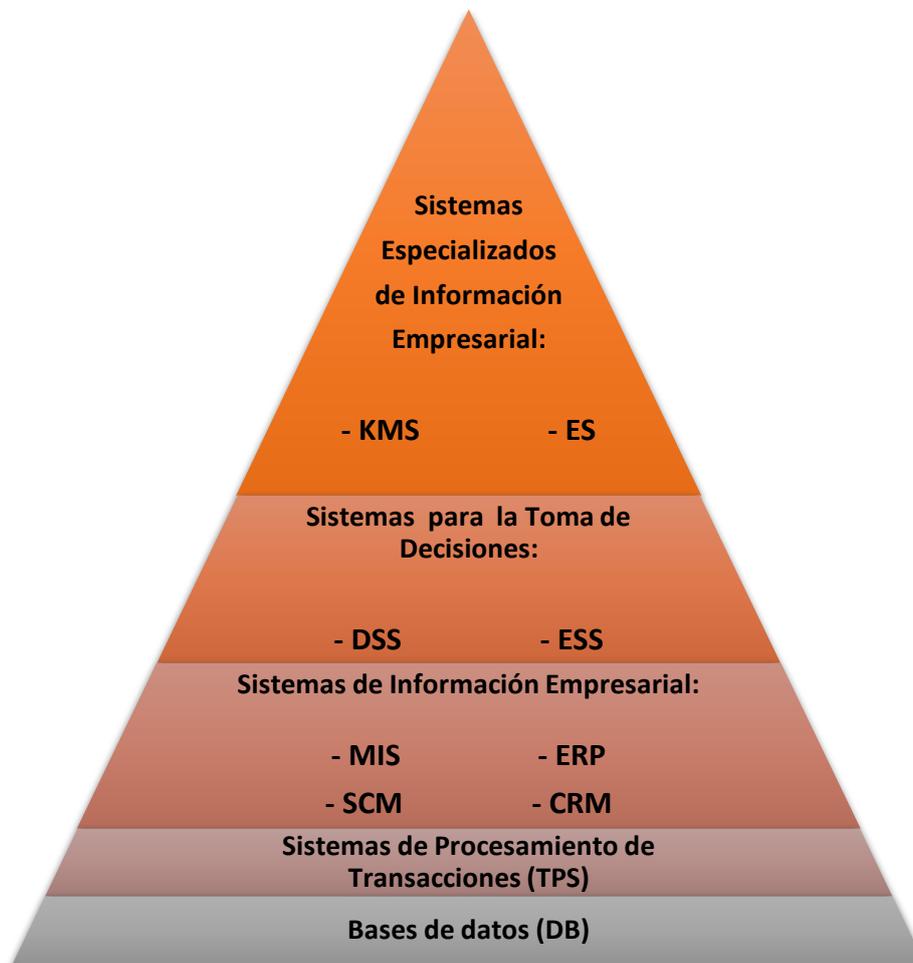
Como veremos en secciones posteriores, la investigación en temas de SI y TI sigue vigente. El desarrollo constante de innovaciones en este campo ha llevado a este tema a un alto interés por parte de investigadores en diversas ramas de la ciencia, y seguirá siendo una tendencia en el desarrollo de nuevas teorías en el ámbito organizativo. Establecido los conceptos e importancia, procederemos a determinar los diferentes tipos de SI y establecer el de mayor interés para esta investigación.

1.1.3 Tipos de Sistemas de Información

Establecidos los conceptos asociados a los SI y las TI como parte fundamental de estos, es importante destacar los diferentes tipos de SI que se manejan en las organizaciones. Basándonos en los planteamientos de Cohen y Asín (2013), González (1994), Lapiedra y

Devece (2012), Laudon y Laudon (2011) y Stair y Reynolds (2010) sobre los diferentes tipos de SI que pueden tener las organizaciones, hemos elaborado la siguiente categorización para comprender la utilidad que tienen los SI más importantes y sus aplicaciones a nivel organizacional.

Figura 1-4: Categorización de los Sistemas de Información



Fuente: elaboración propia

1.1.3.1 Bases de Datos

Tal y como se observa en la figura anterior, las DB (acrónimo en inglés de “Data Bases”) son el nivel más básico que puede tener una organización en términos de SI, y corresponde a una forma organizada de registrar, almacenar y consultar la información que posee y requiere la organización. Las DB pueden manejarse en diferentes niveles de la organización, ya sea a nivel individual, departamental o en toda la empresa, conocida como base de datos corporativa. Adicionalmente, las organizaciones tienen la posibilidad de consultar bases de datos externas con información de interés provistas por entidades

gubernamentales o privadas, como información tecnológica, económica, legislativa y social (González, 1994).

1.1.3.2 Sistemas de Procesamiento de Transacciones

Los TPS (acrónimo en inglés de “Transaction Processing Systems”) fueron una de las primeras aplicaciones informáticas desarrolladas para las organizaciones, y permitieron organizar a personas, procedimientos, software y bases de datos (Stair & Reynolds, 2010) con el fin de recolectar, procesar y almacenar las transacciones que se realizan diariamente en una organización (Lapiedra & Devece, 2012). Los TPS permiten a los gerentes supervisar las operaciones internas y externas de la empresa, como el pago a empleados, cobro a clientes, pago a proveedores, flujo de materiales, etc., además de ser una fuente de información para otros sistemas y funciones de la organización (Laudon & Laudon, 2011).

1.1.3.3 Sistemas de Información Empresarial

Los TPS han venido evolucionando y por muchos años han sido una solución para el manejo de las transacciones en las organizaciones. Sin embargo, el uso de estos sistemas ha venido disminuyendo debido al creciente uso de sistemas más avanzados y que proporcionan una mayor cantidad de soluciones (Stair & Reynolds, 2010). Estos sistemas son los siguientes:

- ✓ **Sistemas de Información Gerencial:** los MIS (acrónimo en inglés de “Management Information Systems”) son sistemas que sintetizan datos e información provista por los TPS con el fin de generar reportes a la alta gerencia y de nivel medio sobre el desempeño en diferentes áreas de la organización, como marketing, finanzas, producción y demás áreas funcionales (Laudon & Laudon, 2011; Stair & Reynolds, 2010). Adicionalmente, los reportes pueden incluir una comparación de los resultados obtenidos con estándares preestablecidos en el sistema o con periodos anteriores, generando una alerta en caso de que se presente una disminución en el desempeño (Lapiedra & Devece, 2012).
- ✓ **Sistema de Planificación de Recursos:** los ERP (acrónimo en inglés de “Enterprise Resource Planning”) son sistemas que integran varios procesos de negocio de la organización en un solo paquete de software, almacenando toda la

información en una base de datos central que permite el acceso a distintas áreas de la organización (Laudon & Laudon, 2011; Stair & Reynolds, 2010). Los ERP funcionan creando transacciones de manera automática entre las diferentes áreas que necesiten procesar un pedido o requerimiento de un cliente, así como generar requerimientos de material o información entre proveedores internos y externos de la organización. La principal ventaja de los ERP consiste en que mejoran los procesos y actividades de la organización y facilitan el acceso oportuno de la información para la posterior toma de decisiones (Stair & Reynolds, 2010).

- ✓ **Sistemas de Administración de la Cadena de Suministros:** Los SCM (acrónimo en inglés de “Supply Chain Management”) son sistemas interorganizacionales (Laudon & Laudon, 2011) que ayudan a la organización a administrar las relaciones con sus proveedores y clientes para el manejo de los bienes y servicios, ya que permiten el flujo de información automáticamente, determinando los suministros que se necesitan a lo largo de toda la cadena de valor (Stair & Reynolds, 2010). Las principales ventajas de los SCM radican en la disminución de costos operacionales en transporte y fabricación, y permiten a los gerentes mejorar la toma de decisiones para organizar y programar el suministro, la producción y la distribución de sus productos (Laudon & Laudon, 2011).

- ✓ **Sistemas de administración de relaciones con el cliente:** los CRM (acrónimo en inglés de “Customer Relationship Management”) son sistemas estratégicos de información que manejan todos los procesos relacionados con el cliente, entre los que se encuentran el marketing, publicidad, ventas, servicio al cliente y programas de fidelidad (D. Cohen & Asín, 2013; Stair & Reynolds, 2010). Los CRM buscan aprender sobre las necesidades y comportamientos del mercado con el fin de generar una satisfacción, fidelización y confianza de los clientes con la organización, mejorando la rentabilidad y las ventas de sus productos (D. Cohen & Asín, 2013).

1.1.3.4 Sistemas para la Toma de Decisiones

Desde la década de 1980, los SI sufrieron una evolución que les permitió ser más poderosos, económicos y accesibles a todas las personas de la organización, llegando así a apoyar en actividades mucho más importantes como la toma de decisiones (Stair &

Reynolds, 2010). A continuación, presentamos los dos tipos de sistemas más importantes que surgieron con esta funcionalidad.

- ✓ **Sistemas de Apoyo para la Toma de Decisiones:** Los DSS (acrónimo en inglés de “Decision Support Systems”) son sistemas interactivos que permiten al usuario expandir sus capacidades para resolver problemas no estructurados o semi-estructurados a la hora de tomar decisiones (Lapiedra & Devece, 2012). Los DSS se utilizan en casos donde es requerida una gran capacidad de procesamiento de información (Stair & Reynolds, 2010), ya que se manejan diversas fuentes internas (como de los sistemas MIS, TPS y bases de datos) y externas (como información de los mercados) (Laudon & Laudon, 2011), para luego ser analizados por medio de modelos diseñados para proponer soluciones, y acompañar el proceso de toma de decisiones en conjunto con el criterio de las personas encargadas (Lapiedra & Devece, 2012). Los gerentes y tomadores de decisiones de una organización juegan un papel muy importante en el desarrollo e implementación de los DSS, debido a que son sistemas que se adaptan al estilo de administración y necesidades particulares de información que requieren estas personas (Stair & Reynolds, 2010).

- ✓ **Sistemas de Apoyo a Ejecutivos:** los ESS (acrónimo en inglés de “Executive Support Systems”) son sistemas que permiten a los altos mandos de una organización tomar decisiones en temas de planeación estratégica, organización del personal, control estratégico y gestión de crisis (Stair & Reynolds, 2010). Los ESS utilizan fuentes de información similar a los DSS. Sin embargo, estos sistemas llevan a cabo un mayor análisis de la información con mayor importancia para la organización con el fin de presentar tendencias, pronósticos y reportes con un alto nivel de detalle y simplicidad en su manejo (Laudon & Laudon, 2011).

1.1.3.5 Sistemas Especializados de Información Empresarial

Las innovaciones en el campo de los SI no se han detenido por décadas, y no lo hará en el futuro próximo. Es por este motivo que la evolución ha llevado a los SI a ser capaces de procesar cantidades de información antes inimaginables, llegando a ser hoy una realidad los sistemas provistos de su propia inteligencia y capaces de proponer soluciones que emulen los niveles de entendimiento de una persona. Hoy en día las organizaciones,

además de los otros tipos de sistemas, confían en los sistemas especializados para encontrar soluciones a problemas en la creciente complejidad y cantidad de la información, siendo los más comunes los presentados a continuación.

- ✓ **Sistemas de Gestión del Conocimiento:** los KMS (acrónimos en inglés de “Knowledge Management Systems”) “son sistemas que organizan a las personas, procedimientos, software, bases de datos y dispositivos para crear, almacenar, compartir y usar el conocimiento y la experiencia de la organización” (Stair & Reynolds, 2010, p. 24). Los DSS y los ESS pueden ser considerados como parte de un KMS (Jashapara, 2004), debido a que son sistemas que recolectan conocimiento y experiencia de la organización para la toma de decisiones (Laudon & Laudon, 2011). El principal objetivo de un KMS se centra en la mejora de los procesos de gestión de la calidad asociados a la capacidad de una organización de capturar, compartir y aplicar nuevo conocimiento, y su éxito dependerá de la interconexión entre el sistema y la estrategia de la organización, enfocado siempre en el cliente final (Jashapara, 2004).

- ✓ **Sistemas Expertos:** los ES (acrónimo en inglés de “Expert Systems”) son sistemas inteligentes capaces de capturar, almacenar y usar el conocimiento de los empleados expertos en un área por medio de un conjunto de reglas, con el fin de utilizar este conocimiento para proponer sugerencias para la toma de decisiones (Laudon & Laudon, 2011). Estos sistemas se encuentran intrínsecamente ligados a los KMS, debido a su naturaleza de trabajar con el conocimiento. Sin embargo, los ES proveen soluciones que se asemejan a las de un experto humano y, en conjunto con otras tecnologías, se han convertido en sistemas capaces de monitorear en tiempo real los procesos de la organización, permitiendo una toma de decisiones más ágil por parte de los gerentes (Stair & Reynolds, 2010).

Como hemos observado anteriormente, la innovación a lo largo de varias décadas en temas de SI ha llevado al desarrollo de tecnologías que permiten una mayor capacidad para la toma de decisiones. Es por este motivo que en la siguiente sección exploramos los nuevos SI y TI que han sido desarrollados hoy en día, enfocados en sacar provecho a la creciente oferta y demanda de información que requieren las organizaciones para satisfacer las, cada vez más exigentes, necesidades de los clientes.

1.2 Big Data y su relación con los sistemas y tecnologías de la información

En las organizaciones, los datos y la información han dejado de ser activos estáticos almacenados en bases de datos que solo se consultan cuando se requiere algún tipo de análisis o para tomar decisiones. La evolución del manejo de los datos y la información ha llevado a modificar la manera en cómo se toman las decisiones basadas en datos, dado el incremento en la robustez y complejidad con la que se generan los datos en esta era gobernada por las redes sociales, dispositivos inteligentes, automatización, asistentes virtuales y demás tecnologías que han cambiado el paradigma del cómo nos comunicamos, trabajamos y vivimos, fenómeno que hoy conocemos como Big Data.

Para afrontar estos cambios, las organizaciones requieren de una comprensión del impacto que tiene el Big Data en sus procesos, cómo afectan la toma de decisiones y qué métodos y técnicas han surgido para enfrentar los cambios que conlleva el Big Data. Por ello, a continuación, presentamos los conceptos relacionados con el Big Data, los métodos, técnicas y herramientas del análisis de datos para el Big Data, la relación del Big Data los SI y las TI, y su aporte en lo que hoy conocemos como la cuarta revolución industrial, esto es, la industria 4.0.

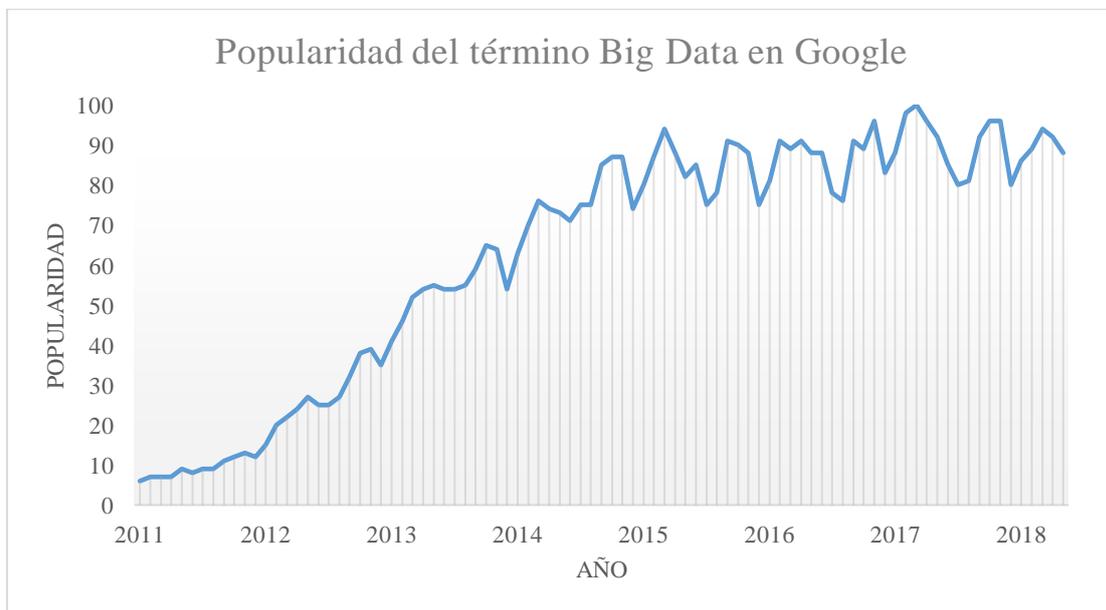
1.2.1 Concepto de Big Data: las 10 V's

El “Big Data” es un término que hoy día se encuentra omnipresente en diversas ramas de la ciencia y las organizaciones. No obstante, su origen no se conoce con certeza (Gandomi & Haider, 2015; Mishra et. al., 2017). Diebold (2012), considerado uno de los primeros en mencionar el término en el año 2000 (Press, 2013), encontró que el término pudo tener su origen a mediados de los 90 en una conversación en la empresa “Silicon Graphics”, e incluso el término hizo presencia en una publicidad de la empresa en la revista “Black Enterprise” en marzo de 1996. Por su parte, Press (2013) en su artículo para la revista Forbes afirma que la primera vez que una publicación menciona la palabra Big Data es en octubre de 1997 con el trabajo de Cox & Ellsworth (1997,p. 235) en donde se establece que “la visualización presenta retos importantes para los sistemas computacionales: los

conjuntos de datos son generalmente bastante grandes, agotando la capacidad de la memoria principal, de discos locales, e incluso de discos remotos. A este problema lo llamamos Big Data”.

Aunque el término tiene ya dos décadas, no se hizo popular hasta el año 2011, debido a que en ese año empresas como IBM presentaban soluciones computacionales a los problemas que se venían manifestando en el manejo de grandes cantidades de datos (Gandomi & Haider, 2015). Sin embargo, su verdadero despegue lo consigue a partir del año 2012 donde fue un tema central en el Foro Económico Mundial y en donde Estados Unidos anunciaba una inversión en investigación sobre Big Data por \$200 millones de dólares (Lake & Drake, 2014). Como podemos notar en las tendencias de Google (ver figura 1.5), el término adquirió un interés exponencial desde entonces y ha llegado, inclusive, al nivel máximo de popularidad a nivel mundial (Google, 2018).

Figura 1-5: Popularidad mundial del término Big Data de 2011 a 2018



Fuente: elaboración propia a partir de Google (2018)

A pesar de lo anterior, aún no se tiene un consenso universal sobre la definición del término Big Data, e incluso ha acarreado confusión su verdadero significado (Frizzo-Barker et. al., 2016; Gandomi & Haider, 2015; Lake & Drake, 2014; Ward & Barker, 2013). No obstante, la literatura especializada (Diebold, 2012; Frizzo-Barker et al., 2016; Gandomi & Haider, 2015; Günther et. al., 2017) concuerdan que la base del concepto de

Big Data recae en el trabajo de Laney (2001) y, aunque no es un artículo sobre Big Data, se definen las tres dimensiones para afrontar los retos que presenta la gestión de datos, conocidas hoy en día como los tres principales atributos (las 3 V's: volumen, velocidad y variedad) del Big Data:

- ✓ **Volumen:** se refiere a la creciente magnitud de los datos generados o recolectados por las transacciones de una organización, las personas o a nivel mundial (Gandomi & Haider, 2015; Laney, 2001), teniendo un incremento exponencial desde el año 2000 a raíz del despliegue mundial del internet (Bumblauskas *et al.*, 2017; Russom, 2011). En la actualidad, casos como el de la empresa Wal-Mart son un ejemplo de la cantidad de datos que se pueden generar, llegando a los 2.5 petabytes por hora (2500 terabytes) como resultado de las transacciones con sus clientes. En el año 2012 se estimó que a nivel mundial se generan 2.5 exabytes (2500 petabytes) de datos cada día, y se prevé que esta cantidad se duplicará cada 40 meses (McAfee & Brynjolfsson, 2012).
- ✓ **Velocidad:** se refiere al incremento en la frecuencia con la que los datos son generados o recolectados (Katal *et al.*, 2013; Laney, 2001; I. Lee, 2017; Russom, 2011). La creciente popularidad de los smartphones, dispositivos y sensores conectados a internet, son la principal razón de este fenómeno, y ha requerido de nuevos sistemas capaces de recolectar y analizar en tiempo real este aumento en la velocidad en la generación de datos (Gandomi & Haider, 2015; I. Lee, 2017; Russom, 2011). Además, se estima que para el año 2020 la cantidad de dispositivos conectados a internet llegue a 13.5 billones, tres veces más que en el año 2016 (Gartner Inc., 2015).
- ✓ **Variedad:** se refiere a la heterogeneidad (o falta de la misma) de un conjunto de datos y a la creciente diversidad de fuentes de las que se puede conseguir (Gandomi & Haider, 2015; Russom, 2011). La variedad es uno de los principales problemas de la gestión de datos debido a la incompatibilidad de formatos, estructuras de datos no alienados e inconsistencia en la semántica de los datos que se presentan en los grandes volúmenes de información (Laney, 2001).

Los datos se pueden dividir en tres tipos de acuerdo a su nivel de estructuración (Gandomi & Haider, 2015; I. Lee, 2017; Russom, 2011):

- ✓ Estructurados: son datos que poseen una estructura determinada y son los más fáciles de comprender y analizar. Por ejemplo, los datos tabulados en una hoja de cálculo o en una base de datos. Sin embargo, son las estructuras menos frecuentes en las que podemos encontrar los datos, llegando solamente al 5% de todos los datos.

- ✓ Semi-estructurados: son datos que poseen ciertos estándares para su organización, pero que necesitan de ciertas reglas para poder ser leídos e interpretados. El ejemplo más común de datos semi-estructurados es el lenguaje XML, el cuál utiliza un conjunto de etiquetas para los datos para poder ser procesados y compartidos a través de la web.

- ✓ No estructurados: son las formas más frecuentes en las que se encuentran los datos, debido a que la mayoría de las fuentes proveen los datos en forma de texto, imágenes, video o audio, las cuales deben ser procesadas para poder ser interpretadas. Este tipo de datos son generados con mayor rapidez que los otros tipos. Sin embargo, poseen una falta de estructura que impide su procesamiento directo, y requieren ser organizados antes de ser leídos e interpretados.

A raíz de los atributos antes descritos, las definiciones más comunes y reconocidas sobre el Big Data se han centrado en estos tres principales atributos para intentar establecer cuál es el significado del término. La literatura (Gandomi & Haider, 2015; Lake & Drake, 2014; Mishra et al., 2017) destaca dos de estas definiciones:

- ✓ “Big Data es alto volumen, alta velocidad y/o alta variedad de los activos de información, los cuales demandan formas innovadoras y económicas de procesar información y permiten mejorar el conocimiento, la toma de decisiones y los procesos de automatización” (Gartner Inc., s. f.).

- ✓ “Big Data es un término que describe los grandes volúmenes de datos de alta velocidad, complejos y variables que requieren técnicas y tecnologías avanzadas para permitir la captura, almacenamiento, distribución, administración y análisis

de la información” (TechAmerica Foundation: Federal Big Data Commission, 2012, p. 10).

En base a los atributos anteriores, empresas importantes del sector del Big Data incluyeron atributos adicionales para describir fenómenos que ocurrían con los grandes volúmenes de datos y que no se describían dentro de los tres primeros atributos (Gandomi & Haider, 2015; Kshetri, 2014; I. Lee, 2017; Mishra et al., 2017). Los atributos incluidos son los siguientes:

- ✓ **Veracidad:** introducido por la empresa IBM, se refiere a los problemas de calidad o integridad de los datos provenientes de las diversas fuentes, conocido como ruido en los datos, lo cual genera una falta de confianza e inconvenientes en su procesamiento (Zikopoulos et al., 2013).

- ✓ **Variabilidad (y complejidad):** introducido por la empresa SAS, se refiere a la inconsistencia en la velocidad de los flujos de datos, desencadenado por eventos puntuales temporales que afectan la cantidad de datos recolectados. SAS incluye además el atributo complejidad, el cual se refiere a la dificultad que conlleva conectar y correlacionar los datos provenientes de las diversas fuentes (SAS Inc., 2013). Sin embargo, este atributo no ha contado con popularidad en la literatura, y autores como Russom (2011) tratan el problema de la diversidad de fuentes en el atributo variedad.

- ✓ **Valor:** introducido por la empresa Oracle, se refiere a la variación en el valor económico que adquieren los datos. El reto de los datos está en extraer la información valiosa dentro de la gran densidad de datos con poco o sin valor, logrado a través de las técnicas de extracción, transformación y análisis de los datos (Dijcks, 2013).

Como consecuencia de la introducción de estos atributos a la teoría de Big Data, nuevas definiciones surgieron con el fin de tener una mejor descripción del significado del término. Entre ellas destacamos las siguientes:

- ✓ “Big Data consiste en extensos conjuntos de datos – principalmente en las características de volumen, variedad, velocidad y/o variabilidad – que requiere una arquitectura escalable para el almacenamiento, manipulación y análisis eficiente” (NIST Big Data Public Working Group, 2015, p. 5).
- ✓ “Big Data es un enfoque holístico para gestionar, procesar y analizar las 5 Vs (es decir, volumen, variedad, velocidad, veracidad y valor) con el fin de crear conocimiento accionable para obtener valor sostenible, medir el rendimiento y establecer ventajas competitivas” (Fosso Wamba *et al.*, 2015, p. 235).

No obstante, la literatura en el tema de Big Data no se ha detenido en la intención de establecer más atributos que expliquen el significado de Big Data y, según Firican (2017), son necesarios para afrontar los fenómenos y desafíos que traen las iniciativas de Big Data. Los cuatro atributos adicionales que se describen son (Firican, 2017):

- ✓ **Validez:** se refiere a qué tan precisos y correctos son los datos para el fin que se piensan utilizar. A pesar de que es similar a veracidad, se diferencia en que la validez se enfoca en los datos en sí más no en la fiabilidad de las fuentes.
- ✓ **Vulnerabilidad:** se refiere a los problemas de seguridad que conlleva el almacenamiento de grandes cantidades de datos, ya que una violación de seguridad en Big Data conlleva a una gran filtración de datos.
- ✓ **Volatilidad:** se refiere a la disminución en la duración del ciclo de vida de los datos, ya que, debido al volumen y la velocidad, se necesitan reglas que determinen la circulación y disponibilidad de los datos, y que se ajusten a las necesidades del negocio y costos asociados a los procesos de recolección y almacenamiento de Big Data.
- ✓ **Visualización:** se refiere a los problemas que conlleva la representación de datos en Big Data debido a su volumen, variedad y velocidad. Las técnicas tradicionales, como los gráficos o las tablas, fallan al momento de representar la realidad de millones de datos, por lo que técnicas como el clustering o diagramas circulares de red son más idóneos para representar las relaciones complejas de los datos.

Teniendo en cuenta lo expresado anteriormente, hoy en día se siguen construyendo nuevas definiciones e inclusive nuevos atributos para el término Big Data. Sin embargo, no se ha llegado a un consenso en su significado o qué atributos son los que mejor describen el término. Así, a efectos de esta investigación se han seleccionado los atributos de mayor importancia en la literatura y aquellos atributos que expresen mejor la esencia del Big Data (ver figura 1-6).

Figura 1-6: Las 10 Vs (Atributos) del Big Data



Fuente: elaboración propia

Por lo tanto, y teniendo en cuenta las definiciones y atributos anteriores, podemos concluir que el Big Data es un término asociado al crecimiento en la cantidad y complejidad de los datos obtenidos mediante una gran diversidad de fuentes, cuyos atributos conllevan a replantear la manera en cómo se recolectan, almacenan, procesan y gestionan los datos en una organización, con el fin de adquirir un mayor conocimiento, mejorar el rendimiento y lograr ventajas competitivas.

Teniendo claro el concepto de Big Data, a continuación, desarrollaremos la descripción de las técnicas o herramientas encargadas de realizar el análisis del Big Data, estudiando los diferentes tipos de análisis que se pueden llevar a cabo y el proceso por el cuál se puede obtener el valor de la información.

1.2.2 Concepto, Métodos, Técnicas y Herramientas de Big Data Analytics

El Big Data, por sí solo, no conlleva una ventaja competitiva para las empresas. Su verdadero valor y potencial está en el análisis de los datos que lleva a tomar mejores decisiones (Gandomi & Haider, 2015). A este análisis de datos se le conoce en la actualidad como “Big Data Analytics” (BDA), el cual es una técnica avanzada para la captura, almacenamiento, transferencia, intercambio, búsqueda, análisis y visualización de grandes volúmenes de datos que se transmiten a una gran velocidad con el objetivo de obtener información útil para la organización (Duan & Xiong, 2015; Gandomi & Haider, 2015; Russom, 2011). No obstante, el término solo es la evolución de las diferentes formas con las que se ha llamado a través de los años al uso y análisis de los datos, por lo que es común encontrar que en la literatura se utilizan otros términos como “Data Analytics”, “Business Intelligence”, o “Advanced Analytics” (H. Chen *et al.*, 2012; Davenport, 2014; Larson & Chang, 2016; Russom, 2011; Tien, 2013). La siguiente tabla nos muestra, de manera sintética, la evolución en la terminología utilizada para el uso y análisis de datos (ver tabla 1-2 en la página siguiente).

Tabla 1-2: Evolución de la terminología para el uso y análisis de datos

Término	Periodo de Tiempo	Significado
Soporte de Decisiones	1970 – 1985	Usar el análisis de datos como soporte para la toma de decisiones
Soporte de Decisiones Ejecutivas	1980 – 1990	Enfoque en el análisis de datos para la toma de decisiones a nivel ejecutivo
Procesamiento Analítico en Línea (OLAP)	1990 – 2000	Software para analizar tablas de datos multidimensionales
Inteligencia de Negocios (Business Intelligence)	1989 – 2005	Herramientas para apoyar la toma de decisiones basada en datos, con énfasis en la presentación de informes
Analytics	2005 – 2010	Enfoque en el análisis matemático y estadístico para tomar decisiones
Big Data (Big Data Analytics)	2010 – Presente	Enfoque en datos de gran volumen, no estructurados y de alta velocidad

Fuente: Davenport (2014, p. 10)

La diferencia del BDA con respecto a las técnicas tradicionales de análisis de datos por métodos estadísticos y matemáticos (técnicas de Analytics) reside en que se requiere de sistemas de información capaces de manejar los atributos del Big Data para llevar a cabo un análisis de datos. Sin embargo, las técnicas estadísticas y matemáticas tradicionales siguen siendo parte fundamental de las herramientas de BDA (Tien, 2013). En este sentido, Russom (2011) establece que la unión del Big Data con las técnicas y herramientas de Analytics ha dado como resultado muchas ventajas para el análisis de datos en las organizaciones, entre los que destacamos:

- ✓ Big Data proporciona muestras estadísticas gigantes, que mejoran los resultados de las herramientas analíticas.
- ✓ Las herramientas y bases de datos analíticas ahora pueden manejar Big Data.
- ✓ Las herramientas de Analytics son ahora más accesibles que nunca.

- ✓ Las herramientas y técnicas modernas para análisis avanzados y de Big Data son muy tolerantes con los datos sin procesar, no estructurados y de baja calidad.
- ✓ Big data y Analytics juntos son un activo especial que merece ser aprovechado.
- ✓ Los análisis de datos basados en muestras de gran tamaño revelan y sacan provecho del cambio organizativo.

Por lo tanto, el BDA tiene un gran potencial de beneficios, donde las organizaciones pueden extraer ideas y conocimiento de su negocio. Por ejemplo, en la encuesta realizada por Russom (2011) a 325 personas entre profesionales de TI, patrocinadores de negocios y consultores en el área, destacaron los siguientes beneficios:

- ✓ El BDA beneficia a todo lo que involucre la relación con el cliente: mejoras en el marketing orientado, mejorar la segmentación, reconocimiento de oportunidades en ventas y mercados, y mejor análisis del comportamiento de los clientes.
- ✓ La inteligencia de negocios se beneficia del BDA: la información empresarial es más numerosa y precisa, mejora la comprensión del cambio organizativo, una mejor planificación y previsión, y la identificación de las causas raíz en los costos.
- ✓ Las aplicaciones específicas de analítica se benefician del BDA: mejoras en la detección de fraudes, cuantificación de riesgos, análisis de la tendencia en los mercados y automatización de las decisiones en tiempo real.

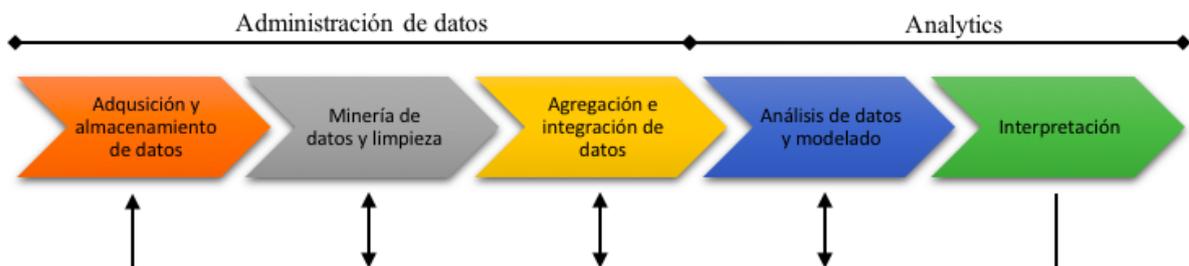
No obstante, para lograr obtener los beneficios que trae el uso del BDA en la organización y poder conseguir la información útil necesaria para la toma de decisiones del Big Data, se tiene que llevar a cabo el análisis de datos por medio de un proceso constituido por cinco etapas, dividido a su vez en dos sub-procesos: administración de datos y Analytics (Gandomi & Haider, 2015; Jagadish, 2015). Estas cinco etapas son las siguientes (Gandomi & Haider, 2015; Sivarajah *et al.*,2017):

- ✓ **Administración de datos:**
 - ✓ Etapa 1 – Adquisición y almacenamiento de datos: consiste en la recolección y almacenamiento de datos provenientes de las diversas fuentes con el objetivo de obtener valor en los pasos posteriores.

- ✓ Etapa 2 – Minería de datos y limpieza: consiste en la extracción y limpieza de los datos útiles dentro de una gran cantidad de datos no estructurados recolectados.
- ✓ Etapa 3 – Agregación e integración de datos: consiste en la agrupación y unión de los datos limpios con características similares, estableciendo una estructura.
- ✓ **Analytics:**
 - ✓ Etapa 4 – Análisis de datos y modelado: consiste en la elaboración de modelos que permitan describir el comportamiento de los datos, con el fin de elaborar un análisis del que se extraiga la información valiosa.
 - ✓ Etapa 5 – Interpretación: consiste en la visualización de los resultados obtenidos de la etapa de análisis y modelado, con el fin de presentar la información de una manera que sea entendible por los tomadores de decisiones y lograr extraer conocimiento.

La siguiente figura muestra gráficamente cada una de las cinco etapas antes mencionadas (ver figura 1-7).

Figura 1-7: Etapas del proceso de Big Data Analytics



Fuente: elaboración propia a partir de Gandomi & Haider (2015) y Jagadish (2015)

Como podemos notar en la figura anterior, todas las etapas son susceptibles de presentar ciclos repetitivos (Jagadish, 2015), debido a que se pueden detectar errores con los datos dentro de la etapa o de errores en etapas anteriores, lo que afectaría a los resultados obtenidos en la etapa de interpretación y acarrear problemas para la toma de decisiones. Otro motivo para presentarse ciclos repetitivos en el proceso de BDA consiste en la constante reevaluación de los métodos y técnicas utilizadas en cada etapa, ya que los

atributos del Big Data obligan a replantear constantemente la manera en cómo se extrae el conocimiento y el valor de la información resultante.

El proceso de BDA descrito previamente muestra una imagen general de cómo debe llevarse a cabo la transformación de los datos en conocimiento. Sin embargo, existen diferentes métodos para lograr este objetivo, dependiendo del nivel de detalle al que se quiera llegar con el análisis de datos. Sivarajah *et al.* (2017), en un análisis de 115 artículos que discutían los métodos de BDA, identificaron y clasificaron estos métodos analíticos en tres grupos principales:

- ✓ **Analítica descriptiva:** es la forma más simple de los métodos de BDA, y lleva a cabo un análisis de datos recolectados en el pasado para obtener una descripción de patrones y relaciones por medio de métodos estadísticos simples como la media, moda, frecuencia y varianza. A pesar de que es considerado como un método para analizar datos históricos, actualmente es común utilizar este método para analizar los resultados obtenidos del método de analítica predictiva como, por ejemplo, los pronósticos de ingresos futuros. El método de analítica descriptiva sigue siendo el más común de los métodos de BDA, ya que permite descubrir patrones o correlaciones que no se habían identificado antes, siendo de mucha utilidad para la toma de decisiones (Duan & Xiong, 2015; Sivarajah *et al.*, 2017).
- ✓ **Analítica predictiva:** es un método de BDA que incluye varias técnicas y modelos estadísticos para pronosticar futuros resultados basado en datos históricos y/o actuales. Este método puede clasificarse en dos grupos: técnicas de regresión (como la regresión logística multinomial) y técnicas de aprendizaje automático (como las redes neuronales). Sin embargo, los resultados de otras técnicas más simples como la media móvil o la regresión lineal, usualmente utilizados para identificar patrones o correlaciones, pueden ser extrapolados para obtener resultados futuros (Duan & Xiong, 2015; Gandomi & Haider, 2015).
- ✓ **Analítica prescriptiva:** es un método de BDA que busca determinar las relaciones causa-efecto entre los resultados obtenidos de los métodos de analítica predictiva, con el fin de optimizar los modelos de procesos de negocio. El método

de analítica prescriptiva es el menos común de todos, ya que se requieren técnicas complejas de simulación y optimización que aún se encuentran en desarrollo. No obstante, es una de las soluciones más completas para la toma de decisiones empresariales, determinando acciones y evaluando su impacto con respecto a los objetivos, requisitos y limitaciones de la organización, permitiendo una evolución continua en los modelos de procesos de negocio (Duan & Xiong, 2015; Sivarajah et al., 2017).

Junto a los métodos anteriores, los autores H. Chen *et al.* (2012) y (Gandomi & Haider, 2015) exponen técnicas de BDA que permiten llevar a cabo un análisis de datos en base al tipo de datos recolectados o la fuente de donde provienen, y cumplen la función de extraer información específica que permita detectar patrones, obtener relaciones y apoyar la toma de decisiones. Estas técnicas son:

- ✓ **Analítica de textos:** técnicas que extraen y analizan información de datos no estructurados que se encuentran en formato de texto. Entre las técnicas de analítica de textos destacamos: técnica de extracción de información, técnica del resumen de información, técnica de respuesta a preguntas y la minería de opiniones.
- ✓ **Analítica de audio:** técnicas que extraen y analizan información de datos no estructurados en formato de audio. La analítica de audio utiliza dos enfoques para el análisis de los datos: los sistemas basados en transcripción y los sistemas basados en fonética.
- ✓ **Analítica de video:** técnicas que extraen y analizan información de datos no estructurados en formato de video. Su principal aplicación se encuentra en los sistemas automatizados de vigilancia y seguridad, analizando los datos (videos) provenientes de los circuitos cerrados de televisión.
- ✓ **Analítica web:** técnicas que extraen y analizan información de datos estructurados y no estructurados de los sitios web con el objetivo de medir su tráfico e impacto y optimizar el uso de la página. La analítica web tiene sus bases en las técnicas de minería de datos, análisis estadístico y en otros tipos de analíticas como la de textos.

- ✓ **Analítica de redes sociales (opiniones):** técnicas que extraen y analizan información de datos estructurados y no estructurados (imágenes, videos, reacciones, etc.) generados por los contenidos y relaciones entre personas, organizaciones y productos en los diferentes canales de redes sociales. La analítica de redes sociales puede dividirse en dos grupos (analítica basada en contenidos y la analítica basada en estructura), y utilizar una serie de técnicas diferentes para extraer información, entre los que destacamos: detección de comunidades, análisis de influencia social y predicción de links.

Como consecuencia de las técnicas y métodos descritos, muchas herramientas de BDA han surgido para suplir las necesidades organizativas para sacar provecho, valor y conocimiento del Big Data, ya que las infraestructuras computacionales tradicionales no pueden manejar los atributos del Big Data (Fan *et al.*, 2014). Dos de las herramientas de BDA que más se han destacado, y comúnmente utilizadas en conjunto, son el Hadoop y el MapReduce (Chen & Zhang, 2014; Frizzo-Barker *et al.*, 2016; Khan *et al.*, 2014), y en conjunto con otros componentes de BDA, constituyen una distribución de código abierto conocido como el ecosistema Hadoop (Fan *et al.*, 2014; Khan *et al.*, 2014).

El sistema Hadoop es un software creado en el año 2006 por la fundación Apache, y esta basado en el lenguaje de programación Java para la gestión y procesamiento de grandes cantidades de datos de una manera escalable a través de su propio sistema distribuido de archivos (HDFS por sus siglas en inglés) (Fan *et al.*, 2014; Frizzo-Barker *et al.*, 2016; Khan *et al.*, 2014). Como mencionamos anteriormente, el sistema Hadoop viene acompañado por el MapReduce, el cual es un modelo de programación desarrollado por empresas como Google y Yahoo! para el procesamiento en paralelo de grandes volúmenes de datos. El Hadoop/MapReduce forman parte de la clase de herramientas de procesamiento por lotes (Batch processing), el cual funciona fraccionando un problema complejo en pequeños sub-problemas hasta que sean escalables para poder ser resueltos (Chen & Zhang, 2014). Los demás componentes que acompañan el ecosistema Hadoop aportan funcionalidades necesarias para todo el procesamiento de los datos, los cuales describimos a continuación (Apache Software Foundation, 2014):

Tabla 1-3: Componentes adicionales del ecosistema Hadoop

Componente	Funcionalidad	Componente	Funcionalidad
HDFS	Sistema de archivos distribuido para el acceso a los datos a una alta tasa de transferencia	Hive	Infraestructura de almacenamiento de datos que proporciona un resumen de datos y consultas ad hoc
YARN	Programación de trabajos y gestión de recursos de clústeres	Mahout	Librería escalable de aprendizaje automático y minería de datos
Ambari	Aprovisionamiento, administración y supervisión de clústeres de Hadoop	Pig	Lenguaje de flujo de datos de alto nivel y de ejecución para computación paralela
Avro	Sistema de serialización de datos	Spark	Motor de cálculo rápido y general para datos de Hadoop
Cassandra	Base de datos multimaestra escalable	Tez	Programación de flujo de datos para procesar datos tanto por lotes como interactivos
Chukwa	Sistema de recolección de datos para administrar grandes sistemas distribuidos	Zookeeper	Servicio de coordinación de alto rendimiento para aplicaciones distribuidas
HBase	Base de datos escalable y distribuida que admite el almacenamiento de datos estructurados para tablas grandes.		

Fuente: Apache Software Foundation (2014)

La herramienta de la empresa Apache es una de las más populares dada su naturaleza de código abierto, ya que ha permitido a los desarrolladores adaptar mejor estas aplicaciones a las necesidades particulares de las organizaciones y convertirse en una herramienta de mucha utilidad para manejar Big Data.

Sin embargo, H. Wang *et al.* (2016) consideran que las herramientas basadas en procesamiento por lotes hacen parte de la primera generación de las herramientas de Big Data, y exponen que este tipo de herramientas tienen un enfoque prioritariamente hacia el volumen, procesando únicamente datos que se encuentran ya almacenados en bases de

datos. Esta primera generación, a pesar de ser herramientas multipropósito, no es capaz de llevar a cabo un procesamiento de datos en tiempo real ni de alto desempeño, características necesarias en aplicaciones actuales como los sensores, comunicación máquina a máquina y la telemática (Chen & Zhang, 2014), por lo que se considera que esta primera generación ya ha concluido (H. Wang, Xu, Fujita, et al., 2016).

Tras la primera generación, surge la necesidad de herramientas capaces de procesar datos con alto volumen, alta velocidad y la complejidad de los diferentes tipos de datos del Big Data, por lo que aparecen las herramientas de procesamiento de flujo (Stream Processing) o de procesamiento en tiempo real (Real-time Processing) (H. Wang, Xu, Fujita, et al., 2016). Esta clase de herramientas lleva a cabo un análisis de datos continuo mediante el procesamiento de una secuencia infinita de pequeños lotes de datos, haciendo que no se procesen grandes cantidades de datos al mismo tiempo, logrando así una latencia de respuesta muy baja (Chen & Zhang, 2014).

En los últimos años, han surgido herramientas que permiten a los usuarios llevar a cabo su propio análisis de los datos por medio de un ambiente interactivo, conocidos como análisis interactivo (Interactive Analysis) o procesamiento híbrido (Hybrid Processing) (Chen & Zhang, 2014). Esta clase de herramientas funciona mediante la conexión directa del usuario con el computador para revisar, comparar y analizar los datos en formato tabular, gráfico o ambos al mismo tiempo, utilizando en conjunto los principios del procesamiento por lotes y del procesamiento de flujo, basado en la arquitectura lambda (C. L. Chen & Zhang, 2014; H. Wang, Xu, Fujita, et al., 2016). La arquitectura lambda consiste en tres capas de procesamiento: la capa de procesamiento por lotes administra el conjunto de datos que se ha almacenado en un sistema distribuido y no es modificable, la capa de servicio visualiza y expone las vistas de la capa de procesamiento por lotes en un almacén de datos para su consulta, y la capa de velocidad trata con los datos con una baja latencia (H. Wang, Xu, Fujita, et al., 2016). La siguiente tabla nos muestra la evolución en las clases de herramientas de Big Data Analytics (ver tabla 1-4 en la página siguiente).

Tabla 1-4: Clases de herramientas de Big Data Analytics

Clase (Generación)	Principales Características	Herramientas
Primera Generación Batch Processing (Procesamiento por lotes) (2006 – 2010)	✓ Grandes cantidades de datos estáticos ✓ Soluciones escalables ✓ Volumen	MapReduce Hadoop (HDFS, Hive, HBase) Flume Scribe Dryad Mahout Jaspersoft BI Suite Pentaho Skytree Server Cascading Spark Tableau Karmasphere Pig Sqoop Kafka
Segunda Generación Stream Processing (Procesamiento en tiempo real) (2010 – 2014)	✓ Flujo de datos en tiempo real ✓ Baja latencia ✓ Velocidad	Kafka Flume Kestrel Strom S4 SQL stream Splunk SAP Hana Spark Streaming
Tercera Generación Hybrid Processing (Análisis Interactivo) (2014 – Presente)	✓ Arquitectura Lambda ✓ Volumen + Velocidad	Lambdooop SummingBird Google Dremel Apache Drill

Fuente: H. Wang *et al.* (2016)

No obstante, la mejor herramienta de BDA es aquella que se acople a las necesidades, recursos y capacidades de la organización, ya que cada una presenta soluciones útiles para obtener valor, conocimiento y ventajas competitivas del Big Data en las organizaciones. Una vez expuestos los conceptos, métodos, técnicas y herramientas de Big Data y BDA, analizaremos la relación de estos conceptos con la teoría de los SI, además de explorar las TI que están llevando al Big Data a ser un fenómeno creciente en las organizaciones.

1.2.3 Big Data y los Sistemas de Información para la Toma de Decisiones y el Conocimiento

Los datos han sido utilizados desde la década de 1950 para apoyar la toma de decisiones (Intezari & Gressel, 2017) y, como hemos observado en secciones anteriores, en la literatura de Big Data y de BDA encontramos que la mención a la toma de decisiones y al conocimiento es algo recurrente. El principal motivo para esta relación es que el Big Data, al ser una evolución del término para referirse al uso y análisis de datos (Davenport, 2014), hereda todos los principios y beneficios que conlleva el análisis de datos para la toma de decisiones y la creación de conocimiento de la era de la inteligencia de negocios y el Analytics (H. Chen et al., 2012).

La literatura especializada (Intezari & Gressel, 2017) establece que el aporte del Big Data y los métodos y herramientas de BDA a la toma de decisiones es innegable, y es debido a que la generación de datos a alta velocidad permite incorporar datos e información en tiempo real a la toma de decisiones. Además, explica que la incorporación del Big Data en las decisiones estratégicas requiere una colaboración entre los que formulan las estrategias de la organización y aquellas personas encargadas de realizar los análisis de datos de Big Data. Por otra parte, Kaivo-oja *et al.* (2015) consideran que la verdadera contribución del Big Data para la toma de decisiones radica en su capacidad para proporcionar la mejor información posible y, simultáneamente, limitar el exceso de información que no tiene valor y, adicionalmente, permitir construir la estrategia de la organización basado en el conocimiento, llevando a comprender mejor los cambios organizativos en base a su entorno.

Sin embargo, existen consideraciones que afectan a la calidad de la toma de decisiones dependiendo de cómo las organizaciones adoptan y utilizan el Big Data. Por lo tanto, Janssen et al. (2017) exponen que los principales factores que afectan a la calidad de la toma de decisiones basados en Big Data son:

- ✓ La calidad y características de las fuentes de Big Data.
- ✓ La calidad y capacidad de los procesos de BDA.
- ✓ Las capacidades de las personas involucradas en el proceso de recolección y procesamiento del Big Data.

- ✓ La disponibilidad de una infraestructura de Big Data.
- ✓ La habilidad de los tomadores de decisiones para comprender los datos y la relación de las variables de un problema.
- ✓ La habilidad de los tomadores de decisiones para colaborar con otras personas involucradas en la cadena de Big Data.

El aporte del Big Data a la toma de decisiones en las organizaciones ha visto un crecimiento en importancia, de tal manera que en una encuesta realizada entre la empresa consultora Capgemini y The Economist Intelligence Unit, en el año 2012, a 607 ejecutivos de empresas, se puso en manifiesto que para el momento el Big Data era utilizado para tomar decisiones el 58% del tiempo, y que en un 29% del tiempo era utilizado para tomar decisiones automatizadas. Adicionalmente, la encuesta arrojó resultados prometedores del Big Data y la toma de decisiones, entre los cuales se destacaba que el 40% de los encuestados consideraba que las decisiones tomadas en los tres últimos años hubieran sido significativamente mejor si se hubieran tenido la cantidad de datos estructurados y no estructurados que posee el Big Data. Además, se espera que la mejora en la toma de decisiones incremente el desempeño de la organización en un 41%, y el 85% de los encuestados consideró que el problema del Big Data no está en el volumen, sino en poder analizar los datos y tomar las decisiones en tiempo real (The Economist Intelligence Unit, 2012).

Por lo tanto, el Big Data puede ser considerado una de las herramientas más importantes para llevar a cabo la toma de decisiones basada en datos, y como han establecido Davenport *et al.* (2012, p. 46), “a medida que el Big Data vaya evolucionando se convertirá en un ecosistema de información de servicios internos y externos que comparten información continuamente, optimiza las decisiones, comunica resultados y genera nuevo conocimiento para los negocios”. Por consiguiente, el incremento de los datos de alto volumen, variedad y velocidad disponibles en tiempo real tendrá un mayor impacto a futuro en la toma de decisiones (Power, 2014).

El conocimiento también ha sido directamente relacionado con el Big Data, mencionado en secciones anteriores por la literatura relevante en el tema (Davenport *et al.*, 2012; Fosso-Wamba *et al.*, 2015; Gartner Inc., n.d.) como un objetivo del Big Data, de tal manera que autores como Rothberg & Erickson (2017) consideran que la conexión entre

ambos es clara, principalmente en términos del intercambio y transferencia de intangibles, y que los sistemas diseñados para ambos tienen como prioridad la comunicación de datos, información o conocimiento crítico para los procesos de monitoreo u operaciones.

Dadas las relaciones existentes entre Big data, la toma de decisiones y el conocimiento, es natural relacionar al Big Data con los SI para la toma de decisiones y los sistemas de gestión del conocimiento (Elgendy & Elragal, 2016; Intezari & Gressel, 2017; Rothberg & Erickson, 2017; Vera-Baquero *et al.*, 2015). En el caso de los DSS, Vera-Baquero *et al.* (2015) establecen que estos sistemas son valiosos en la transformación de los datos en información y la información en conocimiento para el apoyo a la toma de decisiones. Sin embargo, con las implicaciones que han traído el Big Data, el monitoreo y análisis de los datos en tiempo real y con baja latencia, se ha convertido en algo clave para la toma de decisiones, por lo que los DSS tradicionales no son suficientes. Dada esta necesidad, autores como Elgendy & Elragal (2016) y Vera-Baquero *et al.* (2015) han propuesto metodologías y marcos de referencia que permiten manejar los atributos del Big Data, mediante la inclusión de las técnicas, métodos y herramientas de BDA en los DSS tradicionales.

En el caso de los sistemas de gestión del conocimiento, Intezari & Gressel (2017) consideran que la relación del Big Data con los KMS es beneficiosa, y establecen que las organizaciones que integran el Big Data a sus repositorios de conocimiento generan un mayor valor a sus ya existentes sistemas de gestión del conocimiento, a lo que ellos denominan como “Sistemas Avanzados de Gestión del Conocimiento”. Estos sistemas avanzados de gestión del conocimiento permiten a la organización manejar la velocidad con la que se recopilan, agregan y comparten los datos y el conocimiento, y facilitan la incorporación del Big Data y el conocimiento en las organizaciones para la toma de decisiones estratégicas, por lo que van más allá de los sistemas de gestión de conocimiento tradicionales (Intezari & Gressel, 2017).

En esta misma línea, Rothberg y Erickson (2017) consideran que los sistemas de gestión del conocimiento aportan a los sistemas de Big Data una guía para buscar identificar, transferir y mostrar indicadores clave de rendimiento, y establecen que aquellas organizaciones que requieran un uso de Big Data solo para el monitoreo de procesos,

transacciones y el sentimiento de los clientes, pueden utilizar un sistema de gestión del conocimiento con estructuras tradicionales.

Una vez establecida la relación entre Big Data y los SI para la toma de decisiones y la gestión del conocimiento, pasamos en el siguiente apartado a enumerar aquellas TI que han favorecido el crecimiento del Big Data en las organizaciones, en especial, aquellas que han adquirido un gran interés debido a su influencia en la nueva era de la transformación digital de las organizaciones.

1.2.4 Big Data y las Tecnologías Claves de la Transformación Digital

Las organizaciones actuales se encuentran viviendo un cambio en las formas de cómo crean competencias y capacidades a raíz de la llegada de la cuarta revolución industrial, la industria 4.0, el cual corresponde a una era dominada por los sistemas ciberfísicos de manufactura y la innovación en los servicios ofrecidos por una organización (J. Lee *et al.*, 2014). La integración de diferentes TI es una capacidad importante en las organizaciones para lograr una transformación digital, capacidades necesarias para crear valor para los clientes (Del Giudice, 2016), lo cual se refleja en una encuesta realizada por Gartner Inc. (2017) a 388 ejecutivos, donde el 42% de los encuestados han comenzado con la transformación digital de sus negocios, y las TI se han convertido en la segunda prioridad para las organizaciones, llegándose a considerar que ha nacido una nueva era en la construcción de habilidades y capacidades tecnológicas.

No obstante, la nueva era de la transformación digital de las organizaciones se encuentra aún en proceso. Sin embargo, la aparición y/o crecimiento de las tecnologías que hacen parte de la industria 4.0, como la Inteligencia Artificial, el Big Data, la Realidad Virtual o Aumentada, la Computación en la Nube y el Internet de las Cosas, demuestran que estamos en una revolución digital (Roblek *et al.*, 2016; S. Wang *et al.*, 2016), de tal forma que la empresa International Data Corporation estima que el gasto en estas tecnologías de transformación digital alcanzaría los \$1.2 trillones de dólares en 2017, y pronostica que para el 2020 esta cifra alcanzará los \$2 trillones de dólares (International Data Corporation, 2017).

Teniendo en cuenta los argumentos anteriores, a continuación, presentamos las tecnologías de la información que han adquirido mayor importancia e interés tanto para las organizaciones como en la investigación en transformación digital e industria 4.0. Estas TI son la Computación en la Nube, la Inteligencia Artificial, el Internet de las Cosas y la Realidad Virtual o Aumentada (S. Wang *et al.*, 2016), así como su relación con el Big Data y el BDA.

1.2.4.1 Computación en la nube

En inglés “Cloud Computing” (CC), son recursos computacionales, de redes y almacenamiento ofrecidos como servicios a través de Internet, con el objetivo de disminuir la carga operativa y financiera de una organización que requiera un sistema de información con una capacidad superior a la instalada (Rehman *et al.*, 2016). Los servicios de CC brindan ventajas competitivas a las organizaciones al permitir una mayor agilidad en el procesamiento y almacenamiento computacional con un costo bajo demanda, logrando así nuevas oportunidades para los negocios (Tien, 2015). Existe una relación fuerte entre el CC y el Big Data (Hashem *et al.*, 2015), ya que los servicios de CC proveen una mayor flexibilidad y capacidad para la gestión, procesamiento y análisis del Big Data (Al-fuqaha *et al.*, 2015; J. Lee *et al.*, 2014), lo que ha llevado al CC a un incremento en su uso para gestionar, procesar y almacenar el Big Data (Rehman *et al.*, 2016).

La infraestructura que poseen los proveedores de CC es la más adecuada para el análisis del Big Data, ya que provee una capacidad de procesamiento que no es equiparable al de un sistema “in-house”, ya que utilizan tecnologías de almacenamiento y procesamiento distribuido en paralelo para poder ejecutar los algoritmos y modelos de análisis de los grandes volúmenes de datos (Hashem *et al.*, 2015). A pesar de lo anterior, Rehman *et al.* (2016) consideran que la utilización de servicios de CC sigue siendo un reto para las pequeñas y medianas empresas debido al costo que esta conlleva.

1.2.4.2 Inteligencia Artificial

En inglés “Artificial Intelligence” (AI), son sistemas computacionales y máquinas que demuestran características de inteligencia, entre las que se destacan el aprender y aplicar conocimiento de la experiencia, comprender, procesar y manipular información, manejar situaciones complejas, solucionar problemas a pesar de tener información faltante,

establecer prioridades, reaccionar rápida y correctamente frente a una nueva situación y utilizar heurísticas (Stair & Reynolds, 2010). Las tecnologías de AI son utilizadas en procesos que requieren de una capacidad de procesamiento y habilidades superiores a las humanas (Loebbecke & Picot, 2015), o en el caso de tareas que consuman mucho tiempo por medio de técnicas tradicionales, en especial en aquellas en las que se requiera una mayor capacidad y habilidad para recolectar información de múltiples fuentes externas, realizar cálculos complejos de forma rápida y precisa, y procesar y transferir información entre máquinas y personas (Stair & Reynolds, 2010). Algunas técnicas de AI son ampliamente utilizados para llevar a cabo BDA, como es el caso del aprendizaje automático (en inglés “Machine Learning”) que es utilizado para reconocer patrones y tomar decisiones inteligentes basadas en datos, y las redes neuronales (en inglés “Neural Networks”) utilizadas para el reconocimiento de patrones en los datos y la optimización (Manyika et al., 2011; O’Leary, 2013; Rehman et al., 2016). Los avances que se logran en el campo de la AI afectan positivamente el desarrollo de los métodos y técnicas de BDA (Loebbecke & Picot, 2015) y viceversa, lo que permite que ambos campos beneficien a las organizaciones en las capacidades para la obtención de valor del Big Data.

1.2.4.3 Internet de las Cosas

En inglés “Internet of Things” (IoT), también conocido como el Internet de Todo o el Internet Industrial, se refiere a la interconexión y comunicación entre máquinas y dispositivos capaces de interactuar entre sí a través de Internet (Dijkman *et al.*, 2015; Lee & Lee, 2015). Además, se relaciona al IoT con el término “Smart”, debido a que las aplicaciones son tan numerosas que van desde aplicaciones industriales, en nuevos servicios que surgen de esta tecnología, los electrodomésticos en los hogares y las llamadas “ciudades inteligentes” (Kaivo-oja et al., 2015).

Una de las principales aplicaciones del IoT para las organizaciones es en el campo del Big Data y BDA (Lee & Lee, 2015), llegándose a considerar como tecnologías interdependientes (M. Chen *et al.*, 2014), ya que el IoT genera grandes cantidades de datos e información en tiempo real proveniente de todos los dispositivos y máquinas conectadas que hacen parte del ecosistema IoT (dentro y fuera de la organización), con el fin de transmitirlos a sistemas o herramientas capaces almacenar, analizar, procesar y tomar decisiones (Del Giudice, 2016; Lee & Lee, 2015). Como consecuencia del gran

incremento que aporta el IoT, principalmente, en volumen, velocidad y variedad del Big Data, las herramientas tradicionales, e inclusive algunas propias de Big Data, no son idóneas para manejar el procesamiento de estos datos en tiempo real (Al-Fuqaha et al., 2015; M. Chen et al., 2014), siendo necesarias nuevas herramientas computacionales para esta labor. Por ello, Lee & Lee (2015) han señalado al CC como una posible solución para el análisis de datos en tiempo real del IoT. Por su parte, Al-fuqaha *et al.* (2015) destacan a la herramienta “TSaaaS”, la cual utiliza análisis de datos de series de tiempo para poder llevar a cabo un análisis en tiempo real de los datos del IoT.

1.2.4.4 Realidad Virtual y Aumentada

En inglés “Virtual and Augmented Reality” (VR y AR), son tecnologías y dispositivos de visualización de datos e información, utilizados para que los usuarios interactúen con un ambiente simulado de la realidad (caso de la VR) o interactúen con una superposición de la realidad con información generada por el computador (caso de la AR) (Nee & Ong, 2013; Stair & Reynolds, 2010). La VR y AR han surgido como una de las soluciones más importantes para llevar a cabo los métodos de visualización de grandes volúmenes de datos complejos. Por este motivo, son tecnologías utilizadas para llevar a cabo análisis de Big Data con el fin superar las limitaciones humanas en la percepción e interpretación de grandes volúmenes de datos por medio de métodos avanzados de visualización de imágenes, diagramas o animaciones, mejorando así la calidad en la toma de decisiones (Olshannikova *et al.*, 2015).

Una vez descrita la relación existente entre Big Data, los SI y las TI, estableceremos la necesidad en las organizaciones de un personal altamente capacitado en competencias digitales en la era del Big Data y la Industria 4.0. Además, estudiaremos cómo el Big Data ha penetrado los diferentes tipos de industrias y cómo se han beneficiado del Big Data para mejorar sus procesos, la toma de decisiones y crear valor. Por último, plantearemos una propuesta de los niveles de implementación de ecosistemas de Big Data en las organizaciones. Todas estas cuestiones serán abordadas en el siguiente apartado de este capítulo.

1.3 Big Data en las Organizaciones

Las organizaciones, al abordar el fenómeno del Big Data en sus procesos internos y externos, han tenido que realizar cambios que les permitan afrontar los retos que conlleva el adoptar las nuevas tecnologías que comprenden el ecosistema del Big Data. Es por este motivo que en esta sección expondremos algunos factores que las organizaciones deben tener en cuenta para conseguir las capacidades y habilidades requeridas en sus empleados para lograr extraer valor del Big Data. Además, estudiaremos las distintas aplicaciones del Big Data en algunos sectores industriales y cómo ha impactado en los procesos específicos de su sector. Para finalizar, y como conclusión a este capítulo, recopilaremos lo visto en las distintas secciones de este capítulo para presentar una escala que permita identificar el nivel de implementación de ecosistemas de Big Data y de sus tecnologías asociadas en una organización. Esta propuesta nos permitirá después seleccionar los casos de estudio que harán parte de esta investigación.

1.3.1 Big Data, la Ciencia de Datos y el Científico de Datos

La ciencia de datos (en inglés “Data Science”) es un campo interdisciplinario que ha surgido para afrontar los retos y oportunidades que ha traído consigo el creciente valor de la información en las organizaciones, y busca adoptar un enfoque holístico frente a estos retos (Ahalt *et al.*, 2012; George *et al.*, 2014). Por lo tanto, podemos definir la ciencia de datos como un conjunto de principios para la extracción de información y conocimiento de los datos, por medio de métodos, procesos y técnicas cualitativas y cuantitativas de observación, desarrollo de teorías, análisis sistemático, evaluación de hipótesis y validación rigurosa, con el objetivo de resolver problemas relevantes, tomar decisiones y predecir resultados (Ahalt *et al.*, 2012; Provost & Fawcett, 2013; Waller & Fawcett, 2013). Provost y Fawcett (2013) establecen que los conceptos fundamentales que subyacen a la ciencia de datos son los siguientes:

- ✓ Extraer conocimiento útil de los datos para resolver problemas de la empresa puede ser tratado sistemáticamente siguiendo etapas razonablemente bien definidas.

- ✓ Evaluar los resultados de la ciencia de datos requiere consideraciones cuidadosas del contexto en que van a ser usados.
- ✓ La relación entre el problema de la empresa y la solución analítica frecuentemente puede ser descompuesta en sub-problemas más manejables, a través de un marco de análisis del valor esperado.
- ✓ Las TI pueden ser usadas para encontrar datos informativos dentro de una gran cantidad de datos.
- ✓ Las entidades que son similares con respecto a las características o atributos conocidos, frecuentemente son similares con respecto a características o atributos desconocidos.
- ✓ Si miras demasiado a un conjunto de datos encontrarás algo, pero puede que no se generalice más allá de los datos que estas observando (Sobreajuste).
- ✓ Para sacar conclusiones casuales, se debe prestar mucha atención a la presencia de factores de confusión; posiblemente de aquellos que no se pueden ver.

Así mismo, Ahalt *et al.* (2012) establecen que los propósitos fundamentales de la ciencia de datos se centran en mejorar las habilidades para describir, predecir, y explicar los procesos naturales y sociales, a través de los siguientes puntos:

- ✓ Crear conocimiento en base a las propiedades de las grandes cantidades de datos dinámicos.
- ✓ Desarrollar métodos para compartir, administrar y analizar datos digitales.
- ✓ Optimizar los procesos de datos en factores como la precisión, latencia y costo.

Con la llegada del Big Data, la ciencia de datos ha adquirido un mayor interés debido a que en la actualidad las organizaciones se encuentran en el proceso de construir capacidades para poder extraer valor del Big Data, lo que Provost & Fawcett (2013) denominan la era del Big Data 1.0. Además, consideran que cuando las organizaciones logren manejar grandes volúmenes de datos en tiempo real de una manera flexible llegará la era del Big Data 2.0, y será cuando la ciencia de datos alcanzará su época dorada e, inclusive, se considera que empresas avanzadas en este campo como Amazon ya han llegado a esta era (*ibíd.*).

Como hemos visto anteriormente, la relación existente entre Big Data y la ciencia de datos es clara, lo que ha provocado una confusión entre ambos términos e, incluso, se han considerado como sinónimos (George et al., 2014), convirtiéndose en uno de los principales mitos del Big Data (Jagadish, 2015). La diferencia principal entre ambos términos está en que el Big Data se enfoca más en la recolección y gestión de las grandes cantidades de datos y sus atributos, mientras que la ciencia de datos se enfoca en los modelos que capturan, analizan y visualizan los datos (George et al., 2014; Jagadish, 2015). El punto de intersección entre ambos es innegable, en especial que ciertas técnicas y métodos de BDA son manejadas por la ciencia de datos. No obstante, la ciencia de datos no solamente es utilizada para llevar a cabo BDA, sino que su campo de acción puede abarcar diferentes áreas relacionadas con los datos que no impliquen Big Data, y viceversa, por lo que concentrarse únicamente en la intersección entre ambos puede ser limitante, y es mejor verlos como complementarios (Jagadish, 2015).

A raíz del creciente interés de las organizaciones por adquirir las capacidades y habilidades que les permitan aprovechar la ciencia de datos y el Big Data, surge un nuevo tipo de trabajador digital, el científico de datos, y dada la importancia que ha adquirido en el mercado laboral, por lo que Davenport y Patil (2012) lo han catalogado como el trabajo más “sexy” de siglo 21. El científico de datos se puede definir como un profesional altamente capacitado que posee un profundo conocimiento y habilidades analíticas en el campo de la ciencia de datos y el Big Data (Davenport & Patil, 2012; Waller & Fawcett, 2013), con capacidades para estructurar grandes cantidades de datos no estructurados, identificar fuentes de datos útiles, y aplicar sus descubrimientos y soluciones a los problemas de una organización desde una perspectiva de los datos, para evaluar sus implicaciones y las futuras direcciones (Davenport & Patil, 2012; Provost & Fawcett, 2013).

