FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

DOBLE GRADO EN ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS Y DERECHO

Implicaciones éticas y legales en el uso del big data

Trabajo Fin de Grado presentado por Alejandro Durán Domínguez, siendo el tutor del mismo el profesor D. José Luis Roldán Salgueiro.

Vº. Bº. del tutor: Alumno:

D. José Luis Roldán Salgueiro D. Alejandro Durán Domínguez

Sevilla, Noviembre de 2019

FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

DOBLE GRADO EN ADMINISTRACIÓN DE EMPRESAS Y DERECHO

TRABAJO FIN DE GRADO

CURSO ACADÉMICO [2018-2019]

TÍTULO: **Implicaciones éticas y legales del uso del big data.**

AUTOR: **ALEJANDRO DURÁN DOMÍNGUEZ**

TUTOR: **JOSÉ LUIS ROLDÁN SALGUEIRO**

DEPARTAMENTO: **Administración de Empresas e Investigación de Mercados**

ÁREA DE CONOCIMIENTO: **Organización de empresas**

RESUMEN:

El auge de la tecnología *big data* en los últimos años junto con la aparición del internet de las cosas ha provocado un aumento de la creación de datos de todo tipo y a su vez un aumento de interés por parte de la sociedad sobre el tratamiento que reciben sus datos personales. Este trabajo trata sobre los aspectos éticos que se derivan del uso de la tecnología *big data* y cómo está regulado el tratamiento que reciben los datos usados durante el proceso. Además se estudia cómo pueden organizarse y establecerse las responsabilidades en las empresas ante esta innovación para evitar posibles vulneraciones y pérdidas de reputación. Por último, se analiza la percepción que tienen los profesionales del sector sobre el cumplimiento de los aspectos éticos y legales que analizamos en este trabajo.

PALABRAS CLAVE:

Big data; governance; ética; datos; privacidad.

ÍNDICE

PARTE I: MARCO TEÓRICO

1. PREÁMBULO

1.1 Justificación del tema y objetivos

El título de este trabajo es "Implicaciones éticas y legales en el uso del *big data"*. He seleccionado este tema debido a la gran importancia que ha adquirido en los últimos años el uso de la tecnología *big data*.

A pesar de que su uso ha aumentado en los últimos años, actualmente no hay una palabra en el diccionario español que defina este concepto. Además, como ocurre en los últimos años con muchos conceptos novedosos o tecnológicos de procedencia anglosajona, el castellano va adoptando muchos anglicismos y este es uno de esos casos. La Real Academia Española aún no se ha decantado por incluir el término *big data* en el diccionario, aunque *big data* es la expresión más usada junto a macrodatos e inteligencia de datos.

Aunque la inteligencia de datos es un tema principalmente de ámbito tecnológico, en este trabajo nos hemos centrado en las consecuencias éticas y legales debido a que el fenómeno *big data* tiene multitud de aplicaciones en el ámbito de las ciencias sociales.

Con este trabajo se busca entender qué es el fenómeno *big data* y por qué es tan importante, sus utilidades y desventajas y las consecuencias éticas y legales que se derivan de su uso. Con ello queremos dar a conocer a las personas el concepto de la inteligencia de datos y los aspectos lo rodean además de dar respuesta a las implicaciones éticas y legales del uso del *big data* en España.

La finalidad que tiene este trabajo es para que las personas ajenas a este fenómeno conozcan un poco más de él y de lo que ocurre con todos los datos que ofrecemos o publicamos en internet. También me gustaría que este trabajo sirva para que, aquellas personas que son profesionales de los macrodatos, inicien un debate acerca de los dilemas éticos y legales que vamos a plantear en este trabajo, dar respuesta a las inquietudes que pueden surgir de esta innovación tecnológica y mostrar que la regulación de los macrodatos no es del todo adecuada o completa como debería ser.

Otra finalidad de este trabajo es dar a conocer el término de *governance* o gobernanza respecto de los datos y plantear cómo se debe gestionar en una empresa el cumplimiento de los estándares éticos y legales cuando se usan los macrodatos.

1.2 Metodología y fuentes de información

El método de investigación seleccionado es el de la revisión bibliográfica para la primera parte y la investigación descriptiva cuantitativa para la segunda parte. La primera parte se corresponde con el marco teórico y la segunda es la investigación empírica.

Inicialmente, se ha realizado una investigación documental para recopilar información sobre los macrodatos y los dilemas éticos que se pueden llegar a plantear debido a su uso y sus consecuencias. También se ha recopilado información sobre el marco regulatorio que existe sobre la materia.

Los documentos y libros consultados han sido muchos, pero debido a la novedad del fenómeno *big data*, nos hemos encontrado con una mayor parte de artículos en revistas científicas ya que como hemos dicho anteriormente se trata de un tema novedoso. Además, el hecho de que se trate de un tema muy específico, las consecuencias éticas y legales del uso del *big data*, hace que resulte más fácil encontrar artículos de revistas que libros dedicados a este tema tan característico.

La búsqueda de la información se ha realizado especialmente en el catálogo de la biblioteca de la Universidad de Sevilla (https://fama.us.es), el portal de información de revistas publicadas en castellano Dialnet (https://dialnet.unirioja.es), las bases de datos de ABI/INFORM y Scopus y por último el buscador de Google Académico (https://scholar.google.es).

1.3 Estructura del trabajo

Este trabajo se divide en dos partes diferenciadas. La primera parte es el marco teórico y está compuesto por los cinco primeros epígrafes. La segunda parte es la investigación empírica y se compone de un epígrafe más las conclusiones finales del trabajo.

El primer epígrafe es el preámbulo en el que nos encontramos donde explico los motivos y objetivos de este trabajo, la metodología empleada y la estructura.

El segundo epígrafe se centra en hacer una introducción al *big data*, empezando por acercarnos al concepto y su definición para, a continuación, hablar de los desafíos que conlleva realizar un proyecto utilizando la tecnología *big data*, destacando en último lugar los desafíos éticos y legales.

El tercer epígrafe entra a analizar aquellas consecuencias éticas que conlleva la adopción de proyectos *big data* y la posterior responsabilidad que se puede alcanzar debido a una mala implantación.

El cuarto epígrafe analiza el marco regulatorio del fenómeno *big data* en España, empezando por un breve resumen histórico del tratamiento legal de los datos hasta analizar la actual normativa española.

El quinto epígrafe trata de cómo se debe gestionar en el ámbito empresarial el uso de la inteligencia de los datos y cómo cumplir con los marcos legales y éticos.

El sexto epígrafe corresponde a la segunda parte del trabajo y se centra en la investigación empírica. Se basa en analizar el resultado de encuestas realizadas a profesionales que usan la inteligencia de datos para conocer su percepción sobre los aspectos éticos y legales que conlleva el uso de los macrodatos.

Y, por último, el séptimo epígrafe corresponde a las conclusiones personales a las que he llegado tras elaborar este trabajo de fin de grado sobre las implicaciones éticas y legales del fenómeno *big data* y todo lo que conlleva su uso sobre las personas.

2. INTRODUCCIÓN AL BIG DATA

2.1 Aspectos generales

Hoy en día hay muchas empresas, entidades y administraciones públicas que recogen datos de todo tipo. Estos datos de carácter personal son obtenidos de múltiples fuentes de recursos y recogen información sobre lo que haces, dónde vas, gustos, preferencias y muchos datos más.

En los últimos años, el rápido desarrollo de internet y su uso extensivo, el internet de las cosas, la computación en la nube o *cloud computing* y la aparición de nuevas tecnologías ha conllevado a un crecimiento exponencial de los datos en casi todas las industrias y áreas de la economía (Jin, Wah, Cheng y Wang, 2015).

Estos millones de datos provienen de múltiples dispositivos tales como teléfonos móviles, GPS, electrodomésticos, contadores de luz, agua, pulseras deportivas, etc. Estos datos son

almacenados y analizados para mejorar la sociedad a través de innovaciones introducidas mediante el uso de los macrodatos (Monleon-Getino, 2015).

Las empresas y los gobiernos realizan muchas actividades, entre ellas, la recopilación y análisis de datos y la aparición del *big data* ha supuesto un mayor desafío para éstas debido a su complejidad. A su vez también supone un beneficio porque les permite analizar una mayor cantidad de datos y obtener mejores resultados (Dinh, Karmakar, Kamruzzaman y Stranieri, 2018).

La inteligencia de datos o *big data* es importante ya que gracias a él se ha ido cambiando y transformando la forma en la que vivimos y trabajamos (Mayer-Schönberger y Cukier, 2013). Un uso adecuado de él puede suponer una ventaja competitiva para una empresa ya que mediante el análisis de los datos puede encontrar al mejor segmento de cliente al que dirigir su producto, puede mejorar el servicio que realice o reducir sus costes debido a un mejor aprovechamiento de sus recursos.

Además de ser una innovación tecnológica, los macrodatos se usan en sectores muy diferentes de la economía y en empresas de todo tipo como por ejemplo el sector bancario y financiero, empresas del sector salud o empresas tecnológicas como Google, Apple o Microsoft. También lo usan empresas distribuidoras de todo tipo como Amazon (comercio electrónico), Inditex (textil) o Mercadona (alimentación) (Cotino, 2017).

Asimismo, se puede usar en el sector primario para mejorar las cosechas, en el sector secundario, lo pueden usar en casi cualquier tipo de industria para ser más eficiente y en el sector terciario su principal aplicación es en el turismo, aunque también se usa en compañías de seguros y hasta en los deportes.

Ya no sólo hablamos de la importancia del *big data* en el ámbito empresarial, sino la importancia que tiene para la sociedad. Puede mejorar la regulación del tráfico de las ciudades suponiendo un ahorro de tiempo y energético. También sirve para identificar áreas turísticas en las ciudades según la afluencia de personas a los monumentos de cada ciudad. También puede servir para mejorar servicios públicos como el suministro de agua, limpieza o transporte urbano. Por todo esto, el uso de la inteligencia de datos tiene mucho interés económico y académico (Galimany, 2014).

2.2 Concepto de *big data*

Inicialmente vamos a definir qué son los macrodatos o *big data.* Según McKinsey, (citado en Sun, Strang y Li, 2014) define al *big data* como "los conjuntos de datos cuyo tamaño supera la capacidad de las herramientas típicas de software de bases de datos para capturar, almacenar, gestionar y analizar" (p. 56).

Según Gartner, (citado en Sun et al., 2018), los macrodatos son "activos de información de alto volumen, alta velocidad y gran variedad que exigen formas rentables e innovadoras de procesamiento de la información para mejorar el conocimiento y la toma de decisiones" (p. 56).

Se puede definir al fenómeno *big data* a partir de su traducción literal que significa grandes datos o macrodatos. Según Galimany (2014) se puede definir a los macrodatos como la capacidad para capturar, agregar y procesar grandes cantidades de datos.

Lo que queda claro, es que actualmente no hay una definición clara sobre el concepto de *big data* (Jin et al., 2015). Si bien una de las definiciones que más se usan es la que aporta Gartner. Para entender mejor el concepto de *big data* debemos analizar sus características.