Las habilidades que debe tener un científico de datos son muy variadas ya que abarcan muchos campos del conocimiento, por lo que Waller y Fawcett (2013) consideran que es casi imposible que un individuo posea todas las habilidades posibles de un científico de datos. Por este motivo, Schoenherr y Speier-Pero (2015) han establecido, en una encuesta realizada a 531 trabajadores en el sector de cadenas de suministros, que las principales habilidades que necesita un científico de datos provienen de los campos del pronóstico (cualitativo y cuantitativo), optimización, estadística (estimación y muestreo), economía

(determinar el costo de oportunidad), modelación matemática y la probabilidad aplicada. Adicional a las anteriores, los autores además hacen énfasis en que una de las habilidades necesarias para el científico de datos debe ser la habilidad de comunicación e interpersonales, debido a que es de suma importancia saber comunicar el conocimiento adquirido por el proceso de análisis de datos (Schoenherr & Speier-Peró, 2015).

Por lo tanto, el científico de datos debe ser una persona cercana al resto de la organización para poder comprender los problemas y comunicar las soluciones. Sin embargo, el pensamiento analítico basado en datos no debe ser exclusivo del científico de datos, ya que el resto de empleados deben tener noción de los principios de la ciencia de datos y el Big Data para poder aprovechar los recursos y soluciones que proveen estos campos. En otras palabras, en organizaciones que manejan ciencia de datos y/o Big Data, el pensamiento analítico basado en datos debe convertirse en parte de la cultura organizativa (Carillo, 2017; Provost & Fawcett, 2013).

La conexión existente entre el científico de datos y el resto de la organización debe darse en especial con los ejecutivos y tomadores de decisiones a cargo de los productos y servicios de la misma (Davenport & Patil, 2012). Así, grandes empresas a nivel mundial han visto la necesidad de crear un nuevo grupo de ejecutivos capaces de comprender y aprovechar los beneficios de la ciencia de datos y el Big Data. Dentro de este nuevo grupo de ejecutivos, destacan el “Chief Data Officer” (CDO), encargado de gestionar el Big Data a nivel ejecutivo y alinearlos con la estrategia del negocio, el “Chief Technology Officer” (CTO), encargado de identificar aplicaciones tecnológicas a los productos, servicios y procesos, y el “Chief Marketing Technologist” (CMT), encargado de todas las actividades de marketing de la organización relacionadas con las nuevas tecnologías (Carillo, 2017). Este nuevo grupo de ejecutivos está obligando a modificar las relaciones de poder en varias de las funciones de una organización, y junto con los científicos de datos, son muestra de que la ciencia de datos y el Big Data están impactando a las organizaciones en sus procesos y roles internos (Carillo, 2017), así como en su estructura organizativa. De ahí el interés de esta tesis doctoral.

Teniendo en cuenta cómo impactan la ciencia de datos, el Big Data, y los trabajadores digitales de esta nueva era a las organizaciones, a continuación, pasamos a estudiar cómo dichos aspectos han impactado en algunos sectores e industrias, analizando de que forma

benefician sus aplicaciones al desarrollo de las mismas y cómo será su evolución en base a esta nueva forma de extraer valor a los datos.

1.3.2 Aplicaciones del Big Data en Diferentes Sectores

Las aplicaciones de Big Data y de los SI y las TI asociadas a ésta han sido numerosas, donde diversos sectores industriales y de servicios han visto la necesidad de incorporar los nuevos tipos de análisis de grandes volúmenes de datos para poder obtener ventajas competitivas en su sector. En un estudio del McKinsey Global Institute en 2011, destaca que los sectores que más han mostrado beneficios en el uso e implementación de sistemas y tecnologías asociadas al Big Data son el sector salud, el sector público y el de cadenas de suministros (Manyika et al., 2011). A continuación, describiremos como el Big Data ha impactado a cada uno de estos sectores.

1.3.2.1 Sector salud

Es un sector que se caracteriza por tener problemas y retos para poder mejorar su productividad, en especial en la salud pública, debido a la variedad que actores que intervienen en ella, como industrias farmacéuticas, proveedores, centros de salud y pacientes (Manyika et al., 2011). Las fuentes de datos del sector salud son muy variadas y con un gran volumen, proviniendo de los datos de investigación y desarrollo en salud, los datos clínicos de pacientes, datos de costos y proveedores, y el comportamiento y preferencia de los pacientes (*ibíd.*). Para Murdoch y Detsky (2013), uno de los principales motivos para el aumento de la cantidad de datos en el sector salud ha sido la llegada y expansión en el uso de registros electrónicos, los cuáles han generado grandes cantidades de datos cuantitativos (p.ej., exámenes de laboratorio), cualitativos (p.ej., recetas médicas en texto) y transaccionales (p.ej., registros de entrega de medicamentos), convirtiéndose en un potencial para la mejora de la eficiencia y la productividad del sector salud.

Así mismo, Murdoch y Detsky (2013) establecen que el Big Data aportará grandes beneficios en el sector salud en cuatro puntos claves:

- ✓ Big Data puede expandir la capacidad para generar nuevo conocimiento mediante el análisis de datos no estructurados de los registros electrónicos a través de técnicas de BDA, permitiendo una recolección de datos automatizada.

- ✓ Big Data ayuda a la diseminación del conocimiento al facilitar el acceso a los análisis de los datos de los registros electrónicos para la toma de decisiones clínicas en tiempo real, consiguiendo así diagnósticos más acertados y estandarización en la atención a los pacientes.
- ✓ Big Data ayuda convertir a las iniciativas de medicina personalizada en una práctica clínica, al ofrecer capacidades para la integración de la biología de sistemas con los datos de los registros electrónicos.
- ✓ Por último, Big Data permite a los pacientes estar más involucrados en sus tratamientos médicos, por medio de la entrega de información relevante para ellos. Además, Big Data permite vincular los registros médicos de los pacientes (como la historia clínica y familiar) con información personal (ingresos, educación, etc.), hábitos alimenticios y regímenes de ejercicios, con el fin de relacionar los factores sociales de salud con el paciente.

Por lo tanto, el análisis de Big Data en el sector salud será vital para la mejora en la calidad de los servicios en salud y para poder responder efectivamente a los nuevos retos que enfrenta este sector (Archenaa & Anita, 2015). Por ello, la aplicación del Big Data ha adquirido un gran potencial para la mejora de los procesos, productividad y la eficiencia de entidades dedicadas a esta labor, tanto en el sector público como privado, direccionándose hacia una atención personalizada y poder mejorar los tratamientos a los pacientes por medio de toma de decisiones clínicas en tiempo real.

1.3.2.2 Sector Público

El Big Data ha tenido un gran impacto en el sector público, en especial en el sector de la salud pública y las aplicaciones asociadas a Big Data como es el caso del gobierno electrónico (Sheng *et al.*, 2017). Estos sectores han visto como el uso del Big Data impacta a la gobernabilidad al permitir a la ciudadanía una mayor participación y transparencia en los programas de administración pública (*ibíd.*). En el estudio realizado por Manyika *et al.* (2011) se han establecido cinco puntos en los que el sector público se beneficia del Big Data:

- ✓ Crear transparencia: el Big Data promueve la accesibilidad a las bases de datos de las diferentes agencias gubernamentales.

- ✓ Descubrir necesidades, exponer la variabilidad y mejorar el desempeño: Big Data permite detectar las variaciones en el desempeño de las diferentes agencias del gobierno con funciones similares, cuya variabilidad no es visible cuando se ve de forma agregada.
- ✓ Segmentar la población para tomar acciones personalizadas: el Big Data permitirá detectar las necesidades especiales de cada sector de la población para poder atender específicamente sus prioridades, mejorando la eficiencia de los servicios prestados y la satisfacción de los ciudadanos.
- ✓ Reemplazar y/o apoyar la toma de decisiones humanas con decisiones automatizadas: las técnicas de Big Data pueden proveer un apoyo para detectar anomalías en diferentes agencias gubernamentales, como en el pago de impuestos o en seguridad social, para poder tomar acciones en tiempo real.
- ✓ Innovar en nuevos modelos de negocio, productos, y servicios: el Big Data puede promover políticas y programas en el sector público para la mejora en el desempeño y la innovación, permitiendo que las herramientas de BDA sean accesibles a todos los actores internos y externos que forman parte del sector público.

Por consiguiente, las políticas públicas harán parte importante para que muchos sectores se beneficien del análisis del Big Data y sus tecnologías asociadas, logrando que empresas internas y externas asociadas al sector público, como las de servicios públicos (Agua, Luz, Gas, etc.), logren un desarrollo y mejoras en la eficiencia y desempeño en la prestación de servicios que favorezcan a todos los ciudadanos.

1.3.2.3 Cadena de suministros

Uno de los sectores que más se ha visto influenciado por el fenómeno del Big Data y de las tecnologías asociadas al Big Data, en especial el IoT, es el de la cadena de suministros (Atzori *et al.*, 2010). La aparición de tecnologías como el RFID (Identificación por radio frecuencia) para el manejo y seguimiento de productos y la integración de otras tecnologías (p.ej., los sensores inteligentes) en los procesos de fabricación, distribución y venta, ha llevado a un mayor nivel de interconexión entre las diferentes empresas que integran la cadena de suministros, lo que genera una gran cantidad de datos proveniente de múltiples fuentes (Ferretti & Schiavone, 2016). A raíz de esta generación de grandes cantidades de datos de diversas fuentes Gunasekaran *et al.* (2017) establecen que el Big

Data y el BDA puede mejorar el desempeño de toda la cadena de suministros utilizando el análisis de los datos para mejorar la visibilidad, la resiliencia, la robustez y el desempeño organizativo.

En esta misma línea, Manyika *et al.* (2011) establecen que son tres los puntos principales en los que las empresas de la cadena de suministros se pueden beneficiar del Big Data y el BDA:

- ✓ Administración de inventario: Big Data permite mejorar el pronóstico de los inventarios al combinar y analizar múltiples fuentes de datos como el historial de venta o los ciclos de temporadas de ventas.
- ✓ Optimización de la distribución y logística: el uso de tecnologías asociadas al Big Data como la localización GPS en tiempo real de los transportes de mercancías, permite generar grandes cantidades de datos para la optimización de rutas, mejorar la eficiencia de combustible y mejorar la productividad de los transportes.
- ✓ Relaciones con el proveedor: Big data permite un mejor análisis de los datos provenientes de las transacciones y precios manejados por los minoristas para establecer una mejor comunicación entre el minorista y el proveedor sobre las preferencias y comportamientos de los clientes.

La cadena de suministros y la venta al por menor vienen presentando grandes cambios en la forma en cómo se distribuyen los productos y la manera en la que se interactúa con el cliente, siendo el ejemplo más destacado la empresa Amazon con sus nuevos supermercados Amazon Go, donde no existen cajeros ni colas, el cliente solo escoge los productos que necesita, sale de la tienda y se cobra automáticamente en la tienda de Amazon virtual (Amazon, 2018), siendo un claro ejemplo de la integración de tecnologías como el IoT y el Big Data, permitiendo adquirir datos para la toma de decisiones en tiempo real y una transparencia de la información a lo largo de toda la cadena de suministros.

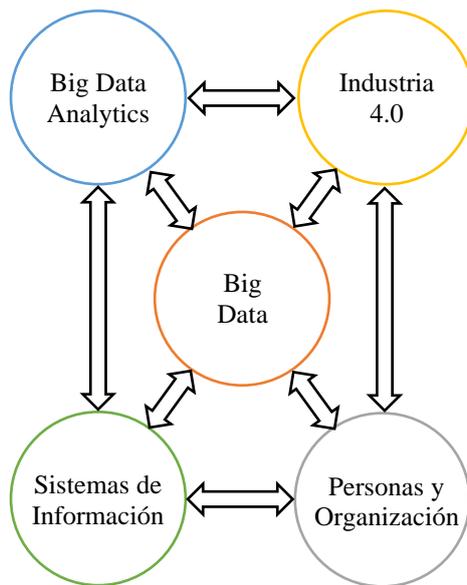
Teniendo en cuenta cómo el Big Data y las TI asociadas han impactado en diferentes sectores, pasaremos a proponer las diferentes etapas que comprenden la implementación de ecosistemas de Big Data, estableciendo qué características debe tener una organización en cada uno de sus niveles.

1.3.3 Niveles de Implementación de Ecosistemas de Big Data en las Organizaciones

Como hemos visto en secciones anteriores, el Big Data y BDA se han convertido en casi una necesidad para las organizaciones, donde los datos se han convertido en protagonistas para la creación de valor, obtener ventajas competitivas y mejorar el desempeño. Sin embargo, muchas organizaciones no poseen las capacidades, habilidades o tecnologías para sacar el mayor provecho al Big Data, en especial las pequeñas y medianas empresas, en donde el nivel de inversión en tecnologías, sistemas y capital humano para aprovechar el Big Data es más reducido en comparación con empresas grandes. A pesar de esta situación, y como sucede con todas las aplicaciones de este tipo, los sistemas y tecnologías asociadas al Big Data irán siendo cada vez más accesibles y tendrán una mejor adaptación a las necesidades particulares de cada organización, lo que permitirá que empresas en todos los niveles puedan aprovechar las ventajas que conlleva el Big Data.

Por lo tanto, a medida que la accesibilidad de los sistemas y tecnologías asociadas a Big Data sea mayor, las organizaciones experimentarán una adopción progresiva de los mismos. Así, como interés en esta investigación estableceremos los niveles de implementación de los sistemas, tecnologías, métodos, técnicas, herramientas y factores humanos que se relacionan con el Big Data en las organizaciones, a lo que, para efectos de simplificación, hemos denominado “Ecosistema de Big Data” (ver figura 1-8 en la página siguiente).

Figura 1-8: Ecosistema de Big Data

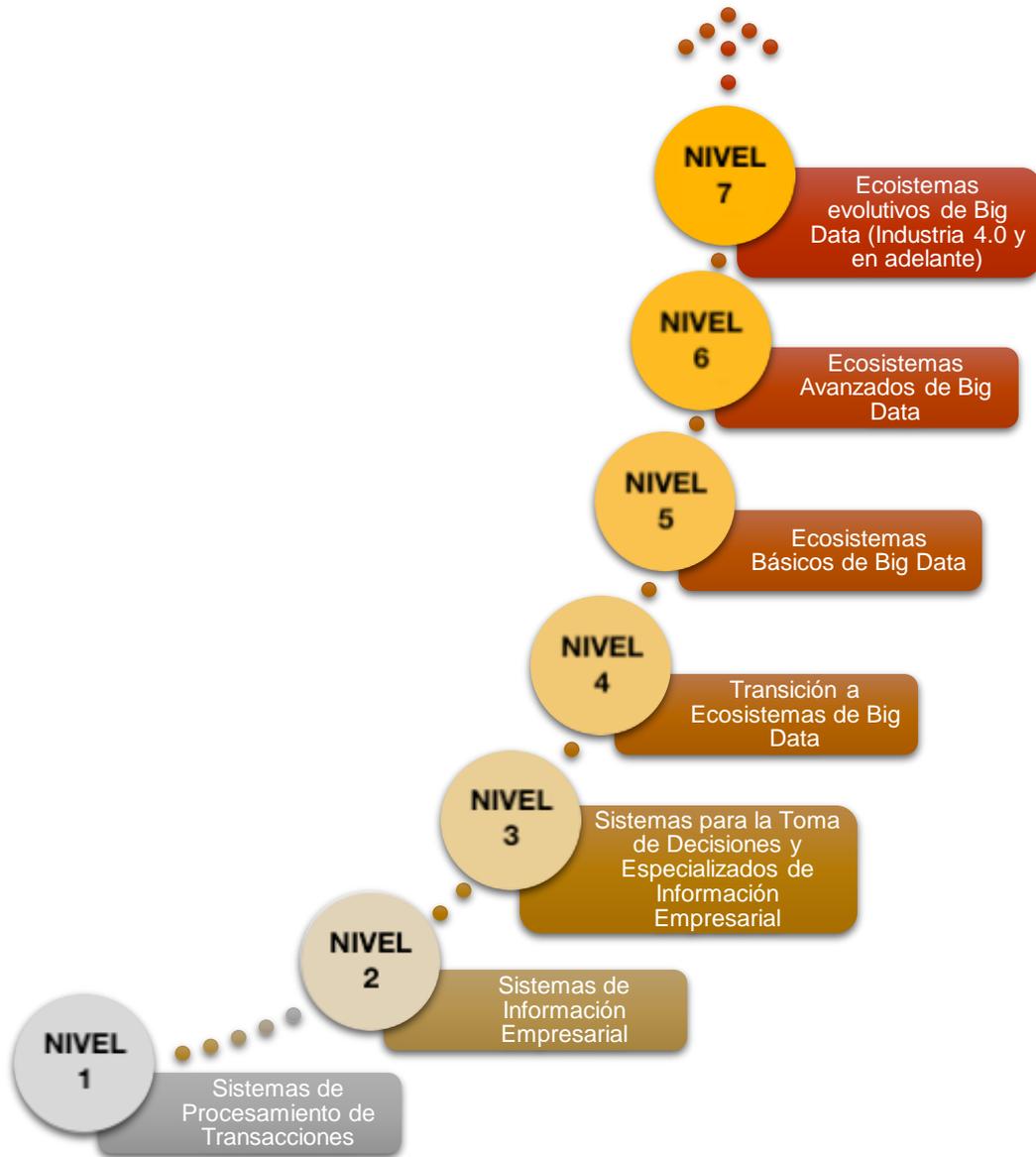


Fuente: elaboración propia

Teniendo en cuenta lo anterior, proponemos siete niveles que definen el grado de implementación de ecosistemas de Big Data en las organizaciones (ver figura 1-9), partiendo de la clasificación de los SI descrita en la sección 1.1.3 y los distintos niveles en los que se puede encontrar la organización de acuerdo al nivel de adopción de los diferentes sistemas y tecnologías asociadas al Big Data. Así, en base a la literatura especializada hemos identificado cuatro niveles de implementación de ecosistemas de Big Data: un nivel de transición al Big Data, un nivel de Big Data Básico, un nivel de Big Data Avanzado y un nivel superior al que hemos denominado Big Data Evolutivo.

Para cada uno de estos niveles, hemos definido cinco categorías o dimensiones, las cuales describen las características que debe tener una organización para corresponder con un nivel. Estas categorías son: implicación organizativa de Big Data, factores humanos y su relación máquinas-sistemas, método de BDA utilizado, tipo de herramienta de BDA e integración de tecnologías de información. Por lo tanto, a continuación, presentamos nuestra propuesta para establecer el nivel de implementación de ecosistemas de Big data en las organizaciones, la cual nos servirá más adelante como una herramienta para la selección de casos de estudio para esta investigación e investigaciones similares en el campo. La figura 1-9 nos muestra gráficamente la propuesta, mientras que la tabla 1-5 sintetiza el contenido de cada uno de los niveles.

Figura 1-9: Niveles en la implementación de ecosistemas de Big data



Fuente: elaboración propia

Tabla 1-5: Niveles y dimensiones de implementación de ecosistemas de Big Data en las organizaciones

Nivel	Implicación Organizativa de Big Data	Factores Humanos y su Relación Máquinas-Sistemas	Método Utilizado de Analytics	Clase de Herramienta de BDA	Integración de Tecnologías de Información
1	La organización aun no adopta ninguna forma de Big Data y solo formas básicas de análisis de datos	Los empleados y directivos se encuentran capacitados para registrar y gestionar las transacciones en un SI	Métodos estadísticos y matemáticos simples	N/A	Integración de ordenadores y servidores
2	La organización realiza análisis de datos internos para realizar reportes, sin llegar a niveles de Big Data	Los empleados y directivos se encuentran capacitados para gestionar y compartir datos e información mediante el SI con actores internos y externos	Métodos estadísticos y matemáticos avanzados	N/A	Integración de tecnologías de seguimiento como el código de barras o el RFID
3	La organización comprende la importancia y realiza análisis de datos para la toma de decisiones y crear conocimiento, sin llegar a niveles de Big Data	La organización posee personal y departamentos especializados en la gestión y mantenimiento de los SI y las TI. Los empleados y directivos se encuentran capacitados para la toma de decisiones y gestión del conocimiento mediante el uso de los SI	Método de analítica descriptiva con un volumen bajo o medio de datos	N/A	Integración de tecnologías para la mejora en el procesamiento de datos, como Hardware de alto desempeño y alta capacidad de almacenamiento, y el seguimiento y monitoreo de procesos, como sensores y sistemas de vigilancia
4	La organización establece los recursos y capacidades necesarios para la adopción de sistemas y tecnologías asociadas a Big Data, seleccionando las prioridades y necesidades de la organización	La organización incorpora o se encuentra en proceso de selección o capacitación de personal competente en métodos, técnicas y herramientas de BDA y ciencia de datos. Se establece un plan de capacitación en toda la organización para uso del Big Data en los procesos de la organización	Método de analítica descriptiva con volumen medio o alto de datos	Procesamiento por lotes	Integración (en proceso) de algunas tecnologías para el procesamiento de Big Data, como algunas herramientas de Cloud Computing o Inteligencia Artificial

<i>Nivel</i>	Implicación Organizativa de Big Data	Factores Humanos y su Relación Máquinas-Sistemas	Método Utilizado de Analytics	Clase de Herramienta de BDA	Integración de Tecnologías de Información
5	La organización posee los recursos y capacidades básicas para utilizar y gestionar las técnicas, métodos y herramientas de BDA. Se empieza a extender por la organización la cultura del pensamiento analítico y se toman más decisiones en mandos medios	Se crea en la organización un departamento especializado encargado de llevar a cabo el análisis y extracción de valor de los datos. El personal y ejecutivos comprenden e incorporan el nuevo conocimiento extraído del Big Data en sus procesos y actividades diarias	Método de analítica descriptiva o predictiva con gran volumen de datos (dependiendo de la organización)	Procesamiento por lotes o Procesamiento en tiempo real. (Dependiendo de la organización)	Integración de soluciones de Big Data basados en Cloud Computing o Inteligencia Artificial para el análisis de datos
6	La organización adopta sistemas avanzados de KM y DSS que integren las técnicas, métodos y herramientas de BDA. Se comienza el proceso de automatización de ciertas decisiones y el pensamiento analítico ya hace parte de la cultura organizativa	La organización incorpora ejecutivos con concomimientos específicos en Big Data y ciencia de datos, como los CDO y CTO, Todos empleados y ejecutivos están capacitados en el manejo de algunos métodos, técnicas y herramientas de BDA	Método de analítica descriptiva o prescriptiva con grandes volúmenes de datos en tiempo real (dependiendo de la organización)	Procesamiento en tiempo real o Procesamiento Híbrido (dependiendo de la organización)	Integración de soluciones personalizadas mediante la unión de tecnologías de Cloud Computing, Inteligencia Artificial, entre otros
7	Los sistemas tienen la capacidad de evolucionar por si solos para adaptarse a los cambios de la organización, automatizándose decisiones operativas en tiempo real. La descentralización de las decisiones es completa y la organización se encuentra completamente conectada con su entorno	Los empleados y ejecutivos tienen capacidades para controlar e interactuar con los datos en tiempo real a través de sistemas avanzados de visualización de datos. El personal de la organización utiliza los sistemas y tecnologías para realzar sus capacidades operacionales y cognitivas	Método de analítica prescriptiva con grandes volúmenes de datos en tiempo real	Procesamiento Híbrido con arquitectura Lambda	Integración de diversas tecnologías que componen la Industria 4.0, como el IoT, Cloud Computing, Inteligencia Artificial, Realidad Virtual o Aumentada, entre otros

Fuente: elaboración propia

Las dimensiones descritas en la tabla anterior brindan una explicación de las características que debe poseer una organización para pertenecer a un nivel dado. En

primer lugar, la implicación organizativa de Big Data establece que tipo de SI o herramienta de Big Data ha adoptado la organización para la toma de decisiones basada en análisis de datos, el nivel de descentralización de esas decisiones y el nivel de implicación del pensamiento analítico en la cultura organizativa. En segundo lugar, los factores humanos y su relación máquinas-sistemas establece el nivel de capacitación e implicación que poseen los empleados y directivos en el manejo y gestión del SI, en el uso de métodos básicos, intermedios o avanzados de análisis de datos o de BDA y en la creación de áreas o departamentos especializados de Big Data y ciencia de datos. En tercer lugar, el método utilizado de Analytics establece que tipo de métodos estadísticos, matemáticos o de BDA utiliza la organización (ver sección 1.2.2). En cuarto lugar, la clase de herramienta de BDA establece cuál de las herramientas de BDA descritas en la sección 1.2.2 utiliza la organización. En quinto lugar, la integración de las TI establece que posibles tecnologías básicas, intermedias o avanzadas ha adoptado la organización.

En la tabla presentada anteriormente, se denota como cada una de las dimensiones va incrementando en intensidad a medida que aumentan los niveles. Esto es debido a que nuestra propuesta implica que un nivel superior abarca las características descritas en la dimensión del nivel anterior, por lo que las dimensiones y niveles tienen carácter acumulativo. Cabe resaltar, que cada uno de los niveles y dimensiones son susceptibles de no cumplirse en su totalidad, dada la gran variedad de formas para implementar este tipo de sistemas y tecnologías, por lo que se pueden presentar casos en que una organización pueda caer en diferentes niveles dependiendo de la dimensión. Sin embargo, proponemos que para que una organización pueda considerarse en un nivel concreto, debe cumplir gran parte de las dimensiones descritas.

Una vez descrito el marco teórico correspondiente a los SI, TI y Big Data, y concluir con nuestra propuesta para establecer el nivel de implementación de ecosistemas de Big Data en las organizaciones, podemos afirmar que esta nueva era del Big Data presenta retos y beneficios para las organizaciones que implican replantear la forma en cómo se estructura la organización para la adopción de los sistemas y tecnologías asociadas al Big Data y la forma en que ésta gestiona sus procesos vinculados al conocimiento. Por ello, el siguiente capítulo está destinado a introducir los fundamentos teóricos básicos de la gestión del conocimiento y del diseño organizativo, ya que nuestro objetivo de investigación tiene

que ver con estudiar cómo el Big Data va a afectar estos dos elementos organizativos: la estructura y los procesos de gestión del conocimiento.

Capítulo 2:

Procesos de Gestión del Conocimiento y Mecanismos de Coordinación y Control en la Organización

En el capítulo anterior, hemos explorado los aspectos más relevantes del Big Data en las organizaciones y el cómo, de forma general, han afectado a muchas áreas, actividades y procesos que son llevados a cabo en una organización de acuerdo al nivel de implementación que se tenga del mismo. Dado que en esta tesis doctoral pretendemos analizar cómo la aplicación del Big Data afecta a los procesos de gestión del conocimiento y a la forma de coordinación del trabajo en las organizaciones, en el presente capítulo expandiremos aquellos conceptos que consideramos relevantes para esta investigación. Así, en la primera parte abordaremos el estudio de la gestión del conocimiento, en donde describimos la relación existente entre datos, información y conocimiento, el concepto general de gestión del conocimiento y los procesos que subyacen a la misma. En la segunda parte del capítulo, exploramos las formas de coordinación del trabajo en las organizaciones, describiendo en primer lugar el concepto general de diseño organizativo y presentando, después, los mecanismos de coordinación y control.

2.1 Concepto y Procesos de la Gestión del Conocimiento

En el capítulo anterior, hemos encontrado referencias con las palabras conocimiento, gestión del conocimiento y sistemas de gestión del conocimiento, esto debido a que el Big Data es considerado como un creador de conocimiento (Fosso-Wamba et al., 2015), por lo que es natural relacionar ambos términos entre sí. Con el fin de establecer claramente la relación existente entre Big Data y la gestión del conocimiento en el capítulo posterior, a continuación, describimos los conceptos más importantes a tener en cuenta en relación a la gestión del conocimiento.

2.1.1 Información, Conocimiento y Sabiduría

Como establecimos en la sección 1.1.1 de esta tesis doctoral, los datos y la información son términos intrínsecamente relacionados, pudiéndose producir como consecuencia una confusión de ambos términos. Sin embargo, al revisar la literatura especializada encontramos que el término información adicionalmente es relacionado con el término conocimiento, siendo términos que en ocasiones pueden ser difíciles de distinguir (Davenport *et. al.* 1998). Para Huber (1991) y Nonaka (1994), los términos información y conocimiento pueden encontrarse en la literatura indistintamente; no obstante, ambos

autores declaran que es necesario establecer una diferencia entre ambos términos, situación que es de igual importancia para los objetivos de esta investigación.

Una de las formas más comunes de describir el significado de conocimiento es estableciendo una diferencia con los datos y la información (Alavi & Leidner, 2001; Nonaka & Peltokorpi, 2006), siendo de las más comunes definir al conocimiento como una forma más rica de información (Gurteen, 1998) o como información procesable (Jashapara, 2004). Estas definiciones colocan al conocimiento un escalón por encima de la información, conociéndose en la literatura como la visión jerárquica de los datos, la información y el conocimiento (ver Figura 2-1) (Alavi & Leidner, 2001; Jashapara, 2004).

Figura 2-1: Visión jerárquica de los datos, información, conocimiento y sabiduría



Fuente: Jashapara (2004, p. 17)

Teniendo en cuenta lo anterior, autores como Alavi & Leidner (2001), Gurteen (1998) y Jashapara (2004) declaran que las definiciones anteriormente expuestas y la visión jerárquica de los datos, la información y el conocimiento es una forma muy simple para lograr definir correctamente al conocimiento. Sin embargo, esto no implica que la información no tenga una relación directa con el conocimiento ya que, como establece Nonaka (1994, p. 16), “la información es el medio o material para inicializar y formalizar el conocimiento”. Por lo tanto, la información es necesaria para que se pueda crear conocimiento y para que el conocimiento sea productivo (Gurteen, 1998).

Como podemos observar, aunque la relación entre información y el conocimiento sea indiscutible, no es la única manera en la que se puede definir al conocimiento. Así, Alavi & Leidner (2001) recopilaron las diferentes perspectivas que han surgido en la literatura para definir al conocimiento, las cuales son presentadas a continuación (ver tabla 2-1).

Tabla 2-1: Perspectivas para definir el conocimiento

PERSPECTIVA	DEFINICIÓN
Conocimiento vis a vis con los datos y la información	Los datos son hechos sin manipular. Información son datos procesados e interpretados. Conocimiento es información personalizada poseída por los individuos
Conocimiento como estado mental	El conocimiento es un estado de saber y comprender mediante la experiencia y el estudio
Conocimiento como objeto	El conocimiento es un objeto que debe ser almacenado y manipulado.
Conocimiento como proceso	El conocimiento es un proceso para la aplicación de la experiencia
Conocimiento como acceso a la información	El conocimiento debe ser organizado para facilitar el acceso y recuperación de la información
Conocimiento como capacidad	El conocimiento es la capacidad para usar información; el aprendizaje y la experiencia resultará en habilidades para interpretar y comprobar la información necesaria para tomar decisiones

Fuente: Alavi & Leidner (2001)

Adicional a las perspectivas expuestas anteriormente, autores como Alavi & Leidner (2001), Jashapara (2004) y Sabherwal & Becerra-Fernandez (2003) consideran como importante para la definición del conocimiento el trabajo de Nonaka (1994), autor que se basa en los aportes de Polanyi (1967). El trabajo de Nonaka (1994) ha servido de base para muchos autores en temas relacionados con el conocimiento, siendo catalogado como la investigación más influyente en este campo (Gaviria-Marin et. al., 2019).

Teniendo en cuenta las anteriores afirmaciones, en esta investigación ampliaremos la definición del conocimiento en base a las observaciones de Nonaka (1994), donde se establece que existen dos dimensiones del conocimiento: la primera corresponde al “conocimiento explícito”, también conocido en la literatura como “know-what”, y la segunda corresponde al “conocimiento tácito”, también conocido en la literatura como “know-how” (Jashapara, 2004; Nonaka, 1994). Estas dos dimensiones son los tipos de conocimiento más citados en la literatura (Alavi & Leidner, 2001), donde el conocimiento

explícito corresponde al conocimiento codificado que es transmisible mediante un lenguaje formal y sistemático (Nonaka, 1994), por ejemplo, en forma de datos, información, especificaciones, manuales, etc. (Sabherwal & Becerra-Fernandez, 2003), y el conocimiento tácito corresponde al “conocimiento arraigado a las acciones, la experiencia e implicaciones en un contexto específico” (Nonaka, 1994 p. 16), por ejemplo, el conocimiento que surge de las ideas, intuiciones y corazonadas (Sabherwal & Becerra-Fernandez, 2003).

A pesar de existir diferencias entre ambas dimensiones, Nonaka (1994) establece que ambas dimensiones no son exclusivas sino complementarias. Sin embargo, queda claro para autores como Alavi & Leidner (2001), Grant (1996) y Nonaka (1994) que el conocimiento tácito es más difícil de comunicar y formalizar que el conocimiento explícito debido a que tiene una cualidad más personal y, por ende, más difícil de transmitir entre los miembros de una organización (Grant, 1996; Nonaka, 1994).

Adicional a las dimensiones anteriormente expuestas, diversos autores en la literatura establecen que la taxonomía del conocimiento es diversa, apareciendo propuestas en las que se separa al conocimiento en categorías subjetivas u objetivas, en aquellas que promueve el positivismo y/o pensamiento, o de acuerdo a la naturaleza procesal y multinivel del conocimiento (Nonaka & Peltokorpi, 2006). Por ejemplo, Nonaka (1994) establece que las dimensiones tácita y explícita del conocimiento pueden existir en múltiples niveles, ya que se pueden presentar de forma individual, grupal u organizacional. En esta línea, el autor (*ibíd.*) expone que el conocimiento tácito se divide en dos elementos, el primero corresponde al cognitivo, también llamado “modelos mentales”, en donde los individuos crean y manipulan analogías mentalmente resultando en creencias, paradigmas y puntos de vista. El segundo corresponde al técnico, que se refiere a las habilidades y destrezas (“Know-how”) concretas que aplican a un determinado contexto (Alavi & Leidner, 2001; Nonaka, 1994).

Así mismo, Alavi & Leidner (2001) resaltan la importancia de estas clasificaciones del conocimiento, dado que los desarrollos teóricos en el campo de la gestión del conocimiento surgen de la distinción entre los diferentes tipos de conocimiento. Por este motivo, recopilan los tipos de conocimiento principales que han surgido en la literatura a lo largo de los años, los cuáles presentamos en la tabla 2-2.

Tabla 2-2: Taxonomías del conocimiento

TIPO DE CONOCIMIENTO	DEFINICIÓN
Tácito	Conocimiento arraigado a las acciones, la experiencia e implicaciones en un contexto específico
Cognitivo:	Modelos mentales
Técnico:	Know-how aplicado a un trabajo específico
Explícito	Conocimiento articulado y generalizado
Individual	Creado por e inherente al individuo
Social	Creado por e inherente a las acciones colectivas de un grupo
Declarativo	Saber sobre algo (“know-about”)
Procesal	Saber cómo hacer algo (“know-how”)
Causal	Saber por qué sucede algo (“know-why”)
Condicional	Saber cuándo aplicar algún conocimiento (“know-when”)
Relacional	Saber cómo interactúa o se relaciona algo (“know-with”)
Pragmático	Identificar el conocimiento útil para una organización

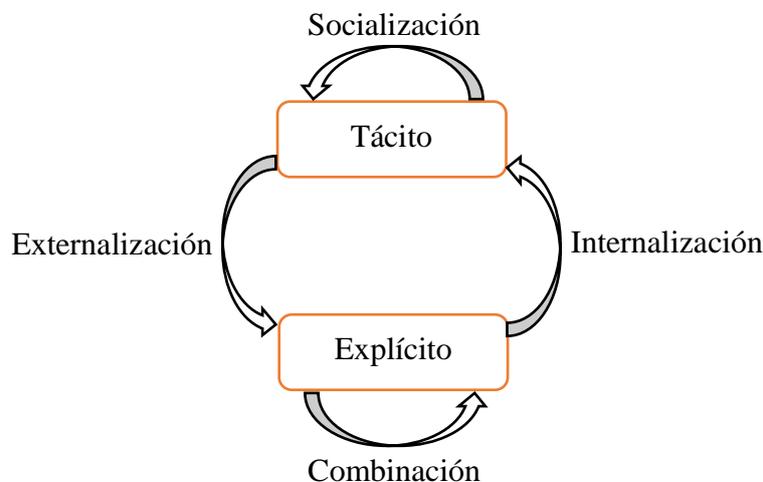
Fuente: Alavi & Leidner (2001, p. 113)

Teniendo en cuenta la taxonomía anterior, la literatura se ha centrado principalmente en comprender la creación de conocimiento en las organizaciones a través del proceso de conversión entre el conocimiento tácito y el conocimiento explícito, conocido como la teoría de la creación del conocimiento (Nonaka, 1994). Esta teoría expone cuatro modos de conversión para crear conocimiento (Cairó-Battistutti & Bork, 2017; Nonaka, 1994; Nonaka & Takeuchi, 1995; Nonaka & von Krogh, 2009), tal y como se representa en la figura 2-2 más abajo:

- 1) De conocimiento tácito a conocimiento tácito: este modo de conversión del conocimiento se presenta mediante la interacción de los individuos y la compartición de su experiencia. Este modo de conversión también se conoce como “socialización”.
- 2) De conocimiento explícito a conocimiento explícito: este modo de conversión del conocimiento se presenta cuando se realiza una reconfiguración de información existente mediante la clasificación, recategorización o recontextualización del conocimiento explícito, con el fin de crear conocimiento explícito superior. Este modo de conversión también se conoce como “combinación”.

- 3) De conocimiento tácito a conocimiento explícito: este modo de conversión del conocimiento se presenta cuando se articula el conocimiento tácito a través de técnicas como la inducción o la deducción para crear conceptos explícitos que puedan ser entendidos y compartidos por otros individuos. Este modo de conversión también se conoce como “externalización”.
- 4) De conocimiento explícito a conocimiento tácito: este modo de conversión del conocimiento se relaciona directamente con el concepto de “aprendizaje”, y consiste en la incorporación individual del conocimiento explícito de un grupo o de la organización mediante la acción y la práctica. Este modo de conversión también se conoce como “internalización”.

Figura 2-2: Modos de conversión del conocimiento



Fuente: Balbastre (2003) a partir de Nonaka y Takeuchi (1995)

Estos cuatro modos de conversión del conocimiento no funcionan de manera estática o aislada, sino que existe una interacción dinámica y constante entre ellos formando lo que se conoce como la “espiral de creación del conocimiento organizativo”. De esta manera, en primer lugar, un individuo adquiere de alguna fuente o de otro individuo un conocimiento tácito mediante la práctica u observación (socialización); luego, ese conocimiento tácito es convertido en conocimiento explícito para poder ser compartido (externalización); después, el conocimiento explícito es consignado en manuales o procedimientos para ser utilizados en procesos, productos y/o tecnologías (combinación); y finalmente, este nuevo conocimiento explícito será incorporado por los miembros de la organización creando conocimiento tácito (internalización) (Balbastre, 2003; Nonaka,

1994; Nonaka & Takeuchi, 1995). Al momento de finalizar un ciclo en la creación del conocimiento, el proceso de conversión vuelve a comenzar, formando una espiral en donde la base del conocimiento de la organización crece de forma ascendente (Balbastre, 2003).

Tomando como base la teoría de la creación del conocimiento, Nonaka y Takeuchi (2011) sugieren que se debe tener en cuenta un tercer, y frecuentemente olvidado, tipo de conocimiento llamado “frónesis” o “sabiduría práctica”. Como pudimos observar en la figura 2-1, la sabiduría se encuentra un escalón por arriba del conocimiento en la visión jerárquica, y es definido por Jashapara (2004, p. 17) como “la habilidad de actuar de manera crítica o práctica en una situación dada, y se basa en el juicio ético relacionado con las creencias de un individuo”. La definición de sabiduría como la expuesta por Jashapara (2004) guarda relación con el concepto de “frónesis”, y es una de las tres formas de conocimiento descritas por Aristóteles, definido en su libro *Nicomachean Ethics* VI.6 como “un estado verdadero y razonado de la capacidad de actuar con respecto a las cosas que son buenas o malas para el hombre” (Nonaka & Takeuchi, 2011, p. 4), y definido por los autores como “un conocimiento que surge de la experiencia y que permite a las personas hacer juicios éticamente sólidos”.

En base el concepto de frónesis, Nonaka *et al.* (2016) proponen que la frónesis es el factor que promueve la espiral de creación del conocimiento, llamando a la síntesis (relación) de estos tres tipos de conocimiento como “la triada del conocimiento”. Así, estos autores (*ibíd.*) proponen que la frónesis conduce los modos de conversión del conocimiento mediante la inclusión de juicios de valor, los cuales permiten interpretar contextos, captar la esencia y crear significado por fuera de un contexto. Por ende, la inclusión de la frónesis en la espiral de la creación del conocimiento añade valor al incluir las creencias, el compromiso, la pasión y el juicio de las personas en el proceso (Nonaka *et al.*, 2016).

Este concepto de frónesis, en términos organizativos, está normalmente relacionado con las capacidades de liderazgo. No obstante, para lograr promover correctamente los juicios de valor en la espiral del conocimiento, la frónesis debe convertirse en una capacidad en todos los niveles de la organización ya que permitirá a las personas hacer juicios prudentes y tomar acciones basadas en la situación real, siempre guiados por sus valores y su moral (Nonaka *et al.*, 2016; Nonaka & Takeuchi, 2011).

Teniendo en cuenta la forma en cómo se crea el conocimiento, las organizaciones tienen un interés prioritario en gestionar la creación de conocimiento de sus empleados, debido a que el conocimiento es el que promueve la innovación dentro de la organización y evita que el conocimiento necesario para el desarrollo de las operaciones y actividades se pierda cuando empleados clave decidan dejar la organización (Alavi & Leidner, 2001). Por este motivo, en la siguiente sección presentaremos el concepto de gestión del conocimiento, qué aspectos se deben tener en cuenta para llevarse a cabo adecuadamente y qué procesos subyacen a la misma.

2.1.2 Concepto de Gestión del Conocimiento

Como notamos en la sección anterior, el conocimiento y su creación juegan un papel importante en todas las organizaciones, al punto de ser considerado como el recurso más importante de una organización (Becerra-Fernandez & Sabherwal, 2014). Es por este motivo que la gestión del conocimiento ha tenido un creciente interés de directivos e investigadores en el campo para mejorar el desempeño de las organizaciones en base al conocimiento (Alavi & Leidner, 2001; Nonaka & Peltokorpi, 2006). Investigaciones en el área han arrojado que las organizaciones consideran que la mayoría del conocimiento que necesitan se encuentra dentro de la organización. Sin embargo, el verdadero problema del conocimiento en las organizaciones se encuentra en identificar qué conocimiento es requerido, encontrar ese conocimiento y aprovecharlo (Alavi & Leidner, 2001).

A pesar de que las organizaciones reconocen que necesitan gestionar el conocimiento para sacar provecho de él, muchos proyectos para implementar la gestión del conocimiento fracasan debido a que son llevados a cabo como proyectos de gestión de la información, en el sentido de que se centran únicamente en la consolidación de datos en información y en el procesamiento de los mismos, actividades que no aportan a la innovación de productos o de los servicios de la organización (Gold *et al.*, 2001). Los autores (*ibíd.*) sugieren que para lograr ir más allá de los proyectos de información se deben crear estructuras que permitan a la organización facilitar el reconocimiento, la creación, la transformación y la distribución del conocimiento. Como consecuencia de esto, muchas definiciones en la literatura sobre gestión del conocimiento

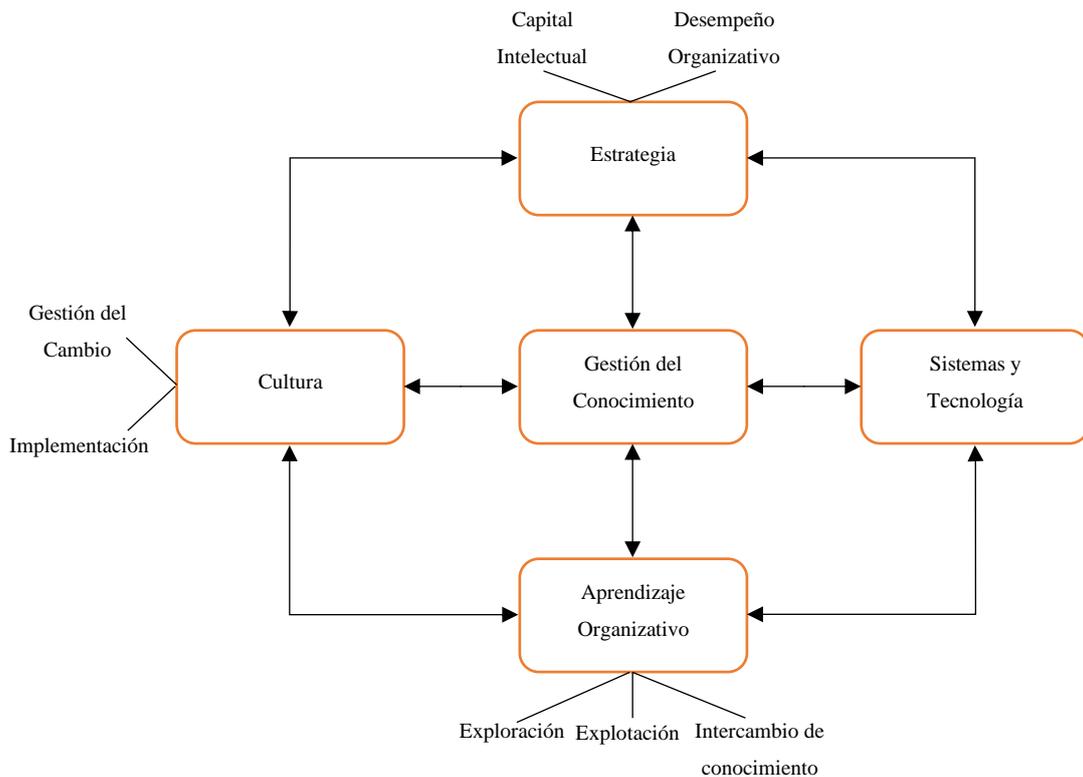
equivocadamente se han basado en los conceptos de gestión de la información y en los sistemas de información (Jashapara, 2004).

Por otro lado, una de las definiciones más comunes que se pueden encontrar en la literatura es ver a la gestión del conocimiento como una disciplina que promueve la creación, adquisición, almacenamiento, transformación, transferencia y aplicación del conocimiento organizativo con el fin de incrementar la innovación y el desempeño en la organización (Becerra-Fernandez & Sabherwal, 2014; Nonaka & Peltokorpi, 2006; Shujahat et al., 2019). Desde este enfoque, la gestión del conocimiento es considerada como un proceso de recursos humanos (Jashapara, 2004).

En contraste con lo anterior, Gurteen (1998) y Jashapara (2004) exponen que la gestión del conocimiento no puede verse desde una única perspectiva, por lo que se debe adoptar una visión más integradora y holística para poder definirla. En esta línea, Gurteen (1998, p. 6) propone que “la gestión del conocimiento es un conjunto emergente de principios de diseño organizativo y operativos, procesos, estructuras organizativas, aplicaciones y tecnologías que ayudan a los trabajadores del conocimiento a aprovechar su creatividad y capacidad para entregar valor a la organización”. Como podemos observar, la visión de Gurteen (1998) implica que la gestión del conocimiento involucra muchas áreas funcionales de la organización, ya que no se centra en único proceso o solo sobre las tecnologías que lo involucran, sino que además se centra en aspectos humanos y organizativos que no son tomados en cuenta por las definiciones tradicionales.

Una de las definiciones más completas que consideramos en esta investigación corresponde a la de Jashapara (2004, p. 12), en donde se consideran cuatro dimensiones que interactúan con la gestión del conocimiento, y que se corresponden con el aprendizaje organizativo, los sistemas y la tecnología, la cultura y la estrategia (ver figura 2-3). En este sentido, Jashapara (2004, p. 12) define la gestión del conocimiento como “los procesos de aprendizaje efectivos asociados con la exploración, explotación e intercambio de conocimiento humano (tácito y explícito) que utilizan la tecnología y los entornos culturales adecuados para mejorar el capital intelectual y el desempeño de una organización” .

Figura 2-3: Dimensiones de la gestión del conocimiento



Fuente: Jashapara (2004, p. 12)

Por lo tanto, el logro de una sinergia entre las cuatro dimensiones para gestionar el conocimiento permitirá a la organización obtener beneficios como mejorar las competencias, acelerar la innovación, empoderar a los empleados, aumentar la calidad de los productos y/o servicios, fortalecer el compromiso organizacional y obtener una ventaja competitiva sostenible (Becerra-Fernandez & Sabherwal, 2014).

Sin embargo, para obtener los beneficios antes mencionados se debe enfocar el proyecto de gestión del conocimiento de acuerdo a unos objetivos específicos que lleven al éxito su implementación. En este sentido, Davenport *et al.* (1998) sugieren que existen cuatro objetivos en los cuales se debe enfocar un proyecto de gestión del conocimiento, los cuales presentamos a continuación:

- 1) *Crear repositorios de conocimiento*: este tipo de proyectos tratan al conocimiento como algo a ser capturado y almacenado, y consiste en una base de datos con documentos, reportes o artículos, donde los individuos o la organización en

general registran sus experiencias sobre un problema o solución en particular, para luego ser comentado y evaluado mediante medios electrónicos por los demás miembros de la organización.

- 2) *Mejorar el acceso al conocimiento*: este tipo de proyectos se basa en la idea de encontrar a la persona indicada con el conocimiento indicado para solucionar problemas específicos que se presenten, permitiendo la conexión de los empleados con una red de expertos internos o externos que brinden soporte y transferencia del conocimiento a la organización.
- 3) *Mejorar el entorno del conocimiento*: en este tipo de proyectos se busca propiciar un ambiente que conduzca a una creación, transferencia y uso más efectivo del conocimiento a través de iniciativas que desarrollen la receptividad cultural de la organización acerca del conocimiento, con el fin de mejorar los procesos de gestión del conocimiento.
- 4) *La gestión del conocimiento como un activo*: este tipo de proyectos tratan al conocimiento como un activo o recurso más de la organización, en donde se suele medir su impacto en el desempeño de la organización, y es parte de los reportes anuales de la misma. Al ser el conocimiento tratado como un activo, la organización debe buscar formas de incrementar el uso efectivo del conocimiento para lograr mejorar sus beneficios.

En base a los cuatro objetivos anteriormente descritos, Davenport *et al.* (1998) establecen que cada objetivo de un proyecto de gestión del conocimiento no es algo aislado, y lo normal es encontrar que una organización se centra principalmente en un objetivo pero con rasgos de otros objetivos que complementen al objetivo principal. El autor (*ibíd.*), además, sugiere que las iniciativas de gestión del conocimiento con múltiples objetivos serán más efectivas que aquellas con uno solo, ya que, como mencionó Jashapara (2004), la gestión del conocimiento se beneficia al utilizar una visión integradora.

Otro aspecto a tener en cuenta para lograr proyectos de gestión del conocimiento exitosos consiste en obtener las capacidades organizacionales necesarias para aprovechar el conocimiento existente y crear nuevo conocimiento que provea ventajas competitivas. En este sentido, Gold *et al.* (2001, p. 187) sugieren que, en primer lugar, la organización debe desarrollar una “capacidad de absorción”, concepto propuesto por Cohen & Levinthal (1990, p. 128) y que definen como “la habilidad de utilizar conocimiento previo para

reconocer el valor de la nueva información, asimilarlo, y aplicarlo para crear nuevos conocimientos y capacidades”.

Sin embargo, existen ciertas capacidades, además de la de absorción, que debe tener la organización para lograr promover el conocimiento y su gestión. Gold *et al.* (2001) las dividen en dos grandes grupos: capacidades para el proceso de gestión del conocimiento, el cual se tratará en la siguiente sección, y las capacidades de infraestructura, las cuáles presentamos a continuación:

- 1) *Tecnología*: corresponde a la infraestructura tecnológica que permite y facilita el enlace de la información y las comunicaciones entre los miembros internos y externos de la organización para crear, transferir y usar conocimiento. Entre las tecnologías que forman parte de una gestión del conocimiento efectiva encontramos: inteligencia de negocios, redes colaborativas, aprendizaje distribuido y aprendizaje profundo.
- 2) *Estructura*: corresponde a la estructura organizativa que posea la organización y a su capacidad de promover la colaboración y el intercambio de conocimiento. Las estructuras organizativas tradicionales tienden a promover un comportamiento individualista respecto al conocimiento, por lo que es deseable una estructura organizativa no jerárquica y auto-organizada, o bien una estructura jerárquica con la inclusión de una dimensión de flexibilidad.
- 3) *Cultura*: es considerado el principal obstáculo para una efectiva gestión del conocimiento, y corresponde a la capacidad de la organización para promover la interacción de sus empleados con el fin de crear y transmitir el conocimiento tácito y explícito, para luego pasarlo de un nivel individual a un nivel organizacional. Por lo tanto, para promover una cultura de interacción y conocimiento la organización debe tener una visión corporativa y un sistema de valores corporativos que incentiven el conocimiento en todas las áreas de la organización.

Como podemos notar, los objetivos y las capacidades que promueven una efectiva gestión del conocimiento son aspectos que pueden ser controlados por la organización. Sin embargo, la gestión del conocimiento también puede verse afectada por aspectos externos e internos no controlables que afectan a su correcto desarrollo. En base a esto, Becerra-Fernandez y Sabherwal (2015) establecen que existen cuatro “fuerzas” que conducen a la

gestión del conocimiento: 1) la primera corresponde al incremento en la complejidad del conocimiento requerido en los procesos, debido al incremento acelerado de la competencia y las tecnologías; 2) la segunda corresponde a la volatilidad acelerada de los mercados que afecta a todos los procesos de las organizaciones; 3) la tercera corresponde a la velocidad intensificada de respuesta, debido al desarrollo acelerado de las tecnologías, donde la toma de decisiones basada en conocimiento debe realizarse en tiempo real; 4) y la cuarta corresponde a la rotación de los empleados por factores internos y externos, donde el conocimiento importante para la organización puede irse con los empleados.

La gestión del conocimiento ha presentado una evolución importante para adaptarse a los tiempos de hoy en día, y desde el año 2010 se considera que la gestión del conocimiento se encuentra en su cuarta generación (Gaviria-Marin et al., 2019), en donde el foco de la investigación actual aborda las complejidades de la gestión del conocimiento en base a los avances en los diferentes aspectos que hemos mencionado anteriormente. Además, la gestión del conocimiento está adquiriendo un enfoque más como un proceso social que como método de gestión, en donde se incentiva la creación e intercambio voluntario de conocimiento dentro de la organización, entre organizaciones y socios externos de la misma (Gaviria-Marin et al., 2019; Tzortzaki & Mihiotis, 2014).

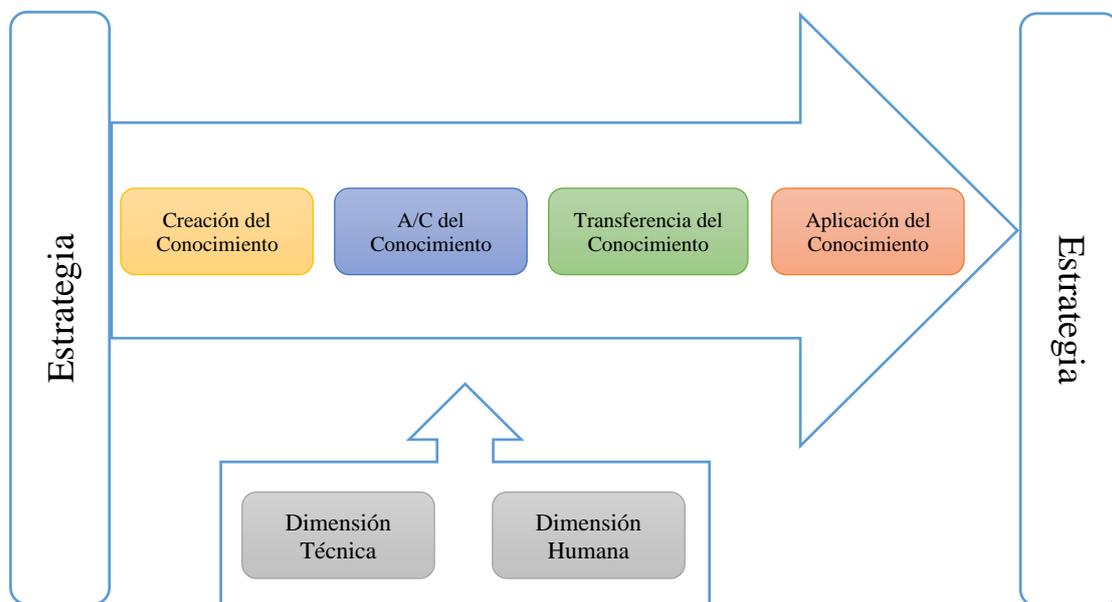
Teniendo claro el concepto, objetivos y aspectos a tener en cuenta de la gestión del conocimiento, y de acuerdo al interés de nuestra investigación, en la siguiente sección exploramos el proceso de gestión del conocimiento, describiendo como es el proceso general para gestionar el conocimiento en una organización.

2.1.3 Procesos de la Gestión del Conocimiento

Como discutimos en la sección anterior, la gestión del conocimiento requiere, además de las capacidades de infraestructura, una serie de procesos que permitan obtener beneficios del conocimiento (Gold et al., 2001). En la literatura es común encontrar muchas propuestas sobre los procesos que componen la gestión del conocimiento. Sin embargo, comúnmente estos procesos se pueden clasificar en cuatro grupos principales correspondientes a los procesos de creación de conocimiento, almacenamiento/consulta de conocimiento, transferencia del conocimiento y aplicación del conocimiento (Alavi &

Leidner, 2001; Jashapara, 2004). Como notaremos en cada uno de estos procesos, la gestión del conocimiento son procesos apoyados desde las dimensiones humanas y técnicas de la organización, siendo importante también la estrategia organizativa para promoverlos del mismo modo que estos procesos apoyan la estrategia organizativa (ver figura 2-4) (Balbastre, 2003). Los procesos de la gestión del conocimiento serán descritos a continuación.

Figura 2-4: Procesos de gestión del conocimiento



Fuente: elaboración propia

2.1.3.1 Creación del conocimiento

El proceso de creación del conocimiento es el proceso más descrito en la literatura, encontrándose en diversos artículos como adquisición, búsqueda, generación o captura del conocimiento (Gold et al., 2001). Este proceso tiene como objetivo crear nuevo contenido o reemplazar el existente para poder acumular conocimiento valioso para la organización, logrado a través de las interacciones sociales de los individuos y los procesos cognitivos individuales (Alavi & Leidner, 2001; Gold et al., 2001; Nonaka, 1994; Yu *et al.* , 2017).

El proceso de creación del conocimiento se centra en las interacciones entre las dimensiones tácita y explícita del conocimiento que forman la espiral del conocimiento en los niveles individual, grupal y organizacional, creándose los cuatro modos de conversión del conocimiento: socialización, combinación, externalización e

internalización (ver sección 2.1.1) (Alavi & Leidner, 2001; Jashapara, 2004; Nonaka, 1994; Rezaei *et al.*, 2018).

Una de las formas más útiles en la literatura de promover un ambiente que facilite y acelere la creación del conocimiento es mediante el concepto de “Ba” (lugar o espacio para crear conocimiento) (Alavi & Leidner, 2001), concepto que fue propuesto por Nonaka & Konno (1998), en donde se establecen cuatro tipos de “Ba” correspondientes a cada uno de los cuatro modos de conversión del conocimiento (Alavi & Leidner, 2001; Balbastre, 2003; Nonaka & Konno, 1998):

- 1) Ba Originador: este Ba se relaciona con la socialización y es donde el proceso de creación del conocimiento comienza, y corresponde a un lugar donde los individuos comparten emociones, experiencias, sentimientos y modelos mentales, y tiene como fin eliminar barreras entre el mismo individuo y los demás miembros de la organización.
- 2) Ba de interacción: este Ba se relaciona con la externalización, y corresponde a un lugar donde el conocimiento tácito se convierte en explícito mediante el diálogo de los modelos mentales entre individuos, donde operan dos procesos en conjunto: los individuos comparten sus modelos mentales con otros y a su vez los individuos reflejan y analizan su propio modelo mental.
- 3) Ciber Ba: este Ba se relaciona con la combinación, y corresponde a un espacio virtual donde se presenta colaboración mediante sistemas de información para el trabajo en red, documentación, bases de datos, repositorios de conocimiento, entre otros.
- 4) Ba de acción: este Ba se relaciona con la internalización, y corresponde a un espacio para el aprendizaje individual activo y continuo.

Como podemos notar, el uso de los Ba permitirá a la organización acelerar la creación del conocimiento. No obstante, se deben tener en cuenta adicionalmente una serie de impulsores que provocan los cambios en los modos de conversión (Nonaka, 1994). En primer lugar, se crea un equipo o campo de interacción para compartir experiencias (socialización); en segundo lugar, se realizan rondas de diálogo para la reflexión colectiva (externalización); en tercer lugar, el conocimiento creado por los equipos puede ser combinado con conocimiento externo o existente para crear un conocimiento más

estructurado (combinación); y en cuarto lugar, el conocimiento creado es concreto e impulsa la experimentación mediante el aprender haciendo (internalización) (Balbastre, 2003; Cairó-Battistutti & Bork, 2017; Jashapara, 2004; Nonaka, 1994; Nonaka & Takeuchi, 1995).

Al tener la organización prácticas que facilitan la creación del conocimiento, adicionalmente esta puede reforzar sus procesos de creación de conocimiento mediante fuentes externas en forma de relaciones y colaboraciones con otras organizaciones, las alianzas estratégicas, la compartición de tecnología, y el intercambio de personal (Gold et al., 2001). En base al concepto del proceso de creación del conocimiento, a continuación, exploramos cómo se lleva a cabo el almacenamiento o recuperación del conocimiento creado.

2.1.3.2 Almacenamiento/consulta del conocimiento

Tal y como vimos en la figura 2-4 anterior, un segundo subproceso de la gestión del conocimiento es el que tiene que ver con su almacenamiento y consulta. Este subproceso está orientado a evitar la pérdida del conocimiento que ha sido creado en la organización. Por este motivo, el proceso de almacenamiento/recuperación del conocimiento (a partir de ahora “A/C-C”) tiene que ver con las actividades de almacenar, organizar y consultar el conocimiento acumulado de la organización, también llamado memoria organizacional (Alavi & Leidner, 2001; Costa & Monteiro, 2018). De no llevarse a cabo un adecuado proceso de A/C-C para mantener la memoria organizacional, se pueden generar grandes pérdidas de conocimiento y, de acuerdo con Huber (1991), la pobre memoria organizacional surge por tres principales motivos: la pérdida de empleados, la no anticipación de necesidades futuras de información y conocimiento, y que los miembros de la organización con necesidades de información y conocimiento no conozcan la existencia del conocimiento dentro de la organización o desconozcan qué miembro de la misma posee el conocimiento que necesitan.

Por ende, podemos considerar que el proceso de A/C-C se encarga de gestionar la memoria organizacional, incluyendo la gestión de la información estructurada en bases de datos, el conocimiento humano codificado en sistemas expertos, procesos y procedimientos documentados de la organización, y el conocimiento táctico de los individuos y de las redes de individuos (Alavi & Leidner, 2001). Adicional a lo anterior,

Alavi y Leidner (2001) añaden que la memoria organizacional puede clasificarse en dos tipos: la primera corresponde a la memoria organizacional semántica, que se refiere al conocimiento explícito y articulado (por ejemplo, un reporte de desempeño anual), y la segunda corresponde a la memoria organizacional episódica, que se refiere al conocimiento específico en un contexto dado (por ejemplo, el conocimiento necesario para tomar decisiones dada una situación o problema específico).

Por lo tanto, el proceso de A/C-C son mecanismos que facilitan a los miembros de la organización mantener y consultar la memoria organizativa (Gold et al., 2001), y por lo general, es un proceso apoyado por sistemas y tecnologías de la información (Jashapara, 2004), siendo utilizado como soporte en la consulta del conocimiento a través de repositorios y bases de datos multimedia, o sistemas de comunicación entre expertos e individuos dentro y fuera de la organización, influenciando positivamente a la memoria organizativa al mantener centralizado y accesible el conocimiento (Alavi & Leidner, 2001; Costa & Monteiro, 2018; Lim *et al.*, 2017).

2.1.3.3 Transferencia del conocimiento

Como hemos notado anteriormente, el conocimiento es omnipresente y se puede encontrar en todos los niveles de la organización, y dada esta naturaleza distribuida del conocimiento, el proceso de transferencia del conocimiento juega un papel importante al encargarse de distribuir el conocimiento de los individuos o grupos a los lugares que se necesita y a las personas que lo necesitan (Alavi & Leidner, 2001; Shujahat et al., 2019). El proceso de transferencia del conocimiento es impulsado por los procesos de comunicación y los flujos de información, y pueden ocurrir en diferentes niveles dentro de la organización: entre individuos, de individuos a fuentes de conocimiento explícito, de individuos a grupos, entre grupos, y de grupos a la organización (Alavi & Leidner, 2001; Gil & Carrillo, 2016).

De acuerdo con Gupta y Govindarajan (2000), el proceso de transferencia de conocimiento tiene su base en los ocho elementos básicos que tiene el proceso de comunicación: un mensaje, un remitente, un esquema de codificación, un canal, transmisión a través del canal, un esquema de decodificación, un destinatario y la asignación de significado al mensaje decodificado. En base a estos elementos, los autores

(*ibíd.*) establecen que el proceso de transferencia del conocimiento se conceptualiza en los cinco factores expuestos a continuación:

- 1) El valor del stock del conocimiento de la fuente: este factor contempla que a medida que el nivel de valor que tiene el conocimiento que posee la fuente sea mayor, mayor será el atractivo de ese conocimiento para otras unidades de la organización.
- 2) Disposición motivacional de la fuente: este factor contempla que a medida que el valor del conocimiento de la fuente sea mayor, habrá una tendencia a acaparar ese conocimiento. Por lo tanto, la organización debe promover la motivación y disposición de los miembros de la organización para mejorar la transferencia de conocimiento.
- 3) Existencia y riqueza de los canales de transmisión: este factor contempla que es necesario canales de transmisión para que el conocimiento pueda ser transferido, y su característica más importante debe ser la riqueza de los enlaces de comunicación.
- 4) Disposición motivacional del receptor: este factor contempla que pueden existir mecanismos de defensa del receptor al nuevo conocimiento al considerarlo poco valioso, actitud usualmente relacionada con el ego o la resistencia al cambio.
- 5) Capacidad de absorción del receptor: este factor contempla que a lo largo de la organización existen diferentes niveles de capacidad de absorción en los receptores, lo que implica que la capacidad para reconocer conocimiento valioso de un receptor puede afectar la capacidad de internalizar y asimilar conocimiento de la organización (Cohen & Levinthal, 1990).

En base a estos factores, los autores Alavi & Leidner (2001) destacan que el tercer factor es el más estudiado en la literatura (Argote *et al.*, 2003; Inkpen & Tsang, 2005; Jashapara, 2004; Khachlouf & Quélin, 2018; Lombardi, 2019; Szulanski, 2000), y que a su vez es común encontrarlo dividido en cuatro tipos de canales de transmisión (o transferencia):

- 1) Los canales informales, como los seminarios y conversaciones informales, se caracterizan por promover la socialización, pero puede impedir la diseminación del conocimiento, y suelen ser más efectivos en pequeñas organizaciones.

- 2) Los canales formales, como las jornadas de capacitación, mejoran la distribución del conocimiento, pero pueden inhibir la creatividad.
- 3) Los canales personales, como las prácticas en empresa o la transferencia en persona, se caracterizan por ser más efectivos para distribuir conocimiento específico en un contexto dado.
- 4) Los canales impersonales, como los repositorios de conocimiento, se caracteriza por ser más efectivo en distribuir conocimiento generalizado.

Teniendo en cuenta lo anterior, el tipo de transferencia de conocimiento utilizado en una organización dependerá de los diferentes factores culturales, tecnológicos y de infraestructura que posean, y a su vez, el valor del proceso de transferencia del conocimiento dependerá del tipo de conocimiento que se esté transfiriendo (Alavi & Leidner, 2001; De Zubielqui *et al.*, 2019), la efectividad con la que se realicen los procesos de creación y de almacenamiento/consulta del conocimiento, ya que de estos procesos dependerá la disponibilidad y accesibilidad del conocimiento en la organización, y la capacidad de absorción individual en las relaciones interorganizacionales (Khachlouf & Quélin, 2018).

2.1.3.4 Aplicación del conocimiento

El último subproceso de gestión del conocimiento es el que tiene que ver con su aplicación (ver figura 2-4 más arriba). La aplicación del conocimiento es considerada como el proceso que permite obtener ventajas competitivas del conocimiento (Abubakar *et al.*, 2019; Alavi & Leidner, 2001), ya que es el proceso encargado de poner en práctica el conocimiento que ha sido creado, almacenado/consultado y transferido en la organización (Shujahat *et al.*, 2019). Por ende, el proceso de aplicación del conocimiento se centra principalmente en el uso y aplicación del conocimiento, y tiene como objetivo convertir el conocimiento en procesos y en comportamiento organizativo para la solución de problemas, ajustar la dirección estratégica de la organización, y mejorar la eficiencia (Cegarra-Navarro & Martelo-Landroguez, 2020; Gold *et al.*, 2001; Shujahat *et al.*, 2017; Teixeira *et al.*, 2018).

De acuerdo Grant (1996), existen cuatro mecanismos para poder aplicar el conocimiento en la organización, los cuales describimos a continuación (Alavi & Leidner, 2001; Grant, 1996):

- 1) Reglas y directrices: corresponde a los estándares, reglas y directrices de la organización que regulan las interacciones entre individuos y proveen los medios para convertir fácilmente el conocimiento tácito en conocimiento explícito.
- 2) Secuenciación: corresponde a una de las formas más simples de aplicar el conocimiento, y consiste en organizar el conocimiento en actividades en una secuencia de tiempo preestablecida, para luego cada individuo recibir el conocimiento en diferentes periodos de tiempo y en forma secuencial.
- 3) Rutinas: corresponde al desarrollo de patrones para coordinar las tareas, así como los protocolos para la interacción y establecer especificaciones que le permiten a las personas aplicar e integrar su conocimiento sin la necesidad de articular y comunicar lo que saben a otros (Alavi & Leidner, 2001; Becker *et al.*, 2005).
- 4) Grupos de resolución de problemas y toma de decisiones: corresponde a una forma de aplicar el conocimiento de una forma más personal, y consiste en crear grupos con una comunicación intensiva, en donde se utilicen reglas, directrices, secuencias y rutinas para integrar el conocimiento con el objetivo de solucionar problemas complejos y lograr tomar decisiones.

Como hemos notado en los procesos de gestión del conocimiento anteriores, el éxito del proceso de aplicación del conocimiento de una organización dependerá de muchos factores organizativos. Sin embargo, para el caso específico de la aplicación de conocimiento, la investigación de Song *et al.* (2005) considera que principalmente la aplicación del conocimiento en una organización dependerá del presupuesto en investigación y desarrollo de la organización, en los procedimientos formales que utiliza la organización para evaluar la efectividad de la creación conocimiento, en orientar la estrategia del conocimiento de la organización a futuro, en la disponibilidad y nivel de inversión en sistemas y tecnologías de información que apoyen el conocimiento, y la superposición de la habilidades y recursos de conocimiento entre los diferentes departamentos y niveles de la organización.

Una vez descrito los conceptos correspondientes al conocimiento, la gestión del conocimiento y el proceso de gestión del conocimiento, en la siguiente sección describimos los conceptos correspondientes al diseño organizativo y los mecanismos de coordinación y control, segundo tema de interés para el desarrollo de esta investigación.

2.2 Diseño Organizativo y los Mecanismos de Coordinación y Control

Un aspecto importante en esta investigación para comprender cómo el Big Data impacta en las organizaciones es evaluar cómo el diseño organizativo y los mecanismos de coordinación y control deben modificarse para adaptarse a los cambios que conlleva el Big Data. Por este motivo, en las siguientes dos secciones describimos estos conceptos clave y aquellos aspectos que se tendrán en cuenta para el desarrollo de esta investigación.

2.2.1 Concepto de Diseño Organizativo

El diseño organizativo es un marco de referencia que surge de la investigación en los campos de la organización y la administración entre los años 1960 y 1970. No obstante, ha adquirido nuevamente interés en los últimos años debido a las nuevas y complejas formas de organización que han surgido debido al auge de ecosistemas complejos, nuevas configuraciones y estructuras, y formas organizativas distribuidas que en ocasiones abarcan a más de una organización (Good *et al.*, 2018). Por ende, con el diseño organizativo se busca la efectividad, la eficiencia y la agilidad que la organización necesita para mantenerse acorde a los entornos dinámicos del mercado, en especial en aquellos sectores intensivos en conocimiento (Fjeldstad & Snow, 2018; Good *et al.*, 2019).

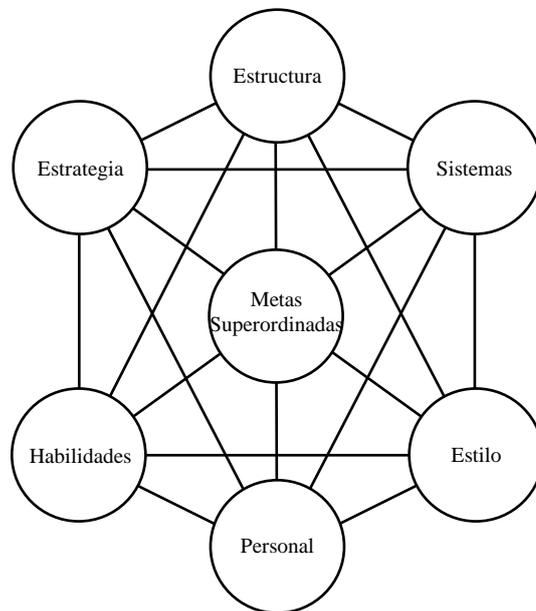
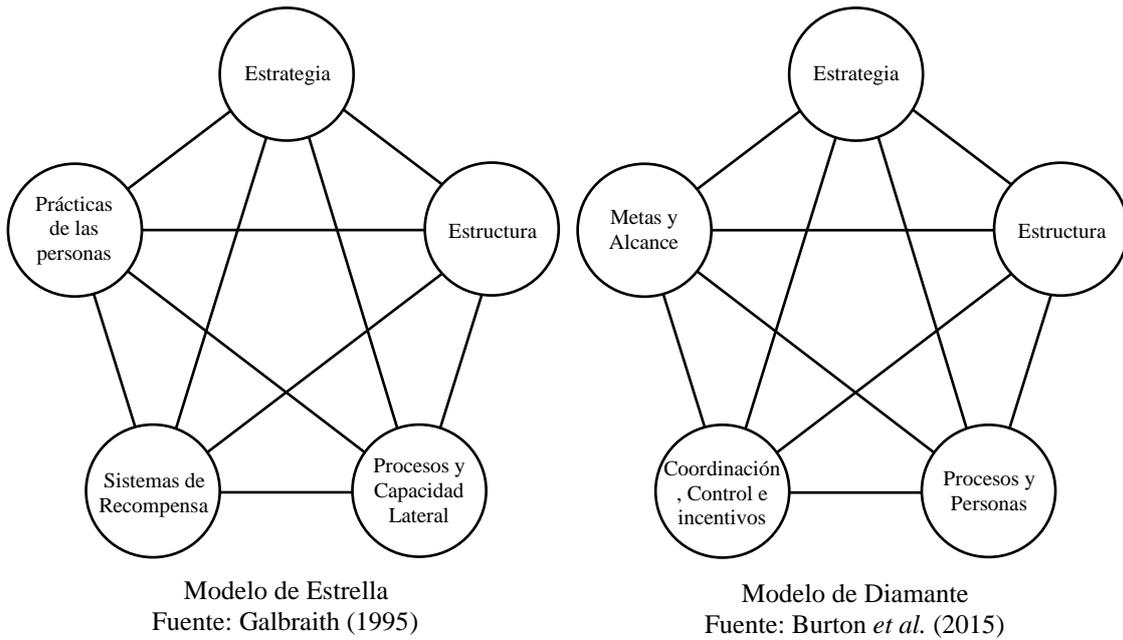
Para lograr lo anterior, el diseño organizativo se enfoca principalmente en resolver dos problemas complementarios (Burton *et al.* 2015): 1) cómo repartir una tarea grande de toda la organización en tareas más pequeñas y en subunidades, y 2) cómo coordinar estas tareas más pequeñas de las subunidades para que se acoplen entre sí y poder cumplir eficientemente la tarea más grande o los objetivos de la organización. Por lo tanto, el diseño organizativo se centra en construir y cambiar factores organizativos para facilitar la construcción y el logro de los objetivos de la organización, construir patrones de división del trabajo y la coordinación y comunicación de las personas que realizan el trabajo (Good *et al.*, 2019; Robbins, 1990).

En ocasiones el término diseño organizativo es relacionado como un símil al término de estructura (Galbraith *et al.*, 2001). Sin embargo, y como notamos anteriormente, el diseño organizativo abarca muchos más factores de la organización, por lo que distintos autores en la literatura han propuesto modelos holísticos para describir los elementos o factores que abarca el diseño organizativo (Miterev *et al.*, 2017).

Para autores como Good *et al.* (2018), Jonker *et al.* (2012) y Miterev *et al.* (2017), entre las propuestas más destacadas se encuentran: a) el trabajo de Galbraith (1995) y su *modelo de estrella*, el cual define al diseño organizativo como “el proceso deliberado de configurar la estructura, procesos, sistemas de recompensa y políticas y prácticas de las personas para crear una organización efectiva capaz de lograr la estrategia de la organización” (Galbraith *et al.*, 2001, p. 2); b) el trabajo de Mintzberg (1979), el cual agrupa el diseño organizativo en diseño de puestos, diseño de la superestructura, diseño de los enlaces laterales y el diseño de sistemas de toma de decisiones (Mintzberg, 1979); y c) el trabajo de Waterman *et al.* (1980) con el modelo de las 7S, donde se establece que un cambio organizativo efectivo proviene de la relación entre la estructura, la estrategia, los sistemas, el estilo de dirección, las habilidades, el personal y las metas superordinadas (ver figura 2-5).

Teniendo en cuenta lo anterior, en la literatura se puede encontrar propuestas de modelos que surgen a partir los trabajos antes mencionados. Por ejemplo, una de estas propuestas es la que realizan Burton *et al.* (2015), con su modelo llamado “modelo de diamante” (ver figura 2-5), en donde se plantea un modelo basado en parte en el *modelo de estrella* de Jay Galbraith y en donde se consideran cinco componentes: metas y alcance, estrategia, estructura, procesos y personas y coordinación, control e incentivos.

Figura 2-5: Comparación entre modelos de diseño organizativo



Fuente: elaboración propia a partir de Burton *et al.* (2015, p. 9), Galbraith (1995, p. 12) y Waterman *et al.* (1980, p. 18)

Como notamos en la figura 2-5, las distintas propuestas de modelos de diseño organizativo sufren de una superposición de los elementos que se consideran importantes, por lo que Good *et al.* (2018) proponen cuatro elementos principales que son consistentes en la literatura: el propósito de la organización, correspondiente a las metas y objetivos

trazados por la organización; las actividades, correspondiente a las tareas que desempeñan los empleados para cumplir el propósito de la organización; la estructura, que corresponde a la forma de gobierno, la estructura organizativa interna, el tamaño de la organización y la ubicación física de la organización; y las personas y la cultura organizacional, que corresponde a los sistemas internos de incentivos y a la cultura interna de la organización. Otro aspecto constante que podemos notar entre los modelos es la interrelación entre los elementos, los cuales representan las conexiones, alineaciones o ajustes entre los elementos, y cuya desalineación o desconexión puede provocar una disminución en el desempeño de la organización (Burton et al., 2015; Miterev et al., 2017).

Por consiguiente, al existir una superposición entre los elementos de los distintos modelos de diseño organizativo y para efectos prácticos, en esta investigación describiremos el *modelo de estrella* de Galbraith para establecer los factores importantes de diseño a considerar en esta investigación, debido a que es uno de los modelos más reconocidos y estudiados de la literatura (Miterev et al., 2017). Además de lo anterior, otra razón que nos mueve a utilizar dicho modelo en esta investigación es que Galbraith es uno de los autores que más ha estudiado el efecto del Big Data sobre el diseño organizativo, convirtiéndose así en un autor de suma relevancia en el ámbito de la presente investigación.

2.2.1.1 Modelo de Estrella de Galbraith

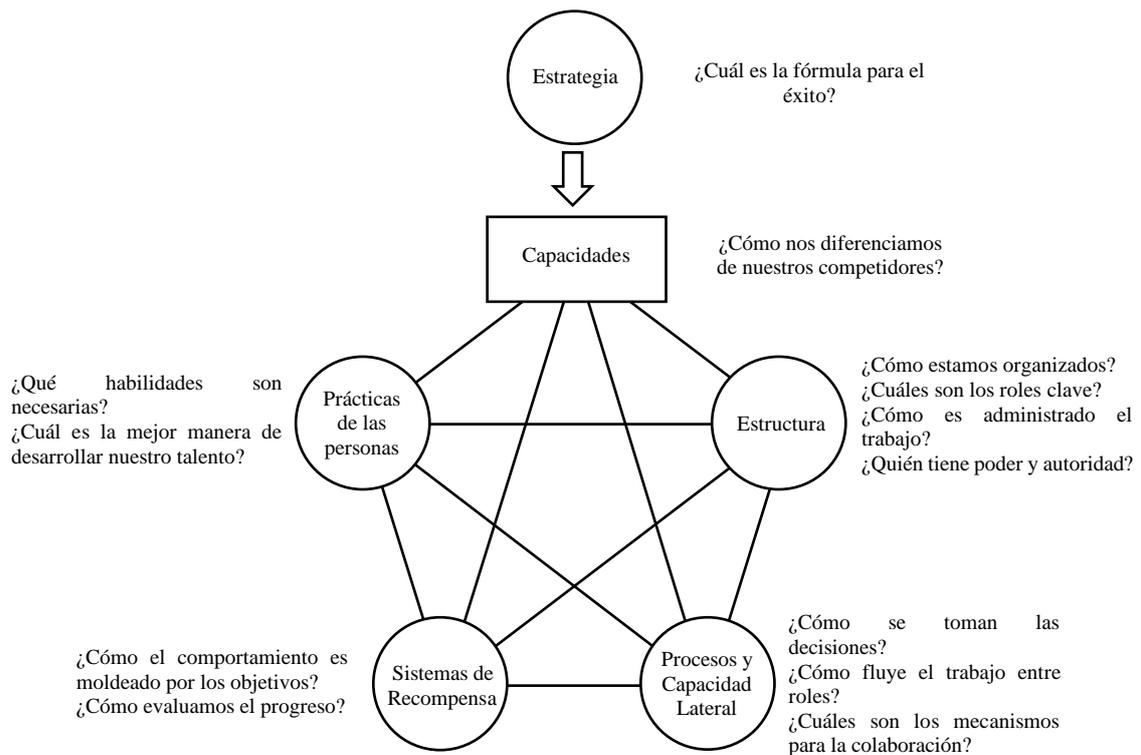
El modelo de estrella es un marco de referencia para el diseño organizativo creado por el investigador y consultor Jay R. Galbraith (1995), entendiéndose como modelo al desarrollo e implementación de procesos para la toma de decisiones basados en información, desarrollo que surge de su experiencia en dichos procesos en cadenas de suministros (Galbraith, 2014a). El modelo surge al presentarse tres problemas al momento de ser implementados los procesos de toma de decisiones en las organizaciones estudiadas (*ibíd.*):

- 1) El desempeño de los gerentes era medido en función de los costos contables y estándares y en esta base se diseñaban los sistemas de recompensa. En cambio, los modelos y procesos de toma de decisiones planteaban para la medición del desempeño los costos económicos, por lo que los gerentes se resistían a utilizar el nuevo modelo al no poseer el mismo tipo de medición.

- 2) Las decisiones eran tomadas al nivel de las plantas de fabricación. En cambio, los modelos y procesos de toma de decisiones planteaban una centralización de las decisiones, por lo que se debía modificar la estructura de la cadena de suministros.
- 3) Para poder implementar los modelos planteados de toma de decisiones, se debía contratar personal con las habilidades cuantitativas necesarias para operar con el nuevo modelo, y pudieran tomar decisiones óptimas para toda la cadena de suministros.

Por lo tanto, analizando aquellos factores o elementos que la organización debía modificar para implementar los nuevos procesos de toma de decisiones basados en información, Galbraith concluyó que se debía hacer cambios sustanciales en las mediciones de desempeño y los sistemas de recompensa, la estructura, y en el conjunto de habilidades y la mentalidad de las personas, por lo que planteó el modelo de estrella para poder ver a la organización desde una perspectiva sistémica u holística (ver figura 2-6) (Galbraith, 1995, 2014a).

Figura 2-6: Modelo de estrella de Galbraith



Fuente: adaptado de Galbraith (1995, p. 12) y Kates & Galbraith (2007, p. 3).

Como notamos en la sección anterior, la cultura organizativa hace parte de muchos modelos o marcos de referencia de diseño organizativo (Good et al., 2019). Sin embargo, el modelo de estrella no lo considera de forma explícita debido a que los líderes de la organización no pueden influir directamente en la cultura. Por lo tanto, consideran que la cultura es el resultado acumulativo de las decisiones de diseño organizativo que se han hecho en el pasado, así como las conductas de liderazgo y de gestión que resultan de esas decisiones (Kates & Galbraith, 2007).

Teniendo en cuenta el modelo propuesto por (Galbraith, 1995), a continuación, describimos cada uno de los elementos que se consideran importantes y aquellos aspectos organizativos que toma en cuenta cada elemento.

Estrategia

La estrategia corresponde a las metas y objetivos de la organización, y establece la dirección que tomará la organización en el corto, mediano y largo plazo, con el propósito de obtener ventajas competitivas y diferenciarse de sus competidores (Galbraith et al., 2001). La estrategia surge por medio de las capacidades organizativas (habilidades, procesos, tecnologías y habilidades del personal inherentes de cada organización) y del análisis y la comprensión de los factores que afectan al desempeño de la organización (e.g., los competidores, proveedores, clientes y tecnologías emergentes), los cuales delimitan los productos, servicios y mercados en los que se enfocará la organización (Galbraith et al., 2001; Kates & Galbraith, 2007). Por ende, la estrategia es considerada la piedra angular del proceso de diseño organizativo ya que, del conocimiento de la estrategia, surgen las decisiones importantes que se deben tomar acerca de la estructura, los sistemas de recompensa, los procesos y las prácticas de las personas de la organización (Galbraith, 1995, 2014a).

Estructura

La estructura determina la distribución de poder y autoridad en la organización, y abarca los componentes de la organización, sus relaciones y la forma de jerarquía utilizada (Galbraith et al., 2001). Por lo tanto, la estructura se puede definir como las diferentes formas de dividir el trabajo en múltiples tareas distribuidas en la organización, y a la vez conseguir una coordinación entre estas tareas (Mintzberg, 1993). En esta línea, las

decisiones de diseño organizativo que debe tomar una organización referentes a la estructura se dividen en cuatro dimensiones (Galbraith, 1995, p. 13):

- ✓ División del trabajo: se refiere al tipo y grado de especialización de los trabajos necesarios para llevar a cabo las tareas de la organización.
- ✓ Forma de la organización: número de personas que constituyen los departamentos en cada nivel, y el número de niveles que posee la estructura.
- ✓ Distribución del poder: abarca dos dimensiones, la dimensión vertical se refiere a la centralización o descentralización de las decisiones en la organización, y la dimensión lateral se refiere al movimiento de poder directamente al departamento que se encuentra lidiando con los problemas críticos para la misión de la organización.
- ✓ Departamentalización: es la base para formar los departamentos en cada nivel de la organización, y se forman en base a las funciones, productos, mercados o regiones que atienden. Los departamentos son normalmente configurados en una jerarquía para la gestión y la toma de decisiones.

Por consiguiente, las decisiones de diseño referentes a la estructura organizativa tienen como objetivo establecer las relaciones jerárquicas, definir la distribución de poder y establecer los canales de comunicación necesarios en la organización (Kates & Galbraith, 2007; Mintzberg, 1979; Miterev et al., 2017). Por lo tanto, la organización debe establecer un tipo de estructura que le permita cumplir sus objetivos y metas de acuerdo con sus capacidades y necesidades. Dentro de los principales tipos de estructuras que puede adoptar una organización destacamos los siguientes (Galbraith, 2014a; Kates & Galbraith, 2007):

- ✓ *Organización funcional*: son más comunes en las organizaciones que poseen un número grande de empleados, y corresponde a una forma de estructura en donde los empleados son organizados en grupos en base a su actividad o tarea (e.g. finanzas, producción o marketing). Este tipo de estructuras promueven la estandarización, una mayor especialización de los empleados, la compartición de conocimiento en un mismo grupo y crean economías de escala.
- ✓ *Organización por productos o unidades de negocio*: son más comunes en organizaciones que se mueven frecuentemente entre diferentes mercados o líneas de producto o servicio cuando el crecimiento del negocio central de la

organización decrece, y corresponde a un tipo de estructura en la que se crean unidades de negocio para cada línea de producto o servicio de la organización, reduciendo los ciclos de desarrollo de productos o servicios, mejorar e innovar en los productos o servicios, y promover la especialización de los empleados en la unidad de negocio a la que pertenecen.

- ✓ *Estructura geográfica:* son más comunes en organizaciones que se expanden hacia nuevos territorios, y corresponde a un tipo de estructura donde las tareas y operaciones se dividen de acuerdo con la ubicación geográfica y donde se necesite una especialización de los empleados en factores como la cultura, el lenguaje, factores políticos y cambios en los patrones de consumo de una región, ciudad o país.
- ✓ *Unidades de negocio basadas en el cliente:* son más comunes en organizaciones que necesiten especializarse en los diferentes segmentos de clientes que poseen, y corresponde a un tipo de estructura similar a la organización por productos con la diferencia de que se crean grupos o unidades de negocio en base a los clientes con necesidades, características o patrones de compra similares.
- ✓ *Canales:* corresponde a un tipo de estructura en donde la organización divide sus actividades y empleados en base a los diferentes canales que poseen para distribuir sus productos (e.g. internet, por catálogo o tiendas físicas).
- ✓ *Estructuras híbridas:* corresponde a un tipo de estructura donde se combinan dos o más tipos de estructura.
- ✓ *Organizaciones matriciales:* corresponde a un tipo de estructura donde la organización es ordenada alrededor de dos dimensiones simultáneamente: funciones (e.g. ingeniería, manufactura o calidad) y centros de ganancia. En este tipo de estructuras los empleados poseen dos o más líderes o jefes, donde cada uno representa un objetivo específico de la organización, y se caracteriza porque cada líder posee una mayor autoridad para establecer objetivos, resolver problemas, administrar el desempeño y tomar decisiones. Este tipo de estructuras se pueden encontrar en organizaciones con estructuras organizativas maduras y con procesos rigurosos, por lo que en la actualidad es más frecuente encontrarlo en grandes organizaciones.

Al momento de seleccionar un tipo de estructura una organización se debe tomar en cuenta las cuatro dimensiones antes mencionadas, ya que la división del trabajo, la forma

de la organización, la distribución de poder y la departamentalización, dependerán del tipo de estructura utilizada (Galbraith, 2014a). En la actualidad, las organizaciones tienen una tendencia hacia las estructuras más planas (i.e. menos niveles jerárquicos) y donde se promueven los cambios en la distribución del poder, mayor grado de especialización, y una descentralización de la toma de decisiones, como es el caso de las organizaciones matriciales (*ibíd.*).

Procesos y capacidad lateral

En las organizaciones existen barreras invisibles que impiden la comunicación y la colaboración efectiva entre sus miembros. Estas barreras son creadas por la propia estructura que posee la organización; es lo que se conoce como “silos organizacionales” (Galbraith et al., 2001; Kates & Galbraith, 2007). Por lo tanto, el diseño organizativo busca romper estas barreras mediante la conexión de actividades que permitan mover información de arriba a abajo y a través de toda la estructura de la organización, a lo que el modelo de estrella denomina “procesos” (Kates & Galbraith, 2007).

Entre los “procesos” que puede manejar una organización se encuentran (Galbraith, 2014a): a) los procesos informales, también denominado procesos auto-organizados, correspondiente a los comportamientos voluntarios de una persona para realizar su trabajo o llevar a cabo una conversación social con los demás miembros; b) los procesos de negocio, correspondientes a procesos programados o que pueden ser automatizados y que involucran a varias funciones de la organización, como el proceso de desarrollo de productos o el proceso de gestión de la relación con el cliente; c) y en tercer lugar, los procesos de gestión, que corresponden a los procesos encargados de asignar los recursos a las oportunidades que la organización presenta, y llevan a cabo la alineación de los objetivos en toda la organización para la ejecución de la estrategia.

Para llevar a cabo los procesos antes mencionados y romper los silos organizacionales, la organización debe desarrollar una capacidad lateral, la cual consiste en los mecanismos para la coordinación y conexión de las actividades necesarias para la ejecución de una tarea a través de distintos departamentos de la estructura organizativa (Galbraith et al., 2001; Kates & Galbraith, 2007). Por ende, la capacidad lateral permite a la organización llevar a cabo procesos interorganizacionales con el objetivo de juntar a las personas adecuadas para la resolución de problemas, crear oportunidades y responder a los retos

que se presenten (Galbraith et al., 2001). El concepto de capacidad lateral (visto como coordinación lateral) será abordado en mayor profundidad en la sección 2.2.2.2.

Sistemas de recompensa

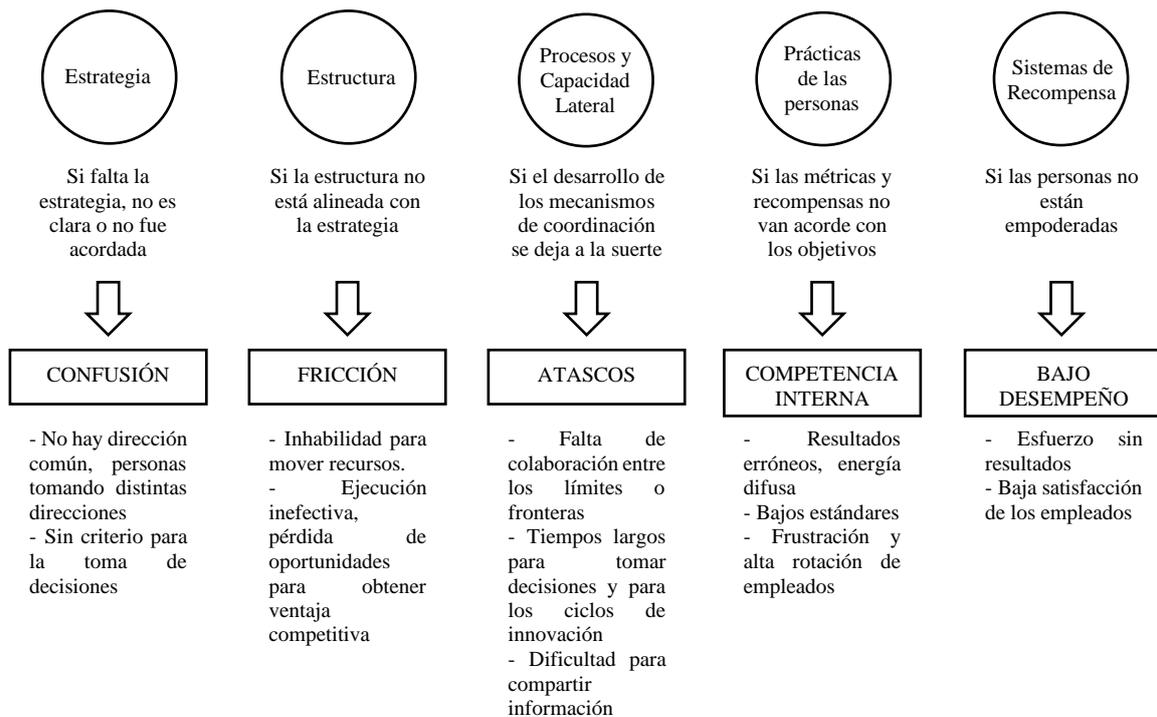
Los sistemas de recompensa son un mecanismo que tiene la organización para alinear sus objetivos con los objetivos de los empleados (Galbraith, 1995), y corresponde a la forma en cómo la organización establece un sistema que le permita comunicar a sus empleados qué es lo que la organización valora y motivar el desempeño, para lo cual se diseñan políticas que regulen los salarios, los bonos, las promociones o los ascensos, las recompensas y los sistemas de reconocimiento por el trabajo (Galbraith, 1995, 2014a). El diseño de los sistemas de recompensa tiene una influencia directa sobre los demás componentes del diseño organizativo, por lo que debe estar acorde con la estructura y los procesos que influyen sobre la dirección estratégica de la organización (Galbraith, 1995; Galbraith et al., 2001).

Prácticas de las personas

Las prácticas de las personas corresponden a todas las prácticas de recursos humanos que facilitan la selección, la dotación de personal y la formación requeridos para desarrollar las capacidades organizativas que surgen a partir de las habilidades y competencias de los empleados (Kates & Galbraith, 2007). Estas prácticas de recursos humanos deben alinearse con la estrategia de la organización lo que permitirá determinar cuáles son las competencias, habilidades y capacidades requeridas por los empleados para poder interactuar entre las barreras organizativas, participar en equipos, tomar decisiones y obtener ventajas competitivas para la organización (Galbraith et al., 2001; Kates & Galbraith, 2007).

El modelo de estrella de Galbraith, al igual que los otros modelos tratados anteriormente, hace énfasis en la alineación que deben tener los elementos que componen las decisiones de diseño organizativo, y cómo la falta de la misma puede provocar problemas en el desempeño general de la organización (Galbraith et al., 2001). En la figura 2-7 se presentan las posibles consecuencias de un diseño organizativo no alienado.

Figura 2-7: Consecuencias de la desalineación del diseño organizativo



Fuente: Galbraith et al. (2001, p. 5)

Una vez descritos los elementos de diseño organizativo que se deben tener en cuenta en una organización, explicados a través del modelo de estrella de Galbraith, a continuación, nos centraremos en el análisis de los mecanismos de coordinación y control, aspecto de interés fundamental para esta tesis doctoral.

2.2.2 Mecanismos de Coordinación y Control

La coordinación y el control son parte fundamental del diseño organizativo, y por ello son considerados por algunos autores (Child, 1973; Galbraith, 1995; Mintzberg, 1979; Van de Ven *et al.*, 1976) como un elemento clave en los modelos de diseño organizativo (ver figura 2-5 más arriba). Estos mecanismos son los encargados de asegurar el correcto funcionamiento y la alineación de las tareas mediante la integración de las diferentes partes de la organización, y tienen como finalidad el cumplimiento de las metas y objetivos que hacen parte de la estrategia de la organización (Burton et al., 2015). Sin embargo, ambos tienen enfoques diferentes para lograr el cumplimiento de la estrategia: la coordinación se encarga principalmente de la cohesión, sinergia y flexibilidad entre los diferentes grupos, unidades o elementos de la estructura de la organización, mientras que

el control busca asegurar la calidad y eficiencia de los flujos de información para regular que las actividades vayan acorde a la estrategia de la organización en todos los niveles de la misma (Burton et al., 2015; Child, 2015).

Íntimamente relacionados con la coordinación y el control se encuentran los conceptos de formalización y centralización/descentralización. Por ello, dedicamos los siguientes párrafos a explicarlos de manera sintética y, con posterioridad, ya nos centraremos en el análisis de los mecanismos organizativos que se pueden utilizar para llevar a cabo las funciones de coordinación y de control.

2.2.2.1 Formalización y centralización/descentralización

La formalización y la centralización/descentralización son decisiones de diseño organizativo que establecen la base para el diseño de los mecanismos de coordinación y control que se utilizan en la organización (Burton et al., 2015). En primer lugar, la formalización define el grado en que la organización tiene estandarizados los trabajos, a través del establecimiento de un conjunto de reglas, normas o regulaciones que determinan cómo se debe realizar el trabajo, quién lo debe realizar, y bajo qué circunstancias o restricciones se debe realizar (Burton et al., 2015; Robbins, 1990).

El grado de formalización lo determina el nivel de detalle y difusión de las reglas o políticas establecidas por la organización, en donde un alto nivel de formalización implica reglas estrictas y una difusión en toda la organización (p.ej., en organizaciones burocráticas) y un bajo nivel de formalización implica que no existen o hay pocas reglas escritas y poco o nada difundidas (Burton *et al.*, 2015; Gibson *et al.*, 2019; Mintzberg, 1979; Moreno-Luzón *et al.*, 2001; Robbins & Coulter, 2018). Cabe resaltar que dentro de una organización el nivel de formalización puede variar, dependiendo de qué roles o circunstancias necesiten más o menos regulación, por lo que a lo largo del tiempo las organizaciones suelen modificar sus niveles de formalización, buscando por lo general un equilibrio (Burton et al., 2015).

Por lo que respecta a la centralización/descentralización, esta decisión de diseño hace referencia a qué tan concentrada están la toma de decisiones y la gestión de los mecanismos de coordinación y control en la organización (Robbins, 1990). Para el caso de la centralización, la toma de decisiones y los mecanismos de coordinación y control

son manejados por una persona o un nivel específico de la organización (usualmente por los altos mandos o directivos), y suele presentarse de forma más habitual en pequeñas empresas u organizaciones burocráticas (Burton et al., 2015; Mintzberg, 1993).

En contraste, la descentralización indica que la autoridad para la toma de decisiones y para la gestión de la coordinación y control recae sobre niveles directivos intermedios, lo que implica una mayor distribución de las responsabilidades en la organización (Burton et al., 2015; Moreno-Luzón et al., 2001). En las organizaciones actuales, la descentralización es una característica deseable, ya que permite un mejor flujo de la información y una mayor flexibilidad de los empleados para desempeñar su labor; no obstante, lo anterior no implica que la centralización no pueda darse en la organización, ya que en determinadas circunstancias es preferible que ciertas decisiones estratégicas sean tomadas por los altos mandos directivos de la empresa (Burton et al., 2015; Mintzberg, 1993).

Tal y como mencionamos anteriormente, la formalización y la centralización/descentralización son conceptos intrínsecamente ligados a los mecanismos de coordinación y control debido a que en estas variables o aspectos del diseño se basan varias de las formas en cómo se coordina o controla el trabajo en las organizaciones. Así, en los siguientes apartados describiremos dichos mecanismos.

2.2.2.2 Mecanismos de coordinación

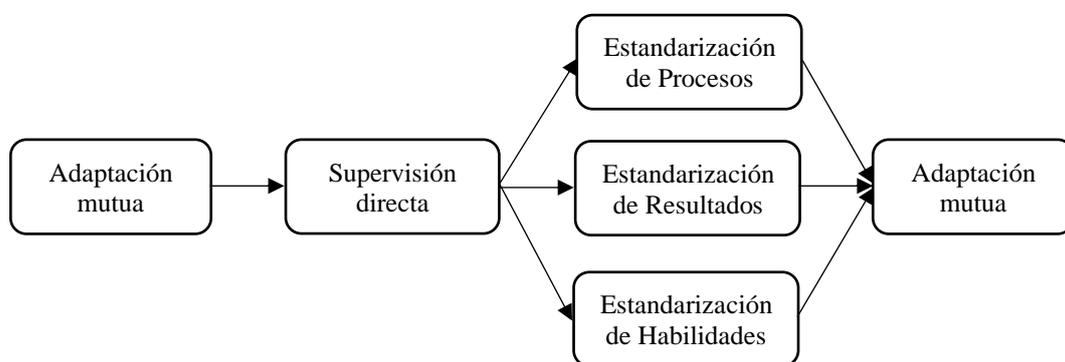
Los mecanismos de coordinación se definen como las herramientas, tecnologías y/o interacciones necesarias para ordenar y juntar a empleados y elementos interdependientes de la organización para conseguir cierto nivel de desempeño (Okhuysen & Bechky, 2009). De acuerdo con Mintzberg (1979), existen cinco tipos de mecanismos de coordinación fundamentales que una organización puede utilizar para coordinar el trabajo:

- 1) Adaptación mutua: es un mecanismo utilizado generalmente por organizaciones simples, y consiste en la coordinación del trabajo mediante un proceso de comunicación informal entre las personas, donde los empleados en principio no conocen bien las actividades que se deben desarrollar para cumplir el trabajo, y se consigue solo cuando las personas se adaptan a las otras.

- 2) Supervisión directa: la coordinación es realizada por una persona que toma la responsabilidad de coordinar el trabajo de los demás, dando instrucciones y monitoreando las actividades.
- 3) Estandarización de procesos de trabajo: la coordinación se consigue cuando el trabajo es especificado o programado en la organización, diseñándose específicamente las instrucciones que se deben seguir para cada trabajo.
- 4) Estandarización de resultados: la coordinación se consigue cuando se especifican los resultados que se deben obtener del trabajo, elaborando estándares de desempeño en la organización.
- 5) Estandarización de habilidades y conocimientos: la coordinación se consigue cuando se estandarizan las habilidades, la formación y el conocimiento que requiere un empleado para desempeñar una actividad o labor específica.

En línea con lo anterior, Mintzberg (1979) establece que a medida que la coordinación del trabajo o las actividades de la organización sean más complicadas, el mecanismo utilizado irá cambiando de forma continua en el modo que se presenta en la figura 2-8, empezando por la adaptación mutua, pasando después a la supervisión directa, luego a la estandarización y, por último, la organización pasaría nuevamente a aplicar el mecanismo de la adaptación mutua.

Figura 2-8: Mecanismos de coordinación en continuo



Fuente: Mintzberg (1979, p. 7).

Con el fin de expandir los mecanismos de coordinación antes mencionados, Child (2015) establece que en la actualidad los principales mecanismos de coordinación se pueden resumir en tres tipos: coordinación por un alto directivo, coordinación mediante procedimientos y planeación formales, y coordinación lateral.

Coordinación por un alto directivo

Corresponden a mecanismos de coordinación centralizados donde la coordinación recae principalmente en los altos directivos de la organización, y suele catalogarse como una forma económica de llevar a cabo la coordinación si se presentan las siguientes tres condiciones (Child, 2015):

- ✓ La necesidad de coordinación no es muy frecuente.
- ✓ El alto directivo debe estar en la posibilidad de comunicarse fácilmente con los empleados.
- ✓ En las situaciones que se necesite coordinación el alto directivo debe ser el más idóneo para coordinar.

Debido a la tendencia hacia organizaciones más descentralizadas, este tipo de mecanismos de coordinación suele ser no aptos para las estructuras organizativas actuales. Además, debido al aumento en la complejidad en las organizaciones, manejar la responsabilidad de la coordinación por una o pocas personas de los altos directivos suele ser muy difícil, por lo que podría no involucrarse lo necesario en la coordinación (Child, 2015).

Coordinación mediante procedimientos y planeación formales

Corresponden a mecanismos de coordinación con un alto nivel de formalización, y se dividen en tres mecanismos (Child, 2015; Mintzberg, 1993):

- ✓ Estandarización: consiste en establecer un conjunto de reglas y procedimientos que definen y programan cómo se debe desarrollar el trabajo en cada departamento o unidad de la organización.
- ✓ Planes y horarios: consiste en establecer un plan y una programación de las actividades a coordinar. Es considerado más flexible que la estandarización al ser más sencilla su modificación.
- ✓ Reuniones de comité: consiste en reunir a miembros de la organización para acordar y coordinar los procedimientos y actividades a desarrollar. En este tipo de comités se ofrece mayor flexibilidad a los empleados al ser partícipes en la toma

de decisiones. Sin embargo, estas decisiones serán luego revisadas en niveles superiores en la jerarquía.

Este tipo de mecanismos de coordinación suele encontrarse en organizaciones burocráticas; pero, a su vez, puede funcionar en organizaciones en donde las actividades son predecibles y cuya complejidad no amerita grandes esfuerzos para la coordinación (Child, 2015). Sin embargo, a medida que las organizaciones se encuentren en condiciones menos estables y requieran un mayor nivel de procesamiento de información, este tipo de mecanismos resultará obsoleto (*ibíd.*).

Coordinación lateral

Por último, la coordinación lateral es, en la actualidad, uno de los mecanismos principales para coordinar las actividades de una organización (Galbraith, 2014a). Este mecanismo consiste en integrar las contribuciones de distintas personas distribuidas en diferentes unidades o departamentos de la organización para llevar a cabo una tarea común (Child, 2015; Galbraith, 2014a; Mintzberg, 1979). Con esta integración, la organización consigue la descentralización de las decisiones mediante la creación de grupos con autonomía en la toma de decisiones, aportando agilidad para atender necesidades o problemas de la organización (Galbraith, 2014a). Según Child (2015), existen cinco tipos de coordinación lateral:

- ✓ Coordinación informal o voluntaria: consiste en la creación espontánea de grupos para la coordinación de las actividades.
- ✓ Coordinación electrónica: consiste en el uso de sistemas y herramientas informáticas para la comunicación y coordinación entre departamentos.
- ✓ Grupos formales: consiste en la creación formal de equipos multifuncionales donde se definen objetivos y se toman decisiones.
- ✓ Integrador: consiste en el uso de personal especializado (p.ej., gerentes de proyectos, gerente de procesos, gerente de marca) dedicado a liderar los grupos formales para lograr una mejor integración de los esfuerzos.
- ✓ Organizaciones matriciales: consiste en organizaciones que manejan dos o más integradores que funcionan paralelamente para coordinar múltiples grupos formales (ver sección 2.2.1.1).

La coordinación lateral, al promover la descentralización y conservando un equilibrio en la formalización, se mantiene como uno de los mecanismos preferidos en organizaciones que se desenvuelven en ambientes dinámicos, debido a que les permite mantener el ritmo de la innovación del mercado al ser mecanismos que promueven la creatividad y la flexibilidad en los miembros del grupo (Galbraith, 2014a). Sin embargo, los mecanismos de coordinación lateral requieren de una mayor inversión por parte de la organización, debido a que es necesaria la contratación de personal que desempeñe roles específicos en la creación, participación y coordinación de los grupos, así como la inversión en sistema de información para la comunicación efectiva de los mismos (Child, 2015).

Una vez hemos analizado distintos mecanismos de coordinación que se pueden aplicar en las organizaciones, vamos a centrarnos en los párrafos siguientes en estudiar las distintas formas de control que se pueden aplicar.

2.2.2.3 Mecanismos de control

Los mecanismos de control se definen como los procesos de gestión en los que se regulan, mediante el poder y la influencia de algunos miembros, las actividades de la organización, de tal manera que los resultados de esas actividades coincidan con los objetivos, la estrategia y expectativas de la organización (Child, 2015). De acuerdo con el autor (*ibíd.*), existen 6 tipos de mecanismos de control (ver tabla 2-3 en la página siguiente).

Tabla 2-3: Tipos de mecanismos de control

Mecanismos de Control	Principales Características
Control Centralizado Personal	<ul style="list-style-type: none"> ✓ Supervisión directa de las actividades de las personas ✓ Toma de decisiones centralizada ✓ Recompensas y castigos acorde a la conformidad con la autoridad personal
Control Burocrático	<ul style="list-style-type: none"> ✓ Dividir la tarea en elementos definidos ✓ Métodos, procedimientos y reglas formales ✓ Tecnología diseñada para limitar la variación en la conducta de las tareas ✓ Recompensas y castigos acorde a la conformidad con las reglas y procedimientos
Control de Resultados	<ul style="list-style-type: none"> ✓ Trabajos y unidades diseñados con responsabilidad para completar objetivos ✓ Definición de resultados estándares y objetivos ✓ Delegación de decisiones operacionales: semi-autónomo ✓ Recompensas y castigos acorde a los resultados
Control Mediante Vigilancia Electrónica	<ul style="list-style-type: none"> ✓ Permite la grabación de las tareas para ser evaluadas remotamente ✓ Desempeño de empleados evaluado contra otros trabajadores ✓ Uso del monitoreo para recompensar y disciplinar a los empleados
Control Mediante Gestión de Recursos Humanos	<ul style="list-style-type: none"> ✓ Uso de métodos de selección para asegurar que el perfil encaje en las capacidades que requiere la empresa ✓ Entrenamiento y capacitación para reforzar los perfiles deseados ✓ Procedimientos de evaluación y sistemas de recompensa
Control Cultural	<ul style="list-style-type: none"> ✓ Desarrollo de la identidad personal del empleado con los objetivos de la organización ✓ Fuerte énfasis en el carácter colectivo y apoyo mutuo en la organización ✓ Empleo caracterizado por la seguridad en la permanencia y progresión en la organización

Fuente: Child (2015, p. 154)

Como podemos observar en la tabla anterior, los tres primeros corresponden a mecanismos caracterizados por una alta formalización y una centralización del control; como hemos visto anteriormente, al poseer estas características son mecanismos que se adaptan mejor a antiguas formas de organización o a organizaciones pequeñas. Por otro lado, el mecanismo de coordinación mediante *vigilancia electrónica* utiliza técnicas de monitoreo que reducen la necesidad de formalización y centralización del control, lo que lo lleva a ser más acorde con estructuras más modernas. En esta línea, los mecanismos de

control mediante *gestión de recursos humanos* y el *control cultural* son los mejores para promover la innovación y la flexibilidad en la organización, al ser mecanismos que fomentan el desarrollo de las capacidades de los empleados potenciando su compromiso y potencial para el cumplimiento de los objetivos de la compañía, características que son requeridas por la filosofía de las nuevas formas organizativas (Child, 2015).

A modo de síntesis

En este capítulo hemos desarrollado varios conceptos clave para esta tesis doctoral, comenzando con el concepto de conocimiento, taxonomía y los modos de conversión del conocimiento, así como la importancia de dichos conceptos en la organización. Luego, definimos los procesos de gestión del conocimiento, los cuales tienen que ver con la creación del conocimiento, el almacenamiento/consulta de conocimiento, la transferencia del conocimiento y su aplicación. Era necesario desarrollar y definir estos conceptos ya que la presente investigación aborda el estudio del efecto que el Big Data tiene sobre los mismos.

Además, en la segunda parte de este capítulo estudiamos el concepto de diseño organizativo y los aspectos más importantes que subyacen al mismo. Con ello pretendíamos explicar las diferentes variables que se deben tomar en cuenta para dividir el trabajo en una organización y coordinar y controlar los esfuerzos de los trabajadores, tomando como referencia principal para ello el modelo de estrella desarrollado por Galbraith (1995). En este contexto, hemos analizado también los conceptos de formalización y centralización/descentralización, conceptos que nos han ayudado a describir y comprender diferentes mecanismos que permiten a las organizaciones llevar a cabo la coordinación y el control de sus actividades y objetivos.

Así pues, los conceptos de conocimiento, procesos de gestión del conocimiento, diseño organizativo y variables subyacentes, formalización, centralización/descentralización y los mecanismos de coordinación y control serán de vital importancia para los capítulos subsecuentes, ya que son la base para el desarrollo del modelo teórico de relaciones propuesto en esta tesis doctoral y sobre el que pivotará todo el proceso de recogida y análisis de datos que se desarrollará con posterioridad.

Una vez concluida la revisión de los fundamentos teóricos relacionados con la gestión del conocimiento, el diseño organizativo y los mecanismos utilizados para la coordinación y el control en el ámbito organizativo, todos ellos de gran interés para nuestra investigación, en el siguiente capítulo analizaremos cómo el ecosistema de Big Data, al ser introducido en una organización, puede tener efecto sobre los procesos de gestión del conocimiento de la misma, sobre los aspectos de la estructura organizativa y sobre las formas de coordinación y control que se utilizan en aquella.

Capítulo 3:

***Big Data, los Procesos de Gestión del
Conocimiento y los Mecanismos de
Coordinación y Control***

Una vez hemos abordado en los capítulos anteriores toda la fundamentación teórica correspondiente a los conceptos de Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación y control, en este capítulo analizaremos cómo se relacionan dichos conceptos y qué factores debe tener en cuenta una organización para adoptar el Big Data en sus procesos. Esta aproximación teórica al estudio de la relación de los tres grandes conceptos en nuestra investigación (i.e., Big Data, procesos de gestión del conocimiento y mecanismos de coordinación y control) tiene por objeto generar un modelo teórico preliminar que sirva como punto de partida para la recogida y el análisis de información empírica en el marco de la presente tesis doctoral. Así mismo, la experiencia y el conocimiento previo del investigador en la temática analizada puede servir de base, en algún momento, para el establecimiento de alguna relación entre las variables identificadas en el modelo preliminar o para la identificación de algún factor que sea relevante para el objeto de investigación de esta tesis y que la revisión de literatura no haya puesto de manifiesto.

Como se verá en el próximo capítulo, esta investigación utiliza una metodología cualitativa con un propósito claramente exploratorio, ya que los fundamentos teóricos existentes sobre la temática de esta tesis doctoral son muy escasos en la literatura especializada. Por ello, necesitamos el modelo preliminar anteriormente mencionado a la hora de guiar nuestras decisiones durante la recogida de información y también durante el análisis de la misma. De esta manera, el modelo que presentaremos al final de este capítulo no tiene por objeto mostrar relaciones entre variables que vayan a ser testadas con una muestra estadísticamente representativa sino, tal y como se ha comentado, servir de guía o mapa de carreteras para recoger, analizar y escribir los resultados finales de la investigación.

En cuanto a la estructura del capítulo, en el primer apartado abordamos la relación teórica entre el concepto de Big Data y los procesos de gestión del conocimiento, describiendo cómo la introducción del Big Data modifica o afecta a cada uno de dichos procesos. En el segundo apartado del capítulo abordamos la relación teórica entre el concepto de Big Data y los mecanismos de coordinación y control, describiendo cómo la introducción del Big Data afecta a la aplicación de cada uno de estos mecanismos en la organización. En el tercer apartado, proponemos nuestro modelo que resume las relaciones teóricas planteadas en los dos primeros apartados y, adicionalmente, planteamos una relación

teórica entre los procesos de gestión del conocimiento y los mecanismos de coordinación y control que surgen a partir de la introducción del Big Data en las organizaciones.

3.1 Big Data y los Procesos de Gestión del Conocimiento

Las organizaciones no pueden ignorar la llegada del Big Data ya que se ha convertido actualmente en una forma de desarrollar mejores soluciones a problemas complejos y para mejorar el desempeño de las organizaciones. Por lo tanto, las organizaciones que plantean objetivos y estrategias para la resolución de problemas y tomar mejores decisiones deben utilizar un proceso sistemático para obtener conocimiento valioso a partir del Big Data (Carayannis *et al.*, 2017).

Como hemos descrito en secciones anteriores (ver sección 1.2), el Big Data se relaciona directamente con el conocimiento dado su potencial para generar gran cantidad de datos e información de diversas fuentes a una alta velocidad. Por este motivo, se ha considerado que el Big Data tiene implicaciones claras sobre la gestión del conocimiento (Tian, 2017), lo que ha llevado a plantear nuevas formas de gestionar el conocimiento teniendo en cuenta el papel que juegan los sistemas de Big Data y BDA para proveer el conocimiento y la inteligencia que se necesita para tomar decisiones efectivas a través del uso del conocimiento generado por estos sistemas (Rothberg & Erickson, 2017; Sumbal *et al.*, 2017).

Esta necesidad de utilizar el Big Data para tomar decisiones basadas en conocimiento deriva del hecho de que, en la actualidad, las organizaciones buscan que las decisiones pasen a estar basadas en evidencias y en conocimiento relevante más que en la intuición de aquellos que toman las decisiones (Acharya *et al.*, 2018). Por ende, el Big Data, en conjunto con sus tecnologías asociadas y la ciencia de datos, surge como la solución para llevar a cabo la toma de decisiones basada en conocimiento, conseguida mediante modelos predictivos que permitan conocer y analizar el comportamiento del entorno y de los procesos de la organización, para así buscar ventajas competitivas (Tian, 2017).

Por lo tanto, las organizaciones han visto la necesidad de adoptar las tecnologías disruptivas asociadas al Big Data (e.g. el IoT o la inteligencia artificial) en sus procesos

para incrementar la eficiencia e innovación, siendo necesario mejorar los procesos de gestión del conocimiento para explorar y explotar el conocimiento externo e interno de la organización y poder manejar, de un manera rápida y efectiva, el incremento en el flujo entrante y saliente de conocimiento (Santoro *et al.*, 2018; Uden & He, 2017).

A pesar de lo anterior, los mayores retos que enfrenta una organización que desea adoptar este tipo de sistemas y tecnologías no están en encontrar las fuentes de donde proviene el Big Data o la infraestructura tecnológica que se requiere para almacenar y procesar el Big Data (O'Connor & Kelly, 2017), sino en desarrollar las capacidades y mecanismos efectivos que se requieren para gestionar el conocimiento generado por el Big Data (O'Connor & Kelly, 2017; Santoro *et al.*, 2018; Sumbal *et al.*, 2017).

En este sentido, Lugmayr *et al.* (2017) han propuesto que esta relación de reciprocidad entre la gestión del conocimiento y el Big Data se le llame “Big Data Cognitivo”, definido como la interdependencia entre el proceso de gestión del conocimiento de las personas con los sistemas de procesamiento de Big Data. Lo anterior implica un modelo dual donde, en primer lugar, se brinda apoyo a las personas para interpretar y relacionar los datos, la información y el conocimiento provisto por los sistemas de Big Data y, en segundo lugar, se promueve el cambio hacia sistemas y tecnologías capaces de aprender, interpretar y sentir las intenciones y los datos creados por las personas.

En un principio, la literatura se ha centrado en estudiar cómo los sistemas avanzados de gestión del conocimiento (ver sección 1.2.3) manejan la velocidad con la que se necesita recolectar, agregar y compartir el conocimiento, y el cómo estos sistemas permiten la integración del conocimiento humano para comprender los datos que provee el Big Data (Intezari & Gressel, 2017). Por lo tanto, la investigación sobre los procesos de gestión del conocimiento y el Big Data ha sido vinculada directamente a la investigación sobre sistemas avanzados de gestión del conocimiento, donde se analiza cómo los SI y las TI avanzadas asociadas a Big Data tienen un impacto en la gestión del conocimiento tradicional (Santoro *et al.*, 2018; Tian, 2017). Este impacto del Big Data sobre los procesos y sistemas de gestión del conocimiento ha llegado a tal magnitud que autores como Lugmayr *et al.* (2017) consideran que esto llevará a que los modelos clásicos de gestión del conocimiento queden obsoletos.

No obstante, aún existen brechas de investigación por cubrir y establecer cómo estos sistemas llevan a cabo todo el proceso de gestión del conocimiento para gestionar los ecosistemas de información creados a partir del Big Data, y cómo se obtienen ventajas competitivas mediante el aprovechamiento del nuevo conocimiento (Santoro et al., 2018).

Debido al carácter incipiente de esta línea de investigación, no hemos encontrado una investigación completa que explore la relación entre todos los procesos de gestión del conocimiento vistos en el capítulo anterior y el Big Data, ya que la mayoría de los trabajos (e.g., Acharya *et al.* (2018) y Sumbal *et al.* (2017)) se han centrado principalmente en el proceso de creación de conocimiento a partir del Big Data, y poco se ha trabajado en los demás procesos. Por lo tanto, a continuación, proponemos los aspectos principales de cómo el Big Data afecta a cada uno de los procesos de gestión del conocimiento vistos en el capítulo precedente (i.e., creación, almacenamiento/consulta, transferencia y aplicación) basándonos, para ello, en nuestra experiencia personal en el tema y en las contribuciones de algunos autores de la literatura.

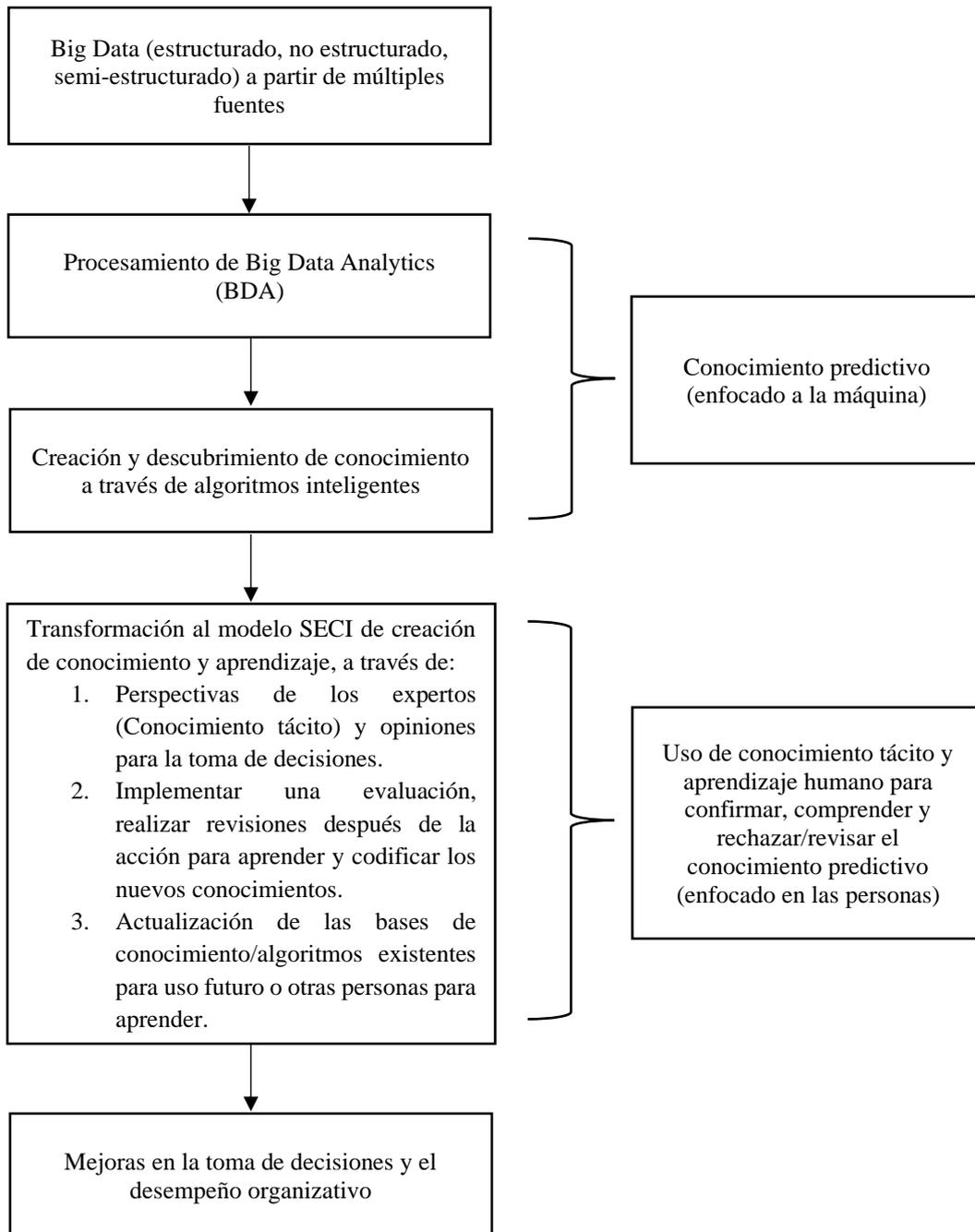
3.1.1 Big Data y el Proceso de Creación del Conocimiento

Como mencionamos anteriormente en la sección 2.1.3.1, el proceso de creación de conocimiento ha sido, en principio, el más estudiado de los procesos de gestión del conocimiento, ya que se considera que es donde el Big Data y el BDA tiene el mayor potencial y donde el nuevo proceso de creación del conocimiento estará apoyado (Lugmayr et al., 2017; Sumbal et al., 2017). Sin embargo, esta relación sigue siendo un campo relativamente nuevo donde apenas se empieza a investigar sobre el mismo (Sumbal et al., 2017).

La llegada del Big Data implica transformar o reformular cómo se lleva a cabo el proceso de creación de conocimiento (Carayannis *et al.*, 2017). Uno de los modelos que consideramos como más importantes hasta el momento es el planteado por Sumbal *et al.* (2017), quienes establecen que el Big Data modificará el proceso de creación de conocimiento por medio de la creación de "conocimiento predictivo", el cual consiste en crear conocimiento que permita predecir y prevenir el comportamiento de los procesos de la organización, lo que genera una mayor capacidad para gestionar el conocimiento.

El modelo de Sumbal *et al.* (2017), en primer lugar, coloca al Big Data como el origen de los datos que provienen de diversas fuentes. En segundo lugar, estos datos pasan a ser procesados mediante métodos, técnicas y herramientas de BDA. En tercer lugar, los resultados del procesamiento son convertidos en conocimiento mediante algoritmos inteligentes que dan como resultado el conocimiento predictivo. Y, en cuarto lugar, el conocimiento predictivo es transformado mediante el modelo SECI tradicional (i.e. socialización, externalización, combinación e internalización) de creación del conocimiento (ver sección 2.1.1), donde se utiliza el conocimiento tácito y el aprendizaje humano para evaluar el conocimiento predictivo. Este modelo plantea, como resultado de este nuevo proceso de creación de conocimiento, que se obtenga una mejora en la toma de decisiones y en el desempeño de la organización (*ibíd.*) (ver figura 3-1, más abajo).

Figura 3-1: Relación entre Big Data y el proceso de creación del conocimiento



Fuente: Sumbal et al. (2017, p. 193)

Así pues, al incluir el conocimiento predictivo en el modelo tradicional para la creación del conocimiento, permite una sinergia entre el conocimiento creado por las personas y el conocimiento automático (o creado por algoritmos y máquinas), donde ambos se retroalimentan con el fin de crear un conocimiento superior al de las técnicas o formas

tradicionales de llevar a cabo el proceso de creación del conocimiento, mejorando así la competitividad de la organización. Sin embargo, la inclusión del Big data en este proceso, y en general en cualquier proceso de gestión del conocimiento, estará limitada por las habilidades analíticas, la confianza en el Big Data y las restricciones de recursos que posea la organización (O'Connor & Kelly, 2017).

3.1.2 Big Data y el Proceso de Almacenamiento/Consulta del Conocimiento

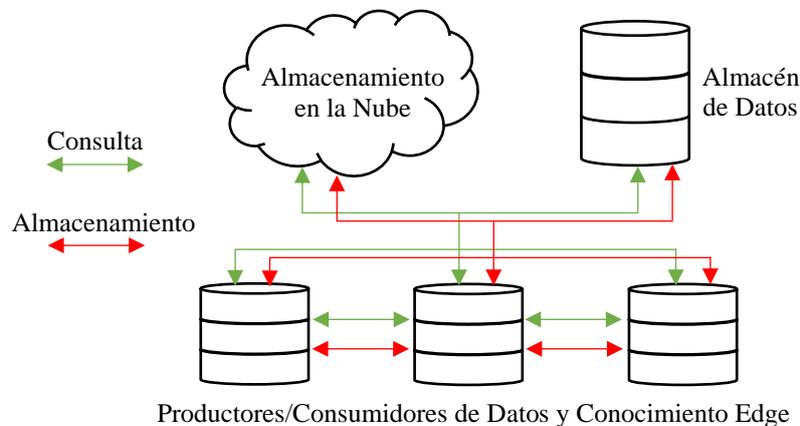
Como mencionamos en la sección 2.1.3.2, el proceso de A/C-C se encarga de gestionar la memoria de la organización, y es un proceso que se apoya en sistemas y tecnologías de la información que facilitan el acceso del conocimiento a todos los miembros de la organización (Alavi & Leidner, 2001). Por ende, la inclusión de sistemas y tecnologías asociadas a Big Data en una organización permitirá una mayor facilidad para la consulta de las bases de datos de conocimiento y una mejora en los sistemas de comunicación entre expertos e individuos dentro y fuera de la organización. Consecuentemente, la introducción de ecosistemas de Big Data en la organización se traducirá en que el almacenamiento centralizado del conocimiento (i.e. enfoque tradicional) experimentará un cambio para convertirse en un sistema descentralizado de conocimiento.

Por lo general, un proceso de A/C-C mantiene la memoria organizativa centralizada para su consulta (Alavi & Leidner, 2001); pero al incluir tecnologías de Big Data, como el IoT, el conocimiento se almacenará de forma distribuida en repositorios de conocimiento dispersos por la organización. Esta distribución del conocimiento ocurre debido a que estas tecnologías utilizan lo que se conoce en el campo como “computación perimetral” (“Edge Computing”, en inglés), y se define como el procesamiento de datos en la fuente, manteniendo el conocimiento resultante almacenado lo más cerca posible al dispositivo, componente o persona que produjo los datos (Linthicum, 2019).

Teniendo en cuenta lo anterior, el uso de la computación perimetral en la organización permitirá una mayor accesibilidad al conocimiento al estar disponible lo más cercano posible a la fuente que lo genera (Linthicum, 2019), yendo acorde con la naturaleza distribuida del conocimiento (Alavi & Leidner, 2001). Sin embargo, a pesar de la naturaleza distribuida del conocimiento, la computación perimetral implica, además, la interconexión digital de todos los dispositivos o sistemas “Edge” que almacenan el

conocimiento, y se realiza mediante la comunicación y cooperación entre ellos para llevar a cabo el procesamiento de los datos y el almacenamiento del conocimiento sin necesidad de terceros, a lo que se le conoce como computación en la niebla (“Fog Computing”, en inglés) (Vaquero & Rodero-Merino, 2014).

Figura 3-2: AS/RS de conocimiento



Fuente: elaboración propia

Por lo tanto, la computación en la niebla define los estándares que establecen el cómo se realiza la comunicación del conocimiento almacenado en dispositivos, máquinas o personas conectadas al sistema, y cómo estos a su vez se relacionan con los centros de almacenamiento de datos y el almacenamiento en la nube que posea la organización (Linthicum, 2019).

Para efectos prácticos, en esta investigación enmarcaremos estos conceptos de *almacenamiento de conocimiento distribuido* mediante la analogía con los sistemas AS/RS (Sistemas automatizados de almacenamiento/recolección, por sus siglas en inglés), proveniente de la teoría de centros de distribución para la recolección y almacenamiento de productos mediante robots y sistemas automatizados (Gagliardi *et al.*, 2012), a lo que denominaremos AS/RS de conocimiento (ver figura 3-2 más arriba). Los AS/RS de conocimiento simbolizan cómo se utilizan las diversas tecnologías asociadas al Big Data para llevar a cabo el almacenamiento y consulta del conocimiento que posee la organización, aplicando aprendizaje automático para determinar las necesidades de conocimiento en tiempo real a lo largo de la organización.

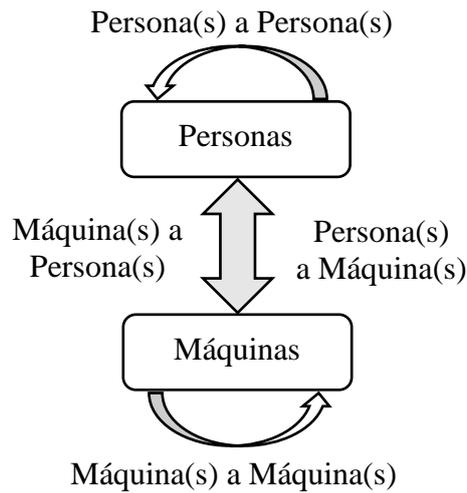
Así pues, a modo de síntesis podemos establecer que la aplicación del Big Data en las organizaciones afectará al almacenamiento y consulta del conocimiento de dos maneras distintas: una, hará que se produzca una descentralización del conocimiento de manera que éste se almacene en repositorios dispersos por la organización como consecuencia de la computación perimetral; y dos, se producirá una interconexión digital entre todos los dispositivos que almacenan el conocimiento como consecuencia de la computación en la niebla.