Las características de los macrodatos nos sirven para entender mejor cómo funciona. En esto también hay divergencias entre los expertos ya que hay algunos autores que señalan que los macrodatos se caracterizan por tres uves (Watson, 2014). En IBM (2012) se determina que existen cuatro uves y hay autores como Colmenarejo (2018) que señalan que son cinco uves.

La gran mayoría de autores coinciden en que los macrodatos se caracterizan, al menos, por las tres uves que se definen en Watson (2014): volumen velocidad y variedad.

El volumen se refiere a que, al día se genera una gran cantidad de datos y ese número no para de crecer exponencialmente. Es tal la cantidad de datos que hasta hace unos años no se pasó de los gigabytes a terabytes, luego se pasaron a los *petabytes* y actualmente a los *zettabytes* (1 *zettabytes* = más de 1 billón de gigas). Esto se debe a que como cada vez se crean y se almacenan una mayor cantidad de datos, se crean nuevas unidades de almacenamiento de información (Galimany, 2014).

La segunda uve que vamos a analizar es la velocidad. Internet, la aparición de los smartphones, los dispositivos GPS y los múltiples dispositivos como pulseras y robots de todo tipo han provocado que se recojan y analicen grandes cantidades de datos simultáneamente y con una frecuencia muy alta. Un ejemplo de la velocidad es cuando Google, a medida que

realizamos una búsqueda va prediciendo las palabras para completar la búsqueda. La localización de un coche que se va actualizando cada pocos segundos mediante GPS es otro ejemplo de la velocidad a la que viajan los datos en la actualidad (Galimany, 2014).

La tercera uve, variedad, se refiere tanto de los lugares de dónde se obtiene la información como de los métodos que se utilizan ya que antes del fenómeno *big data*, el análisis de los datos solía ser de datos estructurados y almacenados en bases de datos relacionadas. Con la aparición de la inteligencia de datos, esto cambia radicalmente y se pueden combinar datos no estructurados de diversas fuentes y con distintos formatos (texto, imágenes, vídeos y audios), lo que provoca que las capacidades de análisis crezcan exponencialmente (Galimany, 2014).

Respecto al resto de uves que mencionan otros autores, actualmente también es bastante aceptada la de veracidad que se refiere a la validez y a la integridad de los datos recogidos. La veracidad es el procedimiento que se dan a los datos no estructurados al combinarlos para el análisis (Watson, 2014). Este procedimiento es complejo debido a la dificultad de discernir entre datos fiables y no fiables, es por eso que se utilizan métodos de limpieza de datos (Jin et al., 2015).

La quinta uve es el valor, esto quiere decir que muchos de estos datos recogidos y almacenados, por sí solos carecen de valor, pero cuando son combinados y analizados cobran valor (Colmenarejo, 2018).

"Los datos no estructurados son recogidos sin un fin específico o distinto al que se aplica en el *big data*. Estos datos son almacenados y agregados a otros con la esperanza de que sean de aplicación" (Nersessian, 2018, p. 846). Pueden ser desde fotografías, audios, mensajes instantáneos, artículos, etc. Cada clic del ratón o el movimiento del cursor en una página web puede ser objeto de recogida, almacenamiento y análisis de *big data* (Monleon-Getino, 2015).

La inteligencia de datos, según Nersessian (2018), no consiste solo en recoger datos de todo tipo y almacenarlos, como hemos dicho, se precisa del análisis. El análisis en la inteligencia de datos es un procedimiento metodológico con técnicas de investigación específicas que analizan los datos con el fin de obtener conocimientos únicos y mejorar la toma de decisiones en un contexto dado (empresarial, científico, político, legal, etc).

Según Hoffman (2018) "el *big data* también puede ayudar en la cura del cáncer, identificar terroristas o proporcionar servicios especializados. También puede llegar a predecir pautas de comportamiento que pueden ser explotadas por las empresas para su beneficio" (p. 8).

Los macrodatos son usados para garantizar la seguridad pública y combatir delitos a la hora de recopilar datos y analizar a sospechosos. En Galimany (2014), se determina que hay comisarías que calculan e identifican en qué lugares y a qué horas es más probable que ocurran hechos delictivos para así establecer una mejor ruta de patrullas y reducir la tasa de criminalidad de las ciudades.

También es usado por científicos para analizar múltiples datos y variables y así predecir o prevenir posibles patrones de enfermedades como la malaria, el dengue o simples brotes de gripe (Nersessian, 2018). Es decir, con un análisis predictivo de ciertos datos en Internet (como pueden ser las búsquedas en Google sobre los síntomas de la gripe), se pueden tomar medidas previas para mejorar la salud pública. (Jin et al., 2015).

Asimismo, se usa la inteligencia de datos para localizar e identificar a los ciudadanos en caso de aglomeraciones y seguir sus pautas de comportamiento mediante su geolocalización por GPS (Goodman, 2014).

En el deporte, el fenómeno *big data,* tiene múltiples usos como puede ser la recopilación de datos para conocer las carencias de un equipo o un deportista en concreto y poder entrenarlas para mejorar, la realización de informes para la prevención de lesiones o mejorar el *ticketing* y los ingresos generados por la venta de entradas y la realización de inversiones o fichajes. La liga de béisbol americana fue la primera en implantar el uso del *big data* en el deporte, pero cada vez más son más deportes los que lo usan como el baloncesto o el fútbol.

Con la inteligencia de datos, las empresas pueden generar ingresos además de clasificar y medir los comportamientos de las personas y sus clientes. Igualmente, venden esos análisis o datos para que sean reutilizados por terceros y así obtener una mayor cantidad de ingresos (Someh, Davern, Breidbach, y Shanks, 2019).

2.3 Desafíos de la adopción del *big data*

Como hemos visto anteriormente, el fenómeno *big data* tiene múltiples aplicaciones tanto para la sociedad como el mundo empresarial. "El *big data* mejora la excelencia operativa, crea una mejor comprensión de las relaciones con los clientes, mejora la gestión de riesgos e impulsan la innovación del modelo de negocio" (Buytendijk y Oestreich, 2015, p. 3).

A partir del auge de la tecnología y del uso de la inteligencia de datos, se ha empezado a reconocer a los datos como un activo más dentro del mundo empresarial. Los datos son un

pilar fundamental de la economía moderna y uno de "los recursos más importantes para mejorar la productividad, desarrollar nuevas tecnologías y mejorar los procesos de información" (Trom y Cronje, 2019, p. 648).

Aun así, los macrodatos, como cualquier innovación, es un arma de doble filo. Puede traer muchos beneficios como la creación de nuevos modelos de negocio o impulsar nuevos servicios o mitigar riesgos y costes empresariales. Pero a su vez también conlleva una serie de riesgos que vamos a ver a continuación (Buytendijk y Heiser, 2013).

Cualquier proyecto de *big data* tiene incluidos una serie de desafíos y oportunidades los cuales vamos a ir desgranando desde una perspectiva más general a una perspectiva cada vez más específica. Estos desafíos van desde la complejidad de llevar a cabo un proyecto de *big data* porque requiere amplios conocimientos y sistemas de procesamiento adecuados hasta la necesidad de una regulación más específica.

Las empresas que deciden usar la inteligencia de datos se enfrentan a desafíos como "la gestión, el procesamiento y la seguridad. Además, tienen que solventar los problemas que surjan de la captura, análisis, almacenamiento, búsqueda, intercambio, visualización, transferencia y violación de la privacidad" (Al-Badi, Tarhini y Khan, 2018, p. 2).

El primer desafío o reto al que se enfrenta una empresa cuando decide usar los macrodatos es la propia dificultad de implantar un proyecto de *big data,* ya que es un proceso complejo, requiere de personal muy cualificado en medios de computación, matemáticas y estadísticas. Además precisa de sistemas de almacenamientos con mucha capacidad para almacenar todos los datos para su posterior análisis (Jin et al., 2015).

Esta complejidad de los macrodatos se debe también a los tipos de datos con los que se trabaja (datos no estructurados), su gran volumen y su dificultad a la hora de analizar los datos. Es decir, lo que caracteriza a los macrodatos, como hemos visto en el epígrafe anterior con las cinco uves, es lo que provoca también que sea un proceso complicado.

Además de la complejidad de recopilar datos por ser datos no estructurados hay que sumar la complejidad de buscar datos de calidad ya que no sirve cualquier tipo de datos. Unos datos de gran calidad son cruciales para un mayor éxito a la hora de utilizar correctamente la inteligencia de datos (Trom y Cronje, 2019).

Otro desafío que hay que afrontar en el uso de la inteligencia de datos es que es una materia novedosa, es decir, es una innovación tecnológica que ha provocado muchos cambios en la sociedad y como es algo novedoso, hasta los últimos años no había muchos estudios ni

tampoco existe una regulación muy extensa. Por tanto las consecuencias éticas de un mal aprovechamiento de los macrodatos no están claramente delimitadas (Saltz y Dewar, 2019).

El tercer desafío al que se enfrenta una organización en el uso de la inteligencia de datos es el que se deriva de la seguridad de los datos y cómo protegerlos de ataques externos para evitar una vulneración de la privacidad de los titulares de dichos datos (Trom y Cronje, 2019).

Aunque el principal reto que plantea la inteligencia de datos es la repercusión ética, ya que el hecho de que sea algo novedoso, provoca que no exista un acuerdo generalizado sobre qué es ético o no (Asadi, Breidbach, Davern, y Shanks, 2016).

Es por ello que el motivo de este trabajo es ir desgranando todos aquellos retos éticos que se plantean del uso de la inteligencia de datos. Por ejemplo, se debe determinar los conceptos éticos en materia de macrodatos (Asadi et al., 2016).

Respecto de la recopilación y del uso de los datos para la inteligencia de datos se pueden producir vulneraciones a la privacidad de las personas durante la recopilación y posterior análisis de los datos como trataremos más adelante.

Por tanto, una empresa, un gobierno o cualquier corporación que decida implantar un proyecto de *big data* su principal reto es plantear y considerar todos los riesgos y consecuencias en materias como privacidad y protección de datos ya que una mala implantación o una interpretación errónea puede generar una pérdida de reputación. Este hecho tendría consecuencias en las ventas, provocaría desaprovechamientos de los recursos internos o que la competencia obtenga ventaja de ello. Incluso si no se cumple con la normativa de protección de datos, pueden llegar a enfrentarse a sanciones y causar rechazo en la sociedad (Nielsen et al., 2015).

Colmenarejo (2018) determina que el uso de los macrodatos provoca que salgan a la luz muchos conflictos sobre ética en las empresas aplicada a la gestión de los macrodatos. La ética en los negocios es una disciplina que sirve para resolver conflictos sobre cómo afectan las decisiones éticas de una empresa a sus grupos de interés o *stakeholders*. Por tanto, las empresas deben plantearse las consecuencias éticas de sus decisiones. También deberían analizar la importancia que se le da a estas decisiones y de si se debe implementar una cultura ética en la empresa.