3.1.3 Big Data y el Proceso de Transferencia del Conocimiento

Como establecimos en la sección 2.1.3.3, el proceso de transferencia de conocimiento tiene lugar a través de los procesos de comunicación y los flujos de información en los diferentes niveles de la organización: entre individuos, de individuos a fuentes de conocimiento explícito, de individuos a grupos, entre grupos, y de grupos a la organización (Alavi & Leidner, 2001). Sin embargo, los procesos tradicionales de transferencia del conocimiento solo tienen en cuenta la interacción entre personas y grupos de personas, cuestión que cambia con la introducción de sistemas y tecnologías asociadas a Big Data.

Al incluir estos sistemas y tecnologías surgen tres tipos nuevos de transferencia de conocimiento: la transferencia máquina a máquina (abreviado como M2M en inglés), transferencia máquina a persona (abreviado como M2H en inglés) y transferencia persona a máquina (abreviado como H2M en inglés) (ver figura 3-3). Estos tipos de transferencia van de la mano de la forma distribuida que adquiere el conocimiento por el uso de la computación perimetral, lo que implica que la transferencia de conocimiento puede ocurrir de manera directa entre las personas, dispositivos y máquinas con conocimiento “Edge” (conocimiento in situ).

Figura 3-3: Transferencia de conocimiento en entornos con tecnologías de Big Data



Fuente: elaboración propia

Adicionalmente, la transferencia de conocimiento no solo ocurrirá en el momento de ser solicitada por un miembro o grupo de personas o máquinas de la organización, ya que los sistemas de Big Data tendrán la habilidad de predecir el comportamiento y las necesidades de conocimiento de cada nivel de la organización mediante la utilización de algoritmos inteligentes (aprendizaje automático e inteligencia artificial), permitiendo una mayor velocidad de respuesta a las necesidades de conocimiento de toda la organización.

Así pues, la aplicación del Big Data hará que la transferencia de conocimiento vaya más allá del ámbito humano y que se generen procesos de transferencia de conocimiento en los que la máquina desempeña un papel fundamental ya que pasa a ocupar uno de los roles centrales en el proceso de comunicación del conocimiento, a saber, el de emisor y/o el de receptor. Así mismo, la velocidad de respuesta a las necesidades organizativas de conocimiento se verá incrementada a través de la introducción de un nuevo proceso de gestión de conocimiento: la predicción del conocimiento necesario en cada momento y en cada nivel de la organización. Estos aspectos son representados en nuestro modelo teórico preliminar (ver figura 3-5, más adelante).

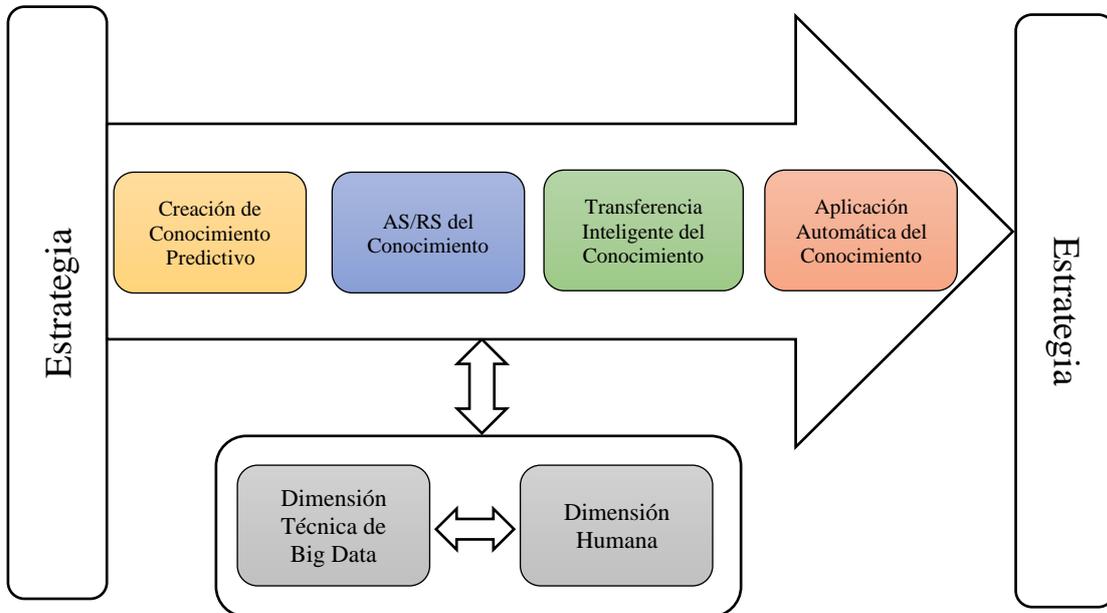
3.1.4 Big Data y el Proceso de Aplicación del Conocimiento

En base a lo expuesto en la sección 2.1.3.4, el proceso de aplicación del conocimiento es el proceso encargado de poner en práctica el conocimiento que ha sido creado, almacenado/consultado y transferido en la organización, y tiene como objetivos el convertir el conocimiento en procesos y comportamiento organizativo para la solución de problemas, ajustar la dirección estratégica y mejorar la eficiencia de la organización (Gold et al., 2001; Shujahat et al., 2019). Como hemos visto anteriormente, estos objetivos son compartidos con los que posee la implementación de ecosistemas de Big Data, por lo que el uso de Big Data en la organización puede considerarse como una forma de mejorar el proceso de aplicación del conocimiento.

En este sentido, Pauleen & Wang (2017) afirman que el conocimiento de las personas es el que decide qué información o conocimiento del Big Data o BDA será aplicado en la organización. Esta afirmación la consideramos parcialmente cierta, ya que el conocimiento y la percepción humana seguirán siendo claves para determinar qué conocimiento es importante para definir estrategias y la toma de decisiones que afecten el rumbo de la organización. No obstante, uno de los objetivos de la implementación de ecosistemas de Big Data consiste en que los dispositivos, máquinas y sistemas sean capaces de tomar decisiones por medio de algoritmos de aprendizaje automático, lo que implica que el conocimiento “Edge” no necesitará de intervención humana para ser aplicado. Esto es, la implementación del Big Data hará que la aplicación del conocimiento en la organización se automatice, lo que puede generar una reducción de los tiempos ciclo a la hora de desarrollar nuevas soluciones a los problemas organizativos (e.g. el desarrollo de un nuevo producto/servicio).

Por lo tanto, podemos concluir que la inclusión del Big Data en la organización tendrá un impacto considerable en los procesos de gestión del conocimiento al facilitar la creación, accesibilidad, transferencia y aplicación del conocimiento mediante personas, dispositivos, máquinas y sistemas que se interrelacionan y comunican para funcionar como productores, consumidores y almacenamiento de conocimiento tácito y explícito, formando sinergias entre el proceso de gestión del conocimiento y las dimensiones técnicas y humanas de la organización (ver figura 3-4).

Figura 3-4: Nuevo proceso de gestión del conocimiento tras la aplicación del Big Data en la organización



Fuente: elaboración propia

Lo anterior conlleva que las organizaciones deben replantear sus procesos de gestión del conocimiento para facilitar la transformación del Big Data en conocimiento de una manera más rápida y efectiva (Uden & He, 2017), ya que en la actualidad el dinamismo de las organizaciones implica ciclos de vida del conocimiento mucho más cortos (García, 2019), por lo que la toma de decisiones y la definición de estrategias deben apoyarse en el conocimiento creado y disponible en tiempo real.

3.2 Big Data y los Mecanismos de Coordinación y Control

Las organizaciones actuales, al encontrarse en mercados altamente dinámicos, han buscado la descentralización de ciertas decisiones operacionales para mejorar el flujo de la información y el conocimiento, así como una mayor flexibilidad de sus empleados (Burton et al., 2015). Esto implica que el Big Data puede proveer la descentralización que se necesita en las decisiones operacionales al ofrecer sistemas y tecnologías que facilitan el manejo de datos, información, conocimiento y procesos de manera automática, y en

ocasiones, sin intervención humana, tal y como se ha comentado en el apartado precedente.

Para lograr lo anterior, el reto de las organizaciones que implementan ecosistemas de Big Data está en desarrollar capacidades de gestión de Big Data y BDA para llevar a cabo la coordinación y el control de las personas, máquinas, dispositivos y sistemas que conforman los grupos interdisciplinarios encargados del procesamiento del Big Data (Colas *et al.*, 2014; Garmaki *et al.*, 2016).

Sin embargo, en la literatura especializada, el estudio de la relación entre los ecosistemas de Big Data y su efecto sobre los mecanismos de coordinación y control está aún sin explorar, encontrándose, en su mayoría, trabajos que exponen conjeturas, opiniones o ideas basados en la experiencia de los autores sobre cómo este nuevo paradigma afecta al diseño organizativo y a la coordinación y el control. Teniendo en cuenta esta consideración, a continuación, presentamos los aspectos relevantes que se ha logrado encontrar en la literatura para explicar los cambios en los mecanismos de coordinación y control tradicionales con la introducción del Big Data en las organizaciones.

3.2.1 Big Data y los Mecanismos de Coordinación

Como describimos en la sección 2.2.2.2, los mecanismos de coordinación son las herramientas, tecnologías y/o interacciones necesarias para ordenar y juntar a empleados y elementos interdependientes de una organización para conseguir los objetivos y cierto nivel de desempeño (Okhuysen & Bechky, 2009). En consecuencia, y en línea con lo visto en la sección anterior, la introducción de ecosistemas de Big Data en la organización permitirá a los gerentes y tomadores de decisiones coordinar las actividades necesarias para sacar provecho del mismo (Gupta & George, 2016), buscando cumplir con los objetivos y estrategias que plantea la organización al ser tecnologías facilitadoras de comunicación efectiva entre personas, máquinas, dispositivos y sistemas.

Susha *et al.* (2017) establecen que las organizaciones enfrentan cinco problemas de coordinación para conseguir crear valor de los datos, lo cual aplica en la implementación de ecosistemas de Big Data:

- 1) Permitir la relación entre proveedores y usuarios de los datos.
- 2) Mantener el control de los datos y evitar usos imprevistos o no adecuados.
- 3) Relacionar los problemas de la organización con los atributos de los datos.
- 4) Asegurar la utilidad de los datos para el usuario.
- 5) Alinear los incentivos de los proveedores de los datos con los objetivos de los usuarios.

Por consiguiente, los mecanismos de coordinación en organizaciones con ecosistemas de Big Data deben asegurar la correcta interconexión y comunicación entre productores y consumidores de datos y conocimiento “Edge”, formando grupos que permitan obtener valor de los datos para el cumplimiento de los objetivos y estrategia individual y de toda la organización.

Una de las ideas que ha surgido para la coordinación en ecosistemas de Big Data en las organizaciones es la creación de centros de excelencia (CoE por sus siglas en inglés), mecanismos encargados de incrementar las capacidades para el análisis avanzado de los datos (BDA) (Pearson & Wegener, 2013). Los CoE consisten en asignar la recolección y la propiedad del Big Data a lo largo de las distintas unidades de la organización, con el fin de planificar cómo se generará el conocimiento y asignar el tiempo de científicos de datos en grupos interdisciplinarios para el desarrollo de actividades relacionadas con Big Data en la organización (*ibíd.*).

Desde nuestra perspectiva personal, la formación de los CoE puede darse de dos formas: a) en primer lugar, la organización asignará roles específicos en base a la idoneidad de la máquina, dispositivo o persona encargada de manejar y almacenar la información o conocimiento requerido para la solución de un problema o el cumplimiento de un objetivo; y b) en segundo lugar, se crearán equipos mediante la asignación automática (utilizando tecnologías de aprendizaje automático e inteligencia artificial) de los miembros de la organización que procesen y almacenen la información o conocimiento requerido para la solución de un problema o el cumplimiento de un objetivo, en especial aquellos con conocimiento “Edge”. Esta última forma de crear los CoE se podrá utilizar dependiendo del nivel de implementación de ecosistemas de Big Data que posea la organización, ya que se requieren sistemas y tecnologías avanzadas de Big Data para la formación de estos grupos.

Otra de las ideas que han surgido para la coordinación en ecosistemas de Big Data en las organizaciones es mediante la coordinación lateral, donde el uso de tecnologías asociadas a Big Data permitirá la automatización de los procesos de negocio, con el fin de integrar los diferentes focos de Big Data que existen a lo largo de la organización (Galbraith, 2012, 2014b). Así, para lograr la utilización eficiente del Big Data, la organización necesita establecer una coordinación lateral inteligente (asignación por gerencia o automática) que integre a expertos en datos, tecnologías y personas con necesidades de conocimiento, brindando una mayor autonomía para tomar decisiones sobre los problemas u objetivos de la organización (Galbraith, 2014a; Garmaki et al., 2016).

Bajo esta premisa, los mecanismos de coordinación que serán más afectados por la aplicación del Big Data son la coordinación electrónica y el uso de estructuras matriciales (ver sección 2.2.2.2), ya que además de ser de los mecanismos principales para coordinar las actividades de una organización en la actualidad (Galbraith, 2014a), permiten a la organización incluir las tecnologías ya adoptadas con la implementación de ecosistemas de Big Data en el seguimiento automático de las actividades desarrolladas por los diferentes grupos formados en la organización (inclusive los CoE) para el desarrollo de habilidades y aplicación del conocimiento necesario para el cumplimiento de los objetivos.

Como consecuencia de lo anterior, una de las principales implicaciones que establece Galbraith (2014a) del Big Data sobre los mecanismos de coordinación y control (y para el diseño organizativo, en general) es que la organización debe sufrir un cambio de poderes en la jerarquía para poder coordinar las actividades, lo que indica un empoderamiento de los empleados y máquinas (i.e. descentralización), dándoles una mayor autonomía para tomar decisiones sobre los beneficios que se pueden obtener del Big Data bajo el criterio de cada grupo, individuo o máquina en poder de los datos, la información y el conocimiento.

Así pues, la aplicación del Big Data afectará a los mecanismos de coordinación de la organización de diversas formas. En primer lugar, se incrementará el uso de grupos de trabajo interdisciplinarios (i.e. centros de excelencia) que faciliten el desarrollo de las actividades relacionadas con el Big Data en la organización. Y, en segundo lugar, se

producirá un incremento en la coordinación lateral inteligente, basada en el uso de tecnologías y técnicas propias del Big Data, a través de coordinación electrónica o la creación de estructuras matriciales. En ambos, casos la descentralización de la toma de decisiones se convierte en una condición necesaria, pero no suficiente.

3.2.2 Big Data y los Mecanismos de Control

Como describimos en la sección 2.2.2.3, los mecanismos de control regulan las actividades de la organización para que los resultados de las mismas se alineen con las estrategias y expectativas de la compañía (Child, 2015). Sin embargo, en las nuevas formas de organización, y en especial con la introducción de ecosistemas de Big Data, las organizaciones buscan cada vez menos los mecanismos tradicionales para estandarizar y controlar el trabajo repetitivo (como, por ejemplo, el control centralizado personal o el control burocrático), pasando a utilizar métodos analíticos para optimizar resultados y reducir el riesgo (McLeod *et al.*, 2017).

Uno de los mayores problemas que presentan las organizaciones para el manejo del Big Data es la falta de talento y personal con las capacidades necesarias para sacar ventajas competitivas del Big Data (Gabel & Tokarski, 2014). Por lo tanto, los mecanismos de control deben promover el desarrollo de las capacidades y la relación y comunicación de las personas con las máquinas, dispositivos y sistemas que forman parte del ecosistema de Big Data, por lo que el enfoque de los mecanismos de control debe ser hacia el desarrollo analítico de los empleados.

En línea con lo anterior, los mecanismos conocidos que se adaptan mejor a estas necesidades son el *control mediante gestión de recursos humanos*, debido a que la organización debe asegurar que las personas que ingresen o se encuentren en la organización poseen o desarrollan las capacidades analíticas necesarias para el Big Data, y el *control cultural*, ya que son mecanismos que permiten al empleado desarrollar sus habilidades colectivas y de apoyo mutuo con todas las áreas de la organización, necesario para la conformación de los grupos de aprovechamiento de Big Data, convirtiéndose en mecanismos que motiven, en vez de controlar, a las personas a través del establecimiento de canales tecnológicos para la interacción (Murawski & Bick, 2017).

Otro de los mecanismos de control que se verá impactado por la introducción de ecosistemas de Big Data será el *control por resultados*, debido a que el control de los objetivos de las personas y máquinas puede medirse mediante el Big Data generado por estos mismos, midiéndose el desempeño de forma automática mediante algoritmos inteligentes, a lo que se le conoce como "gestión algorítmica" o "gestión científica 2.0" (Schildt, 2017). Esto implicará que los mecanismos de control para medir resultados en las organizaciones dejen de ser una práctica meramente humana y se convierta en un proceso en conjunto con, o desarrollado meramente, por tecnologías asociadas a Big Data.

Adicionalmente, la visibilidad de los datos, información y el conocimiento que provee el Big Data permitirá una mayor capacidad para llevar a cabo mecanismos de *vigilancia electrónica*, donde las actividades y resultados de la organización se analizarán mediante técnicas de BDA como la analítica de textos, audio y video (ver sección 1.2.2), manteniendo así un mejor control sobre el cumplimiento de los objetivos de la organización (Berner *et al.*, 2014).

A continuación, en la siguiente tabla presentamos un resumen de los mecanismos de control impactados y las características que conlleva la introducción de ecosistemas de Big Data en las organizaciones.

Tabla 3-1: Mecanismos de control mediante la inclusión de ecosistemas de Big Data

Mecanismos de Control	Características con la introducción de Big Data
Control de Resultados	<ul style="list-style-type: none"> ✓ Generación de Big Data para evaluación de resultados por parte de personas y máquinas ✓ Medición automática de resultados (gestión algorítmica) ✓ Control sinérgico entre máquinas, personas y sistemas
Control Mediante Vigilancia Electrónica	<ul style="list-style-type: none"> ✓ Técnicas de BDA (analítica de textos, audio y video) ✓ Visibilidad de actividades y resultados en la organización ✓ Control sinérgico entre máquinas, personas y sistemas
Control Mediante Gestión de Recursos Humanos	<ul style="list-style-type: none"> ✓ Selección de personal con capacidades en Big Data ✓ Entrenamiento y capacitación en métodos y técnicas analíticas ✓ Control sinérgico entre máquinas, personas y sistemas
Control Cultural	<ul style="list-style-type: none"> ✓ Desarrollo de la cultura de pensamiento analítico ✓ Énfasis en la participación de grupos interdisciplinarios de Big Data ✓ Control sinérgico entre máquinas, personas y sistemas

Fuente: elaboración propia

Como podemos notar en la tabla anterior, se incluyó un factor común entre los mecanismos de control correspondiente al control sinérgico entre máquinas, personas y

sistemas. Esto se debe a que los ecosistemas de Big Data implican por defecto la constante revisión de los resultados obtenidos por máquinas, personas y sistemas, llevando a cabo entre sí una retroalimentación automática, evaluando el cumplimiento de los objetivos planteados por la organización.

En entornos de Big Data será común que las organizaciones seleccionen varios de estos mecanismos de control para su implementación, ya que los ecosistemas de Big Data están diseñados para evaluar en diferentes frentes y de forma automática el desempeño en tiempo real de todos los niveles de organización, usando una combinación de cualquiera de los mecanismos de control antes mencionado. Por lo tanto, a medida que se delegue el control, y a su vez la coordinación, a procesos, máquinas y dispositivos, la organización tenderá hacia una descentralización total de la toma de decisiones debido a la introducción de los ecosistemas de Big Data (Schildt, 2017). Sin embargo, para comprender mejor cómo los mecanismos de coordinación y control funcionan conjuntamente con los ecosistemas de Big Data se necesita desarrollar más investigación en el campo (Colbert et al., 2016).

3.3 Big Data y su relación con el proceso de gestión del conocimiento y los mecanismos de coordinación y control: modelo teórico preliminar

Una vez establecida la relación existente entre el Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación y control, y estudiados los conceptos asociados a cada uno de estos temas, en este último apartado nos disponemos a ilustrar y describir el modelo teórico que explica la relación existente entre estos tres aspectos. La construcción de este modelo está basada en la literatura y en las propuestas planteadas en las secciones anteriores de este mismo capítulo.

En primer lugar, planteamos que las dimensiones expuestas en la sección 1.3.3 y, por ende, el nivel de implementación de ecosistemas Big Data en la organización, son los factores organizativos que determinan el grado en el que se modifican los procesos de gestión del conocimiento y los mecanismos de coordinación y control utilizados por la organización, ya que de estos depende el aprovechamiento del Big Data para la toma de

decisiones basada en conocimiento, la resolución de problemas, el cumplimiento de objetivos y estrategias, la creación de valor y las mejoras en el desempeño.

En segundo lugar, planteamos que mediante la evolución de los procesos de gestión del conocimiento a procesos que involucren a las personas, sistemas, tecnologías y aspectos organizativos del ecosistema de Big Data, el nuevo proceso de gestión del conocimiento resultante modificará el nivel de implementación de ecosistemas de Big Data en la organización mediante la aplicación del conocimiento proveniente de este nuevo proceso. Por tanto, existe una relación bidireccional entre la aplicación del Big Data y los procesos de gestión del conocimiento en la organización, tal y como se ilustra en la figura 3-5 más abajo.

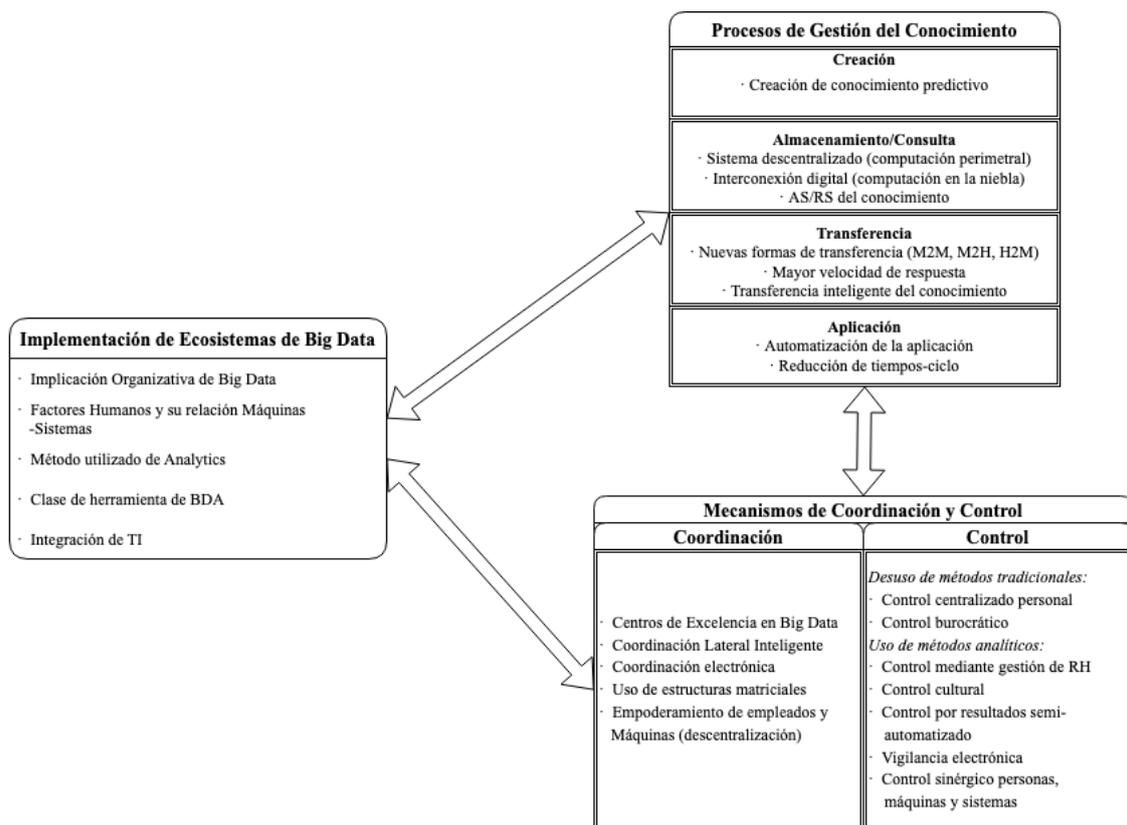
En tercer lugar, y de manera similar al proceso de gestión del conocimiento, planteamos que la evolución de los mecanismos de coordinación y control que involucren a las personas, sistemas, tecnologías y aspectos organizativos del ecosistema de Big Data permitirá a la organización obtener las capacidades y habilidades analíticas requeridas para la toma de decisiones en contextos de Big Data y el cumplimiento de sus objetivos, lo que facilitará el incremento en el nivel de implementación de ecosistemas de Big Data en la organización.

En cuarto lugar, planteamos que los mecanismos de coordinación y control influyen, en la introducción del Big Data, en los procesos de gestión del conocimiento al ser mecanismos que facilitan la creación de grupos interdisciplinarios para el manejo de Big Data y asegurar las capacidades y la cultura analítica necesarias para llevar a cabo todos los procesos que involucren a la gestión del conocimiento.

Por último, y de forma recíproca, planteamos que los procesos de gestión del conocimiento influyen, en la introducción del Big Data, en los mecanismos de coordinación y control al proveer el conocimiento necesario por los algoritmos que llevan a cabo la coordinación de lateral inteligente y, a su vez, el control sinérgico requerirá de este conocimiento para llevar a cabo la retroalimentación automática de personas, máquinas y sistemas.

Partiendo de estas relaciones básicas que se han mencionado en los párrafos precedentes, a continuación presentamos el modelo teórico de relaciones entre la implementación del Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación y control. Como se explicó al inicio de este capítulo, este modelo no pretende mostrar relaciones susceptibles de ser contrastadas estadísticamente; más bien al contrario, su finalidad radica en poder guiar la mirada de este investigador a la hora de recoger y analizar la información empírica que utilizaremos en esta tesis doctoral.

Figura 3-5: Modelo teórico de relaciones entre la implementación de ecosistemas de Big Data, procesos de gestión del conocimiento y mecanismos de coordinación y control



Fuente: elaboración propia

Una vez completado el marco teórico de interés para esta investigación, y tras plantear el modelo teórico de relaciones entre la implementación de ecosistemas de Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación y control, en el capítulo posterior estableceremos y justificaremos las decisiones metodológicas necesarias para llevar a cabo el trabajo de campo requerido para alcanzar los objetivos que nos habíamos propuesto al inicio de esta tesis doctoral.

Capítulo 4:

Metodología de Investigación

Una vez hemos concluido con la descripción de los conceptos correspondientes al marco teórico de esta investigación y planteado el modelo teórico de relaciones entre la implementación de ecosistemas de Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación y control en las organizaciones, en este capítulo pasamos a describir las decisiones metodológicas que hemos tomado para poder llevar a cabo el trabajo de campo de esta investigación y alcanzar, de esta manera, nuestro objetivo inicial propuesto. En este sentido, es necesario recordar que el modelo teórico presentado en el capítulo anterior nos va a servir de base para guiar nuestros procesos de recogida y análisis de información, así como para escribir los resultados de la presente investigación, lo cual es consecuencia del marcado carácter exploratorio de la misma.

Así, el presente capítulo está estructurado del siguiente modo. En primer lugar, presentamos las consideraciones y características propias de la metodología cualitativa y también las de la cuantitativa, lo que nos permite comparar ambos diseños metodológicos y elegir el más apropiado a nuestra investigación. Dentro de la metodología cualitativa explicamos conceptualmente la estrategia del estudio de casos ya que es la que más se adecúa a nuestros objetivos de investigación. Así, hemos explicado y argumentado las principales decisiones que, dentro de éste, se deben tomar, haciendo referencia al tipo de estudio de casos que hemos desarrollado, los criterios seguidos para su elección y el número de casos a estudiar. Con posterioridad, se detallan los métodos y técnicas utilizados para recoger y analizar la información de los casos y, por último, describimos las tácticas que se han aplicado para dotar de calidad a la presente investigación y abordamos los aspectos éticos de la misma.

4.1 Consideraciones y características generales de la metodología cuantitativa y cualitativa

Como establece gran parte de la literatura especializada (Bluhm et al., 2011; Boeije, 2010; Bryman, 1988; Leech & Onwuegbuzie, 2009; Saunders et al., 2009; Yin, 2014), la investigación en ciencias sociales sigue principalmente dos vertientes metodológicas para resolver preguntas de investigación, a saber, la cuantitativa y la cualitativa. La utilización de una metodología u otra dependerá, entre otras cuestiones, de las características y objetivos de la investigación desarrollada, donde la cuantitativa busca resolver problemas

mediante métodos estadísticos y técnicas como los cuestionarios o los estudios demográficos para confirmar una relación entre la teoría y la investigación, mientras que la cualitativa utiliza la participación o la entrevista en profundidad para resolver relaciones emergentes entre la teoría y la investigación (Balbastre, 2003).

Por lo tanto, la investigación cuantitativa se utiliza principalmente para establecer proposiciones e hipótesis que pueden ser comprobadas mediante el análisis de datos numéricos, mientras que la investigación cualitativa tiene la capacidad de descubrir los significados e interpretaciones de un fenómeno mediante el análisis de datos no numéricos (Balbastre, 2003; Saunders et al., 2016).

Teniendo en cuenta ambas vertientes, Bryman (1988) enumera una serie de diferencias entre las metodologías cualitativa y cuantitativa que ayudan decidir qué tipo de metodología es más idónea acorde al tipo de investigación desarrollado. Entre estas diferencias podemos destacar las siguientes (ver tabla 4-1 en la siguiente página).

Tabla 4-1: Diferencias entre diseño metodológico cuantitativo y cualitativo en una investigación

	Diseño Cuantitativo	Diseño Cualitativo
Papel de la investigación cualitativa	Preparatoria	Medio para explorar las interpretaciones de los actores
Relación entre investigador y sujeto a investigar	Distante	Próxima
Posición del investigador respecto al sujeto a investigar	Intruso	Conocido
Relación entre teoría/conceptos e investigación	Confirmación	Emergente
Estrategia de investigación	Estructurada	No estructurada
Alcance de los hallazgos	Nomotético (generalizables)	Ideográfico (no generalizables)
Imagen de la realidad social	Estática y externa al actor	Dinámica y socialmente construida por el actor
Naturaleza de los datos	Tangibles, fiables	Ricos, profundos

Fuente: Bryman (1988, p. 94)

- ✓ En primer lugar, la relación y posición existente entre investigador y el fenómeno (o sujeto) que es el objeto de estudio es cercana y de contacto permanente en la metodología cualitativa, mientras que en la cuantitativa es distante y con una postura de intruso por parte del investigador en relación al fenómeno que se está estudiando.
- ✓ En segundo lugar, la relación entre la teoría y los datos en la metodología cualitativa es de carácter emergente y la investigación busca desarrollar teorías y conceptos, mientras que en la metodología cuantitativa se busca verificar (o no) la veracidad de dichas teorías y conceptos por lo que la teoría existente es el inicio del proceso de investigación.
- ✓ En tercer lugar, la metodología cualitativa tiene un carácter no estructurado (abierto y flexible), mientras que la metodología cuantitativa es estructurada para facilitar que todos los participantes estén sometidos al mismo ‘estímulo’.
- ✓ En cuarto lugar, los resultados de la investigación en la metodología cualitativa no son generalizables desde el punto de vista estadístico (aunque sí que lo son

desde una perspectiva analítica), mientras que en la metodología cuantitativa sí que pueden ser generalizables estadísticamente hablando.

- ✓ En quinto lugar, en la investigación cualitativa se utiliza la participación de los sujetos para el estudio de la realidad social, mientras que en la metodología cuantitativa se ve a la realidad social como algo externo al investigador y estático.
- ✓ Por último, y como mencionamos anteriormente, la metodología cuantitativa se basa en datos tangibles y numéricos, mientras que en la metodología cualitativa los datos son considerados ricos y de carácter profundo.

Teniendo en cuenta las consideraciones anteriores y dada la naturaleza del tema de investigación de esta tesis doctoral y nuestro objetivo de investigación, la metodología que se adapta mejor al mismo es la metodología cualitativa. Nuestro objetivo de investigación tiene que ver con analizar cómo la implementación del Big Data afecta a los procesos de gestión del conocimiento y a la forma de coordinación y control en las organizaciones. Dicho objetivo está relacionado con un fenómeno poco conocido o explorado, presenta un carácter dinámico ya que al mismo subyace un proceso de cambio organizativo y, además, existe un pobre marco teórico que lo explique, lo que dota a nuestra investigación de un marcado carácter exploratorio.

De hecho, y tal como se vio en el capítulo precedente, muchas de las relaciones propuestas en el modelo teórico preliminar provienen de una unión entre la experiencia del autor de esta tesis y la literatura especializada. Además de lo anterior, nuestra investigación requiere de información basada en experiencias, relatos y vivencias de los sujetos que han participado en el fenómeno que estamos investigando aquí. Dicho fenómeno se ha desarrollado en un contexto social (i.e. la organización) que se ha visto afectado por el mismo y que, a su vez, lo ha afectado, lo que nos lleva a no poder separar, a la hora de estudiarlo, el fenómeno del contexto en el que el mismo tiene lugar, todo lo cual nos refuerza la idea de que una metodología de corte cualitativo es la mejor opción para desarrollar esta investigación.

Teniendo claro que la metodología a utilizar ha de ser de corte cualitativo, necesitamos ahora tomar una decisión respecto a qué estrategia de investigación cualitativa es la más apropiada. En este sentido, Saunders *et al.* (2016) establecen que puede suceder que más de una estrategia sea válida para acometer unos mismos objetivos de investigación por lo

que no es demasiado correcto hablar de la estrategia ‘más apropiada’ sino de una estrategia ‘adecuada’ a los objetivos de la investigación. Así, estos autores hacen referencia a las siguientes estrategias de corte más cualitativo:

- ✓ ***Estudio de casos:*** es una estrategia que implica una investigación empírica de un fenómeno contemporáneo en un contexto de la vida real usando múltiples fuentes de evidencia. Típicamente, el estudio de casos se realiza cuando no se controla el contexto en el cual el fenómeno objeto de estudio está ubicado y se quiere ganar un entendimiento mayor de dicho fenómeno.
- ✓ ***Investigación-acción o investigación-acción participante:*** es una estrategia que tiene como objetivo la resolución de problemas organizacionales, y puede involucrar a los empleados, investigadores y consultores externos o internos para el cambio de procesos de la organización. En ella, se rompe con el principio de separación entre investigador e investigado, ya que ambos forman parte de un único equipo de investigación en el que cada uno desarrolla un papel diferente.
- ✓ ***Teoría fundamentada:*** es una estrategia de investigación que consiste en el desarrollo y construcción de nueva teoría a través de la combinación de la deducción y la inducción. La teoría se desarrolla en base a un marco de referencia teórico y una serie de observaciones que llevan a la generación de predicciones que luego son probadas mediante una serie de nuevas observaciones.
- ✓ ***Etnografía:*** es una estrategia que tiene como propósito describir el entorno social en el que se desenvuelve un fenómeno u objeto de estudio, e implica que el investigador debe sumergirse en dicho entorno social para poder describir y explicar los cambios y nuevos patrones que presenta dicho fenómeno u objeto de estudio a lo largo del tiempo.
- ✓ ***Investigación de documentos:*** es una estrategia en donde se utilizan los registros y documentos históricos o actuales de una organización o fenómeno. Esta estrategia puede ser utilizada como complemento en estrategias exploratorias, descriptivas o explicativas.

Teniendo en cuenta las estrategias antes mencionadas, la que consideramos más adecuada para los objetivos de esta investigación es la del estudio de casos. Como veremos en el apartado subsiguiente, esto se explica debido a que nuestro objetivo de investigación para esta tesis doctoral tiene que ver con un fenómeno contemporáneo poco estudiado en la

literatura y que se presenta en ámbitos organizaciones en los que es muy difícil separar el fenómeno de su contexto de desarrollo. A continuación, profundizaremos en las características de esta estrategia de investigación cualitativa y justificaremos mejor las razones que nos han llevado a su elección para desarrollar la presente investigación.

4.2 El estudio de casos como estrategia de investigación cualitativa

Tal y como se vio en el apartado precedente, la metodología de investigación cualitativa ha sido la seleccionada para llevar a cabo la presente investigación debido a que es la metodología adecuada cuando se requiere investigar, documentar y conceptualizar un nuevo fenómeno o que es poco conocido (Bluhm *et al.*, 2011), lo cual se ajusta a nuestro objetivo de investigación que, como vimos en la introducción de esta tesis doctoral, está relacionado con estudiar cómo la implementación de ecosistemas de Big Data en las organizaciones va a afectar a los procesos de gestión del conocimiento y a las formas de coordinación y control del trabajo en la organización.

Dentro de la metodología de corte cualitativo, el estudio de casos, como vimos anteriormente, es una de las opciones metodológicas más utilizadas en el ámbito de las ciencias sociales (Njie & Asimiran, 2014; Stake, 1995). En particular, el estudio de casos es adecuado cuando se quiere comprender el contexto de un fenómeno contemporáneo del que se tiene poco conocimiento y donde el investigador no tiene la capacidad de manipular los eventos que suceden en dicho fenómeno (Saunders *et al.*, 2016; Yin, 2014).

Dado que el interés de esta investigación se centra en el “cómo” y el “por qué” la implementación de ecosistemas de Big Data en las organizaciones influye en los procesos de gestión del conocimiento y en los mecanismos de coordinación y control del trabajo, para llevar a cabo la presente investigación hemos seleccionado el estudio de casos ya que este tipo de preguntas de investigación son susceptibles de ser respondidas mediante esta estrategia de investigación (Stake, 1995; Swanborn, 2010; Yin, 2014). Además, esta estrategia resulta idónea cuando el investigador tiene poco o nada de control sobre los eventos investigados y cuando es fundamental que el estudio del fenómeno se realice en el mismo contexto en el cual éste tiene lugar (Yin, 2014). La siguiente tabla (ver tabla 4-2) sintetiza las condiciones en las que distintas estrategias de investigación pueden ser

utilizadas y pone de manifiesto que el estudio de casos es la estrategia que más se adecúa a nuestros objetivos de investigación.

Tabla 4-2: Condiciones para la selección de estrategias de investigación

Estrategia	Forma de la Pregunta de Investigación	¿Se requiere control sobre los eventos?	¿Enfoque en eventos contemporáneos?
Experimento	Cómo, Por qué	Si	Si
Encuesta	Quién, Qué, Dónde, Cuánto(s)	No	Si
Análisis de Archivos	Quién, Qué, Dónde, Cuánto(s)	No	Si/No
Historia	Cómo, Por qué	No	No
Estudio de casos	Cómo, Por qué	No	Si

Fuente: Yin (2014, p. 6)

Del mismo modo, autores en el campo de la investigación sobre la influencia del Big Data en los procesos y la gestión de las organizaciones (e.g. Acharya *et al.*, 2018; Uden & He, 2017; Zeng & Glaister, 2017 y Zhan *et al.*, 2017) han seleccionado el estudio de casos para investigar sobre el fenómeno del Big Data y sus ecosistemas y para obtener una mejor comprensión de su influencia en distintos ámbitos de las organizaciones como la innovación, el conocimiento o el diseño organizativo, así como en diversos sectores de actividad. Por ello, el uso del estudio de casos en nuestra investigación para analizar una temática similar a la de los autores anteriormente mencionados es una decisión que tiene sentido.

Una vez justificada la elección de la estrategia de investigación a elegir, no podemos continuar nuestra investigación sin definir a qué hacemos referencia cuando hablamos del estudio de casos. En este sentido, Yin (2014, p. 13) establece que el estudio de casos se puede definir como la investigación empírica de fenómenos contemporáneos (el caso) en profundidad y en un contexto del mundo real, en especial cuando los límites entre el fenómeno y el contexto no es claramente evidente, y donde hay más variables de interés que puntos de recolección de datos. De la misma forma, Dul & Hak (2010, p. 4) definen el estudio de casos como “un estudio donde un solo caso o un número pequeño de casos

(estudio de casos comparativo) en su contexto de la vida real son seleccionados, y se analizan de forma cualitativa los resultados obtenidos de dichos casos”.

Por su parte, Swanborn (2010, p. 13) establece que el estudio de casos se refiere al estudio de un fenómeno social con las siguientes características:

- ✓ Es llevado a cabo dentro de los límites de un sistema social (el caso).
- ✓ Es conducido en el contexto natural del caso.
- ✓ El investigador se enfoca en el proceso de describir y explicar el proceso social.
- ✓ El investigador, guiado por una pregunta de investigación inicial amplia, explora los datos para luego formular preguntas de investigación más precisas.
- ✓ Se utilizan múltiples fuentes de datos, como documentos, entrevistas y la observación.

Una vez hemos definido el estudio de casos como estrategia de investigación cualitativa para esta tesis doctoral, a continuación, vamos a clasificar el tipo de estudio de casos que se va a desarrollar en la presente investigación.

4.2.1 Tipos de estudio de casos

Un aspecto a definir cuando se desarrolla un estudio de casos es el relativo a su clasificación. En la literatura especializada podemos encontrar distintas formas de clasificar los estudios de casos. En este sentido, una de las tipologías más conocidas es la propuesta por Stake (1995) quien plantea tres tipos de estudios de casos:

- ✓ ***Estudio de caso intrínseco:*** tiene como objetivo comprender el caso en sí mismo, es decir, el caso es el interés de la investigación.
- ✓ ***Estudio de caso instrumental:*** tiene como objetivo estudiar un problema o teoría donde el caso juega un papel de apoyo para la formulación de afirmaciones, y dicho caso puede ser atípico a otros casos.
- ✓ ***Estudio de caso colectivo:*** tiene como objetivo estudiar las diferencias y semejanzas dentro y entre múltiples casos, con el fin de poder replicar los hallazgos.

Por su parte, Yin (2014) plantea una clasificación de los estudios de casos en función de dos dimensiones distintas, a saber, su propósito y el número de casos a estudiar. En cuanto al primer criterio, Yin (*ibíd.*) establece que, dependiendo del propósito de la investigación, los estudios de casos pueden clasificarse en exploratorios, descriptivos y explicativos (ver tabla 4-3). Los estudios de casos exploratorios se utilizan cuando el investigador pretende explorar fenómenos para los que no existe un conjunto de resultados claros o únicos. Por su parte, los estudios de casos descriptivos se utilizan para describir una intervención o fenómeno y el contexto de la vida real en el que éste ocurre. Por último, los estudios de casos explicativos son adecuados si el investigador está buscando contestar una cuestión que pretende explicar los presuntos vínculos causales en intervenciones de la vida real que son demasiado complejas para otras estrategias de investigación como la encuesta o los experimentos.

Tabla 4-3: Categorías de estudio de casos en función del propósito de la investigación

Categoría	Definición
Exploratorio	Este tipo de estudio de casos se utiliza cuando se desea estudiar situaciones en donde el fenómeno bajo estudio se encuentra poco o nada explorado, por lo que no se han definido resultados únicos o visibles. Las investigaciones exploratorias no tienen un curso fijo y van cambiando a medida que se encuentran nuevos datos.
Descriptivo	Este tipo de estudio de casos busca representar o describir el perfil exacto del fenómeno, persona o evento bajo estudio en un contexto del mundo real
Explicativo	Este tipo de estudio de casos busca establecer relaciones causales entre variables bajo estudio que son muy difíciles de responder mediante técnicas de entrevista o estrategias experimentales.

Fuente: elaboración propia a partir de Baxter y Jack (2008), Saunders et al. (2009) y Yin (2014)

Por lo que respecta al segundo de los criterios, esto es, al que hacía referencia al número de casos a estudiar, Yin (2014) establece una tipología en la que tendríamos, por un lado, el estudio de caso único o sencillo y, por otro, el estudio múltiple de casos, presentándose una variación a los anteriores en función de que existan o no (holístico) unidades de análisis incrustadas en los casos, como, por ejemplo, unidades de negocio o procesos específicos de la(s) organización(es) estudiada(s).

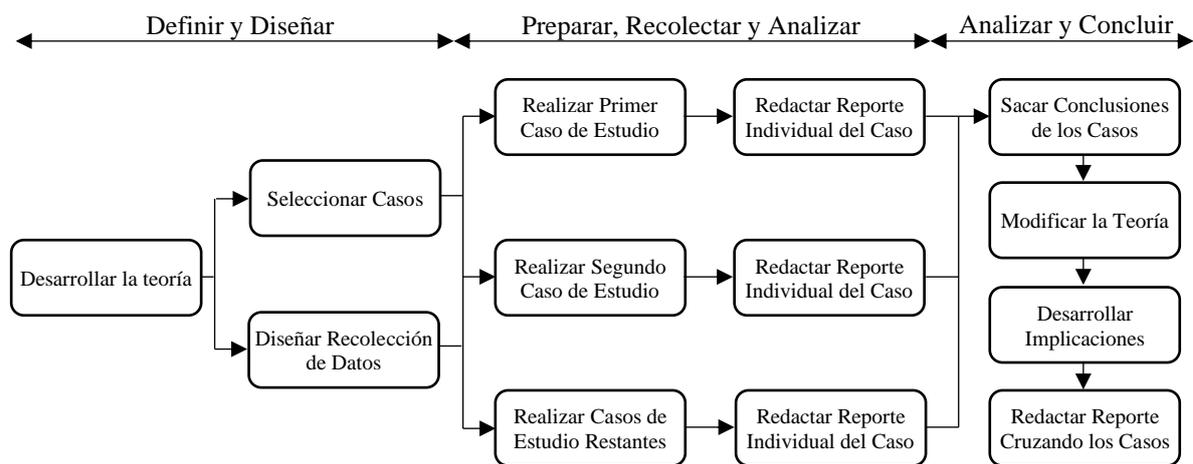
No obstante, es habitual en la literatura (e.g. Baxter & Jack (2008), Swanborn (2010) y Yin (2014) comparar entre el estudio de casos único vs. el estudio múltiple de casos. Así, el estudio de casos único suele usarse cuando el caso que es estudiado representa un caso extremo, único, raro o crítico que permite explorar, observar y analizar un fenómeno que no se había considerado antes, y cuyos datos resultantes son suficientes para responder la pregunta de investigación (Dul & Hak, 2010; Saunders *et al.*, 2009). Esta tipología sería el equivalente a lo que Stake (1995) consideraba un estudio de caso intrínseco. Por otro lado, el estudio de casos múltiple, a veces denominado como estudio de casos comparativo (Dul & Hak, 2010), se utiliza cuando en la investigación se desea comparar las similitudes y/o diferencias entre los casos con el objetivo de predecir resultados similares o predecir resultados contrastantes, donde cada caso responde a una necesidad u objetivo específico de la investigación (Baxter & Jack, 2015; Yin, 2014). En términos generales, los estudios de casos múltiples son considerados como más convincentes y robustos que el estudio de casos único. Sin embargo, la decisión de seleccionar uno u otro tipo de estudio de caso dependerá, en gran medida, de los objetivos de la investigación (Yin, 2014).

Dado que en nuestra investigación pretendemos analizar un fenómeno muy poco investigado hasta la fecha como es el del efecto de la implementación de ecosistemas de Big Data sobre los procesos de gestión del conocimiento y los mecanismos de coordinación y control en las organizaciones, nuestro estudio de casos tendrá un claro propósito exploratorio ya que pretende arrojar algo de luz sobre esta temática que, como se ha dicho, está poco analizada por la literatura académica hasta el momento.

Por lo que se refiere al número de casos a analizar, en esta investigación hemos considerado que la madurez de una organización en el uso de ecosistemas de Big Data puede afectar a los cambios que tendrán lugar en los ámbitos de la gestión del conocimiento y del uso de mecanismos de coordinación y control. Tal y como veíamos en el capítulo 3 de esta tesis doctoral, teorizamos que distintos niveles de implementación del Big Data pueden afectar a los procesos de gestión del conocimiento y a los mecanismos de coordinación y control a través de palancas o mecanismos diferentes. Por ello, es necesario estudiar casos con distintos niveles de madurez en la implantación de ecosistemas de Big Data lo que nos lleva al diseño de un estudio múltiple de casos.

A la hora de desarrollar un estudio múltiple de casos, Yin (2014) establece una sistemática que es la que vamos a seguir en esta investigación. Según esta sistemática, en el estudio de casos múltiples podemos distinguir tres etapas (ver figura 4-1): 1) en primer lugar, se deben elegir los casos a estudiar y diseñar los métodos para la recolección de datos; 2) en segundo lugar, se debe preparar, recolectar y analizar los datos con el fin de presentar reportes individuales; 3) y, en tercer lugar, se debe analizar los casos en conjunto con el fin de modificar la teoría existente, sus implicaciones y redactar un reporte comparando los casos estudiados.

Figura 4-1: Procedimiento para estudio de casos múltiple



Fuente: adaptado de Yin (2014, p. 49)

Habiendo establecido ya el tipo de estudio de casos que vamos a desarrollar en la presente tesis doctoral y el proceso general que vamos a seguir para realizar el trabajo de campo, en el siguiente apartado vamos a abordar las decisiones vinculadas a la configuración de la muestra de nuestro estudio de casos. Con ello estaremos abordando la selección de los casos, parte de esta primera etapa del procedimiento o sistemática propuesta por Yin (2014). Con posterioridad, en el siguiente apartado abordaremos las decisiones relativas a los métodos y técnicas para la recogida de los datos, segundo aspecto dentro de la primera etapa de la sistemática establecida por Yin (*ibíd.*).

4.3 Muestra del estudio de casos

Una de las decisiones más importantes en el diseño de un estudio múltiple de casos es la que tiene que ver con su muestra. De esta manera, en el presente apartado presentamos las decisiones que se tomaron para determinar la unidad de análisis del estudio de casos, el número de casos que se va a estudiar y cuáles fueron los criterios de selección de los mismos. Por lo tanto, dividiremos la sección en dos subapartados, donde el primero abordará la definición de la unidad de análisis y el segundo se centrará en justificar las decisiones relativas al número y criterios de selección de los casos estudiados en esta investigación.

4.3.1 Unidad de análisis

Una vez se ha definido que la presente investigación cualitativa se debe realizar por medio de un estudio de casos, uno de los principales aspectos que se deben definir es la unidad de análisis (Tellis, 1997). De acuerdo con Yin (2014), Dul & Hak (2010) y Swanborn (2010), la unidad de análisis en un estudio de casos está relacionada directamente con la manera en cómo el investigador ha planteado las preguntas de investigación, pudiendo ser dicha unidad de análisis un individuo, un grupo o representar eventos, entidades u organizaciones completas.

Así pues y teniendo en cuenta lo anterior, la unidad de análisis en esta investigación es la organización en su conjunto ya que el interés de nuestra investigación está en estudiar cómo los ecosistemas de Big Data afectan a los procesos de gestión del conocimiento y los mecanismos de coordinación y control en la organización y, consecuentemente, estos aspectos de ámbito organizativo necesitan ser analizados tomando en consideración la totalidad del contexto en el que los mismos se desarrollan.

Una vez tenemos claro qué entendemos por ‘caso’ en esta investigación, las siguientes decisiones concernientes a la muestra del estudio de casos tienen que ver con cuántos casos vamos a estudiar y con qué criterios son elegidos los mismos, a lo cual dedicamos el siguiente subapartado.

4.3.2 Número y criterios de selección de los datos

A diferencia de otras estrategias de investigación como la encuesta, por ejemplo, los casos a ser utilizados en un estudio de casos no son seleccionados basándonos en criterios estadísticos o a través de una selección aleatoria de una población; más bien al contrario, dichos casos deben ser seleccionados intencionalmente por el investigador para apoyar las necesidades, requerimientos y particularidades de la investigación que se está llevando a cabo (representatividad vs. relevancia), ya que el objetivo final de la misma no es la obtención de unos resultados extrapolables a una población mayor sino la obtención de un conocimiento rico y profundo sobre el fenómeno que se está analizando. Bajo esta forma de muestreo, los casos son elegidos de manera deliberada por la información importante que pueden suministrar y que no puede ser obtenida de ninguna otra forma (Maxwell, 2008). Por ello, estos métodos de muestreo se conocen como “muestreo deliberado”, “de selección intencional” o “muestreo cualitativo” (Boeije, 2010; Teddlie & Yu, 2007). Por tanto, los casos serán elegidos en función de su contribución esperada al conocimiento del fenómeno objeto de estudio y los criterios de selección deberán estar alineados con el objetivo general de la investigación.

El criterio básico de selección en un estudio de casos es, según Stake (1995), que los casos elegidos permitan al investigador aprender lo más posible sobre el objeto de investigación. Este criterio de carácter general se concreta en esta investigación, de manera particular, en los siguientes criterios:

- ✓ En primer lugar, necesitamos organizaciones que tengan un cierto nivel de madurez en la implementación de ecosistemas de Big Data, ya que este es el factor impulsor cuyo efecto queremos analizar sobre los procesos de gestión del conocimiento y los mecanismos de coordinación y control de la organización. Como vimos en el capítulo 1 de esta tesis doctoral, la aplicación del Big Data puede caracterizarse a través del uso de diversos sistemas de información y del impacto que tienen sobre la organización, y en este sentido se desarrolló, desde un punto de vista teórico, una propuesta de “Niveles y dimensiones de implementación de ecosistemas de Big Data en las organizaciones” (ver Tabla 1-5 en el capítulo 1). Además, dado que el estudio múltiple de casos tiene un claro enfoque comparativo necesitamos que las organizaciones a estudiar presenten

distintos niveles de intensidad en la variable relevante objeto de estudio. Por tanto, se elegirán distintas organizaciones con distinto nivel de madurez en la aplicación de ecosistemas de Big Data.

- ✓ Debido a que el fenómeno bajo estudio implica que la organización debe poseer algún tipo de método, técnicas o herramientas avanzadas de análisis de datos para poder explorar su efecto en los procesos de gestión del conocimiento y los mecanismos de coordinación y control, consideraremos organizaciones que estén como mínimo en el nivel tres (i.e. Sistemas para la Toma de Decisiones y Especializados de Información Empresarial) de dicha propuesta, que posean personal entrenado en técnicas, métodos y herramientas de BDA y que, en la actualidad, se encuentren al menos en proceso de pasar al nivel cuatro, o tengan varias dimensiones en dicho nivel o superiores (i.e. Transición a Sistemas de Big Data).
- ✓ Otro criterio importante para la selección de los casos en esta investigación tiene que ver con el hecho de que las organizaciones a estudiar deben tener una estrategia y una cultura para el manejo del conocimiento en la organización, siendo deseable que la organización posea un proceso visible de gestión del conocimiento con objetivos y prácticas claramente definidas.
- ✓ Así mismo, las organizaciones a estudiar deben ser empresas de un tamaño relativamente grande (medianas o grandes empresas). Este criterio está íntimamente relacionado con el anterior ya que en las micro y pequeñas empresas se da un pobre desarrollo en temas de sistemas y tecnologías de información, lo que nos lleva a establecer que necesitamos estudiar empresas medianas, grandes y/o multinacionales, donde sea visible su inversión en métodos, técnicas y herramientas para el análisis de datos y/o tecnologías asociadas a Big Data. Este hecho nos garantizará que dichas organizaciones tienen áreas de sistemas de información y de gestión del conocimiento desarrolladas, así como políticas definidas en ese sentido.
- ✓ Por su parte, Rodríguez *et al.* (1996) indican que es importante que exista una elevada probabilidad de que se dé una mezcla de procesos, programas, personas,

interacciones y/o estructuras relacionadas con las cuestiones de investigación, y el hecho de que las organizaciones elegidas tengan un tamaño relativamente elevado nos facilitará el cumplimiento de este criterio.

- ✓ Del mismo modo, las organizaciones a analizar deben pertenecer a sectores de actividad diferentes, ya que eso nos permitirá introducir el criterio de la variedad de contextos en el fenómeno a estudiar.
- ✓ Desde el punto de vista de su ámbito demográfico, las organizaciones a estudiar deben estar ubicadas en Colombia dada la ubicación geográfica del investigador al momento de realizar el estudio de casos facilitando, de esta manera, el acceso del mismo a los datos primarios.
- ✓ Y, por último, el investigador necesita tener acceso al fenómeno objeto de estudio para lo cual es fundamental desarrollar una buena relación con los informantes, todo lo cual redundará en un mejor acceso a la información y, consecuentemente, en un mejor desarrollo del trabajo de campo.

Una vez definidos los criterios de selección de los casos, debemos abordar ahora la decisión relativa al número de casos a analizar. En este sentido, Eisenhardt (1989) sugiere que para un estudio de casos el número ideal de casos a seleccionar podría estar entre cuatro y diez casos. Sin embargo, la dificultad del análisis de un número elevado de casos puede llevar a problemas en el manejo y procesamiento de los datos, lo que acarrearía posteriores inconvenientes en obtener los resultados deseados (*ibíd.*).

Por otro lado, Dyer y Wilkins (1991) consideran que se debe seleccionar un número de casos reducido para poder analizar con mayor profundidad el fenómeno bajo estudio, en particular, sugieren que se seleccione un único caso. Sin embargo, la selección de un único caso puede no garantizar un análisis profundo del fenómeno y se podrían pasar por alto aspectos claves que no toma en cuenta el caso seleccionado, al igual que no se garantiza con un número elevado de casos (Balbastre, 2003).

Por lo tanto, no existe en la literatura un acuerdo respecto al número de casos que se deben seleccionar (Swanborn, 2010); en cambio, se considera que el número de casos depende

más de la confianza y cercanía del investigador a los casos, de tal forma que se permita una mayor confianza en los datos provistos por los mismos, buscando un equilibrio entre la profundidad en el análisis de los casos y una variedad en los aspectos diferenciales de los mismos (Flyvbjerg, 2006; Shakir, 2002).

Teniendo en cuenta lo anterior, y considerando la cercanía del investigador con los casos a seleccionar, la profundidad y variedad que se desea en el análisis de los casos y los demás criterios mencionados anteriormente, en esta tesis doctoral hemos decidido utilizar tres casos para estudiar el fenómeno objeto de esta investigación. Este tamaño muestral: a) nos permitirá realizar comparaciones sobre cómo la aplicación de ecosistemas de Big Data afecta a la gestión del conocimiento y a la coordinación y el control del trabajo en organizaciones con distinto nivel de madurez en Big Data; b) al ser una muestra de tamaño reducido nos permitirá estudiar cada caso con un nivel de profundidad notable; y c) nos permite alejarnos de los riesgos derivados de la elección de un único caso y su difícil justificación en una investigación (i.e. ¿hasta qué punto el caso es “único?”).

Así pues, aplicando las consideraciones anteriores recogidas en los criterios de selección de los casos y ajustándonos al número de casos elegido, las organizaciones que se han analizado en esta tesis doctoral son las siguientes:

- ✓ **Cementos Colombia S.A.:** empresa multinacional colombiana dedicada a la fabricación de cemento, concreto y derivados. Tipo de empresa: Grande. Nivel alcanzado en la implementación de ecosistemas de Big Data: 5.
- ✓ **Fundación Universidad del Norte:** empresa de educación superior privada. Tipo de empresa: Grande. Nivel alcanzado en la implementación de ecosistemas de Big Data: 4.
- ✓ **Eco Fuels Energy S.A.S:** empresa dedicada a la fabricación de biodiesel. Tipo de empresa: Mediana. Nivel alcanzado en la implementación de ecosistemas de Big Data: 3.

Las tres empresas seleccionadas fueron estudiadas en detalle previo a su selección definitiva, comprobando todos los requisitos planteados anteriormente y estudiando si la organización se encontraba en un proceso de implementación de sistemas y tecnologías asociadas a Big Data utilizando el marco de referencia planteado en la sección 1.3.3.

Dicho análisis ha sido validado con uno de los entrevistados en cada empresa para comprobar el análisis previo realizado por el autor de esta tesis doctoral, lo que luego ha sido verificado durante las entrevistas.

El nombre de la empresa Cementos Colombia S.A. es un nombre ficticio que ha sido cambiado por cuestiones de confidencialidad y será utilizado de esta manera en esta tesis doctoral.

Tras presentar los casos que van a ser estudiados en esta investigación, en el siguiente apartado describimos los métodos y técnicas que hemos utilizado para recoger los datos e información relativa a los tres casos que se han analizado.

4.4 Métodos y técnicas para la recogida de datos

En la presente investigación, se han utilizado dos métodos básicos para la recogida de datos e información de los casos a estudiar: las entrevistas en profundidad y la documentación corporativa.

Por lo que respecta a las *entrevista en profundidad* y tal y como establecen Tellis (1997) y Yin (2014), la recogida de datos en un estudio de casos implica el desarrollo de tres tareas principales: la preparación para la recolección de los datos, la distribución de los cuestionarios (guion de la entrevista, en nuestro caso) y llevar a cabo las entrevistas.

En primer lugar, para la preparación de la recolección de los datos se decidió realizar entrevistas de carácter abierto para permitir mayor fluidez de la conversación con el entrevistado, recurriendo a lo que se comúnmente se conoce como “entrevista en profundidad” (Cassell, 2015). Dichas entrevistas son también categorizadas como entrevistas no estructuradas debido a que no se elabora una lista previa de preguntas; en cambio, el entrevistador establece las ideas principales a discutir en la entrevista y el entrevistado tiene la posibilidad de hablar libremente sobre los temas, y sus respuestas van conduciendo el desarrollo de la entrevista a modo de conversación (Saunders et al., 2016)

En este sentido, como instrumento de recogida de información a utilizar en las entrevistas en profundidad se elaboró un guion que contiene las temáticas de interés de la investigación, dividiéndose dicho guion en cinco bloques o secciones cada una de las cuales abordaba una serie de ítems que recogían cuestiones abiertas sobre las que conversar con el entrevistado (ver Anexo B). Una vez elaborado el guion, el mismo fue sometido a un análisis preliminar respecto a su idoneidad para esta investigación por parte de investigadores con experiencia en el desarrollo de estudio de casos, los cuales realizaron correcciones y ajustes que nos permitieron mejorar la versión inicial del guion y, con ello, poder tener una herramienta apropiada que nos lleve a estudiar el fenómeno adecuadamente y a dar respuesta a las preguntas de investigación.

Una vez definido el guion de la entrevista, se preparó una documentación que apoya a la entrevista en caso de que los entrevistados desconocieran algunos términos de las temáticas de la investigación, documentación que exclusivamente fue manejada por el entrevistador y se utilizó únicamente en caso de que el entrevistado así lo requiriera. A modo de ejemplo, Barranco (2012), Firican (2017) y Gandomi y Haider (2015) ha sido bibliografía de apoyo durante las entrevistas para presentar los conceptos asociados a Big Data. Además, se prepararon y probaron los dispositivos de grabación necesarios con el fin de tener digitalmente el audio de las entrevistas para ser, con posterioridad, procesadas, analizadas y revisadas durante la etapa de análisis.

Tras la preparación de la recolección de datos, en segundo lugar, se realizó el contacto con las personas que tenían que ser los informantes en nuestra investigación. Estas personas eran las encargadas de las áreas de la organización que trabajan con datos para la toma de decisiones, Big Data o técnicas de BDA, y las áreas de la organización encargadas de los procesos de gestión del conocimiento. Respecto a los mecanismos de coordinación y control, todos los entrevistados, con independencia de su área, nos podían dar información al respecto ya que dichos mecanismos están inherentes en el trabajo diario de la organización. El contacto se realizó por medio de correo electrónico, presentando la solicitud formal de entrevista y explicando en ella los objetivos de nuestra investigación y la importancia de su aporte para el mismo (ver Anexo C). Una vez recibida la respuesta por parte de los entrevistados, se programó una cita por video llamada en la que se le enviaba el guion de la entrevista para su estudio, de manera que

pudiéramos responder, previo a la entrevista, a las dudas respecto a las temáticas a tratar. En la tabla 4-4 se muestra la planificación de las entrevistas a realizar.

Tabla 4-4: Planificación de las entrevistas a realizar en esta investigación

Organización	Cargo a entrevistar	Duración prevista	Fecha prevista	Bloques del guion a tratar
Cementos Colombia S.A.	Director de Analytics	Una hora y treinta minutos	Junio 2019	1 a 5
	Director de Gestión Humana	Una hora y quince minutos	Junio 2019	1, 2, 4 y 5
Fundación Universidad del Norte	Jefe de Información y Estadística	Una hora y treinta minutos	Junio 2019	1 a 5
	Jefe de Capacitación y Desarrollo	Una hora y quince minutos	Junio 2019	1, 2, 4 y 5
Eco Fuels Energy S.A.S.	Director de Innovación y Tecnología	Una hora y treinta minutos	Junio 2019	1 a 5
	Director Administrativo	Una hora y quince minutos	Junio 2019	1, 2, 4 y 5

Fuente: elaboración propia

En tercer lugar, al momento de llevarse a cabo la entrevista se solicitó a los participantes la autorización para la grabación y uso del audio con fines académicos, así como el consentimiento informado para poder publicar los resultados obtenidos del análisis resultante de la investigación (ver Anexo D). Durante las entrevistas, y dada la naturaleza abierta de la mismas, surgieron cuestiones y temas tratados que no se encontraban en el guion original de la entrevista y que corresponden a características propias de cada organización en la gestión de datos, los procesos de gestión del conocimiento y el uso de mecanismos de coordinación y control. También se recogieron notas de campo para complementar la información obtenida de las entrevistas. Estas notas de campo formaron, junto a las transcripciones y los documentos corporativos, la información primaria que se analizó en cada uno de los casos.

La siguiente tabla (ver tabla 4-5) sintetiza los detalles de las entrevistas realmente realizadas.

Tabla 4-5: Entrevistas realizadas en esta investigación

Organización	Entrevistado	Duración	Fecha	Bloques tratados en la entrevista	Páginas de transcripción
Cementos Colombia S.A.	Director de Analytics	Una hora y cinco minutos	Junio 2019	1 a 5	16
	Director de Gestión Humana	Cuarenta minutos	Julio 2019	1, 2, 4 y 5	10
Fundación Universidad del Norte	Jefe de Información y Estadística	Una hora y cuarenta minutos	Junio 2019	1 a 5	19
	Jefe de Capacitación y Desarrollo	Una hora y veinte minutos	Julio 2019	1 a 5	18
Eco Fuels Energy S.A.S.	Director de Innovación y Tecnología	Una hora y quince minutos	Junio 2019	1 a 5	15
	Director Administrativo	Cincuenta minutos	Agosto 2019	1, 2, 4 y 5	11

Fuente: Elaboración propia.

La segunda fuente de información que hemos utilizado es la *documentación corporativa*, tal y como se ha señalado al inicio de este apartado. El uso de una segunda fuente de evidencias nos permite triangular la información recopilada a través de las entrevistas, lo cual redundará en última instancia en una mejora de la validez interna de la investigación que hemos realizado. En particular, hemos recogido y analizado documentación de diversa índole: información disponible en la web de los casos estudiados, memorias corporativas, informes de gestión, plan estratégico, manual de calidad y procedimientos de trabajo. Para ello, contábamos con el consentimiento explícito de las organizaciones analizadas.

Una vez hemos descrito cómo se llevó a cabo la recolección de datos e información, en el siguiente apartado presentamos las cuestiones relacionadas con el análisis de los datos y la información recolectada de las entrevistas, las notas de campo y la documentación corporativa consultada.

4.5 Análisis de los datos

Tras haber recolectado todos los datos y la información de cada uno de los casos estudiados, el siguiente paso consiste en llevar a cabo el análisis de estos datos e información. Para tal fin, en primer lugar, los audios de las entrevistas deben ser transcritos para poder trabajar los datos con mayor facilidad. Sin embargo, como establece Yin (2014), la transcripción de las entrevistas suele ser una tarea tediosa y que consume bastante tiempo y energía del investigador, por lo que muchos investigadores optan por no realizar la transcripción y únicamente tomar apuntes de los audios en las partes que son de interés.

A pesar de lo anterior, los avances en las técnicas de BDA, como la analítica de audio (técnica descrita en el apartado 1.2.2 de esta tesis doctoral), ha permitido que mediante algoritmos de aprendizaje automático e inteligencia artificial se logre realizar transcripciones del audio mediante un servicio de reconocimiento de discurso automático (ASR, por sus siglas en inglés), permitiendo así convertir el audio en texto en tiempo real para ser usado en los análisis posteriores (Amazon, 2019).

Por lo tanto, y al ser esta una tesis doctoral que trata temas de BDA, vimos pertinente utilizar un servicio de ASR para apoyar la transcripción de los audios en texto, seleccionando para tal fin el servicio ofrecido por la plataforma de Amazon Web Services (AWS) con su programa “Amazon Transcribe”. En línea con lo anterior, Saunders *et al.* (2009) mencionan que el uso de sistemas automáticos de reconocimiento de voz podría tener problemas como la necesidad de entrenar el software para que reconozca la voz.

No obstante, los servicios de ASR, como el seleccionado para esta tesis doctoral, han evolucionado de tal manera que no necesitan de entrenamiento previo para el reconocimiento de la voz y tampoco se necesita dictar completa o parcialmente el audio, ya que los algoritmos de aprendizaje automático están en constante entrenamiento y mejoran a medida que se les ingrese audio de todos los usuarios del programa a nivel mundial. De esta manera, el principal problema que se ha tenido que tener en consideración en esta investigación es el chequeo de la transcripción, debido a que el programa puede fallar al utilizarse acentos o expresiones inherentes a cada persona, tener

partes de la entrevista con volumen bajo o poco claros, y fallas en el reconocimiento de expresiones y frases en el idioma español debido a que, a pesar de que estos algoritmos son compatibles con nuestro idioma, por el momento son algoritmos diseñados y entrenados principalmente para trabajar con el idioma inglés.

Con el objetivo de corregir dichas imprecisiones de las transcripciones automáticas, hemos realizado una revisión minuciosa de la transcripción contrastándola con el audio recolectado de las entrevistas, realizándose las respectivas correcciones que dieran a lugar una mejor comprensión del texto transcrito.

Al momento de tener las entrevistas transcritas y revisadas, la documentación corporativa y las notas de campo organizadas, el siguiente paso consiste en realizar el análisis de los datos cualitativos. El análisis de datos cualitativos, una de cuyas técnicas más conocidas es la “coincidencia de patrones” (Dul & Hak, 2010), está basado en la segmentación de los datos recolectados en partes, unidades o elementos manejables, para luego reorganizarlos de una manera comprensible mediante la búsqueda de patrones y relaciones entre las partes o elementos, con el fin de crear una comprensión del fenómeno desde una perspectiva teórica (Boeije, 2010; Flick, 2018; Miles *et al.* , 2014).

Esta segmentación de los datos lleva implícita la asignación de etiquetas (i.e. códigos) a dichos segmentos. Así, la codificación consiste en asignar un nombre o un texto a cada parte o elemento de acuerdo a los atributos de cada uno de ellos, lo cual permite relacionar con posterioridad grupos de elementos al ser codificados con el mismo nombre (Elliott, 2018). Estos códigos pueden después ser agrupados en categorías que, a su vez, podrán ser agrupadas en un nivel superior dentro del análisis dando lugar a temas o conceptos, los cuales en última instancia serán la base para explicar el fenómeno estudiado (Boeije, 2010; Saldaña, 2015).

De acuerdo con Saunders *et al.* (2009) y Yin (2014), la codificación de los datos cualitativos puede seguir dos enfoques básicos. El primero es un enfoque deductivo, donde se establecen códigos y/o categorías en base a la teoría y/o a la experiencia previa y expectativas del investigador, de manera que los elementos de análisis vienen preestablecidos antes de iniciar este proceso (Miles & Huberman, 1994). Por su parte, el segundo es un enfoque inductivo, donde los códigos y/o categorías surgen a raíz de los

datos recolectados, y se van explorando y creando a medida que se avanza en el análisis (Braun & Clarke, 2006, 2014; Clarke & Braun, 2013). En cualquier caso, el proceso de análisis cualitativo debe ser un proceso iterativo donde constantemente se están revisando la etapas previas (Boeije, 2010; Dul & Hak, 2010), lo que permitirá la construcción de códigos y categorías más complejos y que describan cada vez mejor el fenómeno bajo estudio (Yin, 2014). Teniendo en cuenta lo anterior, en esta investigación tomaremos un enfoque mixto de codificación donde, en primer lugar, definimos algunos códigos de acuerdo a la teoría y experiencia propia del investigador para después, y dada la naturaleza exploratoria de la investigación, generar unos códigos que no se hayan contemplado con anterioridad y que surgirán a raíz de la evidencia observada.

Para apoyar el proceso de construcción de los códigos y categorías utilizadas en esta tesis doctoral se desarrolló una herramienta en Microsoft Excel 2016 (ver figura 4-2) que permitiera la organización de los códigos, subcategorías y categorías que surgieron de forma deductiva e inductiva, junto con su respectiva definición y una descripción de cuando usar dichos códigos. Así, al momento de analizar los fragmentos de las transcripciones de los casos estudiados era más sencillo asociar un código al mismo, y de forma automática se asociaba la subcategoría y categoría que correspondía.

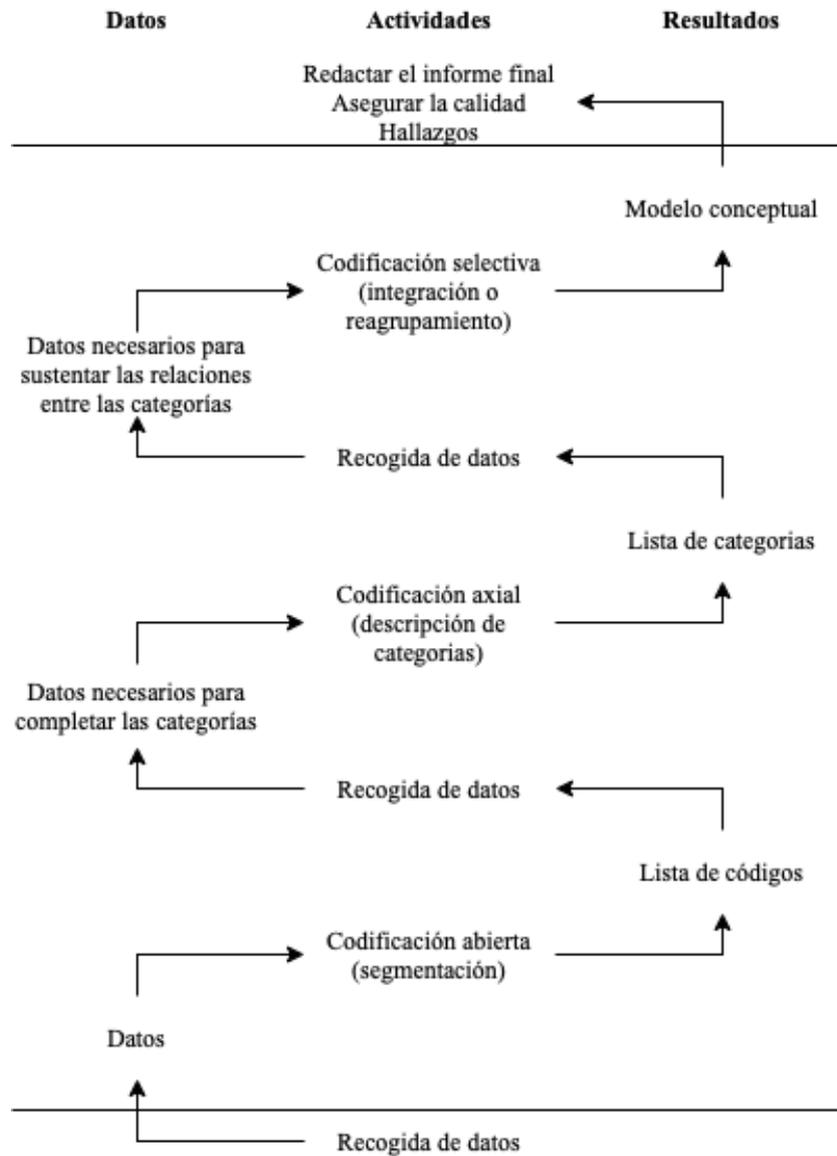
Figura 4-2: Herramienta en Microsoft Excel para la codificación de las entrevistas

CÓDIGO	DEFINICIÓN	CUANDO USARLO	SUBCATEGORÍA	CATEGORÍA	CEMENTOS	UNINORTE	ECO FUELS	
1	Implicación Organizativa de Big Data	Establece el tipo de SI que ha adoptado la organización para la toma de decisiones basadas en análisis de datos, el nivel de descentralización de esas decisiones y el nivel de implicación del pensamiento analítico en la cultura organizativa	Se utiliza cuando se describe como los SI y el Big Data ha influenciado la toma de decisiones, descentralización y el cambio cultural en toda o parte de la organización	Implementación de Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	3	8	3
2	Factores Humanos y su Reducción Máquinas-Sistemas	Establece el nivel de capacitación e implicación que poseen los empleados y directivos en el manejo y gestión del SI, en el uso de métodos básicos, intermedios o avanzados de análisis de datos o de BDA y en la creación de áreas o departamentos especializados de Big Data y ciencia de datos	Se utiliza cuando se describe la implicación de los diferentes ítems de la organización en el desarrollo y uso del análisis de datos, el BDA. Además, el apoyo de la organización para capacitar a sus empleados en técnicas/humanistas y el apoyo de la dirección.	Implementación de Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	6	11	2
3	Método Utilizado de Analytics	Establece que tipo de métodos estadísticos, matemáticos o de BDA utiliza la organización	Se utiliza cuando se presenta los métodos de análisis de datos o BDA utilizados en la organización	Implementación de Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	1	3	4
4	Clase de Herramienta de BDA	Establece las herramientas de BDA utilizadas por la organización	Se utiliza cuando se describen que herramientas de BDA utiliza la organización para llevar a cabo análisis de datos	Implementación de Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	0	2	0
5	Integración de TI	Establece que posibles tecnologías básicas, intermedias o avanzadas que ha adoptado la organización	Se utiliza cuando se describe cómo la organización ha incluido en sus procesos sistemas y tecnologías de asociadas a Big Data para tomar decisiones o llevar a cabo otras actividades	Implementación de Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	3	3	3
6	Creación de Conocimiento Predictivo	Indica que la organización utiliza el BDA para llevar a cabo la creación de conocimiento a partir de algoritmos inteligentes y análisis de datos	Se utiliza cuando se describe cómo se ha utilizado el Big Data y el BDA para crear nuevo conocimiento en la organización	Proceso de Gestión del Conocimiento y Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	2	5	3
7	AS/RS del Conocimiento	Establece las nuevas formas de guardar y almacenar conocimiento a través de la implementación del ecosistema de Big data, relacionado con el conocimiento y consultar el conocimiento a partir de los nuevos sistemas y tecnologías asociadas a Big Data	Se utiliza cuando se evidencian las diferentes formas nuevas de almacenar y consultar el conocimiento a partir de los nuevos sistemas y tecnologías asociadas a Big Data	Proceso de Gestión del Conocimiento y Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	2	5	0
8	Transferencia Inteligente de Conocimiento	Establece la forma de transferir el conocimiento: transferencia máquina a máquina, transferencia máquina a persona y transferencia persona a máquina, que surge a raíz del Big Data	Se utiliza cuando se describe nuevas formas de transferir conocimiento a los individuos que han surgido a raíz del Big Data	Proceso de Gestión del Conocimiento y Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	2	3	0
9	Aplicación Automática del Conocimiento	Establece la forma en como la organización aplica el nuevo conocimiento en la organización cuando a partir del BDA o los cambios de la organización a utilizar y aplicar conocimiento "Edge"	Se utiliza cuando se describe cómo la organización aplica el conocimiento que surge a raíz del Big Data, y/o aplicación in situ o "edge"	Proceso de Gestión del Conocimiento y Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	2	1	4
10	Centros de Excelem en Big Data	Establece la utilización y el apoyo de la organización en la creación de departamentos que impulsen el uso del análisis de datos en la organización y formar departamentos especializados en el análisis de datos para coordinar las iniciativas y proyectos de Big Data	Se utiliza cuando se describen las iniciativas de la organización para formar departamentos especializados en el análisis de datos para coordinar las actividades asociadas a Big Data	Mecanismos de Coordinación y Control y Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	3	2	0
11	Coordinación lateral Inteligente	Establece el uso de la coordinación lateral mediante la integración interdepartamental de forma automática y/o manual para el desarrollo de actividades relacionadas con el Big Data	Se utiliza cuando se describen las iniciativas para conformar grupos interdisciplinarios a raíz de necesidades para la toma de decisiones a través de BDA o análisis de datos	Mecanismos de Coordinación y Control y Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	3	1	1
12	Medición de Desempeño Automático	Establece las técnicas de medición automática que utiliza la organización para controlar los objetivos de la organización	Se utiliza cuando se describe como la organización utiliza formas de medición de desempeño mediante tecnologías o sistemas asociados a Big Data	Mecanismos de Coordinación y Control y Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	2	2	2
13	Control Sinérgico Personas Máquinas, Sistemas	Establece la retroalimentación recíproca entre personas, máquinas y sistemas de la organización	Se utiliza cuando se muestra la forma cómo la organización utiliza los sistemas y tecnologías asociadas a Big Data para evaluar constantemente el cumplimiento de objetivos	Mecanismos de Coordinación y Control y Ecosistemas de Big Data	Efectos asociados a Ecosistemas de Big Data	1	1	2
14	Gestión del Cambio Organizacional	Establece los cambios en procesos que ha tenido que adoptar la organización para poder incluir el Big Data y el BDA	Se utiliza cuando se muestra la necesidad o los cambios que ha tenido que realizar la organización para modificar ciertos procesos y estructuras para obtener ventajas del Big Data	Factores de Diseño Organizativo	Elementos Organizacionales Tradicionales	3	3	0
15	Estructura, Tipo de Coordinación y Control	Establece la distribución de poder y autoridad en la organización, y muestra los componentes de la organización, sus relaciones y la forma de jerarquía utilizada	Se utiliza cuando se expresa el tipo de estructura de la organización y como influye la misma en la toma de decisiones basadas en datos	Factores de Diseño Organizativo	Elementos Organizacionales Tradicionales	0	6	8
16	Medición de Desempeño	Establece como se mide las metas u objetivos en la organización	Se utiliza cuando se describe la forma en como la organización mide el desempeño de sus empleados	Factores de Diseño Organizativo	Elementos Organizacionales Tradicionales	0	2	1
17	Centralización	Establece el nivel de centralización o descentralización que posee la organización en su estado actual	Se utiliza cuando se describen los métodos de tomar las decisiones en la organización y acorde a la forma tradicional de tomarlas	Factores de Diseño Organizativo	Elementos Organizacionales Tradicionales	0	4	1
18	Toma de decisiones	Establece cómo la organización realiza el proceso de toma de decisiones y donde se toman en su estructura organizativa	Se utiliza cuando se menciona algo respecto al proceso de toma de decisiones en la organización y quien toma las decisiones	Factores de Diseño Organizativo	Elementos Organizacionales Tradicionales	0	2	8
19	Gestión de la Calidad	Establece cuando la gestión de la calidad en la organización toma un papel en la implementación de ecosistemas de Big Data, la gestión del conocimiento y/o los mecanismos de coordinación y control	Se utiliza cuando se describe cómo la gestión de la calidad hace parte de la forma cómo se gestiona el conocimiento, se utiliza para los ecosistemas de Big Data o para la coordinación y/o control	Factores de Calidad	Elementos Organizacionales Tradicionales	0	9	0
20	Creación del Conocimiento	Establece como la organización lleva a cabo el proceso de creación del conocimiento	Se utiliza cuando se describe cómo se lleva a cabo este proceso de gestión del conocimiento en la organización de forma tradicional	Proceso de Gestión del Conocimiento	Elementos Organizacionales Tradicionales	0	4	6
21	Almacenamiento del Conocimiento	Establece como la organización lleva a cabo el proceso almacenamiento del conocimiento	Se utiliza cuando se describe cómo se lleva a cabo este proceso de gestión del conocimiento en la organización de forma tradicional	Proceso de Gestión del Conocimiento	Elementos Organizacionales Tradicionales	0	2	2
22								

Fuente: elaboración propia

Con el fin de estructurar la forma en cómo se realiza el análisis de los datos recolectados del estudio de casos, hemos tomado como referencia el modelo de espiral propuesto por Boeije (2010) y que presentamos en la figura 4-3; pero lo hemos adaptado para introducir nuestros códigos deductivos basados en el modelo teórico preliminar expuesto, previamente, en el capítulo 3.

Figura 4-3: Espiral de análisis del proceso de investigación cualitativa



Fuente: adaptado de Boeije (2010, p. 90)

Acorde a lo anterior, en primer lugar, hemos realizado una codificación deductiva a partir de los grandes conceptos que fueron expresados en el modelo teórico preliminar y que fueron introducidos en el capítulo anterior (ver figura 4-4). Dichos códigos fueron agregados a la herramienta en Excel con su respectiva definición y una descripción de cuándo utilizarlos. Posteriormente, hicimos una recolección inicial de los datos en la herramienta desarrollada en Microsoft Excel en donde segmentamos las transcripciones y la documentación corporativa de los casos en fragmentos relevantes para la

investigación. En una primera fase, asignamos códigos deductivos a aquellos fragmentos que se ajustaban a la definición que habíamos hecho del código.

Figura 4-4: Ejemplos de códigos deductivos

#	CÓDIGO	DEFINICIÓN	CUANDO USARLO
6	Creación de Conocimiento Predictivo	Indica que la organización utiliza el BDA para llevar a cabo la creación de conocimiento a partir de algoritmos inteligentes y análisis de datos intermedio/avanzado	Se utiliza cuando se describe como se ha utilizado el Big Data y el BDA para crear nuevo conocimiento en la organización
7	AS/RS del Conocimiento	Establece las nuevas formas de guardar y almacenar conocimiento a través de la implementación del ecosistema de Big data, relacionado con el conocimiento "Edge"	Se utiliza cuando se evidencia las diferentes formas nuevas de almacenar y consultar el conocimiento a partir de los nuevos sistemas y tecnologías asociadas a Big Data
8	Transferencia Inteligente del Conocimiento	Establece la forma de transferir el conocimiento: transferencia máquina a máquina, transferencia máquina a persona y transferencia persona a máquina, que surge a raíz del Big Data	Se utiliza cuando se describe nuevas formas de transferir conocimiento a las tradicionales que han surgido a raíz del Big Data
9	Aplicación Automática del Conocimiento	Establece la forma en como la organización aplica el nuevo conocimiento en la organización creado a partir del BDA o los cambios de la organización a utilizar y aplicar conocimiento "Edge"	Se utiliza cuando se describe cómo la organización aplica el conocimiento que surge a raíz del Big Data, y/o aplicación in situ o edge
10	Centros de Excelencia en Big Data	Establece la utilización y el apoyo de la organización en la creación de departamentos que impulsen el uso del análisis de datos en la organización y para coordinar las iniciativas y proyectos de Big Data	Se utiliza cuando se describen las iniciativas de la organización para formar departamentos especializados para el análisis de datos para coordinar las actividades asociadas a Big Data

Fuente: elaboración propia

Este proceso de codificación deductiva nos dejó algunos segmentos relevantes por codificar ya que los mismos no se ajustaban a la definición de los códigos deductivos. En ese momento, procedimos a crear códigos inductivos que se ajustaran a dichos segmentos de datos primarios que no se habían podido ajustar a los códigos deductivos iniciales. Tras finalizar esta primera etapa de codificación teníamos una lista de códigos en la que algunos de ellos habían sido creados deductivamente mientras que otros lo habían sido de manera inductiva. Esta lista fue el resultado de un proceso en el que se incorporaron nuevos códigos, se reformularon códigos existentes y/o eliminaron códigos redundantes que se habían detectado.

En segundo lugar, realizamos una codificación axial a través de la que se establecieron subcategorías y categorías que integraron dichos códigos generados en la primera fase del proceso de análisis. Algunas de estas subcategorías / categorías procedían de nuestro modelo teórico preliminar mientras que otras fueron generadas de manera específica a los datos de nuestra investigación. Durante todo el proceso de análisis aplicamos el principio de la comparación constante. Como resultado del proceso de análisis establecimos qué códigos, subcategorías y categorías eran relevantes para responder la pregunta de investigación de esta tesis doctoral, lo que permitirá la construcción del modelo conceptual que surge de los casos estudiados. Los códigos definitivos fueron consignados en un informe preliminar para cada caso y enviado para su revisión a uno de los

entrevistados para su verificación, lo que incrementó la confiabilidad en el proceso de análisis cualitativo realizado.

4.6 Criterios de calidad

Al momento de concluir con el análisis en la investigación cualitativa, uno de los problemas que enfrenta el investigador y la investigación es poder demostrar la credibilidad y relevancia de los resultados obtenidos (Saunders et al., 2016). Este problema en la investigación cualitativa, y en particular en el estudio de casos, es recurrente y uno de los puntos más criticados, al considerarse en ocasiones que los resultados obtenidos en este tipo de investigaciones no corresponden o no se ajustan ampliamente a la vida real (Tellis, 1997).

Con el objetivo de aportar credibilidad a la investigación cualitativa en ciencias sociales, incluyendo al estudio de casos dentro de los mismos, la literatura especializada propone cuatro criterios que permiten incrementar la calidad en la investigación cualitativa. Estos criterios son los que describimos a continuación:

- ✓ **Fiabilidad:** consiste en determinar si la investigación realizada es razonablemente estable en el tiempo, y que los métodos, procedimientos y resultados utilizados y obtenidos en la investigación son replicables por otros investigadores (Miles y Huberman, 1994; Saunders *et al.*, 2009). Acorde con Yin (2014), las principales tácticas para aumentar la fiabilidad son: 1) elaborar y aplicar un protocolo de estudio de casos, 2) la construcción de una base de datos del estudio, 3) elaborar un libro de códigos y 4) utilizar citas textuales procedentes de las entrevistas.
- ✓ **Validez del Constructo:** consiste en identificar si las medidas y conceptos seleccionados son adecuados para estudiar el fenómeno de interés en la investigación (Boeije, 2010; Yin, 2014). Acorde con Yin (2014), las principales tácticas para incrementar la validez del constructo son: 1) uso de múltiples fuentes de evidencia, 2) la triangulación, 3) establecer una cadena de evidencia, 4) comprobar las transcripciones con los entrevistados y 5) revisión del borrador del caso por parte del participante clave.

- ✓ **Validez Interna:** consiste en establecer las relaciones causales entre las variables y los hallazgos encontrados en los resultados de la investigación, y cuya relación corresponda con la vida real (Miles & y Huberman, 1994; Saunders *et al.*, 2009). Yin (2014) considera que esta prueba es requerida en estudios explicativos o causales mas no aplica en estudios descriptivos o exploratorios, y establece que las tácticas más utilizadas para aumentar la validez interna son: 1) definir un modelo teórico inicial, 2) la búsqueda de correspondencia con patrones identificados en otros contextos en la literatura, 3) la triangulación teórica durante la construcción del marco teórico, y 4) la aplicación del método de la comparación constante.
- ✓ **Validez Externa:** consiste en determinar si la investigación realizada y sus resultados o hallazgos se pueden generalizar más allá de la propia investigación (Saunders et al., 2016; Yin, 2014). Acorde con Yin (2014), las principales tácticas para aumentar la validez externa son: 1) la lógica de la réplica de los hallazgos en estudios con múltiples casos, 2) la selección de casos con distintas características, y 3) la explicación con detalle del contexto de los casos de estudio.

Teniendo en cuenta lo anterior y excluyendo a la validez interna al no corresponder con esta investigación por ser de carácter exploratorio, aplicando los criterios anteriores a nuestra investigación podemos establecer lo siguiente. En relación a la *fiabilidad*, utilizamos un protocolo de casos de estudio el cual hemos consignado en el anexo A de esta tesis doctoral, y una base de datos de los casos en donde se registran los códigos iniciales y los que han surgido a lo largo del análisis de los datos de las entrevistas y de la documentación corporativa recolectada para la construcción del modelo para cada uno de los casos. Además, durante la elaboración del informa final utilizaremos citas textuales de los entrevistados para sustentar los argumentos presentados en el informe final.

Por lo que se refiere a la *validez de constructo*, hemos utilizado múltiples fuentes de evidencia mediante la documentación corporativa y la realización de las entrevistas con distintas personas que se encontraban relacionadas con los temas de interés de la investigación en cada uno de los casos que aportaban un análisis del fenómeno desde distintos ángulos. A su vez, hemos comparando las entrevistas con la documentación corporativa y los datos recolectados, táctica conocida como triangulación (Boeije, 2010). Como táctica adicional para la validez del constructo, se han validado los resultados

obtenidos en cada caso directamente con los entrevistados. En este sentido, se envió el informe final de cada caso a los entrevistados correspondientes al caso para que modificaran aquello que no se ajustaba a la realidad.

Para el criterio de la *validez externa*, hemos aplicado la lógica de la réplica literal (Yin, 2014) al seleccionar cuidadosamente cada uno de los casos para poder predecir resultados similares o predecir resultados contrastantes, tema que se encuentra relacionado con los criterios que hemos definido para la selección de los casos más idóneos para esta investigación (ver sección 4.2.1). Además, hemos seleccionado organizaciones con distintas características para estudiar el fenómeno desde distintas perspectivas, características que serán explicadas en detalle en la elaboración del informe final de los casos.

A modo de síntesis, en la tabla 4-6 (siguiente página) se resumen las tácticas utilizadas para incrementar la calidad en esta investigación.

Tabla 4-6: Tácticas para cumplir los criterios de calidad de la presente investigación

Criterio	Táctica para casos de estudio	Etapa de la Investigación
Fiabilidad	✓ Usar un protocolo de casos de estudio	Diseño del estudio de casos
	✓ Desarrollar una base de datos de los casos de estudio	Recolección y análisis de datos
	✓ Utilización de citas textuales de los entrevistados	Informe final del caso
Validez del Constructo	✓ Usar múltiples fuentes de evidencia (entrevistas, documentos corporativos)	Recolección de datos
	✓ Triangulación	Análisis de datos
	✓ Revisión de los borradores de los reportes	Informe final del caso
Validez Externa	✓ Usar lógica de la réplica	Diseño del estudio de casos
	✓ Selección de casos con distintas características	Diseño del estudio de casos
	✓ Explicación del contexto de los casos de estudio	Informe final del caso

Fuente: elaboración propia a partir de Yin (2014)

Una vez hemos definido todos los aspectos metodológicos a aplicar en esta investigación, en el siguiente capítulo presentaremos los resultados obtenidos como consecuencia de la aplicación de los mismos. Así, en el siguiente capítulo presentaremos el análisis individual de cada uno de los casos de estudio para posteriormente realizar un análisis conjunto de los casos donde se contrastarán cada uno de los aspectos analizados en los casos individuales y que resultará en el modelo final propuesto en esta tesis doctoral.

Capítulo 5:

Análisis de Resultados

En el presente capítulo procedemos a analizar la información primaria de cada uno de los casos seleccionados para esta investigación y que hemos presentado en el capítulo anterior (ver sección 4.2.1). Al momento de explorar y analizar los datos e información recolectado de cada caso, además de seguir las consideraciones descritas en el capítulo anterior hemos buscado una uniformidad y objetividad en el análisis para que el aporte de cada caso sea de importancia para los objetivos de esta investigación.

Cabe aclarar que durante la realización de las entrevistas se tomaron notas sobre aportes de los entrevistados en aspectos que no fueron considerados durante el diseño del modelo teórico de relaciones inicialmente planteado y en la construcción del guion de la entrevista. Por lo tanto, cada caso tiene aportes específicos que enriquecen el modelo teórico de relaciones, y, además, contribuyen al planteamiento de futuras líneas de investigación que surgen a raíz del análisis de los casos.

Para facilitar el seguimiento del análisis y la comparación entre los casos, todos los casos presentan la misma estructura. Así, hemos dividido cada caso en cinco secciones distintas. En primer lugar, realizamos una descripción del perfil de la organización; en segundo lugar, describimos cómo se desarrolla la gestión del conocimiento en el caso estudiado; en tercer lugar, analizamos cómo se aplican los mecanismos de coordinación y control en la organización; en cuarto lugar, describimos el ecosistema de Big Data presente en la misma; y, en quinto lugar, realizamos un análisis del efecto de la implementación del ecosistema de Big Data existente en la empresa sobre sus procesos de gestión del conocimiento y sus mecanismos de coordinación y control.

Con este último apartado del caso estaremos dando respuesta a la pregunta que dio origen a la presente investigación y que tiene que ver con cómo la introducción del Big Data en una organización afecta a sus procesos de gestión del conocimiento y a la forma de coordinación del trabajo. Así mismo, el análisis de cada caso nos permitirá revelar patrones y relaciones que no estaban contemplados en el modelo teórico preliminar de la tesis, lo que representaremos de manera gráfica en forma de modelo inducido del caso. Por último, una vez presentado el resultado del análisis individual de cada caso, en un último apartado procedemos a desarrollar el análisis conjunto de los mismos lo que nos permitirá establecer patrones comunes de ocurrencia, tal y como sugiere Yin (2014).

5.1 Caso A: Cementos Colombia S.A.

En el presente caso de estudio analizamos la empresa Cementos Colombia, empresa dedicada a la producción y comercialización de cemento, concreto y sus derivados a nivel nacional e internacional. Para tal fin, presentamos el perfil de la organización, su proceso de gestión del conocimiento, los mecanismos de coordinación y control utilizados y el efecto de la implementación del ecosistema de Big Data en la organización en los dos últimos puntos antes mencionados y en otros procesos particulares del caso que se han visto afectados por la implementación del ecosistema.

5.1.1 Perfil de la Organización

La organización Cementos Colombia S.A. fue fundada en el año 1934 y posee como actividad principal la producción de cemento, concreto y agregados para la construcción. En este sector, la organización posee un posicionamiento privilegiado en el mercado colombiano al ser el líder en ventas de cemento y concreto, y ser una de las empresas más importantes para la economía colombiana al ser una de las 20 empresas del índice COLCAP de la bolsa de valores de Colombia (equivalente al IBEX 35 en España).

Además del mercado colombiano, Cementos Colombia S.A. se ha posicionado como una de las empresas más importantes y con mayor liderazgo de la industria del cemento y concreto en toda América como consecuencia de su presencia en otros 14 países, ubicados principalmente en América del Norte, Centroamérica y el Caribe. A continuación, detallamos la presencia de Cementos Colombia S.A. en toda América, divididos entre plantas de cemento, plantas de concreto, plantas de molienda de Clínker (insumo del cemento) y puertos/terminales de empaque y transporte (ver tabla 5-1 página siguiente):

Tabla 5-1 Presencia de Cementos Colombia S.A. en América

PAÍS	Tipo de Propiedad				Total
	Planta de Cemento	Planta de Concreto	Planta de Clíinker	Puertos y/o Terminales	
Estados Unidos	4	236	2	21	263
Colombia	7	65	1	1	74
Panamá		7	1	1	9
República Dominicana		4	1	1	6
Surinam		1	1	1	3
Honduras	1		1		2
Puerto Rico	1			1	2
Guayana Francesa			1	1	2
Haití			1		1
Antigua				1	1
Curazao				1	1
Dominica				1	1
San Martín				1	1
Santo Tomás				1	1
Venezuela*					
Total	13	313	9	32	367

* presencia a través Cemento Andes C.A., en disputa mediante proceso judicial por expropiación.

Fuente: Elaboración propia a partir de (Cementos Colombia S. A., 2019d).

Esta presencia en todo el continente le ha valido a Cementos Colombia S.A. el ser el quinto productor de cemento más grande en toda América Latina, y en un mercado tan importante como es el de Estados Unidos su presencia lo ha llevado a ser el cuarto productor más grande de cemento del país y a ser el segundo productor más grande de concreto del mismo.

En línea con lo anterior, para el año 2018 Cementos Colombia S.A. obtuvo como resultado de su operación en todos los países en donde tiene presencia un total de 16 millones de toneladas de cemento y 10,3 millones de metros cúbicos de concreto vendidos lo cual, en conjunto con sus demás productos y servicios, se tradujo en 8,6 billones de

pesos colombianos (aproximadamente 2.221 millones de Euros) en ingresos totales (52% Estados Unidos, 27% Colombia y 21% Centroamérica y el Caribe), atendiendo a un total de 22.000 clientes y logrando exportaciones a más de 34 países del mundo (Cementos Colombia S. A., 2019e).

En la actualidad, Cementos Colombia S.A. pertenece a una organización llamada Grupo Colombia, grupo de inversiones enfocado principalmente en infraestructura y energía, siendo sus principales empresas: Cementos Colombia S.A., casa matriz del Grupo Colombia; la empresa Celta S.A. E.S.P., enfocada en la generación de energía convencional y renovable; y la empresa Odenza S.A., enfocada en proyectos de infraestructura y concesiones viales y aeroportuarias. El Grupo Colombia se ha consolidado en Colombia como uno de los conglomerados empresariales más importantes del país, siendo parte, al igual que Cementos Colombia, del índice COLCAP de la bolsa de valores de Colombia (Bolsa de Valores de Colombia, 2019).

5.1.2 Proceso de Gestión del Conocimiento en Cementos Colombia S.A.

Tal y como vimos en la sección 2.1.3, la gestión del conocimiento presenta cuatro fases o etapas fundamentales que son la creación del conocimiento, el almacenamiento del conocimiento, la transferencia del conocimiento y la aplicación del conocimiento. A continuación, exponemos como se desarrolla el proceso de gestión del conocimiento en Cementos Colombia S.A.

La empresa Cementos Colombia S.A. es una organización con una visión clara respecto a la gestión del conocimiento y lo enmarca en una estrategia de innovación, con un sólido apoyo por parte de los directivos para su desarrollo a nivel interno de la organización, de sus proveedores y clientes. Para esta organización es importante la articulación entre la gestión del conocimiento y la innovación, tanto que se incluye en su política corporativa (Cementos Colombia S. A., 2019e), y es considerado como fundamental para la generación de valor, promoviéndose el desarrollo integral de los actores que involucran a la organización para que se incorpore nuevo conocimiento, producir mejores prácticas y así poder mejorar el desempeño de la organización.

La importancia de la innovación como parte fundamental de la estrategia del conocimiento de la empresa ha llevado a la creación del Centro Colombia para la Innovación, un centro de investigación con 4.807 m² de construcción ubicado en la ciudad de Medellín, en donde se concentra su operación de innovación y donde se desarrollan proyectos para el fortalecimiento de la industria, la organización y aportes científicos al país mediante un modelo de apoyo entre la empresa, universidades y el Estado Colombiano (Cementos Colombia S. A., 2019a). Por medio de este centro, se busca resolver problemas de la industria del cemento y el concreto y que, a través de la resolución de los mismos, se dé la creación de nuevo conocimiento de valor para la empresa.

En línea con lo anterior, un factor que se considera importante para promover el desarrollo de los proyectos de innovación es el empoderamiento de empleados, investigadores, estudiantes, proveedores y clientes, a la hora de proponer ideas que ayuden a la generación de valor para la organización y que, a su vez, se beneficien las personas, organizaciones, proveedores, clientes y/o el Estado, del conocimiento que surge a raíz de las innovaciones desarrolladas (Cementos Colombia S. A., 2019c).

La generación, validación e implementación de estas ideas que surgen en la organización son gestionadas a través de su sistema de innovación llamado “Ideacción”, que luego de un proceso de evaluación, aprobación y materialización de las mismas, los colaboradores y empleados de la organización pueden interactuar con las ideas y soluciones propuestas, facilitando así su transferencia y aplicación. Por lo tanto, “Ideacción” se ha convertido en una herramienta tecnológica en donde se realizan los procesos de creación, almacenamiento, transferencia y aplicación de conocimiento de la organización, dada la permeabilidad que tiene en todas las áreas funcionales, la visibilidad de la información y conocimiento que provee, y la promoción de una cultura de compartición del conocimiento (Cementos Colombia S. A., 2019c).

Sin embargo, en la organización existen formas tradicionales para gestionar el conocimiento, en especial del conocimiento explícito, como los manuales de operación y de funciones, procedimientos y procesos documentados, que son almacenados de manera local en la organización por medio de su sistema de información, y que son actualizados constantemente de acuerdo con los cambios organizacionales que vayan surgiendo y a las

políticas del sistema de gestión de calidad. Además, la experiencia, habilidades y conocimientos de sus empleados y colaboradores son documentados con el fin de crear un directorio de expertos que sirva en casos de que se requieran sus habilidades y conocimientos en otras áreas de la organización (Cementos Colombia S. A., 2019c).

El área de gestión humana es la encargada de gestionar los esfuerzos de innovación y conocimiento de la organización, y tiene como objetivo el promover la transparencia, creación, transferencia y aplicación del conocimiento de la organización, no solo a través de las plataformas tecnológicas, sino además de la comunicación directa entre las personas, jornadas de capacitaciones entre clientes, proveedores y empleados, y el desarrollo de competencias (Cementos Colombia S. A., 2019b).

La organización reconoce que la innovación y el conocimiento debe ser una de sus prioridades, ya que lo ven como su principal fuente de competitividad sostenible, y es lo que le permitirá el desarrollo de nuevos productos, servicios, nuevos modelos de negocio y nuevas estructuras organizacionales que le faciliten crear valor para sus clientes y seguir siendo una de las empresas líderes en el continente.

Por lo tanto, y a modo de resumen, la empresa Cementos Colombia promueve una sinergia entre el conocimiento y la innovación para que sus empleados, proveedores y clientes hagan parte de la creación de valor de la organización mediante las ideas y proyectos planteados para poder crear nuevo conocimiento, apoyándose en plataformas tecnológicas como “Ideacción” y diversas fuentes de conocimiento tácito y explícito de la organización para poder gestionar los procesos de creación, almacenamiento, transferencia y aplicación de conocimiento. Todas las estrategias de innovación y conocimiento están enmarcadas en políticas claras que hacen que el conocimiento sea un activo accesible en toda la organización, para que de esta forma se convierta en uno de sus principales medios de creación de valor y mejoras en el desempeño.

5.1.3 Mecanismos de Coordinación y Control en Cementos Colombia S.A.

Tal y como vimos en el marco teórico (sección 2.2.2 de esta tesis doctoral) queremos conocer cómo se producen la coordinación y el control en Cementos Colombia S.A. Por lo que respecta a los mecanismos de coordinación estableceremos en primer lugar el nivel

de formalización y de descentralización que posee la organización, para luego exponer los mecanismos de coordinación vistos en el marco teórico, que puede ser por un alto directivo, por procedimientos y planeación formales, lateral o una combinación de las anteriores. Por lo que respecta a los mecanismos de control, estableceremos si el utilizado por la organización corresponde a un control centralizado personal, burocrático, por resultados, mediante vigilancia electrónica, mediante gestión de recursos humanos, cultural o una combinación de las anteriores.

Con respecto a los *mecanismos de coordinación*, la empresa Cementos Colombia S.A. posee una alta formalización de las funciones que desempeñan sus empleados, establecido mediante políticas y procedimientos que están difundidos abiertamente a lo largo de la organización. Estas políticas y procedimientos rigen el proceder de cada una de las áreas de la organización y establecen las reglas, directrices y recomendaciones para desempeñar cualquier labor de la organización acorde a la cultura y los valores que desea promover. Dichos procedimientos, normas y directrices se encuentran consignados como documentación en un estructurado sistema de gestión de la calidad que es de acceso público en toda la organización. El sistema de gestión de la calidad ha promovido la estandarización de los procesos de la organización, lo que ha facilitado la coordinación de las actividades de manera transversal en la compañía.

A pesar de lo anterior, Cementos Colombia S.A. no lleva a cabo una coordinación mediante procedimientos y planeación formales. Al contrario, la empresa ha optado por un equilibrio en la centralización/descentralización de las decisiones, dado que la coordinación del trabajo y las actividades se realiza a niveles medios como las jefaturas y los directores de áreas. Por lo tanto, en los mandos medios se toman las decisiones de carácter operativo, y pueden ser apoyadas por otros empleados que conforman las células o grupos de trabajo interdisciplinarios.

Por ende, los altos mandos como la junta directiva, presidente, vicepresidencias o gerentes, no se dedican a coordinar las actividades de la organización. Esta coordinación del trabajo tiene lugar en el interior de cada célula de trabajo, donde los mandos medios son los encargados de formar los grupos y establecer el propósito y las actividades a desarrollar en cada la célula de trabajo. Tal y como lo reconoce el Director de Gestión Humana:

“La empresa trabaja mucho como si fueran células de trabajo independientes, cada célula de trabajo, o cada grupo de trabajo, coordina el trabajo al interior de esa área”.

Sin embargo, a pesar de haber cierto nivel de descentralización de las decisiones operacionales, las decisiones importantes como cambios en procesos o algunas cuestiones del negocio a nivel estratégico se toman en mandos altos, es decir, desde las gerencias, vicepresidencias, presidencia y junta directiva, encontrándose una mayor autonomía en este tipo de decisiones desde los niveles de vicepresidencia para arriba.

Por lo tanto, y como mencionamos en la sección 2.2.2.1, Cementos Colombia S.A. sigue la tendencia de las organizaciones actuales al haber un equilibrio entre la centralización (decisiones estratégicas) y la descentralización (decisiones operacionales), y en donde la coordinación de las actividades de la organización se lleva a cabo mediante una coordinación lateral en las células de trabajo liderada por los mandos medios, con autonomía en la toma de decisiones operacionales.

En el caso de los mecanismos de control, en Cementos Colombia S.A. se aplican dos mecanismos principalmente: el control de resultados en conjunto con la supervisión directa. Por lo que respecta al control de resultados, los jefes de cada área definen para cada actividad y/o proyecto un indicador o KPI (acrónimo en inglés de “Key Performance Indicator”) que mide el cumplimiento de los objetivos y metas trazadas al inicio de cada actividad y/o proyecto. Sin embargo, la organización ha preestablecido indicadores de rendimiento que miden el nivel de desempeño individual acorde a las responsabilidades de cada puesto de trabajo. Tal y como establece el Director de Analytics:

“En todos los proyectos se define cuáles son los objetivos del proyecto y cuáles son los KPIs con los que se mide el éxito del proyecto...todo proyecto involucra pues un líder, unos recursos, un cronograma, un presupuesto y siempre hay como unas metas al final en cada proyecto”.

Para apoyar la coordinación y el control de actividades, además de los mecanismos tradicionales de coordinación y control en esta empresa se lleva a cabo la coordinación y

el control de actividades a través de otros mecanismos como la aplicación de metodologías ágiles. En este sentido, Cementos Colombia S.A. ha empezado a implementar la metodología SCRUM. Esta metodología consiste en un marco de trabajo para organizar equipos para la ejecución de un proyecto, donde se procura hacer un desarrollo incremental en el proyecto priorizando la creación de valor, realizado mediante una sincronización periódica del equipo (en ocasiones realizadas diariamente), sus resultados y el cumplimiento de los objetivos, dando la autoridad necesaria para la toma de decisiones dentro del equipo para poder completar cada proyecto (Scrum.org, 2019). Por lo cual, estas metodologías ágiles se han adaptado a la forma de trabajar en células de trabajo de la organización, consiguiendo así una mejor coordinación y control de las actividades y proyectos desarrollados. Podríamos establecer, por tanto, que la metodología SCRUM es una forma contemporánea de normalización de resultados.

Por su parte, en relación con la supervisión directa, cada jefe o coordinador de cada área o célula de trabajo es el encargado de realizar el seguimiento a la ejecución de las actividades, apoyándose en la metodología SCRUM que le permite evaluar en tiempo real como está el cumplimiento de los objetivos. A su vez, los jefes o coordinadores son los encargados de solicitar reportes sobre el avance de cada empleado, reportes que son consolidados en cada área para presentar dichos resultados a los altos mandos.

5.1.4 Ecosistema de Big Data en Cementos Colombia S.A.

La empresa Cementos Colombia S.A. ha sido una de las pioneras en Colombia en la introducción de los ecosistemas de Big Data, siendo una de las primeras en considerar la implementación de ecosistemas de Big Data como estrategia de crecimiento. Al ser una de las empresas más grandes del país y de Latinoamérica, ha visto la necesidad de obtener ventajas competitivas en su mercado y una de las estrategias para conseguir esto ha sido mediante el análisis de datos de la organización, estrategia que le ha llevado a implementar varias herramientas y tecnologías que apoyan la toma de decisiones y el conocimiento de la organización.

Para analizar cómo la empresa Cementos Colombia S.A. ha ido realizando la implementación de ecosistemas de Big Data, a continuación, abordaremos el estudio de

cada una de las cinco dimensiones planteadas en la sección 1.3.3 lo que nos permitirá determinar el nivel de implementación de Big Data en el que se encuentra la organización.

Dimensión 1: Implicación Organizativa de Big Data

Como mencionamos anteriormente, la organización ha sido una de las pioneras en Colombia en la implementación de ecosistemas de Big Data. Sin embargo, la organización ha necesitado realizar cambios en su forma de trabajar, principalmente en su estructura y cultura. Tal y como nos comenta el Director de Analytics de la compañía, para implementar con éxito este tipo de iniciativas es crucial el apoyo de los altos mandos de la organización.

“[...] también depende mucho del apoyo. Estamos trabajando en una vicepresidencia que es la vicepresidencia financiera, porque el Centro de Analítica se encuentra dentro de la vicepresidencia financiera, y tenemos el apoyo total del vicepresidente. Entonces, eso hace que sea fácil de que las cosas se den”.

A raíz del apoyo mencionado, la organización optó por crear el Centro de Analítica (considerado como Centro de Excelencia, ver sección 3.2.1), que es el encargado de llevar a cabo el análisis de los datos de la organización para la toma de decisiones. Además, el Centro de Analítica tiene como función crear y difundir la cultura de análisis de datos (cultura analítica) a lo largo de toda la organización, basándose para ello en estrategias de difusión del conocimiento mediante modelos, metodologías y técnicas, que apliquen el análisis de datos para tomar decisiones.

Por lo tanto, Cementos Colombia S. A. ha realizado esfuerzos en todos los niveles de mando para expandir la necesidad del uso del análisis de datos en toda la organización para mejorar las labores diarias y establecer una independencia en los mandos medios para la toma de decisiones operacionales que se encuentren basadas en datos, aportando así a la descentralización de estas decisiones. Para lograr lo anterior, la empresa ha tenido claro que la inversión en nuevas tecnologías tiene que ser clave para el crecimiento económico de la organización. En este sentido, el Director de Analytics reconoce lo siguiente:

“[...] es política de la empresa crecer en este tipo de nuevas tecnologías porque le ve claramente un beneficio...de alguna manera saben que están invirtiendo recursos para conseguir mejoras a nivel de desempeño de la compañía”.

A modo de resumen y por lo que respecta a esta dimensión, la empresa Cementos Colombia S.A. ha destinado los recursos y ha adquirido las capacidades para utilizar y gestionar el ecosistema de Big Data, expandiendo la cultura de pensamiento analítico a lo largo de la organización y descentralizando las decisiones operacionales, buscando así ventajas competitivas y mejorar el desempeño. Consecuentemente, consideramos que la organización, según lo planteado en la sección 1.3.3 de esta tesis doctoral, se encuentra en el nivel cinco en la dimensión de implicación organizativa de Big Data.

Dimensión 2: Factores Humanos y su Relación Máquinas - Sistemas

Como hemos mencionado en secciones anteriores de este caso, la empresa Cementos Colombia S.A. es una de las pocas organizaciones en Colombia en poseer un departamento especializado en análisis de datos. Este departamento se denomina Centro de Analítica. La incorporación del Centro de Analítica a la organización ha implicado adquirir nuevos conocimientos y habilidades mediante la incorporación de nuevo personal capacitado en los temas que requiere la compañía y en la capacitación del personal para el desarrollo y/o uso de las herramientas y reportes que surgen del Centro de Analítica. En este sentido, el Director de Analytics nos manifiesta lo siguiente:

“...el tema es hacer una capacitación con la gente que de alguna forma los llevé a utilizar las herramientas que está desarrollando el Centro de Analítica [...] la idea es aprovechar las grandes cantidades de datos que posee la organización para poder hacer análisis y extraer conclusiones que permitan mejorar el desempeño [...] eso es un cambio de mentalidad, ya que la idea es tratar de usar menos la intuición y que todas las decisiones sean lo más racional posible, sin dejar de lado por supuesto lo que es la experiencia o la experticia de algunas personas que ya llevan mucho tiempo en la organización”.

Por lo tanto, el Centro de Analítica se ha convertido, además del núcleo principal para el análisis de datos de la organización, en el área encargada de ofrecer capacitaciones en

temas y herramientas de Big Data y Analytics, así como en el uso de los desarrollos y reportes que surjan del Centro de Analítica.

Así, la organización completa ha ido incorporando poco a poco estas herramientas, técnicas y métodos de los ecosistemas de Big Data en sus labores diarias, ya que la cantidad de datos e información en los puestos de trabajo ha crecido sustancialmente, teniendo los directivos que hacer énfasis en el cambio de mentalidad en la organización (cultura analítica) para aprovechar estos datos para mejorar el desempeño. En palabras del Director de Analytics:

“[...] la gente en este momento ya sabe que tiene que pasar a algo más interactivo usando herramientas de visualización y nosotros estamos impulsando ese cambio [...] y dado el volumen de información de la empresa, te obliga a que tengas que manejar herramientas e información en la nube”.

Por consiguiente, tal y como hemos visto la organización ya cuenta con un departamento especializado para el análisis de datos, y su personal y sus directivos están incorporando tanto las herramientas, técnicas y métodos de los ecosistemas de Big Data y el conocimiento adquirido mediante los análisis que surgen del Centro de Analítica. Así, consideramos que, según lo planteado en la sección 1.3.3 de esta tesis doctoral, la organización se encuentra en el nivel cinco en la dimensión de factores humanos y su relación máquinas-sistemas.

Dimensión 3: Método Utilizado de Analytics

En la actualidad, el principal método utilizado de Analytics en la empresa Cementos Colombia S.A. es el de la analítica descriptiva, empleado principalmente para manejar y presentar informes e información de la parte operativa y del negocio de concreto y cemento, haciendo análisis de las ventas, de la operación y de sus clientes para tomar decisiones en cualquiera de estas áreas.

Sin embargo, la organización está comenzando a realizar los primeros análisis de datos con analítica predictiva y prescriptiva mediante la aplicación de modelos de optimización y modelos de simulación, permitiendo el análisis de futuros escenarios del negocio y anticipar decisiones que se deben tomar acorde se vaya moviendo el mercado, buscando

utilizar los tres métodos conjuntos a medida que el Centro de Analítica se vaya fortaleciendo. En este sentido, el Director de Analytics nos comenta lo siguiente:

“Vamos avanzando en la medida que hay más madurez en el equipo, digamos de analítica. Ya en algún momento, pues esperamos hacer algo de Machine Learning e inteligencia artificial”.

Por consiguiente, la organización ya se encuentra realizando plenamente analítica descriptiva y aplicando, en ocasiones, analítica predictiva y prescriptiva. De esta manera, consideramos que, según lo planteado en la sección 1.3.3 de esta tesis doctoral, la organización se encuentra en el nivel cinco en la dimensión de método utilizado de Analytics.

Dimensión 4: Clase de Herramienta de BDA

En la empresa Cementos Colombia S.A. se utilizan principalmente grandes cantidades de datos e información histórica para realizar reportes y análisis del negocio, sin necesidad de que hagan parte de este análisis datos en tiempo real, por lo que se requiere para tal fin un procesamiento por lotes.

Sin embargo, la organización también posee, mediante su sistema de información, la posibilidad de analizar ciertos datos del negocio que son recolectados y almacenados en la nube en tiempo real, siendo el sistema capaz de realizar análisis y reportes rápidos a las personas encargadas de tomar decisiones, siendo un ejemplo del uso del procesamiento en tiempo real. Por tanto, dado que la empresa utiliza el procesamiento por lotes y, en ocasiones, el procesamiento en tiempo real, consideramos que, según lo planteado en la sección 1.3.3 de esta tesis doctoral, la organización se encuentra en el nivel cinco en la dimensión de clase de herramienta de BDA.

Dimensión 5: Integración de Tecnologías de Información

La empresa Cementos Colombia S.A., a raíz de la implementación del ecosistema de Big Data en la organización, ha realizado inversiones importantes en tecnología y transformación digital para poder incorporarse de la mejor manera en todos los procesos de trabajo. Como ejemplo de esto, la empresa ha dejado poco a poco de utilizar servidores in situ para trasladar los datos, la información, el conocimiento explícito y el manejo de

ciertos procesos, a servicios que se encuentran en la nube (Cloud Computing), transformación que requirió un apoyo importante del Centro de Analítica para su implementación.

Además de trasladar muchos procesos a servicios en la nube, la organización ha comenzado proyectos para utilizar Internet de las Cosas para manejar datos e información del proceso productivo en tiempo real, utilizando sensores que miden variables específicas del proceso para luego construir un gemelo digital en la nube (representación digital de un proceso o máquina física). El objetivo de toda esta actuación es que la información esté disponible para las personas encargadas y que puedan tomar decisiones de forma ágil y sin necesidad de estar presentes.

De acuerdo con lo anterior, con la implementación de las tecnologías de Cloud Computing e Internet de las Cosas en la organización consideramos que, según lo planteado en la sección 1.3.3 de esta tesis doctoral, la organización se encuentra en el nivel cinco en la dimensión de Integración de Tecnologías de Información.

Como conclusión de esta sección y analizando las cinco dimensiones de los niveles de implementación de ecosistemas de Big Data, nos damos cuenta de que la empresa Cementos Colombia S.A. ha conseguido las habilidades, las tecnologías, el personal y un desarrollo de sus procesos para alcanzar el nivel cinco en cada una de las dimensiones antes expuestas. De esta manera, podemos establecer que la organización estudiada alcanza el nivel cinco en la implementación de ecosistemas de Big Data, lo cual se corresponde, tal y como vimos en el capítulo 1 de esta tesis doctoral, con un *Ecosistema Básico de Big Data* (ver figura 1-9 en el capítulo 1 de la tesis).

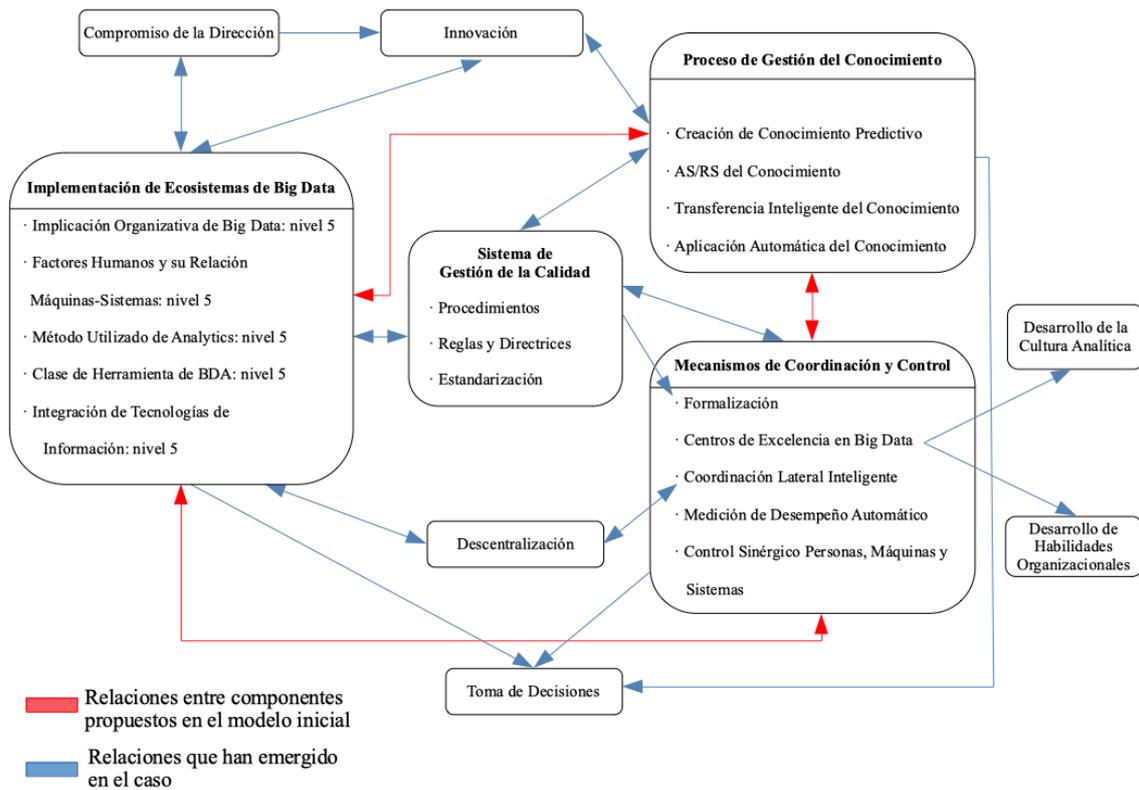
Consideramos que el hecho que la empresa Cementos Colombia S.A. pertenezca al nivel cinco en todas las dimensiones se debe a que es una organización que se ha percatado de la necesidad de utilizar el Big Data para conseguir ventajas competitivas en su mercado, sobre todo por el hecho de que compite en mercados internacionales y el apoyo de sistemas y tecnologías asociadas al Big Data ha sido el camino a seguir de grandes empresas en el continente. Dado lo anterior, la organización ha visto la necesidad de involucrar a todos los niveles de mando y a todas las áreas de la organización en este proceso de implementación para conseguir un proceso de transformación hacia el

ecosistema básico de Big Data, lo cual ha ido consiguiendo mediante la unificación de esfuerzos, criterios y objetivos en toda la organización.

5.1.5 Efecto de la implementación del ecosistema de Big Data en los procesos de gestión del conocimiento y en los mecanismos de coordinación y control en Cementos Colombia S.A.

Una vez hemos explicado y analizado cómo se desarrolla el proceso de gestión del conocimiento, hemos descrito y analizado los mecanismos de coordinación y control, y hemos determinado el nivel de implementación del ecosistema de Big Data en Cementos Colombia S.A., en la presente sección pasamos a estudiar el efecto que ha tenido esta implementación del ecosistema de Big Data en el proceso de gestión del conocimiento y en los mecanismos de coordinación y control de la compañía. A su vez, estos últimos han condicionado también la implementación del ecosistema de Big Data en la empresa lo que nos lleva a plantearnos la existencia de un efecto bidireccional entre ambos. La siguiente figura nos muestra el resultado final de nuestro análisis del caso de la empresa Cementos Colombia S. A. (ver figura 5-1 en página siguiente).

Figura 5-1: Modelo final de relaciones para el caso de Cementos Colombia S.A.



Fuente: elaboración propia

Tal y como puede apreciarse en la figura anterior, el modelo final obtenido tras el análisis de la información del caso de la empresa Cementos Colombia S. A. nos presenta cómo se relaciona la implantación del Ecosistema de Big Data con los procesos de gestión del conocimiento y los mecanismos de coordinación y control en dicha organización. Al compararlo con nuestro modelo inicial (ver figura 3-5 en el capítulo 3 de esta tesis), vemos que tras el estudio del caso han aparecido nuevos elementos, factores organizativos y patrones que nos ayudan a entender mejor cómo tiene lugar dicha relación (están unidos mediante flechas de color azul en la figura 5-1), al igual que también existen otros constructos y elementos que ya habían sido contemplados en el modelo inicial (unidos mediante flechas de color rojo en la figura 5-1). Para facilitar al lector el seguimiento de nuestro análisis, primero presentaremos aquellos componentes del modelo que ya estaban presentes en el modelo inicial y, posteriormente, haremos lo propio con aquellos otros que han emergido durante el análisis del caso de Cementos Colombia S. A.

Componentes propuestos en el modelo inicial

Tal y como se ha descrito en el apartado 5.1.4 anterior, Cementos Colombia S. A. presenta un Ecosistema Básico de Big Data con cinco dimensiones que alcanzan un nivel cinco en todas ellas. Del mismo modo, existen cuatro grandes procesos de gestión del conocimiento que se dan en la organización y que fueron analizados y descritos anteriormente en el apartado 5.1.2 de este capítulo. Sin embargo, y tal y como anticipábamos unas líneas más arriba, ambos constructos no coexisten de forma aislada en el caso estudiado. Como puede verse en la figura 5-1 anterior, existe una relación bidireccional entre ambos. Así, por lo que respecta a la relación entre el Ecosistema de Big Data y los procesos de gestión del conocimiento en Cementos Colombia S. A. nuestro análisis ha puesto de manifiesto lo siguiente:

- ✓ En primer lugar, la creación de conocimiento ha sido una de las políticas más importantes de la organización en los últimos años. Por lo tanto, el impulso a la creación del conocimiento y todo el proceso de gestión del conocimiento es constante. Así, con la implementación del ecosistema de Big Data la empresa ha buscado proveer de las capacidades y habilidades en Analytics a su personal con el fin de obtener nuevo conocimiento basado en datos, generando de esta manera la creación de conocimiento predictivo, y así con la creación de este conocimiento se impulsa, simultáneamente, el uso del ecosistema de Big Data en toda la organización.
- ✓ En segundo lugar, el almacenamiento del conocimiento ha venido evolucionando en la organización con el paso del tiempo, existiendo políticas claras para la documentación formal (conocimiento explícito) y la consolidación del directorio de expertos en diferentes áreas relevantes para la compañía (conocimiento tácito), todo ello actualmente apoyado por su sistema de información. El ecosistema de Big data aporta nuevas herramientas para capturar y almacenar el conocimiento a través de las interacciones de las personas con el ecosistema y con las diferentes áreas, teniendo disponibilidad inmediata a través de la nube para cualquier área que necesite de cierto conocimiento específico, obteniendo de esta manera un AS/RS del conocimiento (ver sección 3.1.2) en un nivel básico lo cual, a su vez, promueve que los empleados utilicen el ecosistema y con ello su implementación.

- ✓ En tercer lugar, la transferencia del conocimiento es, junto a la creación, uno de los ejes de la política de conocimiento de la compañía, debido a que la organización considera un activo importante el conocimiento de sus empleados y, de ahí, el interés de mantener el conocimiento dentro de la organización. Así, la divulgación de conocimiento mediante el directorio de expertos de la organización, la documentación y procedimientos formales y la comunicación, es impulsada mediante sistemas y tecnologías, lo que ha llevado a la organización a ver en el ecosistema de Big Data una forma de hacer llegar el conocimiento necesario a la persona que lo requiera en el momento que lo necesita, obteniendo una transferencia inteligente del conocimiento que, a su vez, promueve la implementación del ecosistema de Big Data.
- ✓ En cuarto lugar, la organización busca que su conocimiento esté disponible y que sea aplicado en las áreas y puestos de trabajo de todas las organizaciones que hacen parte de su grupo empresarial, así como de toda su cadena de suministro. Por tal motivo, los sistemas de la organización buscan apoyar la aplicación inmediata del conocimiento al determinar qué conocimiento debe aplicarse en los diferentes puestos y áreas de trabajo para la creación de nuevos productos, servicios y nuevos modelos de negocio, logrando una accesibilidad automática acorde a las necesidades de los consumidores del conocimiento. Por ello, consideramos que la organización está consiguiendo una aplicación automática del conocimiento que promueve la implementación del ecosistema de Big Data al proveer las tecnologías y sistemas necesarios para este fin.

En cuanto a los mecanismos de coordinación y control, el apartado 5.1.3 anterior contiene nuestro análisis de los mismos para el caso estudiado. Pero de forma similar a lo que sucede respecto a los procesos de gestión del conocimiento, la relación entre el Ecosistema de Big Data de Cementos Colombia S. A. y los mecanismos de coordinación y control de la compañía es bidireccional. En este sentido, nuestro análisis ha revelado que:

- ✓ En primer lugar, el Ecosistema Básico de Big Data ha promovido que las actividades que impliquen decisiones operativas se realicen mediante el análisis de datos, actividades que han sido gestionadas con la ayuda del Centro de Analítica, el cual apoya en la coordinación de actividades que impliquen el uso de datos,

sistemas y tecnologías asociadas a Big Data y Analytics, considerándose así que el Centro de Analítica de la organización hace las funciones de un Centro de Excelencia en Big Data. Del mismo modo, el Centro de Excelencia en Big Data promueve el uso del propio ecosistema al apoyar su uso en las diferentes áreas de la organización, lo que genera una relación de doble dirección entre ambos componentes.

- ✓ En segundo lugar, la organización ha visto la necesidad de coordinar mejor las actividades dado su tamaño y los múltiples procesos que posee, viendo la necesidad de establecer una coordinación lateral mediante células de trabajo. En este sentido, el ecosistema de Big data, en conjunto con el Centro de Excelencia, ha impulsado la coordinación lateral en la compañía mediante la mejora de las comunicaciones y la conformación de las células de trabajo acorde a la relación de los conocimientos requeridos y los procesos de negocio involucrados, consiguiendo de esta manera los principios de una coordinación lateral inteligente que, a su vez, impulsará la implementación del ecosistema al volverse indispensable en la conformación de las células de trabajo y el desarrollo de las actividades dentro del mismo.
- ✓ En tercer lugar, la organización define con claridad los indicadores de desempeño que se deben medir de acuerdo a la actividad o proyecto que se vaya a ejecutar, estableciéndose para tal fin los KPIs que mejor indiquen la ejecución de las actividades de sus empleados, lo cual implica una forma particular de normalización de resultados tal y como vimos en el apartado 5.1.3. Con el ecosistema de Big Data se busca que la medición del desempeño sea de forma automática al concentrar la ejecución de las actividades en sus sistemas de información y que, de esta manera, los KPIs tengan una visibilidad global en toda la organización para el seguimiento de su ejecución. Así, la organización está migrando a una medición automática del desempeño, lo que impulsará a su vez el uso e implementación del ecosistema de Big Data en toda la organización.
- ✓ Así mismo, y basándose en la normalización de resultados como forma de coordinación y control, Cementos Colombia S. A. ha venido utilizado las metodologías ágiles (i.e. SCRUM) para llevar a cabo la supervisión directa y realizar el control de las actividades de sus empleados, realizando una integración con plataformas especializadas en metodologías ágiles para tal fin. Con el ecosistema de Big Data se busca fortalecer el control de las personas a través de las herramientas tecnológicas provistas por el mismo. El propio ecosistema facilita la

vigilancia por parte de directivos y mandos medios del desarrollo de las actividades de los empleados obteniendo un control sinérgico entre personas y sistemas, control que impulsa el uso e implementación del ecosistema de Big Data al ser una herramienta utilizada por cargos de dirección y mandos medios.

Teniendo en cuenta los anteriores puntos, consideramos que el caso de Cementos Colombia S.A. ilustra una relación positiva y recíproca existente entre la implementación del ecosistema de Big Data con los procesos de gestión del conocimiento y los mecanismos de coordinación y control. Adicional a esto, logramos evidenciar en este caso que existe una relación positiva y recíproca entre los procesos de gestión del conocimiento y los mecanismos de coordinación y control, dado que el proceso de gestión del conocimiento es la base para obtener el directorio de expertos de la organización y los procedimientos para el mejor desarrollo de las actividades diarias a ser coordinadas y controladas, y, en el otro sentido, los mecanismos de coordinación y control permiten establecer las relaciones de la organización necesarias para llevar a cabo todo el proceso de gestión del conocimiento, que con el ecosistema de Big Data tiene que ser apoyado por el Centro de Excelencia en Big Data.

A continuación, presentamos las relaciones o patrones adicionales que han surgido inductivamente en el análisis de este caso particular.

Nuevos factores y patrones que han emergido del análisis del caso

Dentro de los factores y patrones específicos que han emergido para el caso de la organización Cementos Colombia S.A. encontramos:

- ✓ En primer lugar, que el *compromiso de la dirección* ha sido clave para la iniciativa del ecosistema de Big Data y para la creación del Centro de Excelencia de Big Data (Centro de Analítica). Con su compromiso, el equipo directivo ha conseguido impulsar el uso del ecosistema y de las herramientas de análisis de datos en las distintas áreas de la organización. Este compromiso de la dirección se verá, a su vez, espoleado por el propio ecosistema de Big Data, ya que su utilización permitirá a los directivos generar resultados y obtener beneficios

derivados del análisis de los datos provisto por el uso de las herramientas y tecnologías disponibles.