Por último, para mejorar estos desafíos éticos en materia de inteligencia de datos debe hacerse frente a otro gran desafío y que también es motivo por el que se realiza este trabajo y es el de la necesidad de elaborar un marco regulatorio que sirva de referencia y que aborde entre

otras cuestiones, todos los dilemas éticos que se planteen y las responsabilidades que pueden derivarse (Saltz y Dewar, 2019).

Este marco regulatorio debe abarcar todo el procedimiento que se lleva a cabo en la implantación de un proyecto con inteligencia de datos, es decir, todo lo que va desde la recopilación de los datos hasta el análisis y el uso de éstos (Dorasamy y Pomazalová, 2018).

3. ÉTICA Y RESPONSABILIDAD SOCIAL EN LA ADOPCIÓN DEL BIG DATA

3.1 Consideraciones generales de ética en *big data*

Para definir la ética en la adopción del *big data,* primero debemos conocer qué es la ética y la diferencia que existe con la ética organizacional ya que la ética en la adopción de la inteligencia de datos es una mezcla entre ambas definiciones con la particularidad del fenómeno *big data*.

Por tanto, la ética podemos definirla como "un conjunto de principios morales internos que nos sirven de guía, los valores y convicciones con que analizamos o interpretamos una situación y decidimos cuál es la conducta correcta o apropiada" (Jones y George, 2014, p. 107).

Mientras que la "ética organizacional son las prácticas e ideas rectoras a través de las cuales una compañía y sus gerentes contemplan su responsabilidad hacia sus grupos de interés" (Jones y George, 2014, p. 128).

Así podemos definir la ética en el uso de tecnología *big data* como "el análisis de la naturaleza y el impacto social de las tecnologías de grandes datos y la correspondiente formulación y justificación de políticas para el uso ético de los grandes datos" (O'Leary, 2016, p. 83).

Algunos valores éticos positivos que han favorecido el auge del *big data* son la libertad de expresión y la democratización de los datos. En cambio, la aparición del *big data* ha conllevado una serie de valores éticos negativos como son la vigilancia masiva y el control que ejercen las empresas sobre nuestros datos.

Como hemos visto, la recogida y posterior análisis de nuestros datos mediante técnicas de *big data* supone una gran oportunidad, aunque conlleva importantes riesgos en diversas materias, especialmente cuando hablamos del tratamiento que se realiza sobre nuestros datos.

En cuanto a estos riesgos, hay que establecer límites legales, que los trataremos en el siguiente punto además de establecer directrices éticas respecto a los datos personales.

Existe hoy día una disyuntiva entre la obtención de unos beneficios, tanto para la empresa que emprende el proyecto de *big data* como para los usuarios que se benefician del producto o servicio que la empresa ofrece, y el de una serie de riesgos ya que pueden verse vulnerados los datos personales de los clientes por el uso y tratamiento que se dan a sus datos (González, 2017). Por ello en los siguientes apartados vamos a tratar los temas éticos más importantes.

3.2 La vulneración de datos personales y medidas de prevención

El principal riesgo que se deriva del tratamiento de los datos en *big data* es el de vulneración de los derechos y libertades de las personas ya que pueden verse perjudicados por un mal uso de estos datos.

Un ejemplo de vulneración de derechos y libertades usando los datos personales fue el registro que se llevó a cabo en los Países Bajos sobre las personas que pertenecían a una determinada religión para así saber qué religión tenía mayor representatividad y otorgar subvenciones en base a la representatividad. Este sistema de recogida de datos para usarlo en beneficio de la sociedad hubiera sido un gran éxito si no fuera porque en la segunda Guerra Mundial este registro fue utilizado por los nazis para dar caza a los judíos que se encontraban en los Países Bajos. Este ejemplo es un caso extremo pero que puede volver a suceder en la actualidad.

Hoffman (2018) determina que, para garantizar los derechos y libertades de las personas en la actualidad, los derechos individuales y libertades civiles como por ejemplo el de la privacidad, debe ser garantizados por los estados democráticos.

A nivel organizacional, para evitar hacer un mal uso de los datos, es importante elaborar un código de ética y conducta y realizar informes de auditorías de ética y responsabilidad social para ver si se están cumpliendo ya que "si los datos se utilizan indebidamente, puede tener efectos perjudiciales para la empresa, tales como daños a la reputación o consecuencias legales" (Trom y Cronje, 2019, p. 651).

Los códigos de ética y conducta son documentos que realiza una empresa y los hace público para demostrar su compromiso ético con la sociedad. También para mostrar que la empresa actúa en base a unos valores morales que plasma dentro del código (Calvo y Osal, 2018).

No sólo las empresas son las que elaboran un código ético de conducta, también "muchas profesiones tienen códigos éticos específicos de sus ámbitos: en medicina, en derecho, militar, científico, ingeniería, en educación, etc" (Duncan, Buytendijk, y Logan, 2016, p. 7).

O'Leary, (2016) establece que la elaboración de un código de ética o códigos de conductas sirven mayoritariamente para demostrar la implicación ética de una empresa por las consecuencias de sus actuaciones. También sirve para indicar al resto el grado de compromiso o ética que tiene la empresa. Por último, también sirve para proporcionar información a los *stakeholders* de cómo va actuar la empresa ante situaciones hipotéticas o dilemas morales.

Estos códigos son elaborados por muchas empresas, pero la aparición del *big data* supone un reto mayor ya que la elaboración de estos códigos es insuficiente y se necesitaría implantar un sistema de vigilancia para comprobar que se cumple con lo establecido legalmente en los proyectos de *big data* y prevenir posibles futuras vulneraciones de datos personales.

Este sistema de vigilancia se puede conocer como *whistleblowing*. Este sistema se utiliza en el mundo empresarial para comprobar que se cumple con todos los aspectos legales, medioambientales, éticos y sociales entre otros (Calvo y Osal, 2018).

El concepto de *whistleblowing* lo debemos entender como la denuncia de irregularidades, es decir, "como un acto llevado a cabo por personas con información privilegiada (empleados o miembros de una empresa) y como un acto por el cual la información sensible obtenida a través de ese acceso privilegiado y se entrega a personas externas de la organización" (Olesen, 2019, p. 280).

Se debe usar este sistema de denuncia de irregularidades de forma preventiva, es decir, promover un sistema de vigilancia en la empresa que aplique *big data* para evitar que se produzcan irregularidades sobre los datos que manejen y así evitar vulneraciones a la privacidad (Calvo y Osal, 2018).

Actualmente, las empresas realizan el control de estos datos unilateralmente, es decir, los que controlan la información que se da al exterior forma parte de la propia empresa, por tanto, puede ocurrir que esa información esté distorsionada o manipulada.

No sólo pueden ser denunciados por este sistema, los actos ilegales, sino también los que siendo legales, son poco éticos ya que la regulación sobre la inteligencia de datos no es del todo completa como vamos a ver en el siguiente punto (Olesen, 2019).

Calvo y Osal (2018), determinan que para mejorar ese control de los datos a través del sistema de *whistleblowing*, la comunicación de las empresas que apliquen las técnicas *big data* con sus

stakeholders debe ser mejorada, al igual que mejorar los planes de actuación y protocolos de control para evitar vulneraciones de datos personales. Este sistema de vigilancia o de denuncia de irregularidades, se usaría en las empresas de forma interna para no dañar la imagen de la empresa (Olesen, 2019).

3.3 La privacidad

En Someh et al. (2019), mencionan "tres grupos de *stakeholders* interrelacionados: individuos, empresas y sociedad". Además, cada grupo de interés tiene sus principales preocupaciones (p. 724).

Las principales preocupaciones de los individuos sobre sus datos son la privacidad y la confianza en las empresas que tienen acceso a ellos. A la sociedad en general le preocupa la regulación existente y la vigilancia masiva. Por último, para las empresas lo más importante es el comercio con los datos, la gobernanza ética, la reputación, la calidad de los datos y los algoritmos de toma de decisiones (Someh et al., 2019).

En cada grupo hay una serie de preocupaciones que están relacionadas entre ellas y que son la privacidad, la vigilancia y las consecuencias éticas y de reputación.

Cuando hablamos de privacidad "nos referimos a la capacidad que tenemos cada uno de restringir y controlar como las empresas usan nuestra información personal" (Someh et al., 2019, p. 725).

Una de las cuestiones que más preocupa dentro de la privacidad es la pérdida de control de los datos personales de cada individuo ya que una vez facilitados los datos, los titulares de ellos pierden poder para controlar lo que la empresa hace con sus datos (Barocas y Nissenbaum, 2014).

Este problema debe ser solucionado facilitando un medio para que los individuos puedan revisar los motivos y a dónde van a parar los datos que se han recogido (Someh et al., 2019).

En la actualidad parece que este problema se va solventando poco a poco ya que cada vez existen más plataformas digitales y empresas que facilitan a los usuarios que lo requieren, toda la información que tienen en su base de datos sobre ellos. Además, los usuarios pueden modificar o eliminar los datos que han facilitado previamente como ocurre en Facebook o Spotify. Por ejemplo, Google te da la opción de solicitar toda la información que tenga de ti, todos los datos sobre tu cuenta de correo electrónico, mensajes, fotos, vídeos, geolocalización,

etc y de restringir el acceso a ciertos datos o incluso a eliminarlos de las bases de datos de Google.

Aun así, la aparición de la inteligencia de datos ha provocado que aumente el uso de los algoritmos y con ello las posibilidades de que se dañe la privacidad de las personas.

Con el uso de los algoritmos se puede ayudar a mejorar la experiencia de los usuarios mediante la personalización de un servicio. También sirve para mejorar la toma de decisiones de las empresas. El problema ético surge cuando estos algoritmos están sesgados porque evalúan a las personas y las clasifican para discriminarlas, vulnerando así su privacidad (Cotino, 2019).

Por ejemplo, una aseguradora puede utilizar un algoritmo usando los datos de sus clientes para evaluar a sus clientes en base al riesgo e imponer una serie de primas según el riesgo de cada cliente. Este es un claro ejemplo de algoritmos sesgados usando esta herramienta con fines poco éticos (Nielsen et al., 2015).

Otro ejemplo del uso de los algoritmos es el que establece Ortiz (2018), en el que según las búsquedas realizadas en internet y los gustos de una persona aparece publicidad de un tipo u otro. Es decir, un algoritmo que ha sido diseñado para mejorar las compras de un usuario puede transformarse debido a un mal uso ético en una herramienta que intenta manipular las preferencias de las personas.

Los algoritmos en sí no discriminan, son una herramienta que se aplican para facilitar la toma de decisiones, el problema surge cuando se decide utilizar algoritmos para usos poco éticos y vulneren los derechos y libertades de ciertos grupos de personas (Chen y Quan-Haase, 2018).

Para intentar proteger la privacidad, muchas bases de datos anonimizan los datos. El problema surge cuando se puede volver a identificar al individuo agregando otros datos y poner a disposición de terceros sin el consentimiento de los individuos (Barocas y Nissenbaum, 2014).