- ✓ En segundo lugar, la organización ha impulsado desde hace muchos años las iniciativas en innovación, siendo una de sus políticas más importantes, promovida principalmente mediante el compromiso que tiene la dirección con las iniciativas de innovación de sus empleados y miembros de toda la cadena de suministro. Como notamos en la sección 5.1.2, la innovación y los procesos de gestión del conocimiento van de la mano (de ahí que en la figura 5-1 coloquemos una flecha bidireccional entre ambos componentes), dado que a través de su política de innovación se promueve la creación, el almacenamiento, la transferencia y la aplicación del conocimiento en toda la organización, conocimiento que es la base para el posterior desarrollo de nuevos productos, servicios y procesos de negocio. Por tanto, el conocimiento y su gestión están, a su vez, en la base de toda la innovación. En línea con lo anterior, la implementación del ecosistema de Big Data y la innovación también presentan una relación de doble dirección en el caso estudiado. Por un lado, la política de innovación afecta al ecosistema de Big Data en el sentido de que todos los desarrollos y la adaptación de la organización al ecosistema corresponden a innovaciones apoyadas por la organización. Pero, al mismo tiempo, el ecosistema de Big Data promueve la innovación en la organización al impulsar, mediante la inclusión de sistemas, tecnologías y el análisis de datos, el desarrollo de nuevos procesos, procedimientos, productos y servicios a la organización.
- ✓ En tercer lugar, el Centro de Excelencia en Big Data será el principal gestor de la implementación del ecosistema de Big Data en las distintas áreas de la organización, dado que será el área responsable de adquirir los conocimientos y habilidades necesarios para aprovechar los beneficios que conlleva el análisis de datos y extender estas ventajas a lo largo de toda la organización, desarrollando de esta manera nuevas habilidades organizacionales. Adicional a esto, el Centro de Excelencia tiene la misión de propagar la necesidad de la organización de usar los datos para la toma de decisiones asertivas y la reducción de la incertidumbre, siendo así el área encargada de promover y desarrollar la cultura analítica en la organización.
- ✓ En cuarto lugar, la organización, al tener un sistema de gestión de la calidad implantado, se rige en gran medida por procedimientos, reglas, directrices,

políticas y la estandarización de sus procesos de negocio. Como consecuencia, la coordinación de gran parte de su trabajo descansa en la documentación del sistema de calidad lo que implica una elevada formalización de los procesos de trabajo como forma de coordinación y control. Por lo tanto, el sistema de gestión de calidad juega un papel fundamental en Cementos Colombia S. A. como nexo de unión entre los procesos de gestión del conocimiento, los mecanismos de coordinación y control y la implementación del ecosistema de Big Data, ya que dicho sistema facilita los procesos básicos de gestión del conocimiento (normaliza su creación, su almacenamiento, su transferencia y su aplicación), impulsa la coordinación y el control del trabajo a través de los propios documentos del sistema y es accesible a toda la compañía por estar integrado en el ecosistema de Big Data.

- ✓ En quinto lugar, la organización ha visto necesaria la descentralización de ciertas decisiones, en especial las operacionales, dadas las necesidades incrementales de coordinación de las diferentes áreas. Como hemos visto, con la implementación del ecosistema de Big Data la organización ha desarrollado nuevas formas de ejecutar y coordinar el trabajo a través de sistemas y tecnologías que impulsan la creación de las células de trabajo, llegando así a una coordinación lateral inteligente. Por lo tanto, la implementación del ecosistema de Big Data y la coordinación lateral inteligente impulsan la descentralización de la toma de decisiones, si bien esta descentralización permitirá, a su vez, una mejor coordinación lateral con el apoyo del ecosistema de Big Data.
- ✓ Por último, el estudio de este caso ha evidenciado que la implementación del ecosistema de Big Data influye en los procesos de gestión del conocimiento y los mecanismos de coordinación y control, y estos últimos, a su vez, promueven la implementación del ecosistema de Big Data, tal y como hemos descrito anteriormente. Dicha relación recíproca entre estos tres factores facilita una mejor toma de decisiones en la compañía, decisiones que tendrán un impacto positivo a nivel operativo y estratégico y que se espera que conlleven a mejoras en el desempeño de la organización y a una mayor rentabilidad.

5.2 Caso B: Fundación Universidad del Norte

En el presente caso de estudio analizamos a la Fundación Universidad del Norte, organización dedicada a la educación superior a niveles de pregrado y posgrado y a la investigación aplicada en diversos campos del conocimiento. Para tal fin, presentamos el perfil de la organización, su proceso de gestión del conocimiento, los mecanismos de coordinación y control utilizados y el efecto de la implementación del ecosistema de Big Data en la organización en los dos últimos puntos antes mencionados y en otros procesos particulares del caso que se han visto afectados por la implementación del ecosistema.

5.2.1 Perfil de la Organización

La Fundación Universidad del Norte es una institución de educación superior privada y sin ánimo de lucro, cuyo proceso de gestación y fundación tuvo lugar en el año de 1966 en la ciudad de Barranquilla, Colombia. La Universidad se constituyó legalmente como centro de educación superior el 24 de enero de 1966 e inició labores el 11 de julio de 1966, empezando con 58 estudiantes y 10 profesores en los programas de pregrado de administración de empresas e ingenierías (Universidad del Norte, 2020).

En la actualidad, la Universidad se encuentra ubicada en el kilómetro 5 de la antigua carretera entre Barranquilla y el municipio de Puerto Colombia, con un campus de aproximadamente 29 hectáreas. Además, la Universidad cuenta (hasta principios del año 2020) con una población de 13.431 estudiantes de pregrado y 2.389 estudiantes de posgrado, repartidos entre 10 divisiones académicas profesionales, con una oferta de 27 programas de pregrado y 123 programas de posgrado (ver tabla 5-2), y con un total de 1.181 profesores (Universidad del Norte, 2019).

Tabla 5-2: Oferta académica de la Fundación Universidad del Norte

Divisiones Académicas	Modalidad			
	Pregrado	Posgrado		
		Especialización	Maestría	Doctorado
Escuela de Negocios	3	8	4	1
División de Derecho, Política y Relaciones Internacionales	3	17	8	1
División de Ciencias de la Salud	3	12	4	1
Especializaciones Clínicas-Médico Quirúrgicas			6	
División de Ingenierías	6	13	11	5
División de Humanidades y Ciencias Sociales	4	5	9	4
Escuela de Arquitectura, Urbanismo y Diseño	3		1	
Instituto de Estudios en Educación (IESE)	1	1	2	1
División de Ciencias Básicas	2		4	2
Instituto de Idiomas	1	2	1	
Música	1			
Total General	27	58	50	15

Fuente: elaboración propia a partir de (Universidad del Norte, 2019)

La Universidad del Norte se ha destacado en los ránquines nacionales como una de las 10 mejores instituciones de educación superior del país, y siendo la mejor de toda la costa caribe colombiana. Además, mediante una clasificación del ministerio de educación nacional de Colombia, la Universidad del Norte ha sido catalogado como una de las 7 universidades en el país con un enfoque doctoral, debido a su enfoque investigativo y su oferta de doctorados.

La Universidad del Norte tiene como objetivo ser un centro de formación de profesionales idóneos y calificados con un enfoque innovador y que sean participes activos en el apoyo al desarrollo de la sociedad, teniendo en cuenta los principales valores de la institución como la excelencia, el liderazgo, la ética y el sentido de pertenencia. Así mismo, la Universidad del Norte está comprometida con el desarrollo social, económico, político, ambiental y cultural de Colombia, con el fin de seguir siendo un referente a nivel nacional, de América Latina y el Caribe en la enseñanza, la creación de conocimiento, innovación, pedagogía, desarrollo científico y desempeño estratégico en sus labores académicas, de investigación y desarrollo empresarial (Universidad del Norte, 2020).

5.2.2 Procesos de Gestión del Conocimiento en Fundación Universidad del Norte.

Tal y como vimos en la sección 2.1.3, la gestión del conocimiento presenta cuatro fases o etapas fundamentales que son la creación del conocimiento, almacenamiento del conocimiento, transferencia del conocimiento y aplicación del conocimiento. A continuación, exponemos como se desarrolla el proceso de gestión del conocimiento en la Fundación Universidad del Norte.

La Fundación Universidad del Norte, al ser un instituto de educación superior, tiene muy claro que la gestión del conocimiento debe ser uno de los pilares de la organización, tanto en la parte académica como en la parte administrativa. Analizando a la Universidad desde la perspectiva administrativa (foco de esta investigación), la Universidad ha establecido en su vigente plan de desarrollo que la institución debe tener un modelo de gestión del conocimiento, esto con el fin de mantener en la organización el conocimiento tácito y explícito que poseen sus empleados, y que estén lo mejor documentados posible (conversión del conocimiento tácito en explícito), y evitar que cualquier puesto de trabajo pierda el conocimiento al momento de irse un trabajador de la organización, así como sucedió en la Universidad años atrás al haber un cambio generacional en varios puestos, tal como lo menciona la Jefa de Capacitación y Desarrollo:

“Hace unos años la Universidad tuvo un relevo generacional bastante grande y fue una preocupación porque se fueron muchos directores que estaban desde la fundación de la Universidad, y estas personas se llevaron mucho conocimiento que había en la organización, no había una herramienta donde dejaran plasmado ese conocimiento,

y de ahí surgió el crear un modelo donde se dejara el conocimiento para que no se pierda... en el plan de desarrollo está determinado que la Universidad debe tener un modelo de gestión del conocimiento como parte de su objetivo, que fuera flexible, adaptable y que fuera diseñado para ser implementado de forma modular”.

En línea con el comentario anterior, la Universidad ha venido desarrollando herramientas tecnológicas para el apoyo de los procesos de gestión del conocimiento para conseguir esa flexibilidad, adaptabilidad y modularidad del modelo planteado en la institución, que consiga que el conocimiento sea propiedad de la Universidad y no se pierda al momento de realizarse cambios de personal.

Por lo tanto, y dada la necesidad expresada por la Universidad, se está trabajando principalmente en estrategias de transferencia de conocimiento que faciliten la conversión del conocimiento tácito en conocimiento explícito (externalización), sobre todo cuando se requiera de la sucesión de un cargo administrativo. En consecuencia, el departamento de gestión humana ha tomado como uno de sus principales objetivos el mantener ese conocimiento dentro de la institución mediante documentación almacenada en las herramientas tecnológicas dispuestas para tal fin, y que dicha documentación haga parte del sistema de gestión de la calidad de la Universidad.

La transferencia del conocimiento en la organización se ha ido convirtiendo en una actividad de obligatorio cumplimiento, no solo al momento de realizarse la sucesión de un cargo, sino para utilizar el conocimiento de un área o departamento para poder crear nuevo conocimiento en otras partes de la organización. En la actualidad, la Universidad considera que ha podido recolectar el 70% del conocimiento de todos los puestos de trabajo y un 60% de las matrices que relacionan el conocimiento de diferentes áreas o departamentos, enfocándose principalmente en los macroprocesos de soporte administrativo, soporte administrativo a la academia, actividades de valor de la institución y extensión de la Universidad. Sin embargo, estas actividades relacionadas con el conocimiento en la organización aún no han sido extendidas a la parte académica y de investigación de la Universidad, pero sigue siendo parte de los objetivos de la institución en el corto y mediano plazo, tal como lo expresa la Jefa de Capacitación y Desarrollo:

“Hemos venido puliendo un proceso para que el conocimiento quede aquí, que sea de obligatorio cumplimiento, no lo aplicamos todavía a la academia, pero si a la parte administrativa, pero si queremos extenderlo a la academia”.

Adicional a lo anterior, la institución también tiene como una de sus prioridades la creación del conocimiento, no solo a través de la transferencia de conocimiento entre departamentos, como mencionamos anteriormente, sino además con actividades que promuevan la socialización, como planes de formación, comités de conocimiento y unidades de aprendizaje, desarrollados entre distintos departamentos y niveles jerárquicos. Tal y como reconoce la Jefa de Capacitación y Desarrollo:

“Hay un comité de conocimiento donde nos reunimos y analizamos toda la información y conocimiento que hemos levantado, y hay que hacer unas directrices que deben incluirse dentro de los compromisos de excelencia de la Universidad en temas de gestión del conocimiento, y ser parte de la política de la Universidad y de los objetivos estratégicos”.

Para la construcción de todo su modelo de gestión del conocimiento y de los comités antes mencionados, la Universidad parte de su estructura organizacional y de los conocimientos tácitos y explícitos que se poseen en esa estructura para determinar qué procesos, personas o áreas de la organización requieren de un cambio o actualización, analizando las brechas existentes en conocimiento y estableciendo las necesidades que surgen a raíz de las brechas. Este análisis se realiza en conjunto con los procedimientos establecidos en el sistema de gestión de la calidad y las personas encargadas de los procesos, buscando dos objetivos: primero, el evitar la duplicidad del conocimiento (conocimiento tácito o explícito ya existente en la institución), y segundo, establecer las estrategias para adquirir el conocimiento faltante. En palabras de la Jefa de Capacitación y Desarrollo:

“Tenemos que identificar los conocimientos, las hojas de vida de los trabajadores, los perfiles de los cargos, los resultados de los informes de auditoría, los resultados de evaluación de desempeño donde identificamos brechas, grupos focales, buzón de sugerencias, todo esto lo tenemos”.

A partir de lo anterior, la Universidad al ser una institución que se rige estrictamente en procedimientos consignados en su sistema de gestión de calidad, ha establecido que la principal forma en que el conocimiento debe ser almacenado en la institución es mediante procedimientos que contenga la matriz de conocimiento de cada puesto de trabajo de la parte administrativa (por el momento), en donde se relaciona el conocimiento que se requiere y que surge al momento de desempeñar el cargo y las relaciones internas y externas que conlleven a un intercambio de conocimiento e información. Este proceso se encuentra actualmente liderado por la responsable de calidad y la Jefa de Capacitación y Desarrollo, encargadas de coordinar la recolección y consolidación de las matrices del conocimiento y los libros de conocimiento de las diferentes áreas y puestos de trabajo, y realizar el mapeo del conocimiento de la organización que surge de los dos anteriores. Tal y como reconoce la Jefa de Capacitación y Desarrollo:

“Muchas cosas del libro de conocimiento que se necesitaban ya estaban en nuestros procedimientos. Invitamos a 300 personas a los talleres con unos consultores en conocimiento para poder llenar los conocimientos, habilidades, riesgos y competencias que no estuvieran documentadas, archivos importantes, relaciones importantes (no lógicas) que tiene el cargo y relaciones externas”.

Por lo tanto, esta estrategia ha permitido identificar plenamente las brechas en conocimiento de su parte administrativa, procedimientos y manuales que no se encontraban bien documentados, incompletos, desactualizados o inexistentes, estandarizar procesos similares en diferentes áreas y establecer los conocimientos claves que posee una persona y que no se encuentran documentados para llevar a cabo una actividad o proceso administrativo.

Por consiguiente, la Universidad considera que mediante su modelo y estrategias de gestión del conocimiento ha conseguido valor para todos sus procesos y tomar acciones que permitan mejorar sus procedimientos, conseguido mediante el aumento de la visibilidad del conocimiento en toda la organización, dando la posibilidad que el conocimiento se convierta en un activo disponible en todos los puestos de trabajo. Al tener una estrategia enfocada en mantener la externalización del conocimiento, el conocimiento explícito se convierte en clave para aplicar ese conocimiento en el desarrollo de competencias y habilidades en otras áreas y puestos de trabajo, y que estas

nuevas competencias y habilidades se conviertan en nuevo conocimiento para un área específica, y, por ende, para toda la organización.

5.2.3 Mecanismos de Coordinación y Control en Fundación Universidad del Norte

Tal y como vimos en el marco teórico, en la sección 2.2.2, queremos conocer cómo se produce la coordinación y el control en Fundación Universidad del Norte. Por lo que respecta a los mecanismos de coordinación estableceremos en primer lugar el nivel de formalización y de descentralización que posee la organización, para luego exponer los mecanismos de coordinación vistos en el marco teórico, que puede ser por un alto directivo, por procedimientos y planeación formales, lateral o una combinación de las anteriores. Por lo que respecta a los mecanismos de control, estableceremos si el utilizado por la organización corresponde a un control centralizado personal, burocrático, por resultados, mediante vigilancia electrónica, mediante gestión de recursos humanos, cultural o una combinación de los anteriores.

La Fundación Universidad del Norte es una organización que posee una estructura jerárquica, presidida principalmente por el Consejo Directivo, conformado por miembros “ad honorem” de empresas y personas fundadoras de la Universidad, y es a quienes se les realiza la rendición de cuentas de los resultados de la Universidad. En el siguiente orden se encuentra el Rector, que determina la visión de la institución y es el encargado de gerenciar la organización, y, por debajo de este cargo, se subdivide en la Vicerrectoría Académica, encargada de gestionar todo lo relacionado con la educación impartida en la Universidad; la Dirección de Investigación, Desarrollo e Innovación, encargada de la gestión investigativa; la Dirección de Extensión, encargada de las consultorías y cursos de educación continua ofrecidas por la Universidad; y por último, la Vicerrectoría Administrativa, encargada de la gestión administrativa de toda la organización.

Teniendo en cuenta lo anterior, la verticalidad de la organización ha implicado una centralización de las decisiones, y cuya jerarquía indica qué persona debe tomar las decisiones. Para el caso de las decisiones operativas, son tomadas principalmente por jefes, coordinadores o directores con el aval de los altos mandos, y en el caso de decisiones estratégicas son tomadas por decanos, vicerrectores o el rector, dependiendo

de la importancia de la misma. Lo anterior es soportado por el Jefe de Información y Estadística:

“A diferencia de otras universidades en el país, la Universidad del norte tiene un modelo centralizado en su gestión, esto quiere decir, que en otras universidades las facultades manejan sus presupuestos, sus ingresos, sus procesos de admisión, su contratación de docentes, y aquí desde sus inicios la Universidad ha optado por un modelo centralizado en donde todos sus procesos están en cabeza de ciertas personas o áreas que sirven a toda la institución”.

Sin embargo, la Universidad ve la centralización como algo positivo ya que le ha favorecido en la implementación de programas y estrategias que se han requerido, manteniendo un control sobre la inversión de los recursos y en el proceso mismo de implementación, coordinado principalmente entre directores y vicerrectores. La Universidad lleva funcionando de esta manera desde sus inicios y ha mantenido su forma de tomar las decisiones desde entonces. No obstante, esta centralización puede ser en ocasiones sectorizada, ya que pueden ser tomadas por solo una parte directiva, dependiendo si afecta a la parte administrativa, la académica, o a ambas. Así lo expresó la Jefa de Capacitación y Desarrollo:

“Esto ha llevado a que cada uno se dedique a lo suyo, la academia a la academia y lo administrativo a lo administrativo, por lo que creo que (la centralización) ha sido un acierto desde la fundación de la Universidad”.

Como mencionamos en la sección anterior, la institución se rige principalmente mediante procedimientos consignados en su sistema de gestión de la calidad y, además, de un conjunto de reglas y directrices establecidas en diferentes reglamentos y códigos de conducta que posee la Universidad para el correcto desempeño de las labores en las diferentes áreas. Esto implica que la Universidad posee un alto nivel de formalización, donde la difusión de estas directrices y reglamentos son de obligatorio cumplimiento.

Por ende, se puede considerar que la Universidad del Norte es una organización con una coordinación mediante procedimientos y planeación formales, en donde existe una combinación entre los tres tipos de mecanismos en los que se subdivide esta forma de

coordinación (ver sección 2.2.2.2). En primer lugar, los procedimientos, reglas y normas rigen el desarrollo de las actividades en los puestos de trabajo y áreas de la institución (estandarización); en segundo lugar, la Universidad posee planes de acción anuales y planes de desarrollo elaborados con una periodicidad quinquenal, elaborados desde todas las áreas estratégicas de la institución (planes y horarios); en tercer lugar, la coordinación de las actividades diarias surge principalmente mediante reuniones de comité conformado principalmente por directores, jefes y actores que hagan parte de la decisión o que esté involucrado en la actividad que se requiera coordinar, y así tener una propuesta consensuada del plan a seguir. Tal y como reconoce el Jefe de Información y Estadística:

“En temas de coordinación, la Universidad funciona mucho con comités para los diferentes temas y son entre diferentes áreas, y ahí se toman las decisiones, ya las decisiones más estratégicas se toman junto con el consejo directivo [...] Hay comités ya definidos y conformados, pero lo que busca la Universidad es que una decisión que se tome o algún cambio de un proceso que se realice, involucre a los actores o procesos que afecte o pueda afectar, en donde una persona encargada identifica los procesos y personas involucradas, y en base a eso, se forman esos grupos, para tener diferentes perspectivas”.

En línea con lo anterior y a raíz de ser una organización con una coordinación mediante procedimientos y planeación formales, los mecanismos de control utilizados por la Universidad van acorde al de una organización centralizada y con una alta formalización. En este sentido, la organización posee, en primer lugar, mecanismos de control burocrático en donde predominan métodos, procedimientos y reglas formales que son de estricto cumplimiento acorde a los planes establecidos en la institución, utilizando principalmente la supervisión directa por un superior que realiza el seguimiento al cumplimiento de las metas y objetivos planteados a nivel operacional. En segundo lugar, la Universidad utiliza mecanismos de control por resultados, donde se hace una trazabilidad del cumplimiento de las metas estratégicas consignadas en los planes de acción y desarrollo, a todos los niveles de la organización, y realizando un análisis y reporte de este cumplimiento de forma anual. En este sentido el Jefe de Información y Estadística reconoce que:

“Respecto al control de resultados hay algo mixto, a niveles más bajos funciona bajo supervisión directa por parte del jefe directo, y ya a niveles estratégicos tenemos un plan de acción que involucra a toda la Universidad, en donde hay unos retos estratégicos y acciones muy específicas y ahí está establecido las metas, cómo se va a medir y los responsables, y en los procesos de autoevaluación que tiene la Universidad que se realizan semestralmente, y al final de año que se hace una rendición de cuentas, se evalúa cómo se está alcanzando esos resultados, acciones o metas a lo largo del año”.

Para realizar el seguimiento de estas metas y objetivos la organización cuenta con un sistema de información denominado “Alena”, en donde se consolida y controla las acciones estratégicas y de gestión, analizando el porcentaje de avance de cumplimientos de tales metas y mostrando informes por área funcional. Este sistema de información ha permitido la sistematización de muchos indicadores de cumplimiento, poder consolidar la autoevaluación de los empleados, analizar el desempeño de los profesores, y proveer datos, información e indicadores solicitados por entes gubernamentales como el Ministerio de Educación Nacional de Colombia o el Departamento Administrativo Nacional de Estadística (conocido por su acrónimo “DANE”) sobre investigación, extensión, bienestar institucional, gestión humana, población estudiantil, docencia y resultados obtenidos, todos estos campos acorde a las metas planteadas.

Además de lo anterior, otro mecanismo de control que se lleva a cabo en la Universidad es el control cultural, debido a que en la organización se promueve el sentido de pertenencia por la institución y que los objetivos personales del empleado se alineen con los objetivos institucionales. En la Universidad se promueve un ambiente de colaboración entre todos los empleados y áreas funcionales para buscar el cumplimiento en conjunto de las metas y objetivos, y se motiva al empleado con mecanismos de promoción y recompensa para incentivar el crecimiento profesional, personal y familiar del mismo.

5.2.4 Ecosistema de Big Data en Fundación Universidad del Norte.

La Fundación Universidad del Norte ha venido trabajando en los últimos años la transición a un ecosistema básico de Big Data que le permita analizar en tiempo real o en un menor tiempo los indicadores de gestión de la organización, en vez de realizarse con

una periodicidad semestral, anual o quinquenal como mencionamos en la sección anterior. Una de las principales razones para que la Universidad haya decidido implementar este tipo de ecosistemas ha sido la falta de confianza del personal en el uso de las herramientas para el análisis de datos y evaluación de los indicadores de gestión, ya que estas herramientas no terminaron de adaptarse a las necesidades y a la forma de trabajar de la organización. En este sentido, el Jefe de Información y Estadística menciona lo siguiente:

“No es la primera vez que la Universidad intenta este tipo de iniciativas. Hubo un proyecto que se llamaba sistema de indicadores académicos “SIA”, demandó muchos recursos, pero al final nadie lo utilizaba y nadie confiaba en esa información, en esos datos”.

Esta transición a un sistema más avanzado está siendo gestionada principalmente por la dirección de gestión humana, la dirección de tecnología informática y de comunicaciones, la oficina de planeación, y otras áreas funcionales interesadas en participar en el proyecto, como la oficina de admisión y la oficina de registro, entre otros. Para analizar cómo se ha ido realizando la implementación de ecosistemas de Big Data en la Fundación Universidad del Norte, a continuación, abordaremos cada una de las cinco dimensiones planteadas en la sección 1.3.3 para determinar el nivel de implementación en la que se encuentra la organización.

Dimensión 1: Implicación Organizativa de Big Data

En la Fundación Universidad del Norte se ha venido creando una necesidad por aprovechar los datos, la información y, en especial, el conocimiento, para determinar su rumbo estratégico, medir indicadores y gestionar la organización, notándose en iniciativas como la estrategia de conocimiento, mencionada en la sección 5.2.2. Sin embargo, la Universidad considera que no es suficiente el esfuerzo realizado hasta el momento para aprovechar todo el valor que pueden brindar los datos que se generan diariamente en la organización, datos que son vitales para la toma de decisiones. En palabras del Jefe de Información y Estadística:

“Actualmente se generan datos; pero no lo estamos aprovechando, porque sí generamos muchísimos datos, pero no lo aprovechamos como debemos ...”.

En consecuencia, la dirección de la institución, al evidenciar la necesidad interna de sacar valor a sus datos y percibir en eventos internacionales y en reuniones con otras universidades que se encuentran desarrollando sus estrategias alrededor del análisis avanzados de datos e implementación de ecosistemas de Big Data, ha decidido tomar la iniciativa de desarrollar e implementar sistemas, herramientas, métodos y tecnologías que le permitan obtener el valor de la información y conocimiento resultante de tal análisis de datos, como se menciona a continuación. En este sentido, el Jefe de Información y Estadística nos comenta lo siguiente:

“Para llevar a cabo todo esto creo que la Universidad debe estar consciente de la problemática de hoy en día, aunque por parte de la alta dirección ya lo están, y de ahí vino el apoyo a esta iniciativa, ver que hay diferentes verdades, diferentes formas de procesar o de obtener un dato, entonces lo que se quiere es unificar y esclarecer eso, y los recursos están dados, pero creo que lo más importante es que las personas de arriba y los que toman las decisiones como directores, vicerrectores y decanos puedan verle la utilidad a esto”.

Para lograr lo anterior, la Universidad ha establecido prioridades para su transición a estos sistemas, destinando recursos y personal para la adaptación de los procesos y procedimientos establecidos, cambios en la gestión de los datos, la información y el conocimiento, analizar las necesidades clave en esta etapa y mantener la unicidad y la colaboración de los esfuerzos, encaminados en sacar adelante la implementación del ecosistema de Big Data en la Universidad. Sin embargo, este proceso ha sido meticuloso y por ende ha demandado mucho tiempo, ya que no se quiere repetir las dificultades que se han presentado en la implementación de otras herramientas (como se mencionó en la sección anterior), sino construir un ecosistema que sirva para toda la institución, que las personas lo utilicen y que crezca progresivamente con la Universidad.

Además de lo anterior, una de las prioridades transversales al cambio de procesos y procedimientos ha sido el trabajo con la calidad de los datos, un problema que ha sido recurrente en la creación de reportes y medición de indicadores, afectando la construcción de los mismos y a caer en reprocesos que afectan la calidad de la toma de decisiones. Esta prioridad ha surgido debido a que las decisiones cada vez se están requiriendo con una

mayor velocidad, y se desea que la veracidad de la información sea eficaz para responder preguntas institucionales que antes no se han podido contestar correctamente o en los tiempos que se han requerido.

Por lo tanto, para llevar a cabo correctamente la implementación de su ecosistema de Big Data, la Universidad está construyendo pruebas piloto en temas prioritarios y sensibles de la institución, como, por ejemplo, analizar los datos de deserción estudiantil, en donde actualmente se tienen preguntas sin responder, y se espera que en los datos encuentren las respuestas para avanzar en este tema y poder tomar acción y decisiones al respecto. La clave de estas pruebas piloto está en mostrar la utilidad del ecosistema y del análisis de los datos, teniendo como objetivo mostrar mejores resultados que los esperados (llamado por el Jefe de Información y Estadística como “victorias tempranas”), y nuevas estrategias e ideas que surjan del análisis, demostrando a los altos mandos (directores, decanos, vicerrectores y rector) que se ha tomado la decisión correcta con la inversión realizada, para así tener el apoyo completo para llevar a cabo el resto de la implementación de una manera más fácil y fluida.

Por consiguiente, debido a que la Fundación Universidad del Norte ya se encuentra estableciendo los recursos y capacidades que se requieren para implementar el ecosistema de Big Data y seleccionando las prioridades y necesidades de la organización para llevar a buen término el proceso de transformación de sus procesos y procedimientos, consideramos que la organización, según lo planteado en la sección 1.3.3 de esta tesis doctoral, se encuentra en el nivel cuatro en la dimensión de implicación organizativa de Big Data.

Dimensión 2: Factores Humanos y su relación Máquinas-Sistemas

La Fundación Universidad del Norte es una institución con múltiples áreas funcionales en donde se realiza análisis de datos para la generación de reportes y medición de indicadores de gestión. Sin embargo, existe un área específica encargada de ser la fuente oficial de estadísticas y de información (luego de procesarse los datos), denominada como la jefatura de información y estadística, contenida en la oficina de planeación, y es el conducto regular para llevar a cabo los procesos de análisis de datos y estadísticos que se requieran para reportes o para la toma de decisiones en toda la organización. Esto es afirmado por el Jefe de Información y Estadística:

“Como departamento de planeación y desde el área de información y estadística tenemos una visión global de la universidad, tanto de la parte administrativa como de la parte académica [...] apoyamos en la toma de decisiones, no las tomamos, pero sí las sugerimos, brindamos el informe o los análisis para que las personas como los directivos, decanos o vicerrectores se les facilite tomar la decisión”.

Por lo tanto, la Universidad ya cuenta con un departamento especializado en el análisis de datos de la organización, además de poseer un departamento de tecnologías y comunicaciones que apoyan en la parte de la infraestructura tecnológica para la implementación del ecosistema de Big Data. A pesar de esto, la Universidad aún se encuentra planteando qué capacidades, habilidades y personal se requieren para llevar a cabo todo el proceso de transición al ecosistema de Big Data, ya que esto implicará cambios que la Universidad aún no se encuentra preparada para asumir.

Dentro de las primeras soluciones que se están planteando está el realizar un plan de capacitación en todas las áreas para que se descentralice todos los análisis del área de información y estadística, y se pase a una autogestión de las áreas en temas de análisis de datos, utilizando un esquema mixto entre el uso de herramientas y modelos que se desarrollen desde el área de información y estadística y el desarrollo propio desde cada una de las áreas. Esto permitirá que la gestión de los datos, la información y el conocimiento se mantenga dentro del área que requiera el análisis o el conocimiento que surja de este análisis, promoviendo una mayor rapidez en la creación, transferencia y aplicación de ese conocimiento. Esto es reafirmado por el Jefe de Información y Estadística:

“Le he dicho a mi equipo que cuando todo esto se termine de implementar ya dejaremos de realizar muchos de los informes que se hacen actualmente, no estaremos tanto tiempo en los informes, sino en crear nuevos análisis y también asesorar a las diferentes áreas a que realicen nuevos análisis con los datos o la información disponible, ser orientador de esos análisis a las nuevas preguntas que se formulen, y formular nuevos modelos”.

Consecuentemente, la Universidad ha empezado con los planes de capacitación en temas interdisciplinarios acorde al área o las personas que serán las encargadas de estos análisis en cada área, temas como: tecnología, econometría, estadística, programación y conocimientos específicos de cada área, en especial en mandos medios y bajos. Además, se está presentando un cambio en los altos mandos para el uso de las tecnologías del ecosistema de Big Data para la toma de decisiones, con el fin analizar las diferentes perspectivas de un problema, pregunta o estrategia que se requiera para el futuro de la Universidad, y se busca que desde estos mandos se promueva una cultura analítica en toda la organización, considerado por los encargados de este proceso de transición como un punto fundamental para esparcir esta cultura analítica por toda la compañía.

Así, tenemos que la Fundación Universidad del Norte ya se encuentra en una etapa de selección de personal y/o capacitación en métodos, técnicas y herramientas de BDA y ciencia de datos, y ha establecido un plan de capacitación para poder llevar a cabo toda la transición a la implementación del ecosistema de Big Data, e iniciar el cambio organizacional y cultural que se requiere para ser adoptado, por lo que consideramos, según lo planteado en la sección 1.3.3 de esta tesis doctoral, que la Universidad se encuentra en el nivel cuatro en la dimensión de Factores Humanos y su Relación Máquinas-Sistemas.

Dimensión 3: Método Utilizado de Analytics

La Fundación Universidad del Norte es una institución que ha utilizado la información histórica recolectada de sus procesos para luego realizar análisis de esos datos, realizar informes, medir los indicadores de gestión y tomar decisiones, utilizando principalmente para ello el método de analítica descriptiva. El volumen de datos manejado por la Universidad puede considerarse de volumen medio, ya que aún solo se manejan, por el momento, datos internos de la organización. Pero, por la cantidad de empleados y de estudiantes que se manejan, la cantidad de datos requiere de una infraestructura tecnológica y herramientas de análisis capaz de almacenar y procesar tales datos, infraestructura y herramientas que se poseen en la actualidad.

La Universidad tiene dentro de su plan estratégico, adicional a los antes mencionados, la selección de su estrategia de Business Intelligence y Analytics para la toma de decisiones institucionales, lo que podría llevarlo a utilizar además el método de analítica predictiva,

como se menciona a continuación para el caso de la prueba piloto para el análisis de la deserción estudiantil. En este sentido, el Jefe de Información y Estadística menciona lo siguiente:

“Ahora estamos en un proceso piloto para determinar con qué herramienta de Business Intelligence nos vamos a ir, y estamos en un piloto de analítica más avanzada que es construir un modelo predictivo de deserción estudiantil, y es una de nuestras metas estratégicas como institución, en cabeza de la dirección de TI con apoyo de planeación y otras áreas involucradas que tienen que ver con los proyectos piloto, pero muchas áreas están interesadas en el proyecto”.

Por lo tanto, como la Universidad se encuentra trabajando con analítica descriptiva desde hace un tiempo con un volumen medio de datos y se encuentra actualmente trabajando en las pruebas piloto utilizando el método de analítica predictiva consideramos que, según lo planteado en la sección 1.3.3 de esta tesis doctoral, la organización se encuentra en el nivel cuatro en la dimensión de Método Utilizado de Analytics.

Dimensión 4: Clase de Herramienta de BDA

La Fundación Universidad del Norte procesa mayoritariamente datos que se encuentran almacenados en sus bases de datos para sus decisiones, análisis, reportes e indicadores de gestión y desempeño, y hasta el momento no ha requerido de procesar datos en tiempo real para llevar a cabo tales labores. Por lo tanto, la Universidad cuenta con herramientas de procesamiento por lotes, algunas de desarrollo propio y otras que han sido adaptadas de proveedores externos.

Con la introducción del ecosistema de Big Data a la organización y una vez se culminen las pruebas piloto, se espera que se analicen algunos datos con herramientas de procesamiento en tiempo real, sobre todo para el seguimiento de indicadores prioritarios y tomar decisiones a problemas que requieran de una pronta respuesta. Sin embargo, no se espera que el uso de este tipo de herramientas sea necesario en el corto plazo.

Por consiguiente, la Universidad al utilizar primordialmente para sus análisis herramientas de procesamiento por lotes consideramos que, según lo planteado en la

sección 1.3.3 de esta tesis doctoral, la organización se encuentra en el nivel cuatro en la dimensión de clase de herramienta de BDA.

Dimensión 5: Integración de Tecnologías de Información

La Fundación Universidad del Norte, a través de la dirección de tecnología informática y de comunicaciones, ha venido liderando la integración de las tecnologías requeridas para la implementación del ecosistema de Big Data, evaluando los diferentes softwares e infraestructura requerida para llevar a cabo los análisis que se van a requerir en la organización. En primera instancia la Universidad ha empezado con un proceso de migración de sus servidores (bases de datos) a la nube (Cloud Computing, ver sección 1.2.4.1), con el fin de mejorar su capacidad y velocidad de procesamiento de datos, tanto de la parte administrativa, académica y de los estudiantes, realizando un estudio costo-beneficio de las herramientas y productos TIC disponibles en el mercado o evaluando la posibilidad de desarrollar ciertas herramientas dentro de la Universidad.

Una de las preocupaciones de la institución y, por ende, uno de los temas principales en los que se está trabajando, es en la estrategia de seguridad de los datos, la información y el conocimiento que posee la organización, ya que, como vimos en el primer capítulo, uno de los 10 principales atributos del Big Data es la vulnerabilidad. En la Universidad toma especial relevancia debido a que los datos manejados dentro de la institución, tanto a nivel de la organización en sí como los datos manejados de los empleados y estudiantes que pertenecen a la institución, pueden llegar a ser datos e información sensible dada la naturaleza personal de los mismos.

Por lo tanto, con la introducción del Cloud Computing en la institución y con el proceso actual de selección de herramientas y tecnologías TIC por parte de la dirección de tecnología informática y de comunicaciones, consideramos que, según lo planteado en la sección 1.3.3 de esta tesis doctoral, la Universidad se encuentra en el nivel cuatro en la dimensión de Integración de Tecnologías de Información.

Teniendo en cuenta lo anterior, al analizar las cinco dimensiones de los niveles de implementación de ecosistemas de Big Data nos damos cuenta de que la Fundación Universidad del Norte ha conseguido las habilidades, las tecnologías, el personal y un desarrollo de sus procesos para alcanzar el nivel cuatro en cada una de las dimensiones

antes expuestas. Por lo tanto, consideramos que la organización posee el nivel cuatro en la implementación de ecosistemas de Big Data, correspondiente al nivel de *Transición a Ecosistema de Big Data* (ver figura 1-9 en el capítulo 1 de la tesis).

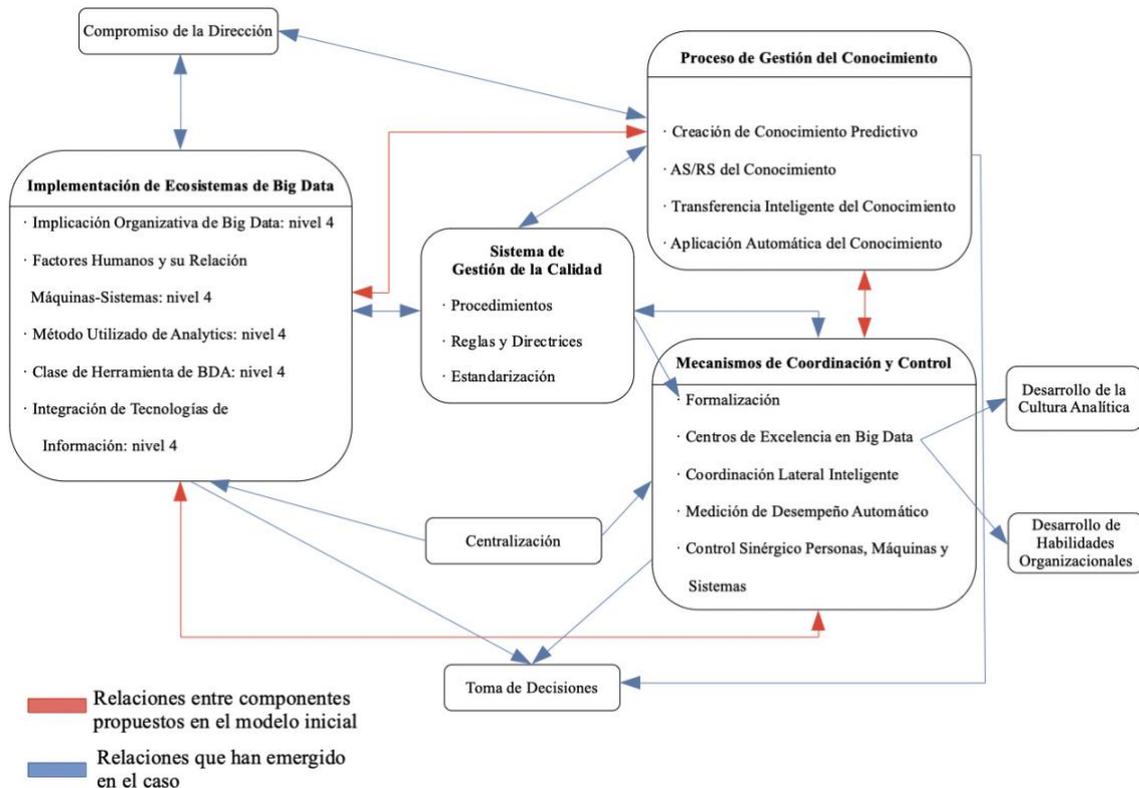
Consideramos que el hecho que la Fundación Universidad del Norte pertenezca al nivel cuatro en todas las dimensiones corresponde a que la organización ha venido manteniendo un esfuerzo sincronizado entre las áreas principales interesadas en el proceso de implementación del ecosistema de Big Data para dar respuesta a las preguntas que requiere la institución y poder tomar mejores decisiones. Para lograr lo anterior, la Universidad ha utilizado su experiencia en la implementación de otros sistemas y tecnologías para determinar las habilidades, capacidades, conocimientos y herramientas que se requieren o que ya se tienen en la organización, con el fin de dar un punto de partida a la transición a un ecosistema de Big Data.

No obstante, consideramos que la organización, a pesar de tener encaminado su transición a un ecosistema de Big Data, ha presentado demoras en esta transición debido a que no quieren un nuevo fracaso en la implementación de un sistema o tecnología como ha ocurrido anteriormente, lo que implica que la organización podría tomarse más tiempo de lo esperado en esta etapa de transición.

5.2.5 Efecto de la implementación del ecosistema de Big Data en los procesos de gestión del conocimiento y en los mecanismos de coordinación y control en Fundación Universidad del Norte.

Una vez hemos explicado y analizado cómo se desarrolla el proceso de gestión del conocimiento, hemos descrito y analizado los mecanismos de coordinación y control, y hemos determinado el nivel de implementación del ecosistema de Big Data en la Fundación Universidad del Norte, en la presente sección pasamos a estudiar el efecto que ha tenido esta implementación del ecosistema de Big Data en el proceso de gestión del conocimiento y en los mecanismos de coordinación y control de la empresa. A su vez, estos últimos han condicionado también la implementación del ecosistema de Big Data en la empresa lo que nos lleva a plantearnos la existencia de un efecto bidireccional entre ambos. La siguiente figura nos muestra el resultado final de nuestro análisis del caso de la empresa Fundación Universidad del Norte (ver figura 5-2).

Figura 5-2: Modelo final de relaciones para el caso de Fundación Universidad del Norte



Fuente: elaboración propia

Tal y como puede apreciarse en la figura anterior, el modelo final obtenido tras el análisis de la información del caso de la empresa Fundación Universidad del Norte nos presenta cómo se relaciona la implantación del Ecosistema de Big Data con los procesos de gestión del conocimiento y los mecanismos de coordinación y control en dicha organización. Al compararlo con nuestro modelo inicial (ver figura 3-5 en el capítulo 3 de esta tesis), vemos que tras el estudio del caso han aparecido nuevos elementos, factores organizativos y patrones que nos ayudan a entender mejor cómo tiene lugar dicha relación (están unidos mediante flechas de color azul en la figura 5-2), al igual que también existen otros constructos y elementos que ya habían sido contemplados en el modelo inicial (unidos mediante flechas de color rojo en la figura 5-2). Para facilitar al lector el seguimiento de nuestro análisis, primero presentaremos aquellos componentes del modelo que ya estaban presentes en el modelo inicial y, posteriormente, haremos lo propio con aquellos otros que han emergido durante el análisis del caso de Fundación Universidad del Norte.

Componentes propuestos en el modelo inicial

Tal y como se ha descrito en el apartado 5.2.4 anterior, la Fundación Universidad del Norte posee un nivel de correspondiente a una transición a ecosistema de Big Data con cinco dimensiones que alcanzan un nivel cuatro en todas ellas. Del mismo modo, existen cuatro grandes procesos de gestión del conocimiento que se dan en la organización y que fueron analizados y descritos anteriormente en el apartado 5.2.2 de este capítulo. Sin embargo, y tal y como anticipábamos unas líneas más arriba, ambos constructos no coexisten de forma aislada en el caso estudiado. Como puede verse en la figura 5-2 anterior, existe una relación bidireccional entre ambos. Así, por lo que respecta a la relación entre el Ecosistema de Big Data y los procesos de gestión del conocimiento en Fundación Universidad del Norte nuestro análisis ha puesto de manifiesto lo siguiente:

- ✓ En primer lugar, consideramos en este caso que la creación del conocimiento ha venido siendo uno de los pilares por el cual se ha iniciado el proceso de implementación (transición) del ecosistema de Big Data debido a la necesidad de la organización de crear conocimiento a través de la colaboración entre cargos, departamentos y distintos niveles jerárquicos (socialización), encontrando relaciones que permitan solucionar problemas institucionales. Con el ecosistema se espera que se cree conocimiento predictivo que permita una relación ágil en la organización para la toma de decisiones, lo que a su vez tendrá un impacto en la implementación del ecosistema al volverse indispensable en esta labor.
- ✓ En segundo lugar, y en línea con lo anterior, otro de los pilares es la transferencia del conocimiento, por lo que se ha buscado que el ecosistema apoye la transferencia entre los diferentes actores de la organización para evitar la pérdida del conocimiento y, además, que se transfiera el conocimiento importante de un área a otra para el desarrollo de nuevo conocimiento en otras áreas, buscando así que el ecosistema se convierta en la principal herramienta para este objetivo institucional. Por lo cual, consideramos que la transferencia del conocimiento y la implementación del ecosistema de Big Data poseen una relación recíproca, resultando en una transferencia inteligente del conocimiento.
- ✓ En tercer lugar, el almacenamiento del conocimiento va a estar respaldado por el ecosistema de Big Data al utilizarse almacenamiento en la nube para que el conocimiento pueda ser consultado desde cualquier área, teniendo una

disponibilidad inmediata en toda la organización. Por lo tanto, la universidad pasará a tener un AS/RS del conocimiento (ver sección 3.1.2) y, además, aportando como un impacto recíproco el facilitar la implementación del ecosistema de Big Data al usarse este tipo de tecnologías para el conocimiento.

- ✓ En cuarto lugar, el ecosistema de Big Data busca utilizar el conocimiento de la organización para apoyar la toma de decisiones, que, al estar disponible en toda la organización mediante el ecosistema de Big Data, permitirá a las personas y las distintas áreas de la organización aplicar el conocimiento y, en ocasiones, se podrán tomar decisiones de forma automática al estar las decisiones preestablecidas en el conocimiento recolectado por la organización (empezando por decisiones administrativas sencillas). Por lo tanto, consideramos que la organización está iniciando el proceso de tener una aplicación automática del conocimiento y que, de forma recíproca, esa aplicación del conocimiento facilitará la implementación del ecosistema de Big Data.

En cuanto a los mecanismos de coordinación y control, el apartado 5.2.3 anterior contiene nuestro análisis de los mismos para el caso estudiado. Pero de forma similar a lo que sucede respecto a los procesos de gestión del conocimiento, la relación entre el Ecosistema de Big Data de Fundación Universidad del Norte y los mecanismos de coordinación y control de la compañía es bidireccional. En este sentido, nuestro análisis ha revelado que:

- ✓ En primer lugar, la Universidad ha venido trabajando en conjunto con el departamento de planeación, en especial con el área de información y estadística, y la dirección de tecnología informática y comunicaciones para adquirir los conocimientos, habilidades y tecnologías asociadas al Big Data necesarios para la implementación del ecosistema, lo que ha llevado a sentar las bases para la creación del centro de excelencia en Big Data entre estas dos áreas. Por lo tanto, consideramos que la implementación de los ecosistemas de Big Data conlleva a la creación de los Centros de Excelencia, parte fundamental para la coordinación de la organización bajo el ecosistema de Big Data y, a su vez, estos centros de excelencia son los encargados de apoyar la implementación del ecosistema de Big Data.

- ✓ En segundo lugar, la coordinación de las actividades en la Universidad es realizada principalmente por mandos altos y medios, conformando comités para solucionar problemas, tomar decisiones y establecer prioridades. Los comités requieren de apoyo tecnológico para mejorar la comunicación y para establecer a quienes afectan las decisiones y los problemas. Por lo tanto, la Universidad ha visto la necesidad de involucrar al ecosistema de Big Data para determinar las conexiones necesarias entre los distintos departamentos y coordinar todo lo necesario para ejecutar las actividades y que, a su vez, los involucrados puedan establecer una mejor comunicación para la coordinación, implicando que se cree una coordinación lateral inteligente, permitiendo que la implementación del ecosistema de Big Data se mejore al ser utilizado por las distintas áreas de la organización.
- ✓ En tercer lugar, la Universidad ha venido trabajando principalmente con mecanismos de control burocrático y por resultados, utilizando indicadores que le permita medir el avance de la ejecución y cumplimiento de las metas y objetivos de su plan de desarrollo, utilizando para ello el apoyo de un sistema de información. Uno de los objetivos de la organización con la implementación del ecosistema de Big Data es reforzar ese seguimiento de las actividades mediante la medición del desempeño de forma automática, lo cual implica que la medición del desempeño promueve la implementación del ecosistema de Big Data y que este, a su vez, permite a la organización evaluar en tiempo real la ejecución de las actividades en todas las áreas.
- ✓ Por último, y como mencionamos en la sección 5.2.3, la Universidad maneja un control cultural de sus empleados al promover el sentido de pertenencia. Adicional a lo anterior, la organización promueve la cultura analítica de todos sus empleados para que exista una relación directa y recíproca entre las personas y el ecosistema de Big Data, lo que permitirá que los empleados puedan tanto vigilar el desarrollo de sus actividades como mejorar la comunicación entre sus empleados, lo que implica que la implementación del ecosistema de Big data impulsa al control sinérgico de personas y sistemas y viceversa.

En consecuencia, consideramos que el caso de la Fundación Universidad del Norte es un ejemplo claro de la relación positiva y recíproca existente entre la implementación del ecosistema de Big Data con los procesos de gestión del conocimiento y los mecanismos

de coordinación y control. Además, pudimos evidenciar en este caso que existe una relación positiva y recíproca entre los procesos de gestión del conocimiento y los mecanismos de coordinación y control, dado que los procesos de gestión del conocimiento proveen las relaciones que deben existir en la organización para llevar a cabo la coordinación y control de las actividades, obtenido mediante las matrices de conocimiento; y, en el otro sentido, los mecanismos de coordinación y control influyen en los procesos de gestión del conocimiento al promover las relaciones y comunicaciones necesarias para crear, almacenar, transferir y aplicar el conocimiento que existe y se crea en esas comunicaciones. Por lo tanto, el caso de la Fundación Universidad del Norte provee evidencias que apoyan las relaciones o conexiones planteadas en el modelo inicial presentado.

A continuación, presentamos las relaciones o patrones adicionales que han surgido inductivamente en el análisis de este caso particular.

Nuevos factores y patrones que han emergido del análisis del caso

Dentro de los factores y patrones que han emergido para el caso de la organización Fundación Universidad del Norte encontramos:

- ✓ En primer lugar, el *compromiso de la dirección*, el cual ha sido una de las claves para iniciar el proceso de implementación del ecosistema de Big Data y reevaluar el proceso de gestión del conocimiento. Esto se debe a los costos que conlleva el cambio tecnológico y organizativo que trae el ecosistema de Big Data, por lo que la dirección necesita ver la utilidad de estos procesos de cambio mediante pruebas y ensayos que den resultados tempranos. Como consecuencia de estos resultados, la dirección impulsará la implementación del ecosistema y promoverá en la organización la gestión del conocimiento y, a su vez, esto lleva a que la dirección se involucre más en dichos procesos de cambio organizativo.
- ✓ En segundo lugar, los centros de excelencia en Big Data se convertirán en una parte importante de la organización dado que será el encargado, además de la dirección, de impulsar la *cultura analítica* y el uso de las herramientas dispuestas para mejorar la comunicación, las relaciones, el conocimiento, la coordinación y control de la organización, así como mantener una mejora continua del ecosistema

de Big Data para responder a las necesidades de la organización. Por lo tanto, consideramos que los centros de excelencia en Big Data serán el eje para el desarrollo de nuevas habilidades organizativas.

- ✓ En tercer lugar, en este caso evidenciamos que la organización, al regirse por procedimientos formales, reglas, directrices y una estandarización de sus procesos, gestiona y formaliza los cambios en su proceso de gestión del conocimiento, los mecanismos de coordinación y control de la organización y, en la actualidad, la implementación del ecosistema de Big Data a través de su sistema de gestión de la calidad. En consecuencia, el sistema de gestión de la calidad de la organización se ha convertido en una de las piezas clave para la implementación del ecosistema de Big Data al gestionar las relaciones que han surgido entre los procesos de gestión del conocimiento, los mecanismos de coordinación y control y el propio ecosistema, relaciones que a su vez suministran nuevos procedimientos, reglas, y directrices al sistema de gestión de la calidad, promoviendo una formalización en la organización.
- ✓ En cuarto lugar, la Universidad posee una marcada centralización desde su fundación, y lo ha llevado a realizar sus actividades, procesos, implementaciones y el cumplimiento de objetivos exitosamente. La centralización ha sido vista en la organización como algo que ha ayudado en la implementación de distintos sistemas y tecnologías, así como la promoción de una mejor coordinación y control de las actividades. Por lo tanto, consideramos que la centralización promueve la implementación del ecosistema de Big Data y facilita los mecanismos de coordinación y control.
- ✓ Por último, consideramos que en este caso se ha puesto de manifiesto que la implementación del ecosistema de Big Data influye positivamente en los procesos de gestión del conocimiento y los mecanismos de coordinación y control, y estos últimos, a su vez, promueven la implementación del ecosistema de Big Data. Dicha relación recíproca entre estos tres factores tiene como fin la toma de mejores decisiones en la organización, y que dichas decisiones tengan un impacto positivo a nivel operativo y estratégico, lo que se espera que conduzca a mejoras en el desempeño de la organización y una mayor rentabilidad.

5.3 Caso C: Eco Fuels Energy S.A.S.

En el presente caso de estudio analizamos a la empresa Eco Fuels Energy S.A.S., dedicada a la producción y comercialización de Biodiesel y derivados a partir de material orgánico. De manera similar a los dos casos anteriores, estructuramos este apartado presentando el perfil de la organización, su proceso de gestión del conocimiento, los mecanismos de coordinación y control utilizados y el efecto de la implementación del ecosistema de Big Data en la organización en los dos últimos puntos antes mencionados y en otros procesos particulares del caso que se han visto afectados por la implementación del ecosistema.

5.3.1 Perfil de la Organización

La empresa Eco Fuels Energy S.A.S. fue fundada el 19 de marzo de 2011 en la ciudad de Barranquilla, Colombia, con una planta de producción ubicada en el municipio de Malambo, departamento del Atlántico. La empresa cuenta con un total de 12 empleados comprendidos entre el presidente, gerente financiero (sin personas a su cargo), gerente de planta (producción), nueve operarios de producción y dos miembros de la junta directiva (director de innovación y tecnología y director administrativo), siendo también parte de esta junta el dueño de la empresa, que actúa a su vez como presidente de la empresa.

Eco Fuels Energy S.A.S. tiene como actividad económica, en primer lugar, la producción de biocombustible producido a partir de la biomasa, es decir, materia orgánica originada en un proceso biológico, utilizable como fuente de energía; en segundo lugar, la producción de biodiesel, derivado de lípidos naturales como aceites vegetales o grasas animales con o sin uso previo, mediante procesos industriales de esterificación y transesterificación; en tercer lugar, la producción de petroquímicos, como el Diésel fuel oil D6, Diésel D2, ULSD (Diésel), Etanol E10, Jet fuel A1/B, combustible colonial grade 54, coque de petróleo, gas natural licuado (GNL) y gas licuado de petróleo (GLP); y por último, la producción de oleoquímicos, como el aceite de crudo de palma, aceite de palma RBD, aceite crudo de palmiste, aceite de palmiste RBD, estearina de palma RBD y la olearina de palma RBD.

A pesar de su completa oferta de productos, el centro de su operación se encuentra en los biocombustibles y el biodiesel, teniendo como objetivo de producción 1.440.000 litros de combustible al mes (2.000 litros por hora), conseguido a través de una planta de procesamiento por medio de reactores de cavitación magnética, innovadora en Colombia, que posee un alto nivel de automatización de su proceso y capaz de procesar insumos sin importar la calidad de los mismos o de insumos más difíciles de procesar como la grasa animal. Con respecto a lo anterior, el Director de Innovación y Tecnología hace el siguiente apunte sobre la calidad de los insumos:

“Como dependas de la calidad de los “palmeros” (cultivadores de palma y productores de aceite de palma) [...] te pueden llevar a la quiebra”.

Por lo tanto, la empresa diseñó e implementó la planta con una inversión de 9.000 millones de pesos (2.324.300 euros aproximadamente), de tamaño mediano y alta rentabilidad para poder suministrar al mercado nacional e internacional la creciente demanda de biocombustibles y biodiesel, mercado cuya demanda de estos dos productos supera a la oferta en Colombia. Los objetivos de Eco Fuels Energy S.A.S. se centran en ser una empresa líder en la utilización e innovación de tecnologías de producción que le permitan proteger el medio ambiente, buscando siempre mejorar la rentabilidad y la calidad de sus productos. Lo anterior es soportado por lo mencionado por el Director de Innovación y Tecnología:

“Por ahora el biodiesel tiene demanda infinita, en teoría, porque la cuota de Biodiesel en Colombia es del 19% y la oferta solo llega al 12%; hay un 7% no cubierta por la oferta”.

5.3.2 Proceso de Gestión del Conocimiento en Eco Fuels Energy S.A.S.

Tal y como vimos en la sección 2.1.3, la gestión del conocimiento presenta cuatro fases o etapas fundamentales que son la creación, almacenamiento, transferencia y aplicación del conocimiento. A continuación, exponemos cómo se desarrolla el proceso de gestión del conocimiento en la empresa analizada.

La gestión del conocimiento en Eco Fuels Energy S.A.S. ha sido uno de sus ejes desde su fundación, debido a que uno de los principales objetivos de la organización ha sido crear un proceso de producción propio y único en Colombia que le permitiera procesar distintos tipos de insumos en distintos grados de calidad. Por lo tanto, al fundarse la empresa se centró en crear el conocimiento requerido para iniciar operaciones bajo su propio modelo de negocio y teniendo en cuenta las condiciones climatológicas de la ciudad, proceso que duró aproximadamente cuatro años. Tal y como reconoce el Director Administrativo:

“La empresa la fundamos en 2011 y empezamos a vender en 2015. Durante unos 4 años solo hicimos investigación y desarrollo porque no podíamos pagar la tecnología de otros competidores, porque cuando adquieres esa tecnología adquieres también el conocimiento que hay detrás de esa tecnología, que son tecnologías ya muy probadas... A nosotros nos tocó esa investigación desde cero porque nuestros reactores son diferentes y no se sabía cómo esos reactores trabajan a las temperaturas que se manejan en la ciudad”.

La investigación y desarrollo que ha realizado la organización mediante diseños experimentales para cada tipo de producto ha dado como resultado unas 250 “recetas” únicas que han sido documentadas para llevar a cabo su operación, recetas que contemplan unas 20 variables como, por ejemplo, la calidad del insumo, la temperatura, calidad final deseada, etc. Así, estas recetas se mantienen en una base de datos para tener almacenado el conocimiento explícito fundamental que posee la organización, donde se almacena todo el conocimiento tácito que pueda surgir de sus empleados y de la operación misma, y, adicional a esto, documentación sobre las pruebas y experimentos que se realizan en el proceso, documentación que se encuentra en constante actualización. En este sentido, el Director de Innovación y Tecnología reconoce lo siguiente:

“El ingeniero químico encargado, que tiene una maestría en procesos de refinación química, evalúa los cambios en las variables del proceso y con su experiencia toma una decisión sobre qué hacer en el proceso. El grupo de operación participa en generación nuevo conocimiento ... por ejemplo, nosotros tenemos un sistema que hace intercambio iónico para filtración de Biodiesel, y en lugar de resinas iónicas, desarrollamos nuestra propia resina que hace lo mismo a un cuarto del costo, estamos

viendo la viabilidad para patentarlo y venderlo, eso es más que todo innovación, y ahí es donde participa el grupo de operaciones”.

Por lo tanto, en Eco Fuels Energy S.A.S. el conocimiento se ha centrado en la parte operativa, en donde se considera que es el eje de su organización y donde se encuentran concentrados la mayoría de los empleados, por lo que principalmente la transferencia del conocimiento se realiza mediante estrategias de externalización para convertir el conocimiento que surge de la operación en el conocimiento explícito de sus recetas y, además, se realiza una socialización de ese conocimiento dentro de la parte operativa teniendo como principal actor al ingeniero encargado de la operación para esta socialización. Como consecuencia de esto, el conocimiento generado ha permitido que la organización impulse la innovación al desarrollar procedimientos y procesos nuevos en su mercado, diferenciándose así de sus competidores.

Adicional a lo anterior, la organización incluye en su gestión del conocimiento a sus otras áreas funcionales mediante un modelo de optimización que incluye, además de las antes mencionadas, variables del negocio que no pertenecen directamente a la operación, en donde se analizan los datos y documentación almacenadas para poder tomar decisiones a nivel operativo y estratégico. Así, esta ha sido la principal aplicación del conocimiento que ha sido creado, almacenado y transferido dentro de la organización desde su fundación, y ha permitido que las decisiones que son apoyadas mediante el modelo sean mejores que aquellas que eran tomadas anteriormente por medio de la intuición del gerente de la empresa. Tal y como lo reconoce el Director Administrativo:

“El conocimiento tácito y explícito en la empresa se maneja con la ayuda del modelo de optimización. Nosotros no tomamos una decisión que no esté modelada, por eso es un modelo que permite tomar decisiones, inclusive decisiones como temas de apalancamiento financiero para poder financiar capital de trabajo... por ejemplo, el modelo nos arrojó que generamos más capital pagando semestral”.

En conclusión, la empresa Eco Fuels Energy S.A.S. tiene una gestión del conocimiento sencilla pero sólida que se centra en los aspectos clave del negocio, y que de este conocimiento ha surgido el modelo de optimización que les permite tomar decisiones estratégicas y operativas de una forma más acertada y que ajuste a las condiciones de su

mercado, su demanda y a las necesidades de desarrollar mejores procesos para obtener un mejor desempeño.

5.3.3 Mecanismos de Coordinación y Control en Eco Fuels Energy S.A.S.

Tal y como vimos en el marco teórico (sección 2.2.2) queremos conocer cómo se producen la coordinación y el control en Eco Fuels Energy S.A.S. Por lo que respecta a los mecanismos de coordinación estableceremos, en primer lugar, el nivel de formalización y de descentralización que posee la organización, para luego exponer los mecanismos de coordinación vistos en el marco teórico, a saber, por un alto directivo, por procedimientos y planeación formales, lateral o una combinación de las anteriores. Por lo que respecta a los mecanismos de control, estableceremos si el utilizado por la organización corresponde a un control centralizado personal, burocrático, por resultados, mediante vigilancia electrónica, mediante gestión de recursos humanos, cultural o una combinación de las anteriores.

La empresa Eco Fuels Energy S.A.S., como mencionamos al inicio de este caso, cuenta con solo 12 empleados incluido el presidente de la organización (dueño de la misma) y dos miembros de la junta directiva. Por lo tanto, al ser de tan pocos empleados la organización posee una estructura jerárquica y centralizada en donde las decisiones estratégicas las toma directamente el presidente. Sin embargo, la participación de la junta directiva y de los demás gerentes es importante para sugerir opciones en temas como estrategias, nuevas ideas de innovación, financiamiento y planeación, apoyando así al presidente para tomar la decisión. En este sentido, el Director Administrativo comenta lo siguiente:

“Cuando hay una decisión de fondo, el presidente, que es el dueño de la compañía, es el que toma las decisiones, tiene el 90% de la compañía, nosotros como parte de la junta apoyamos, pero él toma la decisión, nosotros damos sugerencias”.

A pesar de lo anterior, la empresa Eco Fuels Energy S.A.S. no se consideraría una organización burocrática debido a que posee una baja formalización y una delegación de las decisiones operativas, lo cual se debe a que no existen procedimientos o planes que rigen el trabajo de los empleados. En cambio, el gerente de planta se encuentra a cargo

de su área específica coordinando las actividades de los operarios, utilizando para tal fin, principalmente, la supervisión directa.

En el caso que la decisión operativa sea de relevancia o se requiera solucionar problemas mayores en la producción, se realiza un comité con el presidente, el gerente de planta y los dos miembros de la junta directiva para establecer el impacto del mismo y la adaptación requerida en la organización con respecto a la decisión o ruta a seguir; no obstante, la decisión final la toma el presidente lo que la convierte en una organización con una coordinación por un alto directivo. Con respecto a lo anterior, el Jefe de Innovación y Tecnología nos comenta lo siguiente:

“En la parte de operaciones el que coordina todo el trabajo es el gerente de planta, y es el encargado de informar si fue exitosa una producción, si hubo algún problema o si el producto salió como no era, y él toma las decisiones internas de la operación... Cuando se requiere la resolución de algún problema la coordinación se hace por el presidente, el cual llama a junta con los socios y el gerente, en donde miramos tecnología, riesgos, costos y propuestas sobre alguna compra de maquinaria o tecnología, pero al final el que toma la decisión es él”.

En consecuencia, podemos establecer que la empresa Eco Fuels Energy S.A.S. maneja dos mecanismos de control básicos: en primer lugar, el control centralizado personal debido a que las decisiones se encuentran en gran parte centralizadas en el presidente, el cual realiza una supervisión directa de sus empleados y de todos los aspectos de la organización; en segundo lugar, la organización posee un control de resultados dado que, a pesar de la centralización, cada área posee sus responsabilidades bien definidas sin regirse mediante procedimientos y reglas formales y con unos objetivos y metas claros que deben cumplirse en cada área, y con una delegación semi-autónoma de las decisiones operacionales. En este sentido, el Director Administrativo nos comenta lo siguiente:

“Hacemos una reunión en donde se muestran los estados financieros, los precios que se han manejado, y demás resultados obtenidos en el periodo; luego él nos presenta un cuadro de mando donde muestra cumplimientos o no cumplimientos, y cuando hay no cumplimientos los planes para llegar a cumplir”.

En conclusión, la empresa Eco Fuels Energy S.A.S. puede considerarse una organización con una coordinación centralizada en el presidente y en el gerente de planta, en donde recae principalmente la coordinación de las actividades de la organización, ejerciendo un control mediante supervisión directa y el establecimiento de objetivos y metas claras en cada área, realizando una revisión periódica del cumplimiento de tales objetivos.

5.3.4 Ecosistema de Big Data en Eco Fuels Energy S.A.S.

En línea con lo visto en secciones anteriores, la empresa Eco Fuels Energy S.A.S. es una organización con un enfoque de producción, en donde su parte administrativa y de dirección se encarga de la toma de decisiones estratégicas pero direccionadas a mejorar la producción y sus productos. Por lo tanto, en la implementación de ecosistemas de Big Data se ha centrado en modelos que permitan a la organización tomar mejores decisiones operativas para satisfacer la calidad requerida por los clientes de sus productos.

Para analizar cómo se ha ido realizando la implementación de ecosistemas de Big Data en Eco Fuels Energy S.A.S., a continuación, abordaremos cada una de las cinco dimensiones planteadas en la sección 1.3.3 lo que nos permitirá determinar el nivel de implementación en la que se encuentra la organización.

Dimensión 1: Implicación Organizativa de Big Data

La empresa Eco Fuels Energy S.A.S. ha sentado sus bases desde la experimentación de su proceso productivo, lo que ha llevado a que los datos se conviertan en una parte fundamental para el desarrollo de sus operaciones y sea la causa por la que la organización ha centrado tomar decisiones operacionales basadas en datos. Por tal motivo, la empresa ha desarrollado un modelo de optimización que utiliza la data de múltiples variables de la organización para asegurarse que su operación genere la rentabilidad deseada.

En línea con lo mencionado en secciones anteriores, la experimentación genera datos que contribuyen con la creación de conocimiento tácito y explícito consignado en las recetas que usa la organización para llevar a cabo sus operaciones. Por lo tanto, el modelo desarrollado por la organización tiene como objetivo evaluar las variables consignadas en las recetas para maximizar la rentabilidad, determinando cómo se debe procesar los insumos para cumplir con la calidad pedida por el cliente y tomar decisiones sobre la

negociación del precio en base a la calidad. Tal y como lo reconoce el Director de Innovación y Tecnología:

“Yo creo que tendremos un uso mayor de la data cuando llegue el momento de tener una mayor cantidad de productos, en donde sí tendremos que llegar al mercado a competir, porque hoy solamente competimos en base a calidad, porque hoy si cumple con los parámetros lo vas a vender”.

Por consiguiente, la empresa Eco Fuels Energy S.A.S. ha establecido una importancia en recolectar y analizar los datos de su proceso productivo, permitiendo la creación de conocimiento y el apoyo a la toma de decisiones mediante el modelo de optimización desarrollado por la organización. Así, consideramos que la organización, según lo planteado en la sección 1.3.3 de esta tesis doctoral, se encuentra en el nivel tres en la dimensión de implicación organizativa de Big Data.

Dimensión 2: Factores Humanos y su Relación Máquinas-Sistemas

Como mencionamos anteriormente, la empresa Eco Fuels Energy S.A.S. tiene centralizada la toma de decisiones en el presidente, común en empresas de pocos empleados como es el caso. La toma de decisiones por parte del presidente se apoya principalmente por el modelo de optimización utilizado para el proceso, un sistema de información para temas administrativos, contables y de inventario, y el apoyo, para ciertas decisiones, del gerente de planta y la junta directiva.

En línea con lo anterior, desde la fundación de la empresa ha sido importante que el presidente (dueño de la empresa) comprendiera la importancia del análisis de datos para mejorar la rentabilidad de la organización, por lo que los miembros de la junta directiva propusieron el modelo de optimización que hemos venido mencionando. Así, todos los empleados están capacitados para recolectar los datos requeridos por el modelo y el uso de las tecnologías de información requeridas para ello, y el presidente se encuentra actualmente capacitado para tomar decisiones operativas en base a los resultados del modelo, y ciertas decisiones administrativas en base a su sistema de información administrativo y contable. Tal y como lo reconoce el Director de Innovación y Tecnología:

“Como tenemos 9 operarios es fácil entrenarlos en estos temas (uso de SI y TI), [...] en la empresa para usar datos para tomar decisiones primero tuvimos que convencer al presidente que el manejo de datos era importante”.

Así, tenemos que la empresa Eco Fuels Energy S.A.S. tiene a sus empleados capacitados para el uso de SI y TI asociados a la producción y el presidente plenamente capacitado para tomar decisiones en base al modelo de optimización y el SI administrativo y contable, por lo que consideramos, según lo planteado en la sección 1.3.3 de esta tesis doctoral, que la empresa Eco Fuels Energy S.A.S. se encuentra en el nivel tres en la dimensión de Factores Humanos y su Relación Máquinas-Sistemas.

Dimensión 3: Método Utilizado de Analytics

Como hemos venido mencionando, la empresa estudiada apoya buena parte de las decisiones de la organización mediante su modelo de optimización, el cual consiste en un modelo matemático con unos 20 parámetros de entrada y 29 parámetros de salida que indica, entre otros, en cuánto tiempo, qué presión, qué temperatura, el uso de ultravioleta para matar patógenos, etc., se debe procesar el insumo adquirido del proveedor para obtener cierta calidad en su producto final. Además, el modelo arroja otras decisiones como la programación de turnos, vacaciones y programación de la producción (Scheduling). En este sentido, el Jefe de Innovación y Tecnología opina de la siguiente manera:

“Se desarrolló un modelo matemático de planeación y con este tomamos las decisiones de procurement (aprovisionamiento), producción y de ventas, se programan los turnos, la programación de vacaciones, y programación del proceso. Es un modelo semanal porque se trabaja de lunes a viernes”.

Las principales fuentes de los datos provienen de las mediciones de calidad de los insumos y los datos que arroja el proceso de forma automática, para luego ser procesados y analizados mediante el modelo de optimización y un software de análisis estadístico que recibe los datos para ser almacenados en una base de datos, para posteriormente realizar análisis descriptivos de lo que sucede en el proceso y generar reportes que faciliten la toma de decisiones.

Por lo tanto, como la empresa Eco Fuels Energy S.A.S. se encuentra analizando datos con un bajo volumen (solo dos fuentes de datos) y utilizando para tal fin la analítica descriptiva para generar reportes y tomar decisiones, consideramos que, según lo planteado en la sección 1.3.3 de esta tesis doctoral, la organización se encuentra en el nivel tres en la dimensión de Método Utilizado de Analytics.

Dimensión 4: Clase de Herramienta de BDA

La empresa Eco Fuel Energy posee un proceso de producción automatizado en donde los operarios solo intervienen cuando se presentan factores exógenos, como fluctuaciones de energía eléctrica. De resto, los operarios solo intervienen para el monitoreo del proceso y probar la calidad de los insumos de entrada. Este proceso automatizado genera datos en tiempo real que sirven para dos fines: en primer lugar, el proceso se puede autorregular para ir cambiando variables del proceso en tiempo real para ajustarse y obtener la calidad deseada, guiándose de las recetas; y, en segundo lugar, enviar datos del proceso al modelo de optimización y al software de análisis estadístico y para la toma de decisiones. Tal y como lo menciona el Jefe de Innovación y Tecnología:

“La planta de producción toma decisiones autónomas dentro del propio proceso, lo que si no toma decisiones es si compramos insumos o vendemos el producto...”

Esto implica, que la empresa Eco Fuels Energy S.A.S. tiene un procesamiento en tiempo real con la autorregulación del proceso de producción y un procesamiento por lotes para la generación de reportes y toma de decisiones. Por consiguiente, consideramos que, según lo planteado en la sección 1.3.3 de esta tesis doctoral, la organización se encuentra en el nivel cinco en la dimensión de clase de herramienta de BDA, nivel obtenido por lo particular de su proceso de producción.

Dimensión 5: Integración de Tecnologías de Información

En base a la dimensión anterior, notamos que la empresa Eco Fuels Energy S.A.S. posee una alta integración de sus sistemas de información y tecnologías de captura de datos en tiempo real y una autorregulación del proceso, todo con el propósito de procesar datos para tomar decisiones en tiempo real (para la autorregulación) o por lotes para decisiones como, por ejemplo, establecer los momentos para comprar insumos y a quién se le debe comprar de acuerdo a la calidad deseada de salida.

Para tal fin, su proceso productivo ha integrado ciertas soluciones de Cloud Computing para almacenar los datos generados del proceso y una inteligencia artificial básica que permite al proceso de producción la autorregulación, todo esto integrado con los sistemas de información que maneja la empresa y el modelo de optimización. Sin embargo, aún son aplicaciones puntuales y no integran a las áreas administrativas de la organización. Basándonos en estas consideraciones entendemos que, según lo planteado en la sección 1.3.3 de esta tesis doctoral, la organización se encuentra en el nivel cuatro en la dimensión de clase de herramienta de BDA, nivel obtenido por lo particular de su proceso de producción.

Como observamos, la empresa Eco Fuels Energy S.A.S. difiere respecto a los otros dos casos vistos en este capítulo con respecto a la asignación de un nivel de implementación de ecosistemas de Big Data, en donde observamos que las organizaciones poseían el mismo nivel en todas las dimensiones, situación que no ocurrió de la misma manera para este caso.

En este caso notamos que las primeras tres dimensiones la empresa posee el nivel tres, explicado dado que las personas que hacen parte de la organización aún no se encuentran en el proceso de adaptar los ecosistemas de Big Data en sus actividades diarias (vistas a partir del nivel 4), pero sí se comprenden la importancia del mismo y es clave para la toma de decisiones. En las dos últimas dimensiones notamos que se consiguieron los niveles cinco y cuatro respectivamente, explicado por la automatización de su sistema productivo en donde se requiere integración de tecnologías asociadas a Big Data para el procesamiento por lotes y en tiempo real utilizado en la autorregulación del proceso y el análisis de datos y la toma de decisiones.

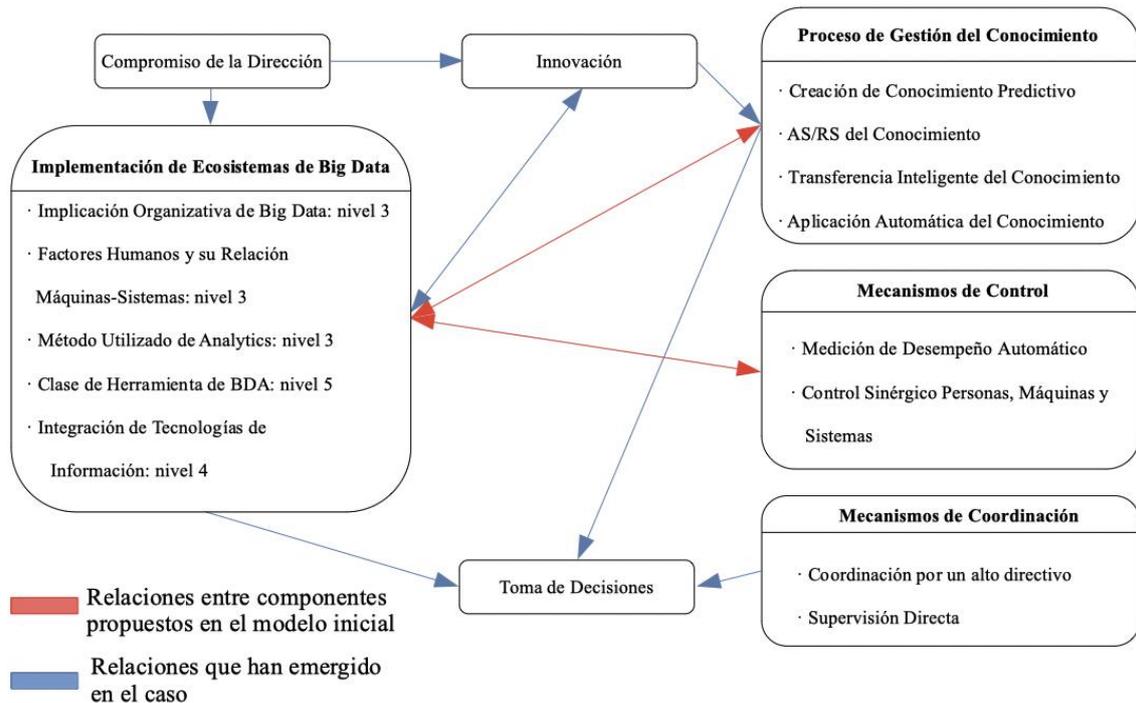
Así, a pesar de tener dos dimensiones en niveles cuatro y cinco, consideramos que la empresa Eco Fuels Energy S.A.S. pertenece al nivel tres en el nivel de implementación de ecosistemas de Big Data, correspondiente a un *Sistema para la Toma de Decisiones y Especializado de Información Empresarial*. La empresa debe trabajar en las dimensiones de implicación organizativa de Big Data, los Factores Humanos y su Relación Máquinas-Sistemas y el Método utilizado de Analytics para ponerse a la par con lo avanzado de su

proceso productivo, y así considerarse (por lo menos) en el nivel de Transición a ecosistemas de Big Data.

5.3.5 Efecto de la implementación del ecosistema de Big Data en los procesos de gestión del conocimiento y en los mecanismos de coordinación y control en Eco Fuels Energy S.A.S.

Una vez hemos explicado y analizado cómo se desarrolla el proceso de gestión del conocimiento, hemos descrito y analizado los mecanismos de coordinación y control, y hemos determinado el nivel de implementación del ecosistema de Big Data en la empresa Eco Fuels Energy S.A.S., en la presente sección pasamos a estudiar el efecto que ha tenido esta implementación del ecosistema de Big Data en el proceso de gestión del conocimiento y en los mecanismos de coordinación y control de la empresa. A su vez, estos últimos han condicionado también la implementación del ecosistema de Big Data en la empresa lo que nos lleva a plantearnos la existencia de un efecto bidireccional entre ambos. La siguiente figura nos muestra el resultado final de nuestro análisis del caso de la empresa Eco Fuels Energy S.A.S. (ver figura 5-3).

Figura 5-3: Modelo final de relaciones para el caso de Eco Fuels Energy S.A.S.



Fuente: elaboración propia

Tal y como puede apreciarse en la figura anterior, el modelo final obtenido tras el análisis de la información del caso de la empresa Eco Fuels Energy S.A.S. nos presenta cómo se relaciona la implantación del Ecosistema de Big Data con los procesos de gestión del conocimiento y los mecanismos de coordinación y control en dicha organización. Al compararlo con nuestro modelo inicial (ver figura 3-5 en el capítulo 3 de esta tesis), vemos que tras el estudio del caso han aparecido nuevos elementos, factores organizativos y patrones que nos ayudan a entender mejor cómo tiene lugar dicha relación (están unidos mediante flechas de color azul en la figura 5-3), al igual que también existen otros constructos y elementos que ya habían sido contemplados en el modelo inicial (unidos mediante flechas de color rojo en la figura 5-3). Para facilitar al lector el seguimiento de nuestro análisis, primero presentaremos aquellos componentes del modelo que ya estaban presentes en el modelo inicial y, posteriormente, haremos lo propio con aquellos otros que han emergido durante el análisis del caso de Eco Fuels Energy S.A.S.

Componentes propuestos en el modelo inicial

Tal y como se ha descrito en el apartado 5.3.4 anterior, la empresa Eco Fuels Energy S.A.S. posee un nivel correspondiente a un sistema para la toma de decisiones y especializado de información empresarial con las tres primeras dimensiones en el nivel tres, la cuarta dimensión en el nivel cinco y la quinta dimensión en el nivel cuatro. Del mismo modo, existen cuatro grandes procesos de gestión del conocimiento que se dan en la organización y que fueron analizados y descritos anteriormente en el apartado 5.3.2 de este capítulo. Sin embargo, y tal y como anticipábamos unas líneas más arriba, ambos constructos no coexisten de forma aislada en el caso estudiado. Como puede verse en la figura 5-3 anterior, existe una relación bidireccional entre algunos de estos componentes; pero, a su vez, otros no cumplen con la relación planteada en el modelo inicial. Así, por lo que respecta a la relación entre el Ecosistema de Big Data y los procesos de gestión del conocimiento en la empresa Eco Fuels Energy S.A.S., nuestro análisis ha puesto de manifiesto lo siguiente:

- ✓ En primer lugar, la creación de conocimiento se ve influenciada por el ecosistema del Big Data dada la autorregulación del proceso de producción en donde se toman decisiones automáticamente en base a los indicadores que se necesitan de una

receta específica. Por lo tanto, el proceso es capaz de medir indicadores necesarios para obtener la calidad deseada del producto final realizando cambios respecto al conocimiento existente en las recetas, resultando en datos que, después de analizados por el gerente de planta (experto en el proceso), se propone cómo mejorar las recetas, obteniendo de esta manera una creación de conocimiento predictivo que surge del análisis automático del proceso. A su vez, consideramos que la creación de conocimiento predictivo promueve la implementación y/o avance del ecosistema de Big Data.

- ✓ En segundo lugar, el almacenamiento del conocimiento se ve influenciado por el ecosistema de Big Data al tener un sistema que permite almacenar datos y decisiones que ha tomado de forma automática el proceso para la consulta en tiempo real de los encargados del proceso y por el presidente para hacer ajustes a las recetas, siendo almacenado de forma automática por el mismo proceso, convirtiéndolo así en AS/RS del conocimiento. A su vez, dados los mismos motivos anteriormente expuestos, consideramos que el AS/RS del conocimiento promueve la implementación y/o avance del ecosistema de Big Data.
- ✓ En tercer lugar, la transferencia del conocimiento en la empresa Eco Fuels Energy S.A.S. sucede de forma recíproca entre el proceso (máquinas) y las personas encargadas del mismo, mediante la comunicación de las recetas (conocimiento explícito) al proceso y la retroalimentación provista por el proceso, que permite mejorar las recetas en conjunto con el modelo de optimización, resultando en una transferencia inteligente de conocimiento.
- ✓ En cuarto lugar, y como mencionamos anteriormente, la toma de decisiones en la organización se lleva a cabo principalmente por el presidente. Sin embargo, el conocimiento plasmado en el modelo de optimización y en las recetas hace que ciertas decisiones operativas se tomen de forma automática. Por lo tanto, al ser estas decisiones automáticas basadas en conocimiento, implica que el ecosistema de Big Data ha llevado que exista una aplicación automática del conocimiento.

En cuanto a los mecanismos de coordinación y control, el apartado 5.3.3 anterior contiene nuestro análisis de los mismos para el caso estudiado. Pero a diferencia de lo que sucede respecto a los procesos de gestión del conocimiento, la relación entre el Ecosistema de Big Data de Eco Fuels Energy S.A.S. y los mecanismos de coordinación y control de la

compañía no es bidireccional en algunos componentes. En este sentido, nuestro análisis ha revelado que:

- ✓ En primer lugar, dada la naturaleza de la organización al ser de pocos empleados, no se notó una influencia significativa del ecosistema de Big Data sobre los mecanismos de coordinación, manteniéndose una coordinación por un alto directivo hasta el momento de escritura de esta tesis doctoral.
- ✓ En segundo lugar, al existir parámetros e indicadores preestablecidos para la producción y que se miden de forma automática, la dirección y el gerente de planta pueden medir en tiempo real los objetivos de calidad de los biocombustibles, y mediante el modelo de optimización establecer la rentabilidad del negocio que se está obteniendo con la operación, implicando que ahora en la organización se realice una medición automática del desempeño.
- ✓ En tercer lugar, con la introducción del ecosistema de Big Data en la organización se está realizando un control de resultados de forma recíproca entre las personas, máquinas y el modelo (sistemas), en donde en ocasiones no se necesita intervención humana para llevar a cabo el control, por lo que consideramos que se está llevando a cabo un control sinérgico entre personas, máquinas y sistemas.

Cabe aclarar que para este caso particular no evidenciamos una relación directa entre los procesos de gestión del conocimiento y los mecanismos de coordinación y control, por lo cual la conexión inicialmente planteada no ha sido representada. A continuación, presentamos las relaciones o patrones adicionales que han surgido inductivamente en el análisis de este caso particular.

Nuevos factores y patrones que han emergido del análisis del caso

Dentro de los factores y patrones que han emergido para el caso de la organización Eco Fuels Energy S.A.S. encontramos:

- ✓ En primer lugar, el compromiso de la dirección ha sido de mucha importancia para el uso y desarrollo de los modelos de optimización y la automatización del proceso, lo que ha llevado a que los datos tomen protagonismo en todas sus

actividades, de tal forma que influye en la implementación del ecosistema de Big Data.

- ✓ En segundo lugar, la innovación ha sido uno de los pilares del desarrollo de la organización desde sus inicios, por lo que ha sido importante el soporte y compromiso de la dirección para impulsar la innovación, en especial para el desarrollo de los sistemas y tecnologías asociadas con el ecosistema de Big Data, teniendo además este ecosistema un efecto en la innovación al impulsar el desarrollo constante de nuevos y mejores sistemas y tecnologías.
- ✓ En tercer lugar, consideramos que la innovación promueve el avance de los procesos de gestión del conocimiento en entornos del ecosistema de Big Data, dado que la innovación impulsará el desarrollo de prácticas ágiles para la creación, almacenamiento, transferencia y la aplicación del conocimiento en la organización.
- ✓ En cuarto lugar, consideramos que la implementación de ecosistemas de Big Data, el proceso de gestión del conocimiento impulsado por el ecosistema, y la coordinación por parte de los directivos, son factores que se juntan entre sí para la toma de mejores decisiones en la organización, utilizando el nuevo conocimiento y las tecnologías que hacen parte del ecosistema, y que esto, a su vez, se espera que derive en mejoras en el desempeño y la rentabilidad de la organización.

5.4 Análisis Conjunto de Casos

Una vez hemos expuesto y analizado individualmente los tres casos de interés para esta investigación, con el objetivo de establecer las relaciones entre la implementación de los ecosistemas de Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación y control, a continuación, procedemos a realizar un análisis conjunto de los tres casos y construir un modelo inducido a partir de los mismos. Para este fin, en esta sección continuaremos con la estructura utilizada en los casos individuales donde realizaremos la comparación de los procesos de gestión del conocimiento, los mecanismos de coordinación y control y, por último, presentaremos el modelo inducido final de esta investigación en donde se relacionarán las conexiones que se presentaron en el modelo inicial planteado y los nuevos factores y patrones que han emergido de los tres casos.

5.4.1 Comparativa de los procesos de gestión del conocimiento

En este apartado realizaremos una comparativa de los procesos de gestión del conocimiento de los casos que hicieron parte de esta investigación, en donde se contrastaran los cuatro procesos: creación, almacenamiento, transferencia y aplicación del conocimiento. A continuación, exponemos la comparación de estos procesos que se encuentran resumidas en la tabla 5-3 (página siguiente):

Tabla 5-3: Comparativa de los procesos de gestión del conocimiento de los casos estudiados

PROCESO DE GESTIÓN DEL CONOCIMIENTO	CASO ESTUDIADO		
	Cementos Colombia S.A.	Fundación Universidad del Norte	Eco Fuels Energy S.A.S.
Creación	- Ba Originador - Ba de Interacción - Ciber Ba - Ba de Acción	- Ba Originador - Ba de Interacción - Ciber Ba - Ba de Acción	- Ba Originador - Ba de Interacción
Almacenamiento	- Sistema de Información - Documentación - Directorio de Expertos	- Sistema de Información - Documentación - Directorio de Expertos	- Documentación
Transferencia	- Canales formales, personales e impersonales	- Canales formales, personales e impersonales	- Canales informales
Aplicación	- Reglas y directrices - Rutinas - Grupos de resolución de problemas y toma de decisiones	- Reglas y directrices - Rutinas - Grupos de resolución de problemas y toma de decisiones	- Reglas y directrices - Rutinas

Fuente: elaboración propia

En primer lugar, para comparar la creación del conocimiento en los tres casos estudiados utilizamos el concepto de “Ba” propuesto por los autores Nonaka y Konno (1998) (espacio utilizado para la creación del conocimiento, ver sección 2.1.3.1), el cual se encuentra relacionado con los modos de conversión del conocimiento (ver sección 2.1.1). Así, nos encontramos que en las tres organizaciones se promueve la socialización entre los empleados, clientes y proveedores (Ba Originador), y la creación de documentos y procedimientos a partir de la externalización del conocimiento (Ba de Interacción).

Adicional a lo anterior, las organizaciones Cementos Colombia S.A. y Fundación Universidad del Norte han desarrollado plataformas y sistemas de información para la colaboración de las personas y mantener el conocimiento en repositorios virtuales (Ciber Ba) y ambas organizaciones promueven la internalización del conocimiento a través de programas de aprendizaje individual (Ba de acción). Así, tenemos que las organizaciones Cementos Colombia S.A. y Fundación Universidad del Norte poseen una combinación de los cuatros tipos de “Ba”, mientras que Eco Fuels Energy S.A.S. solamente utiliza el Ba Originador y el Ba de interacción.

En segundo lugar, respecto al proceso de almacenamiento de conocimiento nos encontramos que las organizaciones Cementos Colombia S.A. y Fundación Universidad del Norte se apoyan en sistemas de información, documentación y un directorio de expertos que les permitan mantener el conocimiento creado dentro de la organización, mientras que la empresa Eco Fuels Energy S.A.S. se centra principalmente en la documentación para mantener las recetas de sus procesos, que es, según los entrevistados, el conocimiento más importante de la organización.

En tercer lugar, para comparar el proceso de transferencia del conocimiento utilizamos los tipos de canales de transferencia del conocimiento propuesto por los autores Alavi & Leidner (2001) (ver sección 2.1.3.3). Por lo tanto, notamos que para las organizaciones Cementos Colombia S.A. y Fundación Universidad del Norte se utilizan principalmente canales formales, como las jornadas de capacitación y una mezcla de canales personales (socialización) e impersonales (internalización) para la distribución del conocimiento en toda la organización. Para el caso de Eco Fuels Energy S.A.S., nos encontramos una organización que recurre a canales informales donde se promueve la socialización del conocimiento de la organización (recetas).

En cuarto lugar, para comparar el proceso de aplicación del conocimiento utilizamos los cuatro mecanismos de aplicación del conocimiento propuestos por Grant (1996) (ver sección 2.1.3.3). Así, nos encontramos que las empresas Cementos Colombia S.A. y la Fundación Universidad del Norte utilizan primordialmente las reglas y directrices, las rutinas y los grupos para la resolución de problemas y la toma de decisiones para aplicar el conocimiento de la organización, grupos donde la comunicación entre personas es vital,

mientras que la empresa Eco Fuels Energy S.A.S. utiliza principalmente las reglas y directrices y las rutinas como las formas para aplicar el conocimiento.

Teniendo en cuenta lo expuesto anteriormente podemos establecer que las organizaciones Cementos Colombia S.A. y la Fundación Universidad del Norte poseen formas muy similares de llevar a cabo los procesos de gestión del conocimiento. Por otro lado, la empresa Eco Fuels Energy es el caso que posee las formas más sencillas de llevar a cabo dichos procesos en comparación con los dos primeros casos, debido a que posee una estrategia enfocada en el conocimiento explícito y en la externalización del conocimiento, situación que consideran, por el momento, suficiente para el desarrollo de sus actividades. Sin embargo, es una estrategia que el Director de Innovación y Tecnología considera que debe cambiar en el corto plazo, y que deben integrar nuevas tecnologías para la gestión del conocimiento.

A pesar de poseer un nivel de madurez similar los primeros dos casos, consideramos que la organización Cementos Colombia S.A. es la que posee una mayor madurez en sus procesos de gestión del conocimiento y en los sistemas y tecnologías utilizadas para tal fin. Esto es debido a que posee procesos más robustos y que integran un mayor número de actores, como lo son los proveedores y clientes, mientras que la Fundación Universidad del Norte posee procesos de gestión del conocimiento relativamente nuevos, lo que implica que aún se encuentra en un constante ajuste de su estrategia de gestión del conocimiento y en el desarrollo constante de nuevas y mejores herramientas tecnológicas para tal fin.

5.4.2 Comparativa de los mecanismos de coordinación y control

En este apartado realizaremos una comparativa de los mecanismos de coordinación y control de los casos que hicieron parte de esta investigación, en donde se contrastaran los distintos mecanismos de coordinación y control vistos en la sección 2.2.2. A continuación, exponemos la comparación de estos mecanismos que se encuentran resumidas en la tabla 5-4:

Tabla 5-4: Comparativa de los mecanismos de coordinación y control de los casos estudiados

MECANISMO	CASO ESTUDIADO		
	Cementos Colombia S.A.	Fundación Universidad del Norte	Eco Fuels Energy S.A.S.
Coordinación	<ul style="list-style-type: none"> - Alta formalización - Descentralización selectiva - Procedimientos y planeación formales - Coordinación lateral - Metodologías ágiles 	<ul style="list-style-type: none"> - Alta formalización - Centralización - Procedimientos y planeación formales - Coordinación lateral 	<ul style="list-style-type: none"> - Baja formalización - Centralización - Coordinación por un alto directivo
Control	<ul style="list-style-type: none"> - Control de resultados - Metodologías ágiles 	<ul style="list-style-type: none"> - Control burocrático - Control de resultados - Control cultural 	<ul style="list-style-type: none"> - Control centralizado personal - Control de resultados

Fuente: elaboración propia

En cuanto, a los mecanismos de coordinación, en primer lugar, analizamos el nivel de formalización de los casos estudiados, en donde nos encontramos que las organizaciones Cementos Colombia S.A. y Fundación Universidad del Norte poseen una alta formalización debido a que se rigen por políticas y reglas de estricta difusión y con una estandarización marcada del trabajo, reforzado por un robusto sistema de gestión de calidad donde se consignan dichas reglas, procedimientos y directrices. Por otro lado, en este aspecto difiere la empresa Eco Fuels Energy S.A.S, debido a que su formalización es baja ya que no existen procedimientos, reglas o políticas redactadas y cada empleado sigue sus funciones acordes a las directrices de los mandos medios y altos.

En segundo lugar, analizamos la centralización o descentralización en los casos estudiados, en donde nos encontramos que la empresa Eco Fuels Energy S.A.S y la Fundación Universidad del Norte poseen un modelo centralizado, donde los altos directivos son los principales encargados de tomar las decisiones estratégicas con cierta

delegación de decisiones operativas en los mandos medios, pero dichas decisiones operativas están sujetas a aprobación por parte de las directivas. Por el contrario, la empresa Cementos Colombia S.A.S. ha optado por migrar a una descentralización selectiva de las decisiones, atribuyendo las decisiones estratégicas a los altos mandos y una autonomía de los mandos medios para tomar las decisiones operativas.

En tercer lugar, analizando los principales mecanismos que utilizan las organizaciones para coordinar las actividades nos encontramos que la empresa Eco Fuels Energy S.A.S. utiliza principalmente una coordinación por un alto directivo, explicado por el número de empleados de la organización y la estructura jerárquica de la misma, donde el presidente es el encargado de realizar la supervisión directa de los empleados. Por otro lado, la empresa Cementos Colombia S.A.S. y la Fundación Universidad del Norte utilizan procedimientos y planeación formales, esto derivado de su alta formalización, donde la estandarización, los planes y horarios, y las reuniones de comité son las principales formas de establecer prioridades y planes de acción en la organización. A su vez, estas dos organizaciones antes mencionadas combinan los procedimientos y planeación formales con una coordinación lateral, donde se utiliza personal especializado para crear y liderar grupos formales para la solución de problemas, integración de esfuerzos y la toma de decisiones.

En cuanto a los mecanismos de control, en primer lugar, hemos visto que las tres organizaciones utilizan una combinación de mecanismos de control. Sin embargo, las tres tienen en común un control de resultados, debido a que en los tres casos se establecen metas, objetivos y los indicadores de desempeño para los empleados y la organización en general. En segundo lugar, se ha evidenciado que la empresa Fundación Universidad del Norte, adicional al control de resultados, aplica un control burocrático dado el uso de métodos, procedimientos y reglas formales que son de estricto cumplimiento, y un control cultural dada la promoción del sentido de pertenencia por la organización. Por otro lado, la empresa Eco Fuels Energy S.A.S. posee un control centralizado personal, la cual se destaca por la supervisión directa y la toma de decisiones centralizada, en este caso ejercida por el presidente de la empresa, explicado por el tamaño de la organización.

Adicional a los puntos expuestos anteriormente, la empresa Cementos Colombia S.A. destaca de los otros casos en que utiliza las metodologías ágiles (SCRUM) como una

forma de llevar a cabo la coordinación y control de las actividades de sus empleados y de los grupos o células de trabajo que se forman para la solución de problemas y la toma de decisiones, destacándose por ser una contemporánea de normalización de resultados y por permitir evaluar en tiempo real como está el cumplimiento de los objetivos.

Acorde a lo visto anteriormente, las tres organizaciones poseen similitudes y diferencias en los mecanismos de coordinación y control lo que ha permitido estudiar el fenómeno de interés a través de una variedad de mecanismos distintos, lo cual permite a esta investigación mirar dicho fenómeno desde diferentes formas de gestionar una organización.

5.4.3 Comparativa de la implementación del ecosistema de Big Data

En este apartado realizaremos una comparativa de los niveles de implementación del ecosistema de Big Data de los casos que hicieron parte de esta investigación. A continuación, la tabla 5-5 (página siguiente) sintetiza los principales aspectos de dicha comparativa.

Tabla 5-5: Comparativa de la implementación del ecosistema de Big Data de los casos estudiados

DIMENSIONES Y NIVEL	CASO ESTUDIADO		
	Cementos Colombia S.A.	Fundación Universidad del Norte	Eco Fuels Energy S.A.S.
Implicación Organizativa de Big Data	5	4	3
Factores Humanos y su Relación Máquinas-Sistemas	5	4	3
Método utilizado de Analytics	5	4	3
Clase de Herramienta de BDA	5	4	5
Integración de Tecnologías de Información	5	4	4
Nivel Conseguido	5 Ecosistema Básico de Big Data	4 Transición a Ecosistema de Big Data	3 Sistema para la Toma de Decisiones y Especializado de Información Empresarial

Fuente: elaboración propia

Como pudimos observar en las secciones 5.1.4, 5.2.4 y 5.3.4, las tres organizaciones estudiadas poseen tres niveles distintos de Big Data, esto debido a que dichas empresas fueron seleccionadas cuidadosamente para poder estudiar el fenómeno desde distintas perspectivas y diferentes etapas de la evolución del ecosistema de Big Data.

Por lo tanto, en primer lugar, nos encontramos que en la dimensión *implicación organizativa de Big Data*, en la empresa Eco Fuels Energy S.A.S. las directivas comprenden la importancia del análisis de datos para la toma de decisiones y la creación de conocimiento (nivel 3). A su vez, la Fundación Universidad del Norte ha establecido las prioridades y necesidades de la organización y ha dispuesto los recursos para la adopción de los sistemas y tecnologías asociadas a Big Data (nivel 4). Por otro lado, la empresa Cementos Colombia S.A. ya posee capacidades y recursos básicos para la utilización de sistemas y tecnologías asociadas a Big Data y, a su vez, se comienza a expandir la cultura analítica en la organización (nivel 5).

En segundo lugar, en la dimensión *factores humanos y su relación máquinas - sistemas* nos encontramos que la empresa Eco Fuels Energy S.A.S. posee directivos capacitados para la toma de decisiones mediante el uso de sistemas de información y herramientas de optimización (nivel 3). A su vez, la Fundación Universidad del Norte se encuentra en proceso de capacitación de su personal en métodos, técnicas y herramientas asociadas a Big Data para poder incorporarlos a sus procesos de negocio (nivel 4). Por otro lado, la empresa Cementos Colombia S.A. ya posee un centro de excelencia (centro de analítica) especializado en el análisis y extracción de valor de los datos, y el conocimiento resultante de dichos análisis es incorporado en los procesos de la organización (nivel 5).

En tercer lugar, en la dimensión *método utilizado de Analytics* nos encontramos que la empresa Eco Fuels Energy S.A.S realiza un análisis descriptivo con un volumen bajo de datos (nivel 3), al igual que la Fundación Universidad del Norte que realiza un análisis descriptivo, pero con un volumen medio de datos (nivel 4), mientras que la empresa Cementos Colombia S.A. utiliza métodos de analítica descriptiva y predictiva con un gran volumen de datos de diversas fuentes (nivel 5).

En cuarto lugar, en la dimensión *clase de herramienta de BDA* nos encontramos que las empresas Eco Fuels Energy S.A.S. y Cementos Colombia S.A. se encuentran realizando procesamiento de datos por lotes y en ocasiones realizan procesamiento de datos en tiempo real debido al nivel de automatización de algunos de sus procesos productivos (nivel 5), mientras que la Fundación Universidad del Norte realiza únicamente procesamiento de datos por lotes (nivel 4).

En quinto lugar, en la dimensión *integración de tecnologías de información* nos encontramos que la empresa Eco Fuels Energy S.A.S. y la Fundación Universidad del Norte han integrado tecnologías de computación en la nube, así como algunas herramientas básicas de inteligencia artificial (nivel 4), mientras que la empresa Cementos Colombia S.A. ha comenzado a utilizar tecnologías del Internet de las Cosas, computación en la nube y herramientas de inteligencia artificial más avanzadas que los casos anteriores para el procesamiento de los datos (nivel 5).

Teniendo en cuenta lo anterior, notamos que los casos estudiados ocupan los tres niveles intermedios de la escala de la figura 1-9 que nos mostraba los niveles de implementación del ecosistema de Big Data en la organización. Como se ha puesto de manifiesto, la empresa Eco Fuels Energy S.A.S. está en un nivel tres (Sistema para la Toma de Decisiones y Especializado de Información Empresarial), en un nivel cuatro se encuentra la Fundación Universidad del Norte (Transición a Ecosistema de Big Data) y en un nivel cinco la empresa Cementos Colombia S.A.

Una vez hemos visto la comparativa de todos los aspectos estudiados en los casos, a continuación, realizamos la comparativa del efecto de la implementación del ecosistema de Big Data en los procesos de gestión del conocimiento y en los mecanismos de coordinación y control en los casos estudiados.

5.4.4 Comparativa del efecto de la implementación del ecosistema de Big Data en los procesos de gestión del conocimiento y en los mecanismos de coordinación y control en los casos estudiados

En este apartado realizaremos una comparativa del efecto recíproco entre la implementación del ecosistema de Big Data, los procesos de gestión del conocimiento, los mecanismos de coordinación y control y los nuevos factores y patrones que han emergido en los casos que hicieron parte de esta investigación, para luego presentar el modelo final de relaciones. Para facilitar la comparación y lectura de las tablas y gráficos presentados a continuación, hemos establecido un color primario y un acrónimo para cada caso de la siguiente manera: Cementos Colombia S.A. (CC), Fundación Universidad del Norte (FUN) y Eco Fuels Energy S.A.S (EFE).

En primer lugar, comparamos el modelo inicial de relaciones planteado en la sección 3.3 el cual propone una relación bidireccional entre la implementación de ecosistemas de Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación y control con los modelos inducidos de cada uno de los casos. Dicha comparación para todos los casos estudiados la podemos encontrar en la tabla 5-6, donde cada casilla representa una relación unidireccional entre los factores y, en caso de darse la misma relación de forma inversa en la tabla, esto indicará que existe una relación bidireccional entre los factores.

Tabla 5-6: Comparativa de las relaciones entre factores propuestos en el modelo inicial de los casos estudiados

COMPONENTES PROPUESTOS EN MODELO INICIAL	COMPONENTES PROPUESTOS EN MODELO INICIAL			
	Implementación de Ecosistemas de Big Data	Proceso de Gestión del Conocimiento	Mecanismos de Coordinación	Mecanismos de Control
Implementación de Ecosistemas de Big Data		CC FUN EFE	CC FUN	CC FUN EFE
Proceso de Gestión del Conocimiento	CC FUN EFE		CC FUN	CC FUN
Mecanismos de Coordinación	CC FUN	CC FUN		CC FUN
Mecanismos de Control	CC FUN EFE	CC FUN	CC FUN	

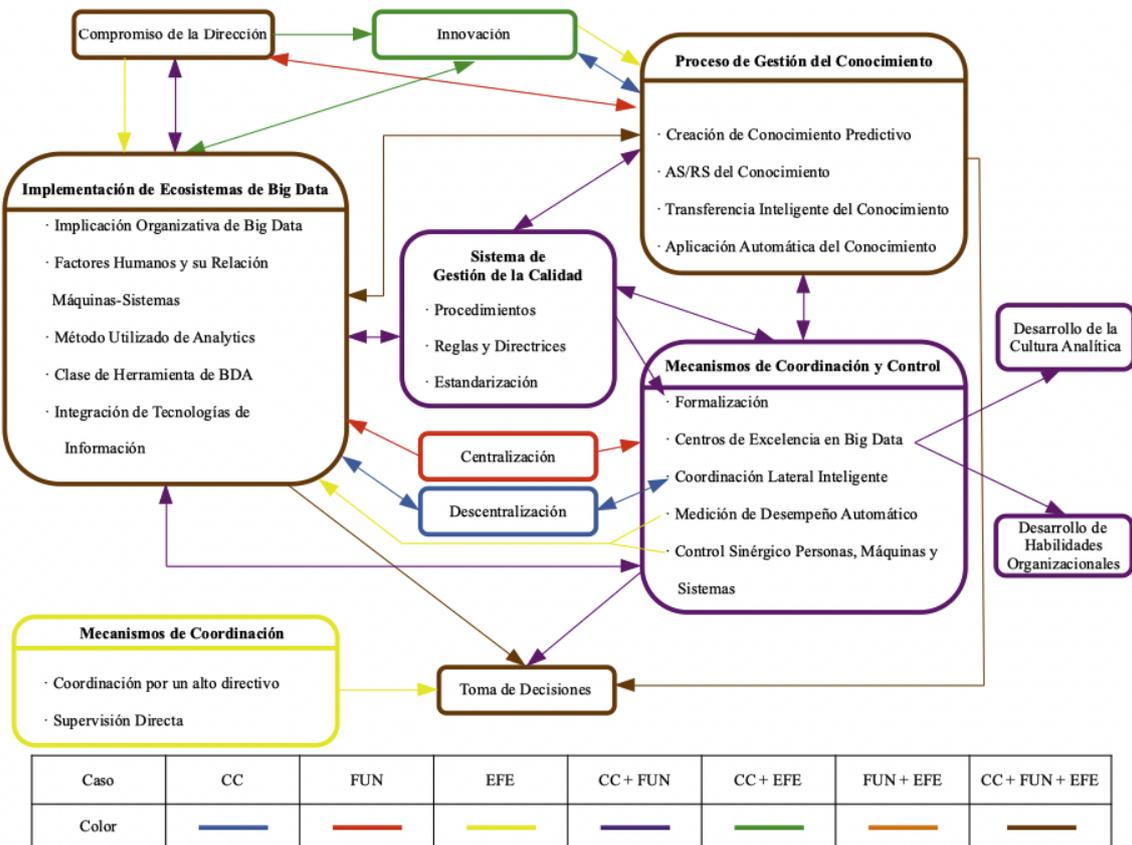
Fuente: elaboración propia

Como podemos observar en la tabla anterior, los casos de la empresa Cementos Colombia S.A. y la Fundación Universidad del Norte cumplen todas las relaciones bidireccionales entre factores planteados en el modelo inicial de la sección 3.3, mientras que para el caso de la empresa Eco Fuels Energy S.A.S. se cumple con las relaciones bidireccionales entre la implementación de ecosistemas de Big Data y los procesos de gestión del conocimiento y entre la implementación de ecosistemas de Big Data y los mecanismos de control. Por

lo tanto y teniendo en cuenta que no se presentaron relaciones unidireccionales, consideramos que los casos estudiados se ajustan al patrón establecido en el modelo inicial planteado debido a que cada relación bidireccional es evidenciada en al menos dos casos.

En segundo lugar, para analizar los nuevos factores y conexiones que han emergido del análisis de los casos (adicional a los planteados en el modelo inicial) hemos optado por presentar de forma gráfica un modelo preliminar de relaciones, donde hemos combinado los modelos finales presentados en las secciones 5.1.5, 5.2.5 y 5.3.5. En la figura 5-4 podemos observar el aporte de cada caso individual con su respectivo color o, en caso de que dos o tres casos coincidan en un factor o conexión, se combinarán los colores primarios y el color resultante representará dicha situación, por lo que se podrán encontrar las combinaciones **CC + FUN**, **CC + EFE**, **FUN + EFE** y **CC + FUN + EFE**, esta última indica que dicha conexión o factor se presentó en los tres casos estudiados.

Figura 5-4: Comparativa preliminar de factores y relaciones planteadas en el modelo inicial y de los nuevos factores y relaciones que han emergido del análisis de los casos



Fuente: elaboración propia

Teniendo en cuenta el gráfico anterior, podemos notar que la mayoría de los factores y conexiones tienen presencia en dos o más casos, con excepción de algunas conexiones y factores como es el caso de los mecanismos de coordinación de la empresa Eco Fuels Energy S.A.S. con los procesos de gestión del conocimiento y la implementación de ecosistemas de Big Data (ver tabla 5-6), así como la conexión unidireccional de este último factor con el compromiso de la dirección.

Además, con respecto a si los casos apoyan la centralización o descentralización de las decisiones en la organización con la implementación de los ecosistemas de Big Data, se puede observar que la empresa Cementos Colombia S.A. (descentralización) y la Fundación Universidad del Norte (centralización) presentan formas contrarias de llevar a cabo esta conexión, por lo que atendiendo a nuestros resultados no se puede afirmar que la descentralización sea un requisito para la implementación de ecosistemas de Big Data, tal como se había expresado en la sección 3.2.1.

Adicional a lo anterior, la figura nos ilustra que la conexión entre el compromiso de la dirección (presente en los tres casos) y los procesos de gestión del conocimiento se da de forma directa y bidireccional para el caso de la Fundación Universidad del Norte, esto debido a que en los otros dos casos vemos que dicha conexión se presenta, pero mediada por el factor de innovación de estos dos casos. En cuanto a la conexión entre la innovación y los procesos de gestión del conocimiento vemos que para la empresa Cementos Colombia S.A. se presenta de forma bidireccional mientras que para la empresa Eco Fuels Energy S.A.S. es de forma unidireccional. Esto último es debido a que en la empresa Eco Fuels S.A.S. la innovación ha sido utilizada como mecanismo para la construcción del conocimiento organizacional, en cambio la empresa Cementos Colombia S.A. lo utiliza como una estrategia organizacional para a partir del conocimiento crear innovación en toda la organización y, a su vez, crear y difundir nuevo conocimiento en la organización.

Estos factores y conexiones que no se han presentado en dos o tres casos no serán tomados en cuenta para la construcción del modelo final de relaciones en esta tesis doctoral. Sin embargo, dichos factores y conexiones pueden ser considerados en futuras investigaciones para explorar más a fondo la presencia o no de los mismos al analizar más casos en diversos sectores y en otros mercados, en especial el análisis sobre el papel de

la descentralización en la relación entre la implementación de ecosistemas de Big Data, los mecanismos de coordinación y control y los procesos de gestión del conocimiento, ya que es posible que dicha descentralización pueda presentarse a partir del nivel cinco de implementación del ecosistema.

Por otro lado, destacamos la importancia del sistema de gestión de la calidad en los casos de la empresa Cementos Colombia S.A. y la Fundación Universidad del Norte, jugando un papel mediador entre los procesos de gestión del conocimiento, los mecanismos de coordinación y control y la implementación de los ecosistemas de Big data, apoyando mediante procedimientos, reglas, directrices y la estandarización de procesos a la correcta implementación del ecosistema, apoyar los procesos de gestión del conocimiento dentro de la organización y apoyar la coordinación y control de las actividades a través de la formalización.

Otro aspecto importante es el surgimiento de los Centros de Excelencia en los casos de la empresa Cementos Colombia S.A. y la Fundación Universidad del Norte, los cuales podemos considerar que surgen a partir del nivel cuatro de implementación del ecosistema de Big Data y se convierten en una parte fundamental para la coordinación y apoyo de las actividades que impliquen soluciones basadas en datos, por lo que dichos centros son los encargados de desarrollar la cultura analítica y de nuevas habilidades en la organización.

Por otro lado, notamos que los entrevistados encargados de los Centro de Excelencia (CC + FUN) o de la innovación y tecnología (EFE) mencionaron que todos los esfuerzos de su área para llevar a cabo la implementación del ecosistema de Big Data serían en vano sin el compromiso de las directivas de la organización, compromiso que se debe mantener con resultados visibles provistos por el ecosistema. Este compromiso, a su vez, promoverá la innovación y dicha innovación tendrá una relación recíproca con la implementación del ecosistema de Big Data (CC + EFE).

Como establecimos en varias secciones anteriores, uno de los principales objetivos que se busca con la implementación de los ecosistemas de Big Data es la toma de mejores decisiones basadas en datos de gran volumen, alta velocidad y gran variedad, factor que estuvo presente en los tres casos, coincidiendo con que dicha implementación, los

procesos de gestión del conocimiento y los mecanismos de coordinación y control (a excepción de los mecanismos de control en Eco Fuels Energy S.A.S.) se combinan para la toma de mejores decisiones.

Teniendo en cuenta lo anterior y eliminando los factores y/o conexiones que solo se han puesto de manifiesto en uno de los casos (exceptuando la conexión entre innovación y los procesos de gestión del conocimiento aportada por Cementos Colombia S.A.), presentamos en la tabla 5-7 el aporte al modelo final de relaciones de cada combinación entre los casos estudiados para analizar la importancia de cada uno de ellos.

Tabla 5-7: Resumen del aporte al modelo final de relaciones de la combinación de los casos estudiados

CASOS ESTUDIADOS	APORTE AL MODELO FINAL	
	FACTORES	CONEXIONES
CC + FUN	4	10
CC + EFE	1	2
FUN + EFE	0	0
CC + FUN + EFE	4	3
CC	0	1
TOTAL	9	16

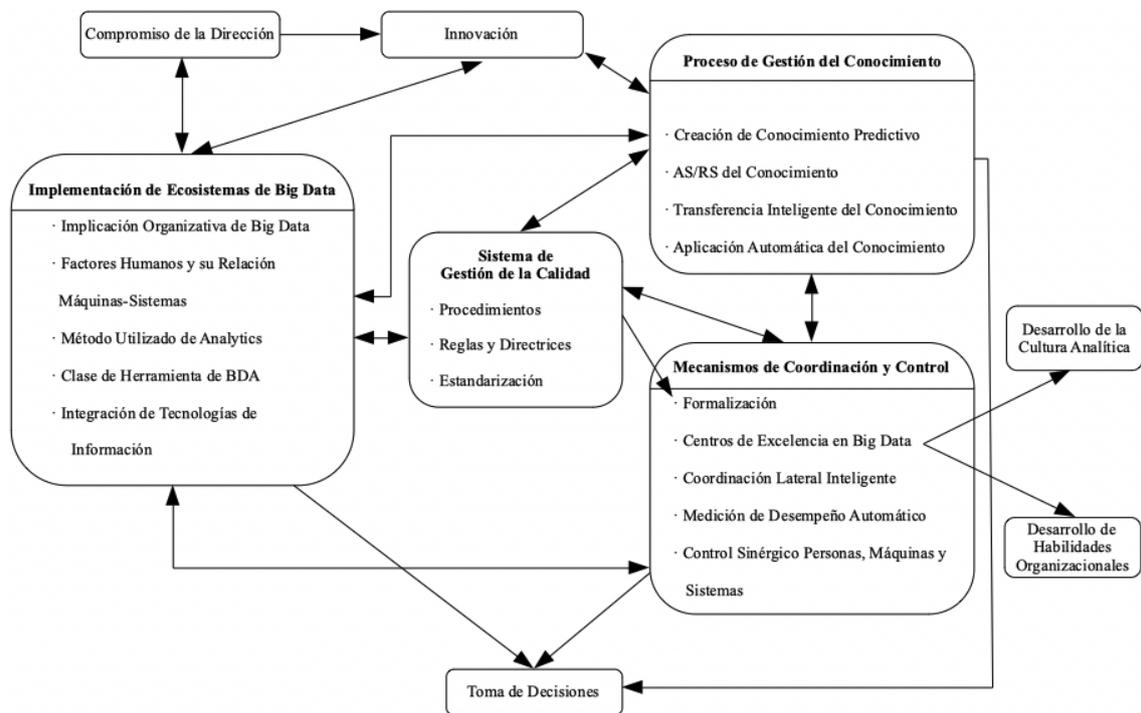
Fuente: elaboración propia

Acorde a la tabla anterior notamos que la combinación entre los casos Cementos Colombia S.A. y Fundación Universidad del Norte son los que más han aportado al modelo final con cuatro factores y 10 conexiones, seguido por la combinación de los tres casos con 4 factores y tres conexiones, luego por la combinación entre Cementos Colombia S.A. y Eco Fuels Energy S.A.S. con un factor y 2 conexiones y, por último, la combinación entre la Fundación Universidad del Norte y la empresa Eco Fuels Energy S.A.S. no aportó ni factores ni conexiones al modelo final de relaciones.

El mayor aporte de las organizaciones Cementos Colombia S.A. y Fundación Universidad del Norte se ha debido a una mayor madurez organizacional y a un mayor avance en el proceso de implementación del ecosistema de Big Data. Sin embargo, consideramos que el caso Eco Fuels Energy S.A.S ha sido de gran aporte debido a que nos brinda una visión de una organización en el inicio del proceso de implementación del ecosistema de Big Data y de la construcción de estrategias de conocimiento y su relación con los mecanismos de coordinación y control.

Por lo tanto, una vez hemos seleccionado los factores y conexiones que harán parte del modelo inducido final de relaciones, presentamos a continuación la representación gráfica de dicho modelo (ver figura 5-5).

Figura 5-5: Modelo inducido final de relaciones a partir del análisis conjunto de los casos



Fuente: elaboración propia

Una vez hemos cumplido con el objetivo de estudiar la relación entre la implementación de ecosistemas de Big Data, los procesos de gestión del conocimiento, los mecanismos de coordinación y control y los nuevos factores y conexiones que han surgido de los casos estudiados, en el siguiente capítulo presentaremos las conclusiones generales de esta tesis

doctoral, así como las futuras investigaciones que pueden surgir a partir de lo planteado a lo largo de este documento.

Capítulo 6:

Conclusiones Generales

Una vez hemos finalizado todos los capítulos que corresponden al desarrollo de esta tesis doctoral, a continuación, presentamos nuestras reflexiones en relación al cumplimiento del objetivo de esta investigación, las principales implicaciones y aportes que se derivan de la presente tesis doctoral, así como las limitaciones que hicieron presencia durante el desarrollo de la misma y posibles futuras líneas de investigación que pueden surgir a partir del trabajo realizado.

En primer lugar y respecto al *cumplimiento de los objetivos* de la presente tesis doctoral, en los capítulos 1, 2 y 3 hemos conseguido establecer un marco teórico amplio y suficiente que describiera los conceptos fundamentales que se requerían para responder a la pregunta de investigación. Como resultado de dicho marco teórico y teniendo en cuenta algunas reflexiones teóricas personales basadas en nuestra experiencia y conocimiento del fenómeno objeto de estudio, al final del capítulo 3 hemos propuesto un modelo que ilustra la relación entre la implementación del ecosistema de Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación y control de la organización, modelo que sirvió como base para poder diseñar las herramientas necesarias para la recogida de datos y llevar a cabo el análisis de la información recopilada.

Mediante el capítulo 4 logramos establecer una metodología de investigación adecuada para dar respuesta a nuestra pregunta de investigación, basándonos para ello en la naturaleza exploratoria de la misma, tomando las decisiones correspondientes a una investigación de tipo cualitativo y seleccionando el estudio de casos como la estrategia de investigación que mejor se adapta al objetivo de esta investigación. Dicho capítulo contiene los detalles y decisiones correspondientes al diseño del estudio de casos efectuado, así como los aspectos relativos a los criterios de calidad de la presente investigación y las consideraciones éticas que se han tenido presente a la hora de llevar a cabo la investigación.

En el capítulo 5 hemos procedido a describir el análisis de los resultados para los casos de estudio seleccionados en esta tesis doctoral. En particular, hemos explicado para cada caso el perfil de la organización, el proceso de gestión del conocimiento llevado a cabo, los mecanismos de coordinación y control utilizados, el ecosistema de Big Data implementado en la organización y elaborado el correspondiente modelo que explica

cómo se produce la relación entre los 3 aspectos antes mencionados. En dichos modelos (uno por cada caso analizado) surgieron nuevos factores y conexiones que ayudaban a entender mejor la relación entre el ecosistema de Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación y control, si bien algunos de los factores o variables iniciales que habíamos considerado en el modelo teórico del capítulo 3 también permanecían en los mismos. En realidad, cada modelo nos muestra un patrón de relaciones que explica cómo la aplicación del ecosistema de Big Data en cada organización estudiada (i.e. los casos) afecta a sus procesos de gestión del conocimiento y a las formas de coordinación y control que se dan en cada organización.

En el apartado final del capítulo 5 se realizó la comparativa de los tres casos de estudio, para cada uno de los tres ejes fundamentales del objetivo de investigación, con el fin de identificar similitudes y diferencias entre los mismos. Nuestra intención era la de identificar un patrón común a los tres casos para lo cual consideramos tan solo aquellos factores y conexiones que hicieron presencia en al menos dos de los casos estudiados, destacando el aporte que hizo cada uno de los casos estudiados al modelo final de relaciones.

Así, el modelo final de relaciones resultado del análisis conjunto de los tres casos estudiados explica cómo se presenta la relación recíproca entre la implementación de los ecosistemas de Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación y control de la organización. Nuestra investigación ha evidenciado que factores como el compromiso de la dirección, la innovación, la existencia de un sistema de gestión de calidad, las características de los procesos de toma de decisiones, la presencia de valores organizativos basados en una cultura analítica y el desarrollo de nuevas habilidades organizacionales como las habilidades para el análisis de los datos, habilidad para la toma de decisiones basada en datos, y habilidad para el uso de sistemas y tecnologías para el desarrollo de las actividades diarias, son fundamentales para entender cómo la implantación de un ecosistema de Big Data modifica la creación, la adquisición, el almacenamiento y la distribución del conocimiento en la compañía, y también cómo se modifican los mecanismos utilizados para coordinar y controlar el trabajo organizativo.

En particular, los casos de estudio nos dejaron lecciones con respecto al fenómeno del Big Data en las organizaciones entre las que destacamos:

- ✓ El apoyo e implicación de la dirección es un aspecto fundamental en todo el proceso de implementación de Big Data, ya que realizar este proceso de implementación para la mejora de las decisiones basada en datos implica cambios estratégicos en la organización, lo cual es responsabilidad de las directivas. Además, las directivas deben apoyar en temas de inversión económica y de coordinación de todo el proceso de implementación para unificar los esfuerzos de la organización en este sentido.
- ✓ A diferencia a lo visto en la literatura, las organizaciones estudiadas están procurando adaptar su estructura tradicional a los ecosistemas de Big Data, evidenciado en los casos de Fundación Universidad del Norte y Cementos Colombia S.A. Esto se debe a que las organizaciones latinoamericanas suelen ser más conservadoras en cuanto su estructura se refiere, y el cambio a nuevas estructuras se realiza de manera progresiva a medida que se consigan mejoras visibles con esta nueva estructura.
- ✓ La adaptación a un ecosistema de Big Data debe empezar por evaluar qué aspectos y procesos clave debe cambiar o crear la organización.
- ✓ La implementación del ecosistema de Big Data debe ser un proceso progresivo donde se vaya escalando en los niveles de implementación, para que de esta manera el ecosistema y los procesos de la organización vayan evolucionando de forma conjunta.
- ✓ Teniendo en cuenta los casos estudiados en la presente tesis, el proceso de implementación de los ecosistemas de Big Data ha sido principalmente mediante prueba y error, y no hay una hoja de ruta definida. Las preocupaciones de los directivos entrevistados se centran en la falta de capital humano para estos nuevos sistemas y tecnologías y el definir cuáles son las mejores prácticas para no fracasar en el intento de llevar a cabo la implementación.
- ✓ La evidencia nos muestra que no es requisito adoptar estructuras de tipo matricial (como sugiere la literatura) al momento de adoptar ecosistemas de Big Data ya que, como mencionamos anteriormente, dicha implementación es un proceso progresivo. Esto lo notamos con la empresa Cementos Colombia S.A. que al estar

en un nivel cinco de implementación ya comenzó a adoptar características de una organización matricial.

- ✓ La implementación del ecosistema de Big Data debe involucrar a todas las personas de la organización y debe convertirse en parte de su cultura, a lo que se le denomina cultura analítica.

De esta manera, tomando en consideración todo lo anterior podemos establecer que nuestro objetivo inicial para esta investigación ha sido alcanzado plenamente.

Respecto a los *aportes realizados y a las implicaciones de esta investigación*, la presente tesis doctoral buscó explorar un terreno poco explorado en el campo de la dirección de las organizaciones. Pero a pesar de esto, es innegable el efecto que ha traído el Big Data a las organizaciones actuales en todo el mundo, fenómeno que tiene mucho camino por recorrer para comprender su influencia sobre las mismas.

Como hemos mencionado anteriormente, la investigación sobre la influencia de los ecosistemas de Big Data se ha centrado principalmente en la relación entre los procesos de gestión del conocimiento y el mismo ecosistema. Sin embargo, estudiar la relación entre diferentes factores organizacionales y el ecosistema de Big Data es fundamental para comprender cómo se modifican dichos procesos en la organización y, a su vez, cómo dichos procesos favorecen la implementación del ecosistema. Por lo tanto, y adicional a los procesos de gestión del conocimiento, un primer aporte ha consistido en el estudio adicional de los mecanismos de coordinación y control, para poder, de esta manera, comprender mejor los cambios que se presentan en este ámbito.

Durante el desarrollo de la presente tesis doctoral se encontraron brechas por cubrir en la literatura especializada en la temática, brechas que era necesario cubrir para poder resolver nuestra pregunta de investigación y que han constituido contribuciones o aportes de la presente investigación. En primer lugar, en el capítulo 1 hemos enmarcado el fenómeno de Big Data en el campo de los sistemas y tecnologías de la información y todos los conceptos asociados a los ecosistemas de Big Data. Sin embargo, hemos notado que en la literatura no existía una clasificación que nos permitiera establecer los niveles de implementación de los ecosistemas de Big Data en las organizaciones (ver sección 1.3.3). En este sentido, un aporte de esta tesis doctoral ha consistido en elaborar una

propuesta de siete niveles de implementación (madurez) del ecosistema de Big Data, niveles que se determinan mediante cinco dimensiones que establecen los requisitos generales que debe cumplir una organización para considerarse en un determinado nivel en cada una de las dimensiones. Dichos niveles de implementación pueden ser utilizados por otros investigadores en posteriores estudios en el campo, al tenerlo como base para la selección de casos de estudio, o en el caso de empresas u organizaciones para establecer qué requisitos se requieren o qué aspectos se deben tomar en cuenta para considerar que una organización se encuentra en determinado nivel de implementación del Big Data.

En segundo lugar, en el capítulo 3 hemos realizado la relación teórica entre la implementación de los ecosistemas de Big Data y los procesos de gestión del conocimiento y la relación teórica entre la implementación de los ecosistemas de Big Data y los mecanismos de coordinación y control. Sin embargo, y como se menciona en las secciones 3.1 y 3.2, las relaciones antes mencionadas no han sido muy exploradas por la literatura lo que ha llevado a que en la presente tesis doctoral se realice una recopilación de la literatura en dichos temas y se complemente con nuevos aportes mediante el análisis de la literatura variada en el área y la experiencia del autor en el campo, y así construir dichas relaciones teóricas y proponer el modelo teórico presentado en dicho capítulo. Las propuestas teóricas planteadas en el capítulo 3 pueden ser un punto de partida importante para investigadores en el área de los ecosistemas de Big Data y su relación con los procesos de gestión del conocimiento y los mecanismos de coordinación y control, o de otros temas de investigación relacionados con el área, para así comprender un poco mejor dicho fenómeno en las organizaciones.

En tercer lugar, el modelo inducido final de relaciones que surge a partir del análisis conjunto de los casos muestra aquellos factores importantes que debe considerarse en una organización para llevar a cabo la implementación del ecosistema de Big Data y pone de manifiesto aquellos aspectos a tener en cuenta para realizar los cambios organizativos que se requieren para dicha implementación. Dicho modelo final puede ser usado como mapa de carreteras por organizaciones que deseen llevar a cabo el proceso de implementación del ecosistema de Big Data y quieran saber, de antemano, su efecto sobre la gestión del conocimiento y la forma de coordinar y controlar el trabajo. Así mismo, este modelo puede ser un referente o punto de partida relevante para la academia y, en particular, para investigadores que deseen profundizar en el estudio de este fenómeno. Las relaciones

puestas de manifiesto en nuestro modelo final inducido pueden constituir una fundamentación teórica de partida para dichos investigadores.

Como cualquier trabajo de investigación, el presente no está exento de *limitaciones*. En este sentido destacamos, en primer lugar, el tamaño reducido de la muestra de casos estudiados. El número limitado de casos que se ha estudiado en la presente investigación (tres) conlleva que los resultados alcanzados no puedan ser extrapolados a otras poblaciones u otros mercados distintos al colombiano, que es el lugar geográfico donde se escogieron los casos. Del mismo modo, los casos han sido escogidos en base al cumplimiento de los requerimientos plasmados en el capítulo 4 por lo que, al considerarse otros aspectos para la selección de los mismos, los resultados podrían variar respecto a los que se presentaron en la presente investigación. En cualquier caso, esta limitación es inherente a la decisión de elegir el estudio de casos como estrategia fundamental de investigación, y es importante considerar que los criterios de diseño de la muestra en un estudio de casos exploratorio como el desarrollado en esta investigación no están basados en la representatividad estadística sino en la relevancia de los casos elegidos en función del objetivo de investigación a alcanzar.

En segundo lugar, el modelo final de relaciones ha surgido mediante un análisis inductivo debido al carácter exploratorio de la presente investigación por lo que no se han probado estadísticamente las relaciones que surgen a raíz de los factores y conexiones del modelo final planteado. Así y tal y como se ha comentado anteriormente, la presente investigación puede servir como punto de partida en futuras investigaciones para comprobar dichas relaciones y factores, lo que puede llevar a que el modelo aquí planteado cambie en un futuro.

En tercer lugar, para cada organización solo se han seleccionado dos entrevistados relacionados con las áreas de interés de esta investigación. Además, los dos entrevistados ocupaban un cargo medio de dirección con lo que pudiera ser que sus perspectivas y opiniones poseyeran algún tipo de sesgo subjetivo. Hubiera sido deseable poder haber obtenido información primaria de cada caso proveniente de individuos ubicados en otros niveles jerárquicos diferentes (e.g. subordinados, superiores) lo que nos hubiera facilitado el proceso de triangulación de la información. En este sentido, y como hemos notado en los modelos que surgieron en los casos, existen más áreas funcionales que hacen parte de

la implementación del ecosistema de Big Data por lo que hubiera sido interesante tener en cuenta a otros entrevistados provenientes de áreas funcionales afectadas por la implementación del ecosistema de Big Data y que no se hayan considerado en esta investigación.

En cuarto lugar, en el capítulo 3 se consideraban una serie de factores limitantes que afectaban a la relación entre la aplicación del Big Data y los procesos de gestión del conocimiento en la organización. Sin embargo, nuestros resultados no han arrojado ninguna luz sobre este tipo de factores lo cual puede constituir una limitación de la presente investigación. En este sentido, futuras investigaciones deberían centrarse en analizar también qué factores limitan el efecto de la aplicación del Big Data sobre los procesos de gestión del conocimiento y cómo lo hacen.

Por último y respecto a las *futuras líneas de investigación*, consideramos que deben centrarse principalmente en la comprobación de los factores y conexiones que se han puesto de manifiesto en el modelo final obtenido en esta tesis doctoral. Dichas relaciones han sido inducidas a partir de las evidencias de los casos analizados. Sin embargo, queda todavía mucho trabajo por realizar si deseamos que estas relaciones puedan ser generalizables a otras poblaciones o colectivos. Así, en el futuro sería necesario diseñar un trabajo de investigación de corte cuantitativo encaminado a contrastar estadísticamente dichas relaciones.

Adicional a lo anterior, consideramos que la implementación de los ecosistemas de Big Data puede ser explorada desde todos los campos de la investigación empresarial ya que, como hemos mencionado antes, es un fenómeno que involucra diversos procesos de la organización. Tomando como punto de partida el modelo final, sería interesante estudiar a profundidad el fenómeno del Big Data y su influencia en los sistemas de gestión de la calidad, la cultura analítica, el desarrollo de habilidades organizacionales, la innovación y el compromiso de la dirección. Esto llevará a plantear nuevas conexiones y factores que no se ha logrado reflejar en la presente investigación y que seguramente enriquecerán el modelo aquí planteado.