También es importante destacar el papel de la ética en la aplicación de la inteligencia de datos porque cada vez hay más dispositivos que recaban información de los usuarios, como son, por ejemplo, los robots domésticos, televisores inteligentes o coches autónomos. En el día a día podemos ver muchas vulneraciones por dispositivos tecnológicos que no paran de captar datos, almacenarlos, analizarlos e incluso compartirlos o venderlos a terceros.

Es importante que todas las situaciones y vulneraciones que se pueden dar al utilizar hayan sido planteadas a la hora de diseñar y lanzar los productos al mercado ya que podemos plantearnos si es lícito que estos dispositivos graben y analicen información y pautas de comportamiento para ser utilizados con fines comerciales.

Por ejemplo, en el caso de los coches autónomos debe analizarse un dilema ético y moral respecto a un hipotético caso de atropello ya que es el coche el que decide a quien salvar porque puede ocurrir que evitando el atropello se produzca un accidente que dañe al conductor o evitar dañar al conductor produciéndose el atropello (Mayer-Schönberger y Cukier, 2013).

Otro de los muchos ejemplos es el de los televisores *smart* de Samsung que registran lo que ve cada persona, la hora a la que ve la televisión e incluso graba sonidos e imágenes de las personas que ven la televisión (Royakkers, Timmer, Kool, y van Est, 2018).

Este hecho viene recogido en su manual de usuario porque es un requisito legal que deben cumplir las empresas. El problema ético es que las empresas informan sobre estos aspectos con un leguaje complejo y se basan en el consentimiento cuando verdaderamente, alguien que se compra una *smart* TV, lo último de lo que se preocupa es de leer los términos y condiciones legales y la letra pequeña de estos manuales (Royakkers et al., 2018).

Otro ejemplo de vulneraciones a la privacidad es el que ocurre con las compañías de seguros que aplican la inteligencia de datos, que se aprovechan de los datos recogidos a través de múltiples fuentes como son dispositivos electrónicos y apps como los smartwatch y pulseras deportivas que monitorizan nuestras constantes vitales cuando hacemos ejercicio o cuenta las calorías que consumimos durante el transcurso del día (Nielsen et al., 2015).

Con estos datos, las aseguradoras médicas elaboran estadísticas demográficas y así realizan modelos predictivos de clientes potenciales y el riesgo que puede tener cada tipo de cliente de sufrir alguna enfermedad (Nielsen et al., 2015).

3.4 La responsabilidad en el empleo del *big data*

Un mayor control de la ética y de la responsabilidad de los datos puede ser percibido por parte de la empresa como un mayor potencial para el mejor desarrollo de su actividad y ésta tenga un mayor prestigio social.

Los casos de filtraciones de datos se producen porque la seguridad que protege esos datos no es lo suficientemente fuerte. También puede ocurrir que una herramienta que ha sido diseñada para mejorar la experiencia del usuario puede volverse en su contra porque no existe esa responsabilidad por parte de la empresa.

Estos casos plantean el riesgo de que se produzca una dictadura *smart,* concepto que acuñó Weltzer (citado en Morte, 2017), ya que las grandes empresas tecnológicas (Alphabet,

Facebook, Microsoft o Apple entre otras) están implantando continuamente sus novedades tecnológicas y un mal uso o aprovechamiento de dichas novedades tiene riesgos para los derechos de las personas (Morte, 2017).

Se puede dar el caso, exagerado, de que "Google proporcione datos a una sistema dictatorial y se use esa información para identificar a las personas para un genocidio" (Nersessian, 2018, p. 849). Por tanto, no vamos a entrar en si estas empresas cumplen o no con la ley, sino en la responsabilidad de cada empresa.

Los casos que hemos mencionado en los apartados anteriores sobre vulneraciones a la privacidad de las personas deben tener un responsable, es decir, debemos identificar al sujeto que toma esas decisiones y atribuirle la responsabilidad sobre ellas. Por eso es básico crear una cultura organizacional, para que el sujeto responsable de tomar estas decisiones lo haga acorde a la cultura de la organización (Colmenarejo, 2018).

La idea de la existencia de un responsable es porque cada vez están apareciendo nuevas máquinas, algoritmos, sistemas de inteligencia artificial que analizan los datos mediante técnicas de *big data* para mejorar la sociedad y es necesario que haya una figura dentro de la organización responsable de coordinar y hacer que todo funcione correctamente (Calvo y Osal, 2018).

Actualmente se está trasladando la responsabilidad sobre el uso de los datos al consumidor. Esto lo vamos a ver más claro con el ejemplo de la implantación y posterior uso de una app para el móvil, pero también es aplicable a cualquier otro ejemplo como el de la compra del televisor Smart TV que comentamos previamente.

Cuando se instala una app en el móvil, es necesario que las condiciones de uso y política de privacidad de la empresa que suministra el servicio app sean aceptadas por el usuario. Lo mismo ocurre con la compra del televisor Smart Tv, en el que se incluye un manual de usuario donde explica su política de privacidad y las condiciones de uso. De hecho, en este caso, el problema es incluso más flagrante ya que el manual se entrega una vez que se compra el producto (Royakkers et al., 2018).

El caso es que, en estas situaciones, esas condiciones de uso y políticas de privacidad se redactan con un leguaje complejo y en la mayoría de los casos son muy extensos. Además, cualquier usuario que descarga una app, lo que quiere es empezar a usarla cuanto antes y no se detiene a leer toda la información y acepta las condiciones sin más para así poder empezar a usar el servicio o producto adquirido (Soto, 2017).

En estos casos, se está obligando al usuario a aceptar las condiciones que impone la empresa y en consecuencia, se está trasladando la responsabilidad al consumidor cuando debería ser una responsabilidad entre ambas partes ya que la empresa debe procurar dicha labor facilitando la información de una manera más asequible y clara el consumidor debe tener una mayor preocupación por su privacidad y los permisos que está concediendo a una empresa para conocer sobre sus datos.

Según Martínez (2019), los usuarios deberían tomar algunas precauciones para proteger mejor sus datos personales como, por ejemplo, evitar publicar todos tus datos en Internet, fotos en redes sociales donde aparezcan datos bancarios, números de teléfono, etc. Otra medida es la de borrar cada cierto tiempo las famosas *cookie*s y el historial de navegación. Por último, la medida más importante es la de leer los términos y condiciones de uso y políticas de privacidad de cada aplicación, producto o servicio que adquiramos.

Martínez (2019) determina que los titulares de los datos, es decir, los usuarios, tienen parte de responsabilidad ya que no es solo responsabilidad de terceros el tratamiento que le den a nuestros datos, sino que los titulares de los datos deben ser más conscientes sobre qué información damos o publicamos en Internet.

Actualmente cualquier persona tiene un correo electrónico, por ejemplo, Gmail, usa redes sociales como Facebook, Twitter o Instagram. Se comunica con sus seres queridos mediante mensajería instantánea (WhatsApp), busca información en Google y usa nuevas formas de entretenimiento como Netflix o YouTube.

A través de todas estas plataformas, recibimos numerosos servicios que tienen muchas ventajas, pero también proporcionamos muchos datos, como páginas visitadas, tiempo de acceso, dispositivos usados, grabaciones de voz y cualquier otro tipo de información sensible. Toda esta información es la que damos permiso al aceptar las condiciones de uso y políticas de privacidad. Estas plataformas almacenan la información en sus bases de datos para prestarnos un mejor servicio, entre otras cosas (Martínez, 2019).

Hammer (2017) establece que estas prácticas son cada vez más comunes y vulneran la privacidad de las personas y esto se produce gracias a que cada vez más las personas y, en especial, las nuevas generaciones usan la tecnología para cualquier aspecto de su vida favoreciendo que se vulnere su privacidad.

Estas vulneraciones a la privacidad hacen que mermen la confianza en las empresas que recopilan datos. Lo que ocurre con Facebook y sus fallos de seguridad o recopilar cualquier dato sobre los gustos de las personas hace que pierda la confianza de las personas y los

últimos escándalos de venta de datos personales a terceras personas supone una gran pérdida de reputación.

La reputación de una empresa está determinada por las percepciones de los *stakeholders*, sobre cómo la organización incorpora la ética en sus prácticas de análisis de inteligencia de datos. Las empresas con una mala reputación por sus prácticas poco éticas de *big data* pueden tener dificultades para desarrollar una cultura ética a nivel interno, lo cual disminuye la confianza de las personas (Someh et al., 2019).

Un problema actual es que, si todos los competidores en un mercado tienen mala reputación, puede provocar que las empresas no se preocupen en cambiar sus actuaciones, aunque sean poco éticas porque no les afecta (Someh et al., 2019).

3.5 Cesión de datos a terceras personas

En la actualidad estamos viviendo lo que algunos autores llaman la revolución tecnológica, lo que provoca que en prácticamente todos los sectores se implementen innovaciones o surjan nuevos modelos de negocio basados en el uso de nuevas tecnologías que usan información y datos. Esto provoca que la competencia aumente y la presión por no quedarse atrás puede provocar que se produzcan abusos en cuanto a las leyes y a la ética.

Hay empresas que aparte de almacenar y usar datos para mejorar sus servicios o productos también ceden o venden nuestros datos para obtener un mayor beneficio. Esto es posible porque previamente el usuario o titular de los datos le ha concedido el permiso para hacerlo. Generalmente se produce cuando se aceptan las condiciones de uso y políticas de privacidad.

Respecto de las ventas de los datos a terceras partes, se plantean los siguientes desafíos: "el consentimiento informado explícito, la transparencia en el suministro y el intercambio de datos, y el mantenimiento del anonimato y la protección de los datos contra el uso no ético a lo largo de la cadena de valor de *big data*" (Someh et al., 2019, p. 727).

Es decir, para evitar mayores vulneraciones de privacidad, cuando una empresa quiera vender los datos de sus usuarios o clientes, no debería bastar con la aceptación generalizada de las condiciones de uso, sino que se debería de volver a preguntar a los usuarios sobre si aceptan que sus datos sean cedidos a terceras personas para ofrecerles un servicio más personalizado o que específicamente haya un casillero de aceptación o rechazo a esta práctica (Someh et al., 2019).

Además, las instituciones deberían obligar a las empresas garantizar que los datos que cedan o vendan a terceros sean anónimos para así proteger los datos y que sean los usuarios los que puedan conocer en todo momento quiénes adquieren sus datos y con qué motivos y poder negarse a ello si han cambiado de opinión.

Uno de los últimos escándalos y que han tenido mayor trascendencia en esta materia es el caso de la empresa Cambridge Analytica, aunque ha habido otros como el de la NSA y Snowden.

En el caso de Cambridge Analytica, Facebook cedió los datos de millones de sus usuarios a esta empresa de analítica y publicidad basada en comportamientos y gustos para así identificar las preferencias, gustos y pasiones de los usuarios en Facebook.

Debido a esto, se ajustaba el contenido y publicidad que veían estos usuarios de Facebook para cambiar la forma de pensar de los usuarios e inducir al voto a favor de Donald Trump en las elecciones presidenciales de Estados Unidos. En este caso se vulneró la privacidad de las personas ya que se usaron datos de los usuarios sin que éstos lo supieran (Martínez, 2019).

Algo parecido es lo que ocurrió con el escándalo de la NSA en Estados Unidos que, en 2013, Edward Snowden filtró documentos de alto secreto del gobierno de los Estados Unidos sobre el sistema de vigilancia masivo que han llevado a cabo las principales agencias de inteligencia estadounidense. Entre dichas vulneraciones se encuentran personajes de las altas esferas políticas mundiales, pero también se estima que han sido afectadas miles de millones de personas.

Esto se puede extrapolar al mundo empresarial y una empresa que adopta implantar un proyecto de *big data* puede verse envuelta en problemas de este tipo (Nielsen et al., 2015).

Estos escándalos han provocado que aumente la preocupación de la sociedad por el uso de sus datos por parte del Gobierno (Nielsen et al., 2015). Empresas como Facebook y Google deben transmitir que no se observa ni monitorea en su vida cotidiana a los usuarios ya que si perciben esto y conocen que se extraen todo tipo de datos con el fin de obtener beneficios, las personas se vuelven recelosas (Someh et al., 2019).

También es éticamente discutible, aunque sea legal, el hecho de dar el consentimiento para el almacenamiento de nuestros datos y gustos y recibir anuncios personalizados o correos electrónicos no deseados con ofertas y promociones porque la empresa encargada del almacenamiento de los datos, ha cedido los datos para su posterior análisis y lucrarse con ello. Empresas como Google, Facebook o Amazon, en base a nuestras búsquedas y preferencias,

realizan un perfil sobre nosotros y lo venden a terceros que son los que contactan con nosotros (Someh et al., 2019).

Un claro ejemplo de estos nuevos modelos de negocios que infringen las reglas es YouTube. La plataforma de vídeos en Internet tiene mucho contenido que vulnera las normas de copyright. Un estudio de YouTube determina que si eliminan todo el material que infringe las normas del copyright, las visitas a su plataforma se reducirían en un 80% y su modelo de negocio dejaría de ser viable (Nielsen et al., 2015).

Por todo lo que hemos visto en este punto, el consentimiento es necesario legalmente para el tratamiento de nuestros datos, pero éticamente es ineficaz ya que aun así se producen vulneraciones a la privacidad con las cesiones a terceras personas y además se intentan ocultar para evitar escándalos como algunos de los que hemos tratado. Es necesario plantear un debate ético y legal para mejorar la privacidad y las consecuencias derivadas de un mal uso de los macrodatos.

4. LOS MARCOS LEGALES Y LA ADOPCIÓN Y USO DEL BIG DATA

4.1 Aspectos generales

Cada día aumenta la cantidad de dispositivos que proporcionan datos que son almacenados. Estos datos necesitan del consentimiento para que puedan ser recogidos y analizados. El problema surge cuando damos el consentimiento a una empresa en concreto y ésta lo extiende a una tercera parte (Hoffman, 2018).

Esta gran cantidad de datos, se generan gracias al Internet de las cosas (Martínez, 2017). Es decir, en la vida cotidiana de una persona existen múltiples dispositivos y sensores que registran datos sobre las personas. Estos dispositivos han multiplicado la generación de datos y la aparición y posterior aumento de la domótica hará que se multiplique de una forma más exponencial a cómo se hace en la actualidad.

Martínez (2017), determina que las redes sociales también han potenciado este fenómeno ya que conectan a las personas a través de la red, se crean comunidades y esas comunidades que tienen unos mismos intereses pueden ser un target en el ámbito de publicidad.

Soto (2017) afirma que:

> Convertir a un individuo en una diana de vigilancia implica hoy día una invasión mucho más extensa de la vida privada, puesto que no solo se suele pretender obtener toda la información posible sobre la persona, sino también, sobre sus relaciones, conexiones e incluso, interacciones. Todo ello supone claramente una amenaza a la privacidad, pero, además, el uso de los datos masivos como modelo predictivo permite la posibilidad de poder juzgar previamente a las personas más allá de su comportamiento (p. 104).

El primer problema que se plantea al tratar un proyecto de *big data* es el de la privacidad. Los principales marcos normativos orientan la privacidad a evitar que se usen los datos para fines distintos de los que fue recopilado. Es decir, la normativa trata de adecuar el fin para el que esos datos fueron recogidos, por tanto, cuando se acomete un proyecto de macrodatos el principal escollo es el de incumplir con la normativa de privacidad.

El uso de *big data* para elaborar modelos predictivos de comportamiento es un claro ejemplo de vulneración de la privacidad. Esto quiere decir que mediante la inteligencia de datos podemos llegar a predecir el comportamiento de un individuo o de un conjunto de individuos que tienen una conducta parecida. Esta situación puede transformar la vida de esos individuos para bien o para mal según el uso que se quiera dar, pero eso queda en el ámbito ético del responsable de tomar esas decisiones (Soto, 2017).

El tema de la privacidad y la protección de datos es importante porque estamos hablando de que en sectores como el bancario, la salud, seguros y la propia Administración llevan a cabo proyectos de macrodatos con nuestros datos (Soto, 2017).

Por eso es importante regular jurídicamente el tema en relación a los datos ya que la tecnología siempre va varios pasos por delante de la legislación y ahora mismo la regulación existente es insuficiente para regular todos los aspectos del fenómeno *big data*. Por ello primero vamos a hacer un repaso de la regulación sobre protección de datos de una forma general para luego concretar sobre aspectos legales de la privacidad.

4.2 Regulación sobre la protección de datos

4.2.1 Regulación histórica

La regulación del derecho a la privacidad se inició en las primeras décadas del S.XX pero hasta la Declaración Universal de Derechos Humanos de las Naciones Unidas en 1948 no se consolidó. En dicha Declaración, el artículo 12 determina que "nadie será objeto de injerencias arbitrarias en su vida privada, su familia, su domicilio o su correspondencia, ni de ataques a su honra o a su reputación" (Gómez-Barroso, Feijóo y Martínez, 2017, p. 114)

En los siguientes años la tecnología se va desarrollando y va aumentando la cantidad de datos almacenados por lo que fue conveniente realizar una actualización de la regulación sobre los datos.

La OCDE en 1980 adoptó las Directrices sobre la protección de la privacidad y flujos fronterizos de datos personales. En ellas se identifica una serie de principios básicos sobre la regulación de datos personales que "a día de hoy sigue presente en muchos ordenamientos jurídicos: finalidad legítima, proporcionalidad, necesidad de consentimiento, transparencia, responsabilidad y restricciones a la transmisión" (Gómez-Barroso et al., 2017, p. 114).

A finales del siglo XX, surgió lo que conocemos actualmente como la protección de datos personales en la Carta de Derechos Fundamentales de la Unión Europea del año 2000. En ella se determina que "los datos se tratarán de forma leal, para fines concretos y con el consentimiento de la persona afectada pudiendo ésta acceder a sus datos y corregirlos" (Gómez-Barroso et al., 2017, p. 115).

Como hemos comentado el auge de internet a principios del Siglo XXI y la aparición del internet de las cosas y todos los dispositivos que recaban información, ha provocado que la regulación existente hasta el momento en cuanto a protección de datos sea insuficiente.

Para solventar dicho problema, la Unión Europea ha incrementado el valor de las sanciones por incumplir con la normativa de protección de datos. Además, incide en que los datos que se recojan, se destinen al fin que inicialmente fueron previstos (Keffer, 2019).

En los últimos años se ha ido completando estas carencias en la regulación europea con el TFUE y otras normas complementarias como la Directiva 2002/58 CE sobre privacidad y comunicaciones electrónicas en las cuales se introducía el concepto de anonimizar los datos o borrarse cuando dejen de ser necesarios (Gómez-Barroso et al., 2017).

4.2.2 Regulación actual: Reglamento General de Protección de Datos (RGPD)

La última gran modificación o actualización de la regulación de protección de datos es la creación del Reglamento General de la Unión Europea sobre Protección de Datos Personales de 2018 (en adelante RGPD). Se hizo debido a la rápida evolución de la tecnología y el aumento exponencial de intercambio de datos (Gómez-Barroso et al., 2017).

Este Reglamento, como hemos venido comentando, era necesario debido al avance de la tecnología y la aparición de la inteligencia de datos y su uso en la economía y en la sociedad. Con este reglamento surge un marco regulatorio más unido, coherente y común para toda la Unión Europea ya que permite a las empresas digitales introducirse en la economía y a los titulares de los datos tener el control de sus datos personales y poder revocar los permisos en cualquier momento (Gómez-Barroso et al., 2017).

Entró en vigor en mayo de 2018 y este reglamento de la Unión Europea es la regulación vigente en España en materia de privacidad y protección de datos junto a la Ley Orgánica de Protección de Datos (en adelante LOPD). Hay que explicar que la LOPD es tan sólo la adaptación de la normativa europea (RGPD) a la normativa española y que la LOPD no puede contradecir lo establecido en el RGPD y tan sólo regula o delimita algunos aspectos técnicos que no son ámbitos de estudio en este trabajo. Por tanto, en este trabajo respecto a la normativa que se aplica en España nos vamos a centrar en el RGPD.

La entrada en vigor del RGPD ha incitado a "que las empresas deben mantener sólo la menor cantidad de datos personales y sólo durante el tiempo estrictamente necesario para un fin específico". (Keffer, 2019, p.178).

De hecho, el artículo 25 del RGPD establece que las empresas sólo deben recopilar la información personal para un propósito específico, guardar la cantidad mínima de información personal necesaria y por el tiempo estrictamente necesario.

Aunque estos principios básicos que establece el RGPD sobre minimización de datos en la recopilación y procesamiento y limitación de fines han supuesto un avance para mejorar la regulación sobre la privacidad, en la inteligencia de datos, "la información no se recopila con un propósito específico y limitado sino que se recopila para descubrir nuevos patrones" (Favaretto, De Clercq y Elger, 2019, p. 22).

Por último, otra característica a mencionar sobre el RGPD es que, aunque sea una norma europea y, por tanto, se aplica en el territorio español, intenta solucionar la problemática de

las empresas globales que tratan con datos personales. Es decir, empresas estadounidenses, por ejemplo, que no tengan presencia física en Europa, si utilizan datos de personas europeas deben cumplir con lo establecido en el RGPD (Keffer, 2019).

4.3 Consideraciones legales sobre la privacidad

4.3.1 Principios jurídicos

Es importante a la hora de regular sobre la privacidad, una serie de principios jurídicos y derechos de la persona que sirvan como base. Esa base "son los derechos y libertades fundamentales de las sociedades democráticas" (Cotino, 2017, p.137).

Uno de esos principios es el de "prohibición con excepción de autorización" (Morte, 2017, p. 230). A partir de ahí surge la interpretación actual que se regula para proteger la privacidad mediante el consentimiento. Según este principio cualquier uso de los datos debe estar autorizado por el sujeto titular o que lo autorice una ley (Morte, 2017).

Según la ONU, (citado en Nersessian, 2018) en sus principios rectores se reflejan tres mandatos que deben respetarse y que considero que afectan al ámbito de aplicación de los macrodatos. Estos mandatos son los siguientes:

1) Los Estados deben protegerse contra los abusos de los derechos humanos independientemente que los realicen personas físicas o jurídicas.
2) Las empresas deben respetar los derechos humanos, tanto como en las políticas de la empresa como en sus actividades empresariales.
3) Los gobiernos y las empresas deben proporcionar un remedio cuando se produzcan abusos de los derechos humanos (Nersessian, 2018, p. 850).

A partir de ahí, Cotino (2017) establece que es necesario un nuevo enfoque jurídico y una nueva interpretación de los derechos fundamentales, no solo desde la perspectiva actual del individuo, sino desde una perspectiva más colectiva ya que el Parlamento Europeo se ha pronunciado sobre la importancia de garantizar la privacidad a través de los derechos fundamentales.

Kemp (2014) elabora un modelo legal para regular el fenómeno *big data* otorgándole gran importancia a la regulación de los datos y la protección de estos.

Kemp (2014) propone que se confieran derechos a las personas e impone obligaciones para el procesamiento de los datos. Actualmente la regulación se basa en poner expresamente los requisitos legales y esperar el consentimiento del usuario para procesar sus datos.

Esto es insuficiente porque cualquier usuario acepta las condiciones de uso sin leerlas ya que suele ser por falta de tiempo, lenguaje complejo e incluir una gran cantidad de texto que todo el mundo evita leer para poder acceder al servicio que ofrece la empresa que va a recoger los datos.

Por eso los Estados deberían elaborar un marco regulatorio que mejore la normativa actual porque además es un tema que cada vez se usa en más sectores y va a acabar afectando prácticamente a cualquier sector que utilice la tecnología.

4.3.2 El consentimiento y la técnica de anonimizar datos

La protección de datos requiere que el titular de esos datos otorgue su consentimiento para que sus datos sean tratados por la otra persona a cambio de que estos derechos sean protegidos y no se conozcan por terceros.

Otra cosa es que el titular de esos datos dé, el consentimiento, o que se informe al titular que sus datos van a ser conocidos por terceras partes y éste acepte. Es decir, no es lo mismo que el titular dé expresamente su consentimiento a que sus datos sean cedidos a terceros que se le informe en un documento extenso o con letra pequeña y lenguaje complejo que sus datos pueden ser usados por terceros y que si no acepta no puede acceder al servicio o producto que desea usar.

Como hemos hablado anteriormente, actualmente el consentimiento es ineficaz ya que según Kamp y Rost (citado en Morte, 2017) "el consentimiento válido debe ser específico, libre e informado".

Actualmente el consentimiento se da por defecto y muchas veces se informa al titular de esos datos que se va a ceder a terceros para que sea analizado con un lenguaje complejo y las empresas se aprovechan del titular de los datos ya que no suele leer los términos y condiciones y las acepta sin más. Por tanto, no hay un equilibrio de poder entre el titular de los datos y la entidad que los recoge.

Soto (2017) afirma que casi cualquier aplicación o plataforma digital venden o ceden los datos que almacenan a otras empresas y por tanto se vulnera la privacidad y confidencialidad ya que el titular de los datos no ha dado su consentimiento o éste se considera que es ineficaz.

Para solucionar esto se está recurriendo a anonimizar los datos. Esto significa que aquellos datos que sean de carácter personal y que puedan identificar a una persona física concreta o a un concreto colectivo, deben eliminarse para que sean anónimos. Es decir, se debe borrar o eliminar cualquier aspecto que haga que dichos datos puedan vincularse con la persona que corresponde. Se suele utilizar en muchos ámbitos, pero el principal es para datos de carácter médico.

Aunque esta técnica tampoco es del todo eficaz ya que con la aplicación de la tecnología *big data* se puede llegar a agrupar los datos y obtener el perfil de una persona a pesar de no tener sus datos personales (Scotti, 2017).

Los proyectos que utilizan estas técnicas para anonimizar datos se aseguran con mayor probabilidad que no van a incumplir con la normativa de protección de datos. Hay diversas técnicas para conseguir este resultado pero básicamente se reducen al aislamiento, la no vinculación y no inferencia de los datos para evitar que se vinculen los datos y llegar a deducir la identidad de la persona a la que corresponde los datos en proyectos de macrodatos (Pérez, 2016).

En cuanto a la regulación sobre la predictibilidad es un riesgo que hay que tener en consideración porque podemos llegar a un caso extremo como el de "una sociedad que establezca un derecho penal preventivo y aplique medidas punitivas a quienes todavía no han cometido un delito" (Martínez, 2017, p.156).

Es decir, con un mal aprovechamiento de las ventajas que tiene la inteligencia de datos podemos llegar a un sistema en el que se castigue a un conjunto de personas por pertenecer a una comunidad.

Al titular de los datos se le reconoce un derecho legal a la privacidad y la expectativa de que sus datos van a ser usados apropiadamente (Spiekerman, Acquisti, Bohme, y Hui, 2015). En base a esto, Hoffman (2018) determina que "hay tres prácticas éticas que necesitan ser abordadas: el derecho a ser olvidado, el derecho a la caducidad de los datos y la propiedad de los gráficos sociales" (p. 6).

4.3.3 La autorregulación

Otra herramienta que se debería implantar para mejorar la regulación y el cumplimiento de los estándares éticos en el uso de *big data* es el de la autorregulación. La autorregulación está relacionada con los códigos de éticos o de conducta que hemos tratado en el epígrafe anterior.

Como hemos dicho, para un mejor cumplimiento es necesario que las empresas y sus directivos realicen normas y directrices en sus proyectos de *big data* y en gestión de datos aparte de cumplir con la normativa estatal e internacional.

Según Martínez (2019):

> La autorregulación sólo redundará en beneficio real de las personas en la medida que sea bien concebida, aplicada y cuente con mecanismos que garanticen su cumplimiento de manera que no se constituyan en meras declaraciones simbólicas de buenas intenciones sin que produzcan efectos concretos en la persona cuyos derechos y libertades pueden ser lesionados o amenazados por el tratamiento indebido de sus datos personales (p. 16).

Para ello, algunas de las medidas que pueden tomar las empresas es la de minimizar los datos que recopilan para que sólo tengan los estrictamente necesarios para el correcto funcionamiento del proyecto y así garantizar la privacidad y demás derechos de los titulares. Por último también existe las técnicas ya mencionadas para anonimizar datos y mejorar el sistema de consentimiento que rige actualmente (Martínez, 2019).

5. CÓMO GESTIONAR EL CUMPLIMIENTO DE LOS ESTÁNDARES ÉTICOS Y LEGALES EN EL USO DEL BIG DATA

La gestión de la inteligencia de datos es un tema muy importante en la actualidad. La inteligencia de datos tiene mucho potencial para mejorar los negocios de las empresas y entender mejor a los clientes pero a su vez conlleva una serie de riesgos que hemos visto y para minimizarlos es necesario una gestión adecuada de los datos (Trom y Cronje, 2019).

La figura que engloba todo lo relacionado a la gestión y toma de decisiones de los datos en una empresa que usan los macrodatos es el gobierno de los datos o *governance*.

Se puede definir el término de *governance* como un proceso de:

> Establecer derechos de decisión y responsabilidad, así como establecer políticas que estén alineadas con el objetivo comercial. Equilibrar las inversiones de acuerdo con las políticas y en apoyo de los objetivos comerciales. Establecimiento de medidas para monitorear el cumplimiento de las decisiones y políticas. Y asegurar que los procesos, comportamientos y procedimientos estén de acuerdo con las políticas y dentro de las tolerancias para respaldar las decisiones (Buytendijk y Oestreich, 2015, p. 9).

Según Soares (citado en Trom y Cronje, 2019) define al término de *governance* como "la gestión eficaz de las personas, los procesos y la tecnología que permite a la organización aprovechar los datos como un activo empresarial valioso" (p. 650).

Es decir, los encargados del gobierno de los datos de una empresa que aplica la inteligencia de datos son los máximos responsables de la toma de decisiones respecto al tratamiento de los datos. Éstos deben establecer una estrategia de acuerdo a los objetivos de la empresa y realizar medidas de control para medir el cumplimiento de las políticas. Además, al ser los máximos responsables, deben tomar decisiones en cuanto a la gestión de riesgos (Trom y Cronje, 2019).

Aparte del *governance* podemos decir que está apareciendo una nueva figura, que es la del director de datos, debido al aumento de empresas que se dedican al análisis de los datos (Chief Data Officer, en adelante CDO) (Duncan et al., 2016).

El CDO, aparte de realizar las tareas propias del gobierno de los datos que hemos descrito anteriormente, también es una figura de influencia. Es decir, el CDO debe ser alguien que tenga en cuenta la cultura de la empresa y motive al resto de trabajadores basándose en los valores de la empresa. Es una nueva forma de rol o liderazgo para impulsar una nueva cultura en la empresa (Duncan et al., 2016).

Es decir, "estos nuevos profesionales creen que la innovación está en los datos, y explorándolos conllevará a nuevas estrategias y oportunidades de negocio" (Buytendijk y Laney, 2014, p. 2).

No sólo se centran en el análisis de los datos, sino que también participan en decisiones comerciales (Duncan et al., 2016). Es decir, el CDO colabora en los objetivos de la empresa con actuaciones para ayudar a conseguirlos (Buytendijk y Laney, 2014).

El CDO o cualquier profesional que forme parte del equipo del gobierno de los datos debe ser una persona que tenga habilidades tecnológicas y de negocio. Igualmente debería tener una

gran capacidad analítica y es por eso, que normalmente el perfil que más se ajusta, es el de un ingeniero, matemático o un estadístico ya que es más fácil que éstos aprendan los aspectos básicos empresariales (Buytendijk y Oestreich, 2015).

Una de las principales líneas de actuación de los responsables del *governance* es que deben tomar las medidas suficientes para garantizar la privacidad de las personas sobre los datos que se han recogido o analizado. Para ello debe tener las herramientas necesarias para facilitar la implementación y realizar un seguimiento para controlar que se cumple con todos los requisitos (Buytendijk y Oestreich, 2015).

Para garantizar la privacidad, es recomendable establecer políticas de seguridad de datos para garantizar la privacidad. Además, la regulación de la privacidad de los datos varía en cada Estado y por tanto es recomendable para las empresas multinacionales, tener un responsable para el gobierno de los datos (Someh et al., 2019).

Aunque muchas empresas tienen la figura de *governance* para garantizar que se cumple con los estándares legales y el cumplimiento de la ley, el mero cumplimiento de la ley es insuficiente para cumplir con los estándares éticos (Buytendijk y Heiser, 2013). Para solucionar este problema, debería plantearse un debate sobre cuáles son los usos correctos y erróneos de los datos y cómo deberíamos mitigar los riesgos.

Para gestionar éticamente el gobierno de los datos, los CDO son los responsables de tener en cuenta todas las consideraciones éticas de los datos y su análisis. Por ello es recomendable que se establezca un código de conducta en el que se defina las pautas éticas según la cultura establecida por el equipo de gobierno de los datos (Duncan et al., 2016).

Así es más difícil que la empresa incurra en vulneraciones de la protección de datos y se evita caer en responsabilidades y que la reputación se vea dañada. De esta forma, aunque sigan existiendo estos riesgos de vulneraciones de la privacidad, mejora la imagen de la empresa y por tanto puede llegar a repercutir en los beneficios.

Mientras un equipo directivo normal se encargaría de elaborar políticas y protocolos de actuación para la empresa, el equipo de *governance* o el responsable de los datos, tiene que ir más allá, y debe tratar de fomentar una cultura dentro de la empresa mediante valores, normas y creencias (Someh et al., 2016).

Además, las diferentes herramientas como estrategias u objetivos para generar valor, minimizar riesgos y cumplir objetivos deber ser utilizados por el equipo de governance para lograr una implementación adecuada y optimizar los recursos (Trom y Cronje, 2019).

Esas medidas o políticas se hacen para garantizar la privacidad y monetizar los datos que van a ser usados con técnicas de *big data*. Se puede decir que estos principios sirven como marco de actuación para el *governance* (Al-Badi et al., 2018).

Según EY (citado en Trom y Cronje, 2019) una buena implementación del *governance* de los datos necesita de unos "procedimientos formales definidos, orientación coherente a la cultura de la empresa y sólida base de toma de decisiones por parte de la administración" (p. 649).

Es decir, las empresas que usan *big data* necesitan realizar un programa de gobierno que sirva como base para la actuación del equipo de *governance*. El *governance* debe rendir cuentas a la dirección de la empresa y encargarse de todo lo relacionado con los datos, desde la gestión hasta el control para evitar riesgos (Al-Badi et al., 2018).

El proceso de *governance* debe ser dinámico ya que la tecnología *big data* y el uso de los datos va variando al igual que la regulación de ésta. Es decir, no basta con elaborar una serie de procesos y adoptar una cultura respecto al uso de los datos en una empresa, sino que a medida que pasa el tiempo, la empresa que ha adoptado un equipo de *governance* de los datos se debe adaptar a los cambios que se van produciendo respecto al uso de los datos.

Según Duncan et al. (2016) "los CDO deben aplicar métodos de ética de la empresa a los datos y los análisis y realizar pruebas periódicas de los impactos éticos de la aplicación que hace la empresa del análisis de los datos" (p. 8).

Es decir, el CDO o el equipo de *governance* es el responsable de monitorizar que se cumple con lo establecido éticamente y medir el impacto ético de las decisiones tomadas en base al análisis de los datos de la empresa. El uso del *big data* requiere de medidas de control y se deben tomar precauciones para proteger los derechos de las personas que son titulares de los datos. Las personas encargadas del equipo de *governance* debe regular todos los aspectos de los macrodatos desde la recogida de datos hasta su procesamiento para evitar filtraciones de datos o robos por parte de hackers. Es decir, los procesos para la gestión de datos deben estar bien definidos y monitorizarse constantemente (Trom y Cronje, 2019).

Cuando se den prácticas poco éticas, los responsables de *governance* deben haber establecido un sistema de sanciones y medidas para educar y entrenar en valores a los infractores dentro de la empresa para no volver a incurrir en esas prácticas poco éticas respecto al uso de los datos (Someh et al., 2019).

PARTE II: INVESTIGACIÓN EMPÍRICA

6. PERCEPCIÓN DEL CUMPLIMIENTO ÉTICO–LEGAL EN EL USO DEL *BIG DATA*

A continuación, vamos a detallar la investigación descriptiva cuantitativa que hemos elaborado mediante encuestas realizadas a profesionales del *big data* y mostrar los resultados de dichas encuestas.

Para la realización de la encuesta se ha optado por encuestas de escalas no comparativas seleccionando entre ellas la escala de Likert. La escala de Likert se utiliza en numerosos cuestionarios y requiere que los encuestados indiquen el grado de conformidad con cada una de las afirmaciones que se realizan más abajo.

Los encuestados deben elegir un número entre 1 y 5 para cada afirmación que se hace donde 1= Totalmente en desacuerdo y 5= Totalmente de acuerdo. Cada afirmación consiste en una descripción del sector o una característica que hemos constatado en los anteriores epígrafes utilizando la revisión bibliográfica. También se han propuesto otras afirmaciones sobre el grado de importancia que los encuestados les otorgan a diversas materias planteadas en el trabajo y tenían que responder con otra escala de Likert donde 1= Muy poco y 5= Mucho. Por último también se han realizado algunas preguntas de opción múltiple.

La encuesta se ha realizado a través de la herramienta de formularios de Google (más conocida como Google forms). Primero se ha elaborado un borrador sobre el tipo de preguntas que se iban a realizar y posteriormente ese borrador se traspasó al cuestionario definitivo realizado en Google forms. En el anexo del trabajo se ha adjuntado la encuesta que se ha realizado.

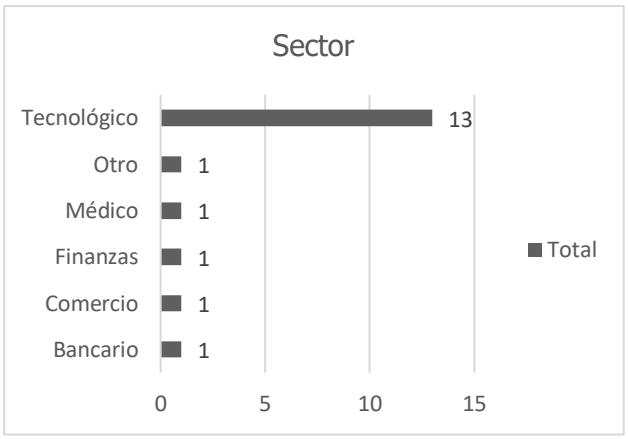
Este cuestionario estaba dirigido a profesionales que se dedican a usar la inteligencia de datos y que tuvieran mucha experiencia. Para aquellos profesionales que tuvieran una experiencia de 2 años o menos, se les ha requerido que tuvieran ciertos conocimientos y es por ello que se ha planteado una pregunta de control de tipo Likert con un baremo de uno a siete y que ha servido para eliminar aquellas respuestas de profesionales que consideraban que no tenían el suficiente conocimiento sobre la materia.

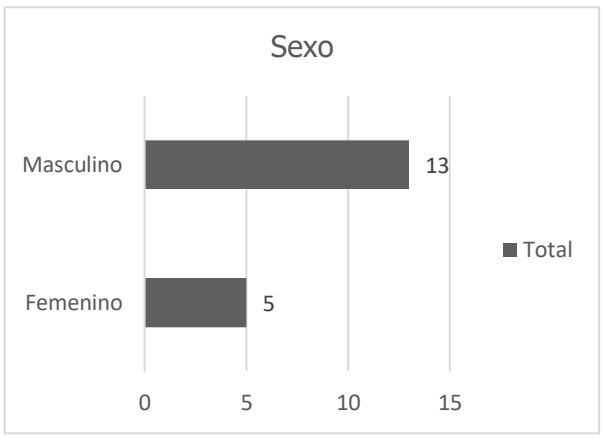
Para realizar la encuesta se ha contactado con profesionales de la inteligencia de datos a través de diversas redes sociales como Linkedin, Facebook y Twitter. En dichas redes sociales se ha buscado grupos y comunidades de profesionales que trabajaran en cualquier sector pero que usaran tecnología *big data.* Además, he podido contactar con algunos amigos y conocidos

que trabajan con macrodatos. En todo caso, se ha contactado individualmente con cada uno de los encuestados para poder enviarle la encuesta por medios telemáticos.

A continuación, vamos a analizar los resultados de dichas encuestas y primero vamos a analizar las estadísticas de los profesionales. Para ello hemos seleccionado las variables de sexo, sector en el que desarrollan su profesión y años de experiencia en el uso del *big data.*

Los resultados de estas variables nos sirven para intentar aclarar cuál es la composición de los profesionales del *big data.*





Figuras 1 y 2: Distribución de los profesionales encuestados por sexo y sector en el que trabajan. Fuente: Elaboración propia.

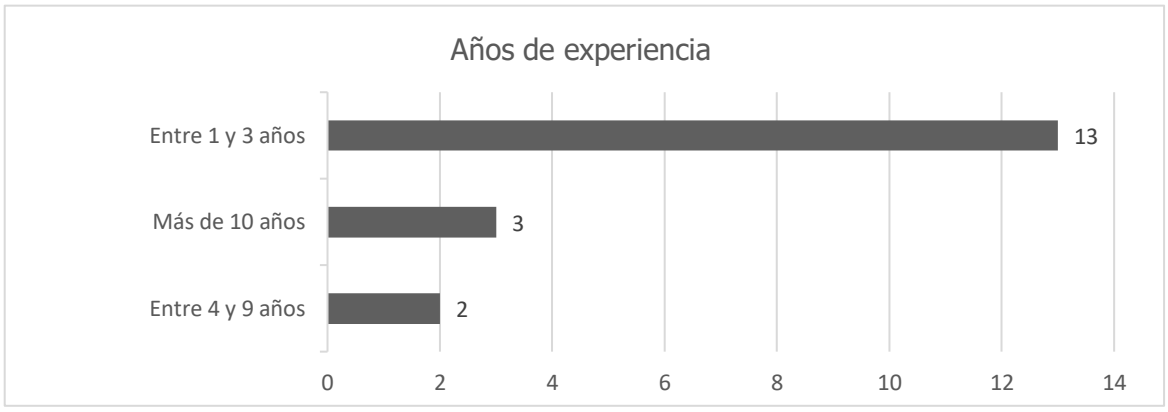


Figura 3: Distribución de los profesionales encuestados según los años de experiencia en el uso de *big data.* Fuente: Elaboración propia.

Estas tres primeras variables nos confirman que el *big data* es una innovación tecnológica y por eso más del 70% de los encuestados trabajan en el sector tecnológico. Además, hay una mayor presencia de hombres que de mujeres ya que tan sólo hay un 27,8% de mujeres frente al 72,2% de hombres.

Este hecho se debe a que históricamente ha habido una mayor presencia de hombre que de mujeres en los estudios universitarios de la rama tecnológica, aunque esta tendencia ya está cambiando y cada año la presencia de las mujeres en estos tipos de estudios y trabajos es mayor.

Respecto a los años de experiencia los resultados obtenidos tienen relación con el estudio realizado en la primera parte del trabajo ya que la inteligencia de datos ha sido una innovación reciente que, aunque se lleve usando hace varios años, ahora ha sido el auge de esta tecnología.

La mayoría de encuestados han comenzado a trabajar con *big data* en los últimos tres años frente a los profesionales que llevan más tiempo usándolo. Esto también se debe, a que la formación y los conocimientos se han ido desarrollando a la par que el auge de este fenómeno.

A partir de ahora vamos a analizar las opiniones que tienen los profesionales encuestados sobre una serie de afirmaciones que les hemos planteamos.

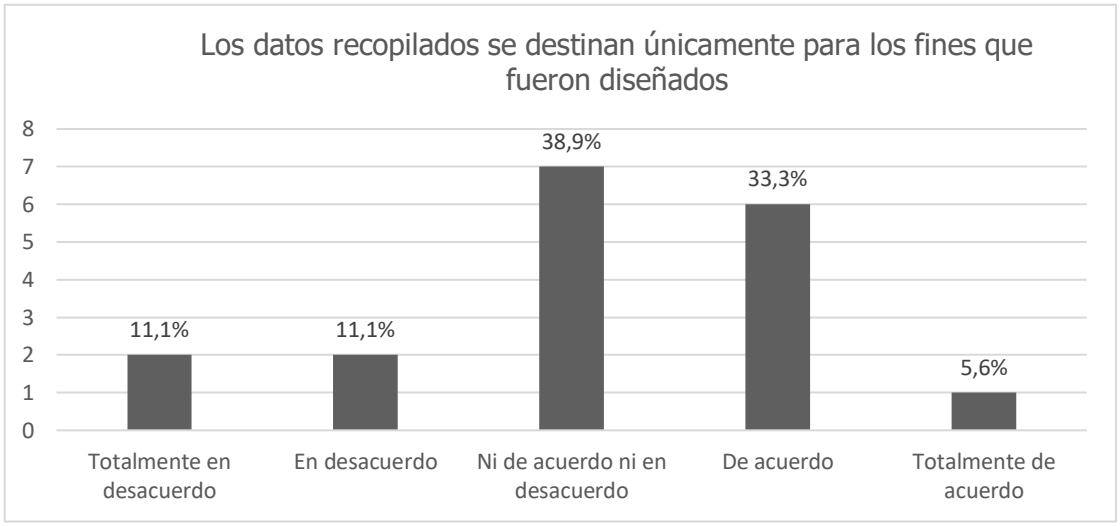


Figura 4: Opinión de los profesionales sobre el destino de los datos recopilados. Fuente: Elaboración propia.

En la figura cuatro nos encontramos que el 22,2% de los encuestados opina que los datos que se recogen para el análisis con técnicas de *big data* no sólo se usan con ese fin, sino que luego se le da un uso diferente mientras que el 39% está de acuerdo con la afirmación planteada.

Se puede confirmar que la normativa actual es insuficiente para garantizar la privacidad de las personas y que el artículo 25 RGPD que se ha estudiado en el epígrafo 4 de este trabajo no cumple la finalidad para la que fue diseñado. Esto es lo que podemos extraer de las opiniones de los profesionales encuestados.

En la figura cinco mostramos que hay bastante consenso entre los profesionales en cuanto a la responsabilidad que tienen los usuarios y clientes por la falta de interés que muestran sobre el destino de sus datos personales.

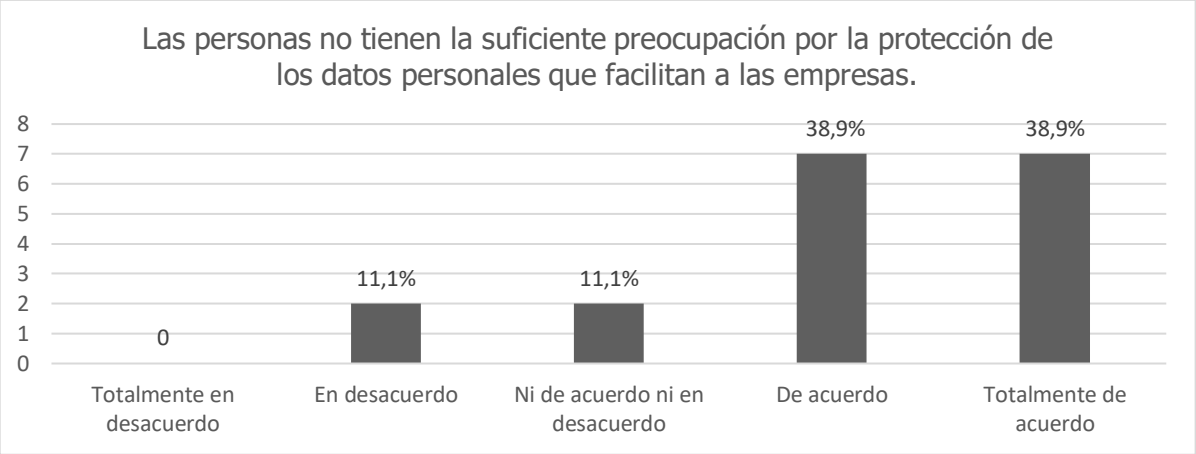


Figura 5: Criterio de los encuestados sobre la responsabilidad por falta de interés de los usuarios sobre sus datos personales. Fuente: Elaboración propia.

Los resultados de la figura cinco nos ratifican las conclusiones a las que se han llegado tras el estudio bibliográfico realizado en la primera parte acerca de la responsabilidad que tienen los propios titulares de datos.

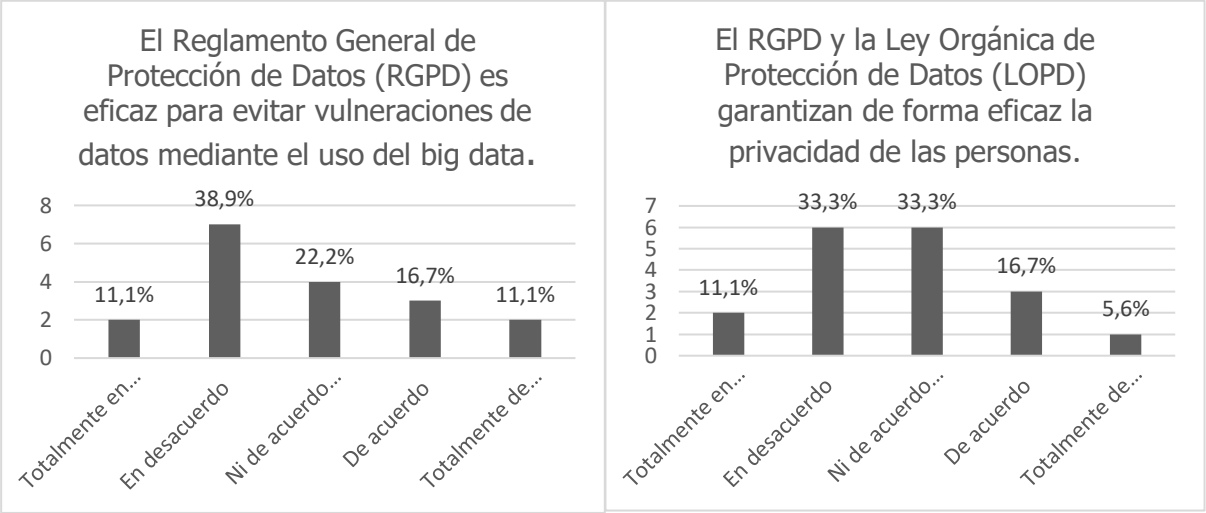
Es cierto que se producen vulneraciones a la privacidad y que continuamente se recopilan datos acerca de las personas, pero también es cierto lo que se ha planteado en la primera parte del estudio que los propios titulares deberían ser más responsables sobre la información que dan.

A continuación, los profesionales tuvieron que comparar, en base a su opinión, si la normativa española que rige actualmente respecto a la protección de datos es eficaz para evitar vulneraciones a la privacidad y si la garantizaban cuando se usa la inteligencia de datos en empresas españolas.

En las figuras seis y siete podemos comprobar que, en ambos casos, alrededor del 50% de los profesionales encuestados no creen que la LOPD y el RGPD sean eficaces para evitar vulneraciones de datos cuando se usen los macrodatos ni que ambas norman garanticen con eficacia la privacidad de las personas. Al frente nos encontramos con apenas un poco más de 22% que si considera que ambas normas cumplen la función para la que fueron elaboradas.

También preguntamos a los profesionales si consideran que debe plantearse un debate para mejorar el sistema existente del consentimiento que otorgan los usuarios para el tratamiento

de sus datos personales y el 77,8% de los encuestados respondieron estar totalmente de acuerdo con tal afirmación frente a un 11,1% en desacuerdo.



Figuras 6 y 7: Opinión de los profesionales sobre la eficacia de la LOPD y el RGPD. Fuente: Elaboración propia.

Las siguientes dos figuras nos muestran el grado de importancia y el grado de implantación que consideran, los profesionales que han realizado la encuesta, que tienen una serie de herramientas para garantizar el cumplimiento de los estándares ético-legales en el uso de la inteligencia de datos en organizaciones en España.

Estas medidas que se han planteado, son algunas de las mencionadas en la primera parte del trabajo. Las medidas planteadas a los profesionales de los macrodatos son los códigos de conducta realizados por las empresas, los protocolos de actuación en materia de protección de datos, el consentimiento del usuario al aceptar las condiciones de uso y políticas de privacidad, la técnica de anonimizar datos, el derecho que tienen los titulares de los datos para modificar o rectificar los datos que ha entregado y por último, el derecho que tienen los usuarios para acceder a la información que han facilitado a las organizaciones.

Todas estas medidas han sido planteadas en la primera parte del trabajo y sirven para intentar garantizar la privacidad y evitar vulneraciones de datos personales. Por eso, primero se ha cuestionado a los expertos sobre el grado de importancia que les otorgan a estas medidas para luego conocer la opinión que tienen respecto al grado de implantación de las mismas medidas en organizaciones españolas que aplican la inteligencia de datos.

En la figura 8 se puede apreciar que los protocolos de actuación, la anonimización y los derechos de revocación y acceso a la información son considerados por más del 66% de los expertos como muy importantes o bastante importante. En cambio, los códigos de conducta

sólo la mitad de los encuestados considera que es muy importante o bastante importante mientras que un 33% considera que es suficiente.

Lo que es más destacable es que tan sólo el 39% de los encuestados considera que el consentimiento de los usuarios es importante, y el 27,8% considera que es suficiente, mientras que el 33% considera que es poco o muy poco importante para garantizar la privacidad de las personas.

Esta cuestión lleva a la conclusión de que una de las medidas que más se aplica para garantizar la privacidad de las personas cuando se hace uso de la inteligencia de datos (figura 9), es insuficiente para cumplir su objetivo según el criterio de los profesionales de los macrodatos.

Esto se debe a que el consentimiento es la medida legal que se obliga a cumplir en todo caso por las empresas españolas y de fácil implantación. En cambio, los códigos de conducta que son voluntarios y de mayor dificultad de implantación que los protocolos de actuación.

En el epígrafe 3.2 se estudió la gran importancia que tiene que la implantación de un código de conducta solventaría muchos problemas y podría evitar posibles vulneraciones a la privacidad. Aun así, los códigos son percibidos de menor importancia.

Por último, también remarcable que entre el 10% y 15% de los encuestados consideran las medidas de la figura 8 muy poco o poco importantes, sí es destacable que el porcentaje crece hasta un 33% y 27% en el consentimiento y el derecho de acceso a la información respectivamente.