Otro aspecto importante para futuras investigaciones es el de explorar el fenómeno aquí estudiado, pero hacerlo desde diversos sectores y geografías con el fin de identificar las

diferencias y semejanzas entre los mismos. Nuestro trabajo ha estado centrado en el ámbito colombiano y en organizaciones de tamaño mediano y grande. ¿Serían similares los resultados obtenidos para organizaciones de otros ámbitos geográficos distintos (e.g. Europa) y/u otros tamaños diferentes (e.g. pequeñas organizaciones y microempresas)? Debido a que la implementación de ecosistemas de Big Data puede aplicarse a empresas de cualquier sector de actividad, tamaño y ámbito geográfico, animamos desde aquí a la comunidad académica a que desarrolle trabajos similares que puedan ser una buena base de comparación con los resultados aquí obtenidos.

Así mismo, futuras investigaciones podrían abordar el estudio de los factores limitantes de los procesos aquí analizados, tal y como antes se ha mencionado, o la inclusión de variables sociales, políticas y económicas, propias del contexto en el que se encuentren las organizaciones analizadas y que, con total seguridad, pueden afectar al modo en que estas aplican el ecosistema de Big Data.

Por último, agradecemos al lector de la presente tesis doctoral su interés por los temas aquí tratados y esperamos que esta investigación sirva de inspiración a futuros investigadores. El estudio del fenómeno del Big Data no ha hecho más que comenzar y, seguramente, generará muchos resultados positivos para el crecimiento de diversas organizaciones a nivel mundial, por lo que invitamos a los investigadores a seguir explorando este fenómeno y a resolver las múltiples preguntas de investigación que están aún sin resolver.

Bibliografía

- Abubakar, A. M., Elrehail, H., Alatailat, M. A., & Elçi, A. (2019). Knowledge management, decision-making style and organizational performance. *Journal of Innovation & Knowledge*, 4(2), 104-114. <https://doi.org/10.1016/j.jik.2017.07.003>
- Acharya, A., Singh, S. K., Pereira, V., & Singh, P. (2018). Big data, knowledge co-creation and decision making in fashion industry. *International Journal of Information Management*, 42, 90-101. <https://doi.org/10.1016/j.ijinfomgt.2018.06.008>
- Ahalt, S., Bedard, D., Carsey, T. M., Crabtree, J., Green, K., Jeffries, C., ... Thakur, S. (2012). Establishing a national consortium for data science. *Renaissance Computing Institute, University of North Carolina at Chapel Hill, USA*. Recuperado de http://data2discovery.org/dev/wp-content/uploads/2012/09/NCDS-Consortium-Roadmap_July.pdf
- Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Communications Surveys & Tutorials*, 17(4), 2347-2376. <https://doi.org/10.1109/COMST.2015.2444095>
- Alavi, M., & Leidner, D. E. (2001). Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, 25(1), 107-136. <https://doi.org/10.2307/3250961>
- Amazon. (2018). Amazon Go. Recuperado 2 de julio de 2018, de <https://www.amazon.com/go>
- Amazon. (2019). Amazon Transcribe. Recuperado 15 de agosto de 2019, de <https://aws.amazon.com/es/transcribe/>
- Apache Software Foundation. (2014). *Welcome to Apache™ Hadoop®!* Recuperado de <http://hadoop.apache.org/index.pdf>
- Archana, J., & Anita, E. A. M. (2015). A Survey of Big Data Analytics in Healthcare and Government. *Procedia Computer Science*, 50, 408-413. <https://doi.org/10.1016/j.procs.2015.04.021>
- Argote, L., McEvily, B., & Reagans, R. (2003). Managing Knowledge in Organizations: An Integrative Framework and Review of Emerging Themes. *Management Science*, 49(4), 571-582. <https://doi.org/10.1287/mnsc.49.4.571.14424>
- Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A survey. *Computer*

- Networks*, 54(15), 2787-2805. <https://doi.org/10.1016/j.comnet.2010.05.010>
- Balbastre, F. (2003). *La autoevaluación según los modelos de gestión de calidad total y el aprendizaje en la organización: una investigación de carácter exploratorio*. Universidad de Valencia.
- Barranco, R. (2012). ¿Qué es Big Data? Recuperado de <https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/>
- Baxter, P., & Jack, S. (2015). Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report*, 13(4), 544-559. <https://doi.org/10.46743/2160-3715/2008.1573>
- Becerra-Fernandez, I., & Sabherwal, R. (2014). *Knowledge Management: Systems and Processes* (Segunda Ed). New York: Routledge. <https://doi.org/10.4324/9781315715117>
- Becker, M. C., Lazaric, N., Nelson, R. R., & Winter, S. G. (2005). Applying organizational routines in understanding organizational change. *Industrial and Corporate Change*, 14(5), 775-791. <https://doi.org/10.1093/icc/dth071>
- Berner, M., Graupner, E., & Maedche, A. (2014). The Information Panopticon in the Big Data Era. *Journal of Organization Design*, 3(1), 14-19. <https://doi.org/10.7146/jod.9736>
- Bluhm, D. J., Harman, W., Lee, T. W., & Mitchell, T. R. (2011). Qualitative Research in Management: A Decade of Progress. *Journal of Management Studies*, 48(8), 1866-1891. <https://doi.org/10.1111/j.1467-6486.2010.00972.x>
- Boeije, H. (2010). *Analysis in Qualitative Research* (Primera Ed). London: Sage Publications.
- Bolsa de Valores de Colombia. (2019). Empresas COLCAP. Recuperado 12 de abril de 2019, de <https://www.bvc.com.co/pps/tibco/portalbvc/Home/Empresas/Ranking+por+Capitalización+Bursátil?action=dummy>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2014). What can “thematic analysis” offer health and wellbeing researchers? *International Journal of Qualitative Studies on Health and Well-being*, 9(1). <https://doi.org/10.3402/qhw.v9.26152>
- Bryman, A. (1988). *Quantity and quality in social research* (Vol. 18). London: Unwin Hyman Ltd.

- Bumblauskas, D., Nold, H., Bumblauskas, P., & Igou, A. (2017). Big data analytics: transforming data to action. *Business Process Management Journal*, 23(3), 703-720. <https://doi.org/10.1108/BPMJ-03-2016-0056>
- Burton, R. M., Obel, B., & Håkansson, D. D. (2015). *Organizational Design: A Step-by-Step Approach* (Third). Cambridge, UK: Cambridge University Press.
- Byrd, T. A., & Turner, D. E. (2000). Measuring the Flexibility of Information Technology Infrastructure: Exploratory Analysis of a Construct. *Journal of Management Information Systems*, 17(1), 167-208. <https://doi.org/10.1080/07421222.2000.11045632>
- Cairó-Battistutti, O., & Bork, D. (2017). Tacit to explicit knowledge conversion. *Cognitive Processing*, 18(4), 461-477. <https://doi.org/10.1007/s10339-017-0825-6>
- Carayannis, E. G., Grigoroudis, E., Del Giudice, M., Della Peruta, M. R., & Sindakis, S. (2017). An exploration of contemporary organizational artifacts and routines in a sustainable excellence context. *Journal of Knowledge Management*, 21(1), 35-56. <https://doi.org/10.1108/JKM-10-2015-0366>
- Carillo, K. D. A. (2017). Let's stop trying to be «sexy» – preparing managers for the (big) data-driven business era. *Business Process Management Journal*, 23(3), 598-622. <https://doi.org/10.1108/BPMJ-09-2016-0188>
- Cassell, C. (2015). *Conducting Research Interviews for Business and Management Students*. London: SAGE Publications Ltd. <https://doi.org/10.4135/9781529716726>
- Cegarra-Navarro, J.-G., & Martelo-Landroguez, S. (2020). The effect of organizational memory on organizational agility. *Journal of Intellectual Capital*, 21(3), 459-479. <https://doi.org/10.1108/JIC-03-2019-0048>
- Cementos Colombia S. A. (2019a). Página Web: centro para la innovación.
- Cementos Colombia S. A. (2019b). *Política de Gestión Humana*.
- Cementos Colombia S. A. (2019c). *Política de Investigación y Desarrollo*.
- Cementos Colombia S. A. (2019d). *Presentación Corporativa*.
- Cementos Colombia S. A. (2019e). *Reporte Integrado 2018*.
- Chen, C. L., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, 275, 314-347. <https://doi.org/10.1016/j.ins.2014.01.015>
- Chen, H., Chiang, R. H. L., & Storey, V. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4), 1165-1188. <https://doi.org/10.2307/41703503>

- Chen, M., Mao, S., & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*, 19(2), 171-209. <https://doi.org/10.1007/s11036-013-0489-0>
- Child, J. (1973). Predicting and Understanding Organization Structure. *Administrative Science Quarterly*, 18(2), 168-185. <https://doi.org/10.2307/2392061>
- Child, J. (2015). *Organization: Contemporary Principles and Practice* (Segunda Ed). Chichester, UK: John Wiley and Sons.
- Clarke, V., & Braun, V. (2013). Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The psychologist*, 26(2), 120-123. Recuperado de <https://uwe-repository.worktribe.com/preview/937606/Teaching>
- Cohen, D., & Asín, E. (2013). *Tecnologías de Informacion en los Negocios* (Quinta Ed). Mexico D.F.: Mc Graw Hill.
- Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1), 128-152. <https://doi.org/10.2307/2393553>
- Colas, M., Finck, I., Buvat, J., Nambiar, R., & Raj Singh, R. (2014). Cracking the Data Conundrum : How Successful Companies Make Big Data Operational. Recuperado de https://www.capgemini-consulting.com/resource-file-access/resource/pdf/cracking_the_data_conundrum-big_data_pov_13-1-15_v2.pdf
- Colbert, A., Yee, N., & George, G. (2016). The Digital Workforce and the Workplace of the Future. *Academy of Management Journal*, 59(3), 731-739. <https://doi.org/10.5465/amj.2016.4003>
- Costa, V., & Monteiro, S. (2018). From Potential Absorptive Capacity to Knowledge Creation in Organisations: The Mediating Role of Knowledge Storage and Realised Absorptive Capacity. *Journal of Information & Knowledge Management*, 17(01). <https://doi.org/10.1142/S0219649218500065>
- Cox, M., & Ellsworth, D. (1997). Application-controlled demand paging for out-of-core visualization. En *Proceedings Visualization '97* (pp. 235-244). IEEE. <https://doi.org/10.1109/VISUAL.1997.663888>
- Daft, R. L., & Lengel, R. H. (1986). Organizational Information Requirements, Media Richness and Structural Design. *Management Science*, 32(5), 554-571. <https://doi.org/10.1287/mnsc.32.5.554>
- Davenport, T. H. (2014). *Big data at work: dispelling the myths, uncovering the opportunities*. Boston: Harvard Business Review Press.
- Davenport, T. H., Barth, P., & Bean, R. (2012). How «Big Data» Is Different. *MIT Sloan*

- Management Review*, (Fall), 43-46.
- Davenport, T. H., Long, D. W. de, & Beers, M. C. (1998). Successful Knowledge Management Projects. *MIT Sloan Management Review*, 39(2), 43-57.
- Davenport, T. H., & Patil, D. J. (2012). Data scientist: the sexiest job of the 21st century: meet the people who can coax treasure out of messy, unstructured data. *Harvard Business Review*, (October), 9.
- De Zubielqui, G. C., Lindsay, N., Lindsay, W., & Jones, J. (2019). Knowledge quality, innovation and firm performance: a study of knowledge transfer in SMEs. *Small Business Economics*, 53(1), 145-164.
- Del Giudice, M. (2016). Discovering the Internet of Things (IoT) within the business process management. *Business Process Management Journal*, 22(2), 263-270. <https://doi.org/10.1108/BPMJ-12-2015-0173>
- Dewett, T. (2001). The role of information technology in the organization: a review, model, and assessment. *Journal of Management*, 27(3), 313-346. [https://doi.org/10.1016/S0149-2063\(01\)00094-0](https://doi.org/10.1016/S0149-2063(01)00094-0)
- Diebold, F. X. (2012). A Personal Perspective on the Origin(s) and Development of «Big Data»: The Phenomenon, the Term, and the Discipline, Second Version. *SSRN Electronic Journal*, 8. <https://doi.org/10.2139/ssrn.2202843>
- Dijcks, J. P. (2013). Oracle: Big data for the enterprise. Recuperado 1 de junio de 2018, de <http://www.oracle.com/us/products/database/big-data-for-enterprise-519135.pdf>
- Dijkman, R. M., Sprenkels, B., Peeters, T., & Janssen, A. (2015). Business models for the Internet of Things. *International Journal of Information Management*, 35(6), 672-678. <https://doi.org/10.1016/j.ijinfomgt.2015.07.008>
- Duan, L., & Xiong, Y. (2015). Big data analytics and business analytics. *Journal of Management Analytics*, 2(1), 1-21. <https://doi.org/10.1080/23270012.2015.1020891>
- Dul, J., & Hak, T. (2010). *Case Study Methodology in Business Research* (Primera Ed). Elsevier Ltd.
- Dyer, W. G., & Wilkins, A. L. (1991). Better stories, not better constructs, to generate better theory: A rejoinder to Eisenhardt. *Academy of management review*, 16(3), 613-619. <https://doi.org/10.5465/amr.1991.4279492>
- Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of management review*, 14(4), 532-550. <https://doi.org/10.5465/amr.1989.4308385>
- Elgendy, N., & Elragal, A. (2016). Big Data Analytics in Support of the Decision Making Process. *Procedia Computer Science*, 100, 1071-1084.

<https://doi.org/10.1016/j.procs.2016.09.251>

- Elliott, V. (2018). Thinking about the coding process in qualitative data analysis. *The Qualitative Report*, 23(11), 2850-2861.
- Fan, J., Han, F., & Liu, H. (2014). Challenges of Big Data analysis. *National Science Review*, 1(2), 293-314. <https://doi.org/10.1093/nsr/nwt032>
- Ferretti, M., & Schiavone, F. (2016). Internet of Things and business processes redesign in seaports: The case of Hamburg. *Business Process Management Journal*, 22(2), 271-284. <https://doi.org/10.1108/BPMJ-05-2015-0079>
- Firican, G. (2017). The 10 Vs of Big Data. Recuperado 1 de junio de 2018, de <https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx>
- Fjeldstad, Ø. D., & Snow, C. C. (2018). Business models and organization design. *Long Range Planning*, 51(1), 32-39. <https://doi.org/10.1016/j.lrp.2017.07.008>
- Flick, U. (2018). *An introduction to qualitative research*. London: SAGE Publications Ltd.
- Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, 12(2), 219-245. <https://doi.org/10.1177/1077800405284363>
- Fosso-Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How «big data» can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, 234-246. <https://doi.org/10.1016/j.ijpe.2014.12.031>
- Frizzo-Barker, J., Chow-White, P. A., Mozafari, M., & Ha, D. (2016). An empirical study of the rise of big data in business scholarship. *International Journal of Information Management*, 36(3), 403-413. <https://doi.org/10.1016/j.ijinfomgt.2016.01.006>
- Gabel, T. J., & Tokarski, C. (2014). Big Data and Organizational Design: Key Challenges Await the Survey Research Firm. *Journal of Organization Design*, 3(1), 37-45. <https://doi.org/10.7146/jod.9753>
- Gagliardi, J., Renaud, J., & Ruiz, A. (2012). Models for automated storage and retrieval systems: a literature review. *International Journal of Production Research*, 50(24), 7110-7125. <https://doi.org/10.1080/00207543.2011.633234>
- Galbraith, J. R. (1974). Organization Design: An Information Processing View. *Interfaces*, 4(3), 28-36. <https://doi.org/10.1287/inte.4.3.28>
- Galbraith, J. R. (1995). *Designing Organizations: an executive briefing on strategy, structure, and process* (First). San Francisco, CA: Jossey-Bass Inc.
- Galbraith, J. R. (2012). The Future of Organization Design. *Journal of Organization*

- Design*, 1(1), 3-6. <https://doi.org/10.7146/jod.6332>
- Galbraith, J. R. (2014a). *Designing Organizations: Strategy, Structure, and Process at the business Unit and Enterprise Levels* (Tercera Ed). San Francisco, CA: Jossey-Bass Inc.
- Galbraith, J. R. (2014b). Organizational Design Challenges Resulting From Big Data. *Journal of Organization Design*, 3(1), 2. <https://doi.org/10.7146/jod.8856>
- Galbraith, J. R., Downey, D., & Kates, A. (2001). *Designing Dynamic Organizations: A Hands-On Guide for Leaders at All Levels* (Primera Ed). American Management Association.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- García, R. (2019). *Siemens Automatización, Foro Regional de Tecnologías Convergentes e industria 4.0: Barreras y Oportunidades para Colombia*. Barranquilla-Colombia.
- Garmaki, M., Boughzala, I., & Fosso-Wamba, S. (2016). The Effect of Big Data Analytics Capability on Firm Performance. *PACIS 2016 Proceedings*, 1-10. Recuperado de <https://aisel.aisnet.org/pacis2016/301/>
- Gartner Inc. (s. f.). Gartner IT Glossary: Big Data. Recuperado de <https://www.gartner.com/it-glossary/big-data>
- Gartner Inc. (2015). Gartner Says 6.4 Billion Connected «Things» Will Be in Use in 2016, Up 30 Percent From 2015. Recuperado de <https://www.gartner.com/newsroom/id/3165317>
- Gartner Inc. (2017). Gartner Survey Shows 42 Percent of CEOs Have Begun Digital Business Transformation. Recuperado de <https://www.gartner.com/newsroom/id/3689017>
- Gaviria-Marin, M., Merigó, J. M., & Baier-Fuentes, H. (2019). Knowledge management: A global examination based on bibliometric analysis. *Technological Forecasting and Social Change*, 140, 194-220. <https://doi.org/10.1016/j.techfore.2018.07.006>
- George, G., Haas, M. R., & Pentland, A. (2014). Big Data and Management. *Academy of Management Journal*, 57(2), 321-326. <https://doi.org/10.5465/amj.2014.4002>
- Gibson, C. B., Dunlop, P. D., & Cordery, J. L. (2019). Managing formalization to increase global team effectiveness and meaningfulness of work in multinational organizations. *Journal of International Business Studies*, 50(6), 1021-1052. <https://doi.org/10.1057/s41267-019-00226-8>

- Gil, A. J., & Carrillo, F. J. (2016). Knowledge transfer and the learning process in Spanish wineries. *Knowledge Management Research & Practice*, 14(1), 60-68. <https://doi.org/10.1057/kmrp.2014.12>
- Glazer, R. (1993). Measuring the value of information: The information-intensive organization. *IBM Systems Journal*, 32(1), 99-110. <https://doi.org/10.1147/sj.321.0099>
- Gold, A. H., Malhotra, A., & Segars, A. H. (2001). Knowledge Management: An Organizational Capabilities Perspective. *Journal of Management Information Systems*, 18(1), 185-214. <https://doi.org/10.1080/07421222.2001.11045669>
- González, M. R. (1994). *Implicaciones Estratégicas de los Sistemas y Tecnologías de la Información: El caso de la provincia alicantina*. Universidad de Alicante.
- Good, M., Knockaert, M., Soppe, B., & Wright, M. (2019). The technology transfer ecosystem in academia. An organizational design perspective. *Technovation*, 82, 35-50. <https://doi.org/10.1016/j.technovation.2018.06.009>
- Google. (2018). Google Trends en Big Data. Recuperado 20 de junio de 2018, de <https://trends.google.es/trends/explore?q=Big Data&geo=ES>
- Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17(S2), 109-122. <https://doi.org/10.1002/smj.4250171110>
- Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*, 70, 308-317. <https://doi.org/10.1016/j.jbusres.2016.08.004>
- Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information Systems*, 26(3), 191-209. <https://doi.org/10.1016/j.jsis.2017.07.003>
- Gupta, A. K., & Govindarajan, V. (2000). Knowledge flows within multinational corporations. *Strategic Management Journal*, 21(4), 473-496. [https://doi.org/10.1002/\(SICI\)1097-0266\(200004\)21:4<473::AID-SMJ84>3.0.CO;2-I](https://doi.org/10.1002/(SICI)1097-0266(200004)21:4<473::AID-SMJ84>3.0.CO;2-I)
- Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, 53(8), 1049-1064. <https://doi.org/10.1016/j.im.2016.07.004>
- Gurteen, D. (1998). Knowledge, Creativity and Innovation. *Journal of Knowledge*

- Management*, 2(1), 5-13. <https://doi.org/10.1108/13673279810800744>
- Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Ullah Khan, S. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, 47, 98-115. <https://doi.org/10.1016/j.is.2014.07.006>
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science In Information Systems Research. *MIS Quarterly*, 28(1), 75-105. <https://doi.org/10.2307/25148625>
- Hilton, R. W. (1981). The Determinants of Information Value: Synthesizing Some General Results. *Management Science*, 27(1), 57-64. <https://doi.org/10.1287/mnsc.27.1.57>
- Huber, G. P. (1991). Organizational Learning: The Contributing Processes and the Literatures. *Organization Science*, 2(1), 88-115. <https://doi.org/10.1287/orsc.2.1.88>
- Inkpen, A. C., & Tsang, E. W. K. (2005). Social Capital, Networks, and Knowledge Transfer. *Academy of Management Review*, 30(1), 146-165. <https://doi.org/10.5465/amr.2005.15281445>
- International Data Corporation. (2017). IDC Forecasts \$1.2 Trillion in Worldwide Spending on Digital Transformation Technologies in 2017. Recuperado de <https://www.idc.com/getdoc.jsp?containerId=prUS42327517>
- Intezari, A., & Gressel, S. (2017). Information and reformation in KM systems: big data and strategic decision-making. *Journal of Knowledge Management*, 21(1), 71-91. <https://doi.org/10.1108/JKM-07-2015-0293>
- Jagdish, H. V. (2015). Big Data and Science: Myths and Reality. *Big Data Research*, 2(2), 49-52. <https://doi.org/10.1016/j.bdr.2015.01.005>
- Janssen, M., Van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, 70, 338-345. <https://doi.org/10.1016/j.jbusres.2016.08.007>
- Jashapara, A. (2004). *Knowledge Management: An Integrated Approach* (Primera Ed). Edinburgh: Pearson Education.
- Jonker, C. M., Popova, V., Sharpanskykh, A., Treur, J., & Yolum, P. (2012). Formal framework to support organizational design. *Knowledge-Based Systems*, 31, 89-105. <https://doi.org/10.1016/j.knosys.2012.02.011>
- Kaivo-oja, J., Virtanen, P., Jalonen, H., & Stenvall, J. (2015). The Effects of the Internet of Things and Big Data to Organizations and Their Knowledge Management Practices. En *10th International Conference on Knowledge Management in Organizations: Lecture Notes in Business Information Processing* (pp. 495-513).

- Maribor, Slovenia: Springer International Publishing. https://doi.org/10.1007/978-3-319-21009-4_38
- Katal, A., Wazid, M., & Goudar, R. H. (2013). Big data: Issues, challenges, tools and Good practices. En *2013 Sixth International Conference on Contemporary Computing (IC3)* (pp. 404-409). Noida, India: IEEE. <https://doi.org/10.1109/IC3.2013.6612229>
- Kates, A., & Galbraith, J. R. (2007). *Designing Your Organization: Using the Star Model to Solve 5 Critical Design Challenges* (First). Jossey-Bass Inc.
- Khachlouf, N., & Quélin, B. V. (2018). Interfirm ties and knowledge transfer: The moderating role of absorptive capacity of managers. *Knowledge and Process Management*, 25(2), 97-107. <https://doi.org/10.1002/kpm.1564>
- Khan, N., Yaqoob, I., Hashem, I. A. T., Inayat, Z., Mahmoud Ali, W. K., Alam, M., ... Gani, A. (2014). Big Data: Survey, Technologies, Opportunities, and Challenges. *The Scientific World Journal*, 2014, 1-18. <https://doi.org/10.1155/2014/712826>
- Kshetri, N. (2014). The emerging role of Big Data in key development issues: Opportunities, challenges, and concerns. *Big Data & Society*, 1(2), 1-20. <https://doi.org/10.1177/2053951714564227>
- Lake, P., & Drake, R. (2014). *Information Systems Management in the Big Data Era* (Primera Ed). Sheffield, UK: Springer International Publishing.
- Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity and Variety. *META Group Research Note*.
- Lapiedra, R., & Devece, C. (2012). *Introduction to Management Information Systems*. Castellón: Universitat Jaume I.
- Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, 36(5), 700-710. <https://doi.org/10.1016/j.ijinfomgt.2016.04.013>
- Laudon, K. C., & Laudon, J. P. (2011). *Essentials of Management Information Systems* (9.^a ed.). Upper Saddle River, NJ: Pearson Education.
- Laudon, K. C., & Laudon, J. P. (2012). *Sistemas de Información Gerencial* (12.^a ed.). Pearson.
- Lee, I. (2017). Big data: Dimensions, evolution, impacts, and challenges. *Business Horizons*, 60(3), 293-303. <https://doi.org/10.1016/j.bushor.2017.01.004>
- Lee, I., & Lee, K. (2015). The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Business Horizons*, 58(4), 431-440.

- <https://doi.org/10.1016/j.bushor.2015.03.008>
- Lee, J., Kao, H.-A., & Yang, S. (2014). Service Innovation and Smart Analytics for Industry 4.0 and Big Data Environment. *Procedia CIRP*, 16, 3-8. <https://doi.org/10.1016/j.procir.2014.02.001>
- Leech, N. L., & Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs. *Quality & Quantity*, 43(2), 265-275. <https://doi.org/10.1007/s11135-007-9105-3>
- Lim, M. K., Tseng, M.-L., Tan, K. H., & Bui, T. D. (2017). Knowledge management in sustainable supply chain management: Improving performance through an interpretive structural modelling approach. *Journal of Cleaner Production*, 162, 806-816. <https://doi.org/10.1016/j.jclepro.2017.06.056>
- Linthicum, D. (2019). Edge computing vs. fog computing: Definitions and enterprise uses. Recuperado 3 de junio de 2019, de <https://www.cisco.com/c/en/us/solutions/enterprise-networks/edge-computing.html>
- Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems*, 24(3), 149-157. <https://doi.org/10.1016/j.jsis.2015.08.002>
- Lombardi, R. (2019). Knowledge transfer and organizational performance and business process: past, present and future researches. *Business Process Management Journal*, 25(1), 2-9. <https://doi.org/10.1108/BPMJ-02-2019-368>
- Lugmayr, A., Stockleben, B., Scheib, C., & Mailaparampil, M. A. (2017). Cognitive big data: survey and review on big data research and its implications. What is really “new” in big data? *Journal of Knowledge Management*, 21(1), 197-212. <https://doi.org/10.1108/JKM-07-2016-0307>
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute. Recuperado de [https://www.mckinsey.com/~media/McKinsey/Business Functions/McKinsey Digital/Our Insights/Big data The next frontier for innovation/MGI_big_data_full_report.ashx](https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_full_report.ashx)
- Maxwell, J. A. (2008). Designing a qualitative study. En *The SAGE handbook of applied social research methods* (Segunda Ed, pp. 214-253). California: Sage Publications.
- McAfee, A., & Brynjolfsson, E. (2012). *Big Data: The Management Revolution*.

- Harvard Business Review*, 90(10), 60-68.
- McLeod, A. J., Bliemel, M., & Jones, N. (2017). Examining the adoption of big data and analytics curriculum. *Business Process Management Journal*, 23(3), 506-517. <https://doi.org/10.1108/BPMJ-12-2015-0174>
- Miles, M. B., & Huberman, M. A. (1994). *Qualitative Data Analysis* (Segunda Ed). London: Sage Publications.
- Miles, M. B., Huberman, M. A., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (Tercera Ed). London: Sage Publications.
- Mintzberg, H. (1979). *The Structuring of Organizations* (Primera Ed). Englewood Cliffs, NJ: Prentice-Hall Inc.
- Mintzberg, H. (1993). *Structure in Fives: Designing Effective Organizations* (Primera Ed). Englewood Cliffs, NJ: Prentice-Hall Inc.
- Mishra, D., Luo, Z., Jiang, S., Papadopoulos, T., & Dubey, R. (2017). A bibliographic study on big data: concepts, trends and challenges. *Business Process Management Journal*, 23(3), 555-573. <https://doi.org/10.1108/BPMJ-10-2015-0149>
- Miterev, M., Turner, J. R., & Mancini, M. (2017). The organization design perspective on the project-based organization: a structured review. *International Journal of Managing Projects in Business*, 10(3), 527-549. <https://doi.org/10.1108/IJMPB-06-2016-0048>
- Moreno-Luzón, M. D., Peris, F., & González, T. (2001). *Gestión de la calidad y diseño de organizaciones: teoría y estudio de casos* (Primera Ed). Pearson Education.
- Murawski, M., & Bick, M. (2017). Digital competences of the workforce ? a research topic? *Business Process Management Journal*, 23(3), 721-734. <https://doi.org/10.1108/BPMJ-06-2016-0126>
- Murdoch, T. B., & Detsky, A. S. (2013). The Inevitable Application of Big Data to Health Care. *Journal of the American Medical Association*, 309(13), 1351-1352.
- Nee, A. Y. C., & Ong, S. K. (2013). Virtual and Augmented Reality Applications in Manufacturing. En *7th IFAC Conference on Manufacturing Modelling, Management and Control* (Vol. 46, pp. 15-26). Springer Science & Business Media. <https://doi.org/10.3182/20130619-3-RU-3018.00637>
- NIST Big Data Public Working Group. (2015). *NIST Special Publication 1500-1 - NIST Big Data Interoperability Framework: Volume 1, Definitions*. NIST Special Publication. Recuperado de <http://dx.doi.org/10.6028/NIST.SP.1500-1>
- Njie, B., & Asimiran, S. (2014). Case Study as a Choice in Qualitative Methodology.

- IOSR Journal of Research & Method in Education*, 4(3), 35-40.
<https://doi.org/10.9790/7388-04313540>
- Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science*, 5(1), 14-37. <https://doi.org/10.1287/orsc.5.1.14>
- Nonaka, I., Hirose, A., & Takeda, Y. (2016). ‘Meso’-Foundations of Dynamic Capabilities: Team-Level Synthesis and Distributed Leadership as the Source of Dynamic Creativity. *Global Strategy Journal*, 6(3), 168-182. <https://doi.org/10.1002/gsj.1125>
- Nonaka, I., & Konno, N. (1998). The Concept of “Ba”: Building a Foundation for Knowledge Creation. *California Management Review*, 40(3), 40-54. <https://doi.org/10.2307/41165942>
- Nonaka, I., & Peltokorpi, V. (2006). Objectivity and subjectivity in knowledge management: a review of 20 top articles. *Knowledge and Process Management*, 13(2), 73-82. <https://doi.org/10.1002/kpm.251>
- Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford: Oxford university press.
- Nonaka, I., & Takeuchi, H. (2011). The Big Idea: The Wise Leader. *Harvard Business Review*, 89(5), 58-67.
- Nonaka, I., & von Krogh, G. (2009). Perspective—Tacit Knowledge and Knowledge Conversion: Controversy and Advancement in Organizational Knowledge Creation Theory. *Organization Science*, 20(3), 635-652. <https://doi.org/10.1287/orsc.1080.0412>
- O’Connor, C., & Kelly, S. (2017). Facilitating knowledge management through filtered big data: SME competitiveness in an agri-food sector. *Journal of Knowledge Management*, 21(1), 156-179. <https://doi.org/10.1108/JKM-08-2016-0357>
- O’Leary, D. E. (2013). Artificial intelligence and big data. *IEEE Intelligent Systems*, 28(2), 96-99. <https://doi.org/10.1109/MIS.2013.39>
- Okhuysen, G. A., & Bechky, B. A. (2009). 10 Coordination in Organizations: An Integrative Perspective. *Academy of Management Annals*, 3(1), 463-502. <https://doi.org/10.5465/19416520903047533>
- Olshannikova, E., Ometov, A., Koucheryavy, Y., & Olsson, T. (2015). Visualizing Big Data with augmented and virtual reality: challenges and research agenda. *Journal of Big Data*, 2(1), 1-27. <https://doi.org/10.1186/s40537-015-0031-2>
- Orlikowski, W. J., & Barley, S. R. (2001). Technology and Institutions: What Can

- Research on Information Technology and Research on Organizations Learn from Each Other? *MIS Quarterly*, 25(2), 145-165. <https://doi.org/10.2307/3250927>
- Orlikowski, W. J., & Iacono, C. S. (2001). Research Commentary: Desperately Seeking the “IT” in IT Research—A Call to Theorizing the IT Artifact. *Information Systems Research*, 12(2), 121-134. <https://doi.org/10.1287/isre.12.2.121.9700>
- Pauleen, D. J., & Wang, W. Y. C. (2017). Does big data mean big knowledge? KM perspectives on big data and analytics. *Journal of Knowledge Management*, 21(1), 1-6. <https://doi.org/10.1108/JKM-08-2016-0339>
- Pearson, T., & Wegener, R. (2013). *Big data: The organizational challenge*. Bain & Company. Recuperado de http://www.bain.com/Images/BAIN_BRIEF_Big_Data_The_organizational_challenge.pdf
- Polanyi, M. (1967). *The Tacit Dimension*. Garden City, N.Y.: Anchor Books.
- Power, D. J. (2014). Using ‘Big Data’ for analytics and decision support. *Journal of Decision Systems*, 23(2), 222-228. <https://doi.org/10.1080/12460125.2014.888848>
- Press, G. (2013). A Very Short History Of Big Data. *Forbes*. Recuperado de <https://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/?sh=6fef997465a1>
- Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, 1(1), 51-59. <https://doi.org/10.1089/big.2013.1508>
- Rehman, M. H. U., Chang, V., Batool, A., & Wah, T. Y. (2016). Big data reduction framework for value creation in sustainable enterprises. *International Journal of Information Management*, 36(6), 917-928. <https://doi.org/10.1016/j.ijinfomgt.2016.05.013>
- Rezaei, A., Allameh, S. M., & Ansari, R. (2018). Impact of knowledge creation and organisational learning on organisational innovation: an empirical investigation. *International Journal of Business Innovation and Research*, 16(1), 117-133.
- Robbins, S. P. (1990). *Organization Theory: Structures, Designs, and Applications* (Tercera Ed). Englewood Cliffs, NJ: Prentice-Hall Inc.
- Robbins, S. P., & Coulter, M. A. (2018). *Management* (14.^a ed.). Pearson Education.
- Roblek, V., Meško, M., & Krapež, A. (2016). A Complex View of Industry 4.0. *SAGE Open*, 6(2), 1-11. <https://doi.org/10.1177/2158244016653987>
- Rodríguez, G., Gil, J., & García, E. (1996). *Metodología de la investigación cualitativa*.

Málaga: Ediciones Aljibe.

- Rothberg, H. N., & Erickson, G. S. (2017). Big data systems: knowledge transfer or intelligence insights? *Journal of Knowledge Management*, 21(1), 92-112. <https://doi.org/10.1108/JKM-07-2015-0300>
- Russom, P. (2011). *Big Data Analytics. TDWI Best Practices Report*. Recuperado de <http://tdwi.org/research/2011/09/best-practices-report-q4-big-data-analytics.aspx>
- Sabherwal, R., & Becerra-Fernandez, I. (2003). An Empirical Study of the Effect of Knowledge Management Processes at Individual, Group, and Organizational Levels. *Decision Sciences*, 34(2), 225-260. <https://doi.org/10.1111/1540-5915.02329>
- Saldaña, J. (2015). *The coding manual for qualitative researchers* (Tercera Ed). California: Sage Publications.
- Santoro, G., Vrontis, D., Thrassou, A., & Dezi, L. (2018). The Internet of Things: Building a knowledge management system for open innovation and knowledge management capacity. *Technological Forecasting and Social Change*, 136, 347-354. <https://doi.org/10.1016/j.techfore.2017.02.034>
- SAS Inc. (2013). Big Data, What it is and why it matters. Recuperado 18 de mayo de 2018, de https://www.sas.com/en_us/insights/big-data/what-is-big-data.html
- Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research Methods for Business Students* (Séptima Ed). New York: Pearson Education.
- Schildt, H. (2017). Big data and organizational design – the brave new world of algorithmic management and computer augmented transparency. *Innovation*, 19(1), 23-30. <https://doi.org/10.1080/14479338.2016.1252043>
- Schoenherr, T., & Speier-Pero, C. (2015). Data Science, Predictive Analytics, and Big Data in Supply Chain Management: Current State and Future Potential. *Journal of Business Logistics*, 36(1), 120-132. <https://doi.org/10.1111/jbl.12082>
- Scrum.org. (2019). What is SCRUM? Recuperado 15 de diciembre de 2019, de <https://www.scrum.org/resources/what-is-scrum>
- Shakir, M. (2002). The selection of case studies: Strategies and their applications to IS implementation cases studies. *Research Letters in the Information and Mathematical Sciences*, (3), 191-198. Recuperado de <http://hdl.handle.net/10179/4373>
- Sheng, J., Amankwah-Amoah, J., & Wang, X. (2017). A multidisciplinary perspective of big data in management research. *International Journal of Production Economics*, 191(June), 97-112. <https://doi.org/10.1016/j.ijpe.2017.06.006>

- Shujahat, M., Sousa, M. J., Hussain, S., Nawaz, F., Wang, M., & Umer, M. (2019). Translating the impact of knowledge management processes into knowledge-based innovation: The neglected and mediating role of knowledge-worker productivity. *Journal of Business Research*, 94, 442-450. <https://doi.org/10.1016/j.jbusres.2017.11.001>
- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263-286. <https://doi.org/10.1016/j.jbusres.2016.08.001>
- Slinger, G., & Morrison, R. (2014). Will Organization Design Be Affected By Big Data? *Journal of Organization Design*, 3(3), 17-26. <https://doi.org/10.7146/jod.9729>
- Song, M., Van Der Bij, H., & Weggeman, M. (2005). Determinants of the Level of Knowledge Application: A Knowledge-Based and Information-Processing Perspective*. *Journal of Product Innovation Management*, 22(5), 430-444. <https://doi.org/10.1111/j.1540-5885.2005.00139.x>
- Stair, R. M., & Reynolds, G. W. (2010). *Principles of Information Systems, A managerial approach* (Novena Edi). Boston, MA: Cengage Learning.
- Stake, R. E. (1995). *The art of case study research* (Primera Ed). California: Sage Publications.
- Sumbal, M. S., Tsui, E., & See-to, E. W. K. (2017). Interrelationship between big data and knowledge management: an exploratory study in the oil and gas sector. *Journal of Knowledge Management*, 21(1), 180-196. <https://doi.org/10.1108/JKM-07-2016-0262>
- Susha, I., Janssen, M., & Verhulst, S. (2017). Data collaboratives as “bazaars”? A review of coordination problems and mechanisms to match demand for data with supply. *Transforming Government: People, Process and Policy*, 11(1), 157-172. <https://doi.org/10.1108/TG-01-2017-0007>
- Swanborn, P. (2010). *Case Study Research: What, Why and How?* (Primera Ed). London: SAGE Publications Ltd. <https://doi.org/10.4135/9781526485168>
- Szulanski, G. (2000). The process of knowledge transfer: A diachronic analysis of stickiness. *Organizational behavior and human decision processes*, 82(1), 9-27.
- TechAmerica Foundation: Federal Big Data Commission. (2012). *Demystifying Big Data: A Practical Guide To Transforming The Business of Government*. Recuperado de https://bigdatawg.nist.gov/_uploadfiles/M0068_v1_3903747095.pdf
- Teddle, C., & Yu, F. (2007). Mixed Methods Sampling: A Typology With Examples.

- Journal of Mixed Methods Research*, 1(1), 77-100.
<https://doi.org/10.1177/1558689806292430>
- Teixeira, E. K., Oliveira, M., & Curado, C. M. M. (2018). Knowledge management process arrangements and their impact on innovation. *Business Information Review*, 35(1), 29-38. <https://doi.org/10.1177/0266382118757771>
- Tellis, W. (1997). Application of a Case Study Methodology. *The Qualitative Report*, 3(3), 1-19. <https://doi.org/10.46743/2160-3715/1997.2015>
- The Economist Intelligence Unit. (2012). *The Deciding Factor: Big Data & Decision Making*. Capgemini Consulting. Recuperado de http://www.capgemini.com/sites/default/files/resource/pdf/The_Deciding_Factor__Big_Data__Decision_Making.pdf
- Tian, X. (2017). Big data and knowledge management: a case of déjà vu or back to the future? *Journal of Knowledge Management*, 21(1), 113-131. <https://doi.org/10.1108/JKM-07-2015-0277>
- Tien, J. M. (2013). Big Data: Unleashing information. *Journal of Systems Science and Systems Engineering*, 22(2), 127-151. <https://doi.org/10.1007/s11518-013-5219-4>
- Tien, J. M. (2015). Internet of connected ServGoods: Considerations, consequences and concerns. *Journal of Systems Science and Systems Engineering*, 24(2), 130-167. <https://doi.org/10.1007/s11518-015-5273-1>
- Tsoukatos, E., Psimani-Voulgaris, F., Vassakis, K., & Lemonakis, C. (2017). The impact of R&D and Information Technology on innovation performance of Greek SMEs. *Global Business and Economics Review*, 19(5), 521-535. <https://doi.org/10.1504/GBER.2017.10004429>
- Tzortzaki, A. M., & Mihiotis, A. (2014). A Review of Knowledge Management Theory and Future Directions. *Knowledge and Process Management*, 21(1), 29-41. <https://doi.org/10.1002/kpm.1429>
- Uden, L., & He, W. (2017). How the Internet of Things can help knowledge management: a case study from the automotive domain. *Journal of Knowledge Management*, 21(1), 57-70. <https://doi.org/10.1108/JKM-07-2015-0291>
- Universidad del Norte. (2019). *Reporte Anual (Boletín Estadístico)*.
- Universidad del Norte. (2020). Página Web: Sobre Nosotros. Recuperado de <https://www.uninorte.edu.co/sobrenosotros>
- Van De Ven, A. H., Delbecq, A. L., & Koenig, R. (1976). Determinants of Coordination Modes within Organizations. *American Sociological Review*, 41(2), 322-338.

<https://doi.org/10.2307/2094477>

- Vaquero, L. M., & Rodero-Merino, L. (2014). Finding your Way in the Fog: Towards a Comprehensive Definition of Fog Computing. *ACM SIGCOMM Computer Communication Review*, 44(5), 27-32. <https://doi.org/10.1145/2677046.2677052>
- Vera-Baquero, A., Colomo-Palacios, R., Stantchev, V., & Molloy, O. (2015). Leveraging big-data for business process analytics. *The Learning Organization*, 22(4), 215-228. <https://doi.org/10.1108/TLO-05-2014-0023>
- Waller, M. A., & Fawcett, S. E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34(2), 77-84. <https://doi.org/10.1111/jbl.12010>
- Wang, H., Xu, Z., Fujita, H., & Liu, S. (2016). Towards felicitous decision making: An overview on challenges and trends of Big Data. *Information Sciences*, 367-368, 747-765. <https://doi.org/10.1016/j.ins.2016.07.007>
- Wang, S., Wan, J., Li, D., & Zhang, C. (2016). Implementing Smart Factory of Industrie 4.0: An Outlook. *International Journal of Distributed Sensor Networks*, 12(1). <https://doi.org/10.1155/2016/3159805>
- Wang, S., Wan, J., Zhang, D., Li, D., & Zhang, C. (2016). Towards smart factory for industry 4.0: A self-organized multi-agent system with big data based feedback and coordination. *Computer Networks*, 101, 158-168. <https://doi.org/10.1016/j.comnet.2015.12.017>
- Ward, J. S., & Barker, A. (2013). Undefined By Data: A Survey of Big Data Definitions. *arXiv preprint arXiv:1309.5821*, September, 2. <https://doi.org/10.48550/arXiv.1309.5821>
- Waterman, R. H., Peters, T. J., & Phillips, J. R. (1980). Structure is not organization. *Business Horizons*, 23(3), 14-26. [https://doi.org/10.1016/0007-6813\(80\)90027-0](https://doi.org/10.1016/0007-6813(80)90027-0)
- Weick, K. E. (1979). *The Social Psychology of Organizing* (Segunda Ed). Reading: Addison-Wesley.
- Yin, R. K. (2014). *Case Study Research: Design and Methods* (Quinta Ed). London: Sage Publications.
- Yu, C., Zhang, Z., Lin, C., & Wu, Y. (2017). Knowledge Creation Process and Sustainable Competitive Advantage: the Role of Technological Innovation Capabilities. *Sustainability*, 9(12), 2280. <https://doi.org/10.3390/su9122280>
- Zammuto, R. F., Griffith, T. L., Majchrzak, A., Dougherty, D. J., & Faraj, S. (2007). Information Technology and the Changing Fabric of Organization. *Organization*

Science, 18(5), 749-762. <https://doi.org/10.1287/orsc.1070.0307>

Zeng, J., & Glaister, K. W. (2018). Value creation from big data: Looking inside the black box. *Strategic Organization*, 16(2), 105-140. <https://doi.org/10.1177/1476127017697510>

Zhan, Y., Tan, K. H., Ji, G., Chung, L., & Tseng, M. (2017). A big data framework for facilitating product innovation processes. *Business Process Management Journal*, 23(3), 518-536. <https://doi.org/10.1108/BPMJ-11-2015-0157>

Zikopoulos, P., DeRoos, D., Parasuraman, D., Deutsch, T., Corrigan, D., & Giles, J. (2013). *Harness the power of big data: the IBM big data platform*. New York: McGraw Hill. Recuperado de ftp://public.dhe.ibm.com/software/pdf/at/SWP10/Harness_the_Power_of_Big_Data.pdf

Anexo A: Protocolo del Estudio de Casos

Objetivos de la investigación

En los últimos años, la recogida y el análisis de datos ha alcanzado niveles que no son manejables a través de los SI y las TI tradicionales, convirtiéndose en lo que hoy conocemos como la era del Big Data. Debido a la introducción de nuevos SI y TI para abordar los problemas de Big Data, éste se convirtió en uno de los principales contribuyentes para la toma de decisiones y la creación de conocimiento, ya que los datos de alta velocidad permiten analizar los datos y la información en tiempo real para la toma de decisiones.

Sin embargo, y a pesar de su creciente importancia para las organizaciones de hoy en día, son pocas las investigaciones en el área de la dirección de empresas que abordan los cambios organizativos que genera la implementación de los ecosistemas de Big Data. Por ende, esta investigación busca analizar la influencia entre la implementación de los ecosistemas de Big Data y factores organizativos importantes como lo son los procesos de gestión del conocimiento y los mecanismos de coordinación y control, relación que tiene origen en la revisión de la literatura pero que hasta el momento no ha sido planteada.

Así pues, acorde a las consideraciones antes expuestas nos planteamos, tal como aparece en la introducción de esta tesis doctoral, como pregunta de investigación básica lo siguiente:

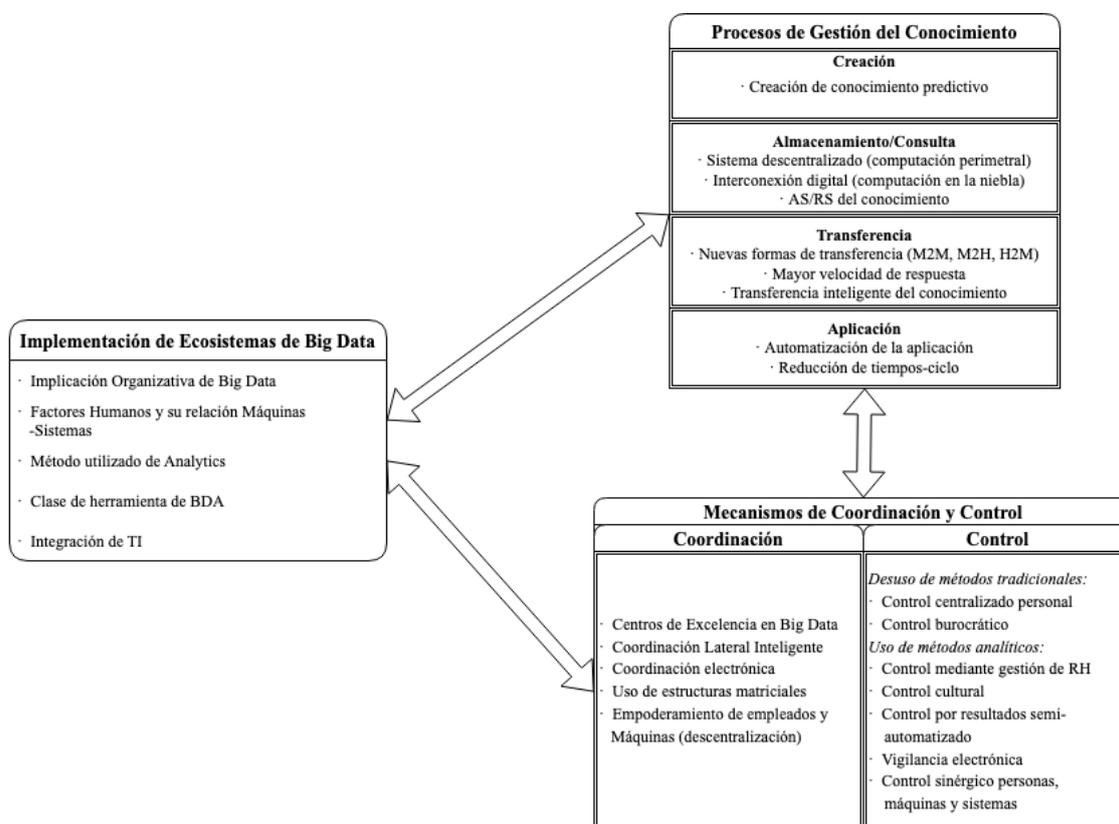
¿Cómo influye la implementación de ecosistemas de Big Data a los procesos de gestión del conocimiento y los mecanismos de coordinación y control en la organización?

Considerando la revisión bibliográfica utilizada para sustentar dicha relación (ver capítulos 1 al 3 de esta tesis doctoral), planteamos una serie de cuestiones específicas a continuación:

- ✓ ¿Cómo influye la implementación del ecosistema de Big Data a los procesos de gestión del conocimiento?
- ✓ ¿Cómo los procesos de gestión del conocimiento impulsan la implementación del ecosistema de Big Data en la organización?
- ✓ ¿Cómo influye la implementación de los ecosistemas de Big Data a los mecanismos de coordinación y control utilizados en la organización?
- ✓ ¿Cómo las nuevas formas de coordinar y controlar las actividades de la organización impulsan la implementación del ecosistema de Big Data?
- ✓ ¿Cómo los procesos de gestión del conocimiento impactan a los mecanismos de coordinación y control en presencia del ecosistema de Big Data?
- ✓ ¿Cómo los mecanismos de coordinación y control impulsan los procesos de gestión del conocimiento en presencia del ecosistema de Big Data?
- ✓ ¿Cómo otros factores organizativos influyen la relación entre la implementación del ecosistema de Big Data, los procesos de gestión del conocimiento y los mecanismos de coordinación?

Estas relaciones se pueden ilustrar mediante la siguiente gráfica (ver figura A-1 más abajo), el cual corresponde al modelo inicial de relaciones planteado para esta tesis doctoral:

Figura A-1: Modelo teórico de relaciones entre la implementación de ecosistemas de Big Data, procesos de gestión del conocimiento y mecanismos de coordinación y control



Fuente: elaboración propia

Por lo tanto, dada la característica novedosa de la presente investigación, se ha seleccionado la estrategia del estudio de casos al ser una investigación de carácter exploratorio. Así, en esta investigación no se plantean hipótesis o proposiciones a ser contrastadas sino elaborar preguntas abiertas para poder construir y estructurar las relaciones planteadas en el modelo teórico presentado anteriormente. Para conseguir lo anterior, se analizarán los casos seleccionados de forma exhaustiva para poder relacionar la implementación de los ecosistemas de Big Data con los procesos de gestión del conocimiento y los mecanismos de coordinación y control, además de otros factores que no hayan sido considerados inicialmente y que surjan de los casos estudiados. De esta manera, obtendremos al finalizar el análisis de los casos un modelo inducido que permitirá validar las relaciones planteadas inicialmente, así como proponer futuras líneas de investigación en el área.

Características de los casos a analizar

Para dar respuesta a la pregunta general de esta investigación, así como a las cuestiones específicas, se propone estudiar tres casos a profundidad. Las características que se han seleccionado en el marco de esta tesis doctoral son las siguientes:

- ✓ En primer lugar, necesitamos organizaciones que tengan un cierto nivel de madurez en la aplicación de ecosistemas de Big Data, ya que este es el factor impulsor cuyo efecto queremos analizar sobre los procesos de gestión del conocimiento y los mecanismos de coordinación y control de la organización.
- ✓ Debido a que el fenómeno bajo estudio implica que la organización debe poseer algún tipo de método, técnicas o herramientas avanzadas de análisis de datos para poder explorar su efecto en los procesos de gestión del conocimiento y los mecanismos de coordinación y control, consideraremos organizaciones que estén como mínimo en el nivel tres (i.e. Sistemas para la Toma de Decisiones y Especializados de Información Empresarial) de dicha propuesta plasmada en capítulo 1 (ver Tabla 1-5 en el capítulo 1).
- ✓ Las organizaciones a estudiar deben tener una estrategia y una cultura para el manejo del conocimiento en la organización, siendo deseable que la organización posea un proceso visible de gestión del conocimiento con objetivos y prácticas claramente definidas.
- ✓ Las organizaciones a estudiar deben ser empresas de un tamaño relativamente grande (medianas o grandes empresas). Este criterio está íntimamente relacionado con el anterior ya que en las micro y pequeñas empresas se da un pobre desarrollo en temas de sistemas y tecnologías de información, lo que nos lleva a establecer que necesitamos estudiar empresas medianas, grandes y/o multinacionales, donde sea visible su inversión en métodos, técnicas y herramientas para el análisis de datos y/o tecnologías asociadas a Big Data.

Procedimientos para la recogida y el análisis de información

Selección de los casos

Adicional a los criterios expuestos anteriormente, los criterios complementarios mencionados a continuación nos permitirán la selección de los candidatos para llevar a cabo el estudio de casos:

- ✓ Las organizaciones a analizar deben pertenecer a sectores de actividad diferentes, ya que eso nos permitirá introducir el criterio de la variedad de contextos en el fenómeno a estudiar.
- ✓ Desde el punto de vista de su ámbito demográfico, las organizaciones a estudiar deben estar ubicadas en Colombia dada la ubicación geográfica del investigador al momento de realizar el estudio de casos facilitando, de esta manera, el acceso del mismo a los datos primarios.
- ✓ El investigador necesita tener acceso al fenómeno objeto de estudio para lo cual es fundamental desarrollar una buena relación con los informantes, todo lo cual redundará en un mejor acceso a la información y, consecuentemente, en un mejor desarrollo del trabajo de campo.
- ✓ La organización debe poseer personal capacitado y/o desempeñarse en cargos asociados a las áreas de interés de la investigación.

Acceso a las organizaciones

Para acceder a las organizaciones establecimos contacto con personal de la empresa para facilitar el nombre y teléfono de las personas encargadas de la implementación del ecosistema de Big Data, los procesos de gestión del conocimiento y conocimiento en los mecanismos de coordinación y control de la empresa. Con estos datos, se contactó a dichas personas para presentar brevemente la investigación, así como comprobar con las personas seleccionadas el cumplimiento de los criterios de selección de los casos.

Una vez confirmada la idoneidad del caso, en una segunda llamada, se procedió con cada persona a presentar con mayor profundidad los objetivos de la investigación, para luego establecer la disponibilidad para llevar a cabo una entrevista y solicitar los permisos necesarios para acceder a la información que se requiere de la empresa.

Recogida de los datos e información

Se utilizará dos métodos básicos para la recogida de datos e información de los casos a estudiar: las entrevistas en profundidad y la documentación corporativa. Adicional a estas dos fuentes, llevaremos a cabo la observación directa del fenómeno para recolectar datos adicionales. Sin embargo, esta última fuente de información solo será de apoyo a las dos primeras, debido a que pueden existir limitaciones de acceso y tiempo para poder observar en detalle el fenómeno bajo estudio.

Entrevistas

Las entrevistas a realizar en el marco de la presente investigación serán de manera abierta y flexible debido al carácter exploratorio del mismo. La tabla A-1 muestra la planificación de las entrevistas a realizar. Sin embargo, durante el desarrollo de las entrevistas pueden surgir personas distintas a las planificadas que aporten a la investigación o puede darse el caso que las entrevistas tengan una duración superior a las aquí planteadas. Estas modificaciones imprevistas podrían modificar la estructura final de las entrevistas.

Tabla A-1: Planificación a priori de las entrevistas

Cargo a entrevistar	Número de entrevistas por persona	Duración prevista	Bloques del guion a tratar
Director/Jefe de Analytics, Tecnología o equivalente	1-2	Una hora y treinta minutos	1 a 5
Director/Jefe de Recursos Humanos o equivalente	1-2	Una hora y treinta minutos	1, 2, 4 y 5
Otras personas	Sin determinar a priori		

Fuente: elaboración propia

Tal y como se muestra en la tabla anterior, las entrevistas se desarrollarán en cinco bloques distintos (ver Anexo B), y dichos bloques se aplicarán en las entrevistas acorde a lo planteado en la planificación. No obstante, durante el desarrollo de las entrevistas se

podrán abordar todos los cinco bloques del guion, así como pueden surgir nuevas preguntas o bloques que sean consecuencia de la naturaleza abierta y flexible de la entrevista.

Documentación

Como mencionamos anteriormente, la segunda fuente de datos e información será la documentación corporativa. Esta documentación corresponde a información de la empresa que se asocie a los objetivos de la investigación y que contribuirá a la triangulación de la información. Entre la documentación corporativa a utilizar encontramos:

- ✓ Documentación elaborada por la empresa: procedimientos, manuales, reportes, página web, organigrama y documentos de interés para la investigación.
- ✓ Documentación no elaborada por la empresa: artículos, noticias o publicaciones en revistas, periódicos y páginas web externas a la empresa.

Para la documentación elaborada por la empresa se solicitará una copia de dicha documentación para su revisión en detalle. En caso que no se facilite la copia, se solicitará observar y analizar el documento dentro de las instalaciones de la empresa y se tomarán los apuntes de interés para la investigación.

Observación directa

La observación directa es nuestra fuente complementaria de datos e información, y para llevarla a cabo se solicitará acceso a áreas, personal y procesos administrativos relacionados con el objeto de estudio de la investigación y observar el fenómeno de interés. De esta forma, se pretende identificar aspectos y factores que no se evidencien de primera mano en las entrevistas o en la documentación corporativa, logrando así una mejor triangulación de los datos.

Anexo B: Guión de la Entrevista

Para la recolección de información a las organizaciones objeto de estudio de esta investigación, hemos dividido la entrevista a realizar en siete bloques en donde se discutirán los distintos aspectos de interés que deseamos que los entrevistados aporten en nuestra investigación, teniendo un carácter amplio y flexible para poder confirmar o enriquecer el modelo inicial planteado. No obstante, durante la realización de las entrevistas pueden surgir preguntas o cuestiones adicionales en base a las particularidades de cada organización, modificándose así la entrevista inicial aquí planteada. A continuación, presentamos los siete bloques que componen la entrevista.

Bloque 1: Información general de la empresa y del entrevistado

- ✓ Nombre de la organización y razón social
- ✓ Sector de actividad
- ✓ Número de empleados
- ✓ Posición en el mercado de acuerdo al sector
- ✓ Estructura organizativa y organigrama
- ✓ Nivel de formación, experiencia en el campo y posición en la organización del entrevistado.

Bloque 2: Proceso de gestión del conocimiento en la organización

- ✓ Políticas de gestión del conocimiento en la organización
- ✓ Cultura del conocimiento en la organización
- ✓ Desarrollo y evolución del proceso de gestión del conocimiento
- ✓ Sistemas y tecnologías de apoyo al proceso de gestión del conocimiento

Bloque 3: Mecanismos de coordinación en la organización

- ✓ Centralización/descentralización para la toma de decisiones
- ✓ Coordinación del trabajo o actividades de la organización
- ✓ Formación de grupos interdisciplinarios en la organización
- ✓ Control de resultados y objetivos de la organización
- ✓ Sistemas y tecnologías de apoyo a la coordinación del trabajo y al control de resultados, personas y máquinas/sistemas

Bloque 4: Factores de adopción de ecosistemas de Big Data en la organización

- ✓ Importancia del análisis de datos en la organización
- ✓ Estrategia para el uso de Big data en la organización
- ✓ Inversión de la organización en BDA u otras técnicas de análisis de datos
- ✓ Disposición de la organización en adoptar el Big Data para la toma de decisiones basadas en datos y conocimiento
- ✓ Asimilación del Big Data en las personas de la organización
- ✓ Métodos y herramientas utilizados para el análisis de datos
- ✓ Integración de tecnologías asociadas a Big Data en sus procesos

Bloque 5: Impacto de ecosistemas de Big Data en los procesos de gestión del conocimiento

- ✓ Impacto del ecosistema de Big Data en los procesos de gestión del conocimiento
- ✓ Nivel de automatización de los procesos de gestión del conocimiento
- ✓ Creación de grupos sobre Big Data para el conocimiento
- ✓ Apoyo de la organización para el uso de Big Data en los procesos de gestión del conocimiento

Bloque 6: Impacto de ecosistemas de Big Data en los mecanismos de coordinación y control

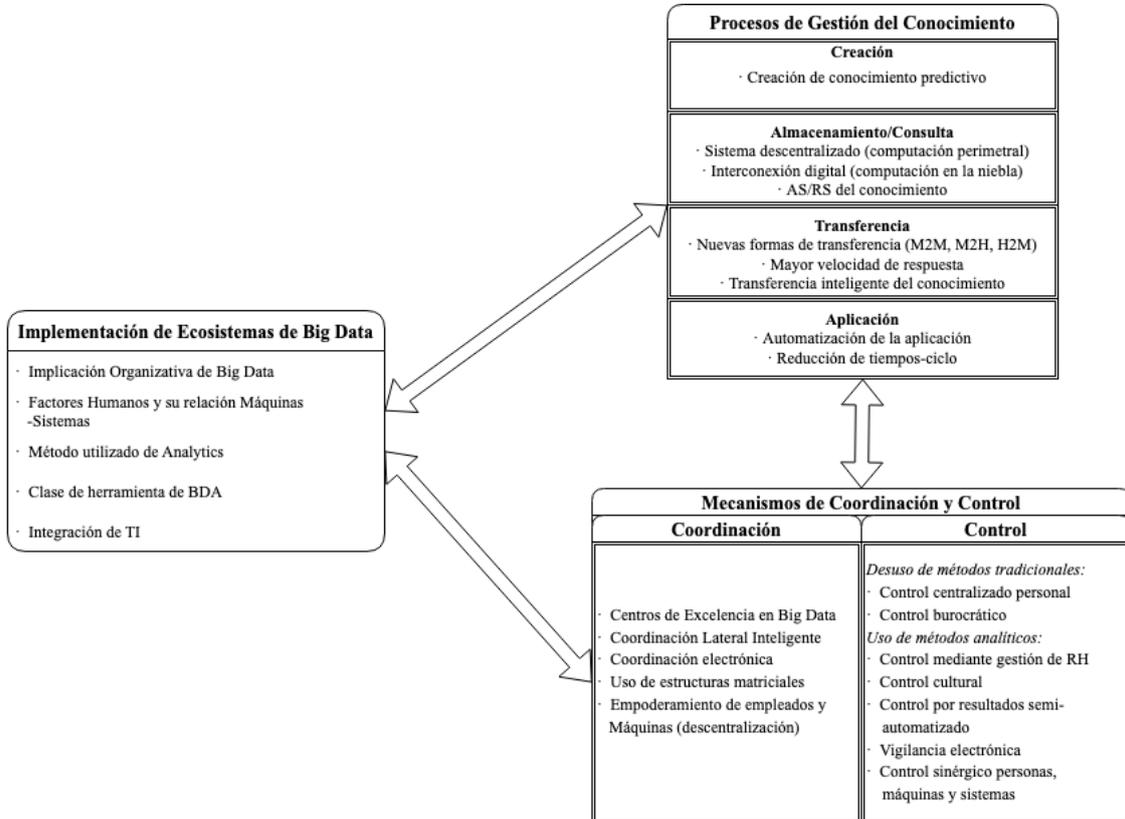
- ✓ Coordinación lateral para proyectos de Big Data
- ✓ Nivel de automatización para la medición de resultados
- ✓ Nivel de automatización para el control de recursos de la organización
- ✓ Apoyo de la organización para el uso de Big Data en la coordinación y control de las actividades y resultados

Bloque 7: Relación entre la implementación de ecosistemas de Big Data, procesos de gestión del conocimiento y los mecanismos de coordinación y control

- ✓ Factores del proceso de gestión del conocimiento que inciden sobre la implementación de ecosistemas de Big Data
- ✓ Factores del proceso de gestión del conocimiento que inciden sobre los mecanismos de coordinación y control

- ✓ Factores de los mecanismos de coordinación y control que inciden sobre el proceso de gestión del conocimiento
- ✓ Factores de los mecanismos de coordinación y control que inciden sobre la implementación de ecosistemas de Big Data

Modelo de relaciones entre la implementación de ecosistemas de Big Data, procesos de gestión del conocimiento y mecanismos de coordinación y control.



Fuente: elaboración propia.

Anexo C: Modelo de Solicitud Formal de Entrevista

(Ciudad), (día) de (mes) de 20__

Estimado (participante),

El profesor Francisco Balbastre Benavent y mi persona, el doctorando Erik Daniel Maldonado Ascanio, ambos pertenecientes al programa de Doctorado en Dirección de Empresas de la Universidad de Valencia en España, nos encontramos realizando una investigación que tiene como objetivo principal estudiar el efecto de la implementación de los ecosistemas de Big Data sobre los procesos de gestión del conocimiento y los mecanismos de coordinación y control.

Dado que la empresa en la que se encuentra laborando reúne los criterios para ser seleccionado como caso de estudio en nuestra investigación, le solicitamos formalmente concedernos una entrevista para abordar las temáticas de interés, estimamos que la entrevista tendrá una duración aproximada de hora y media. Previo a la entrevista, le entregaremos un guion para que esté enterado de los puntos a tratar, entrevista que será de carácter abierto y flexible.

Una vez se haya finalizado la investigación, le remitiremos un informe que desvelará cómo se presenta la relación entre la implementación de los ecosistemas de Big Data sobre los procesos de gestión del conocimiento, los mecanismos de coordinación y control y factores particulares que han surgido en la organización, así como la comparación con otros casos realizados en la investigación. Estamos a su disposición si requiere de información adicional o si se genera alguna duda sobre la entrevista o sobre los objetivos de la misma. Para tal fin puede contactarnos mediante los medios presentados al final de este correo. Agradecemos su valiosa colaboración, cordial saludo.

Erik Daniel Maldonado Ascanio

Doctorando e investigador

Email: emalas@alumni.uv.es

Teléfono: +57 3008027819

Anexo D: Modelo de Consentimiento Informado del Participante

(Ciudad), (día) de (mes) de 20__

Estimado (participante),

Por medio de la presente carta, el profesor Francisco Balbastre Benavent y mi persona, el doctorando Erik Daniel Maldonado Ascanio, ambos pertenecientes al programa de Doctorado en Dirección de Empresas de la Universidad de Valencia en España, le informamos que su participación en la presente investigación es libre y estrictamente voluntaria. Las respuestas brindadas en el marco de la investigación serán de carácter confidencial y serán codificadas para mantener la privacidad. Los audios grabados durante la(s) entrevista(s) serán eliminados.

Teniendo en cuenta la anterior información y lo expresado en el(los) correo(s) y llamada(s) previa(s), mediante la firma de este documento acepta participar de manera voluntaria en la investigación, y reconoce que ha sido informado del alcance y objetivos del mismo y comprendo que cualquier dato o información provista será de carácter confidencial y que no será utilizado para fines distintos al académico.

Además, reconoce que ha sido informado sobre la opción de desistir de su participación en cualquier momento, sin que esto afecte a su persona y que, ante cualquier duda, aclaración o solicitud de resultados parciales o finales, tiene el derecho a requerir respuesta por parte de uno de los investigadores. Así mismo, manifiesta que se encuentra satisfecho con la información recibida sobre el proceso de investigación a realizar, y consiente que su participación es libre y voluntaria.

Firma del participante

Nombre: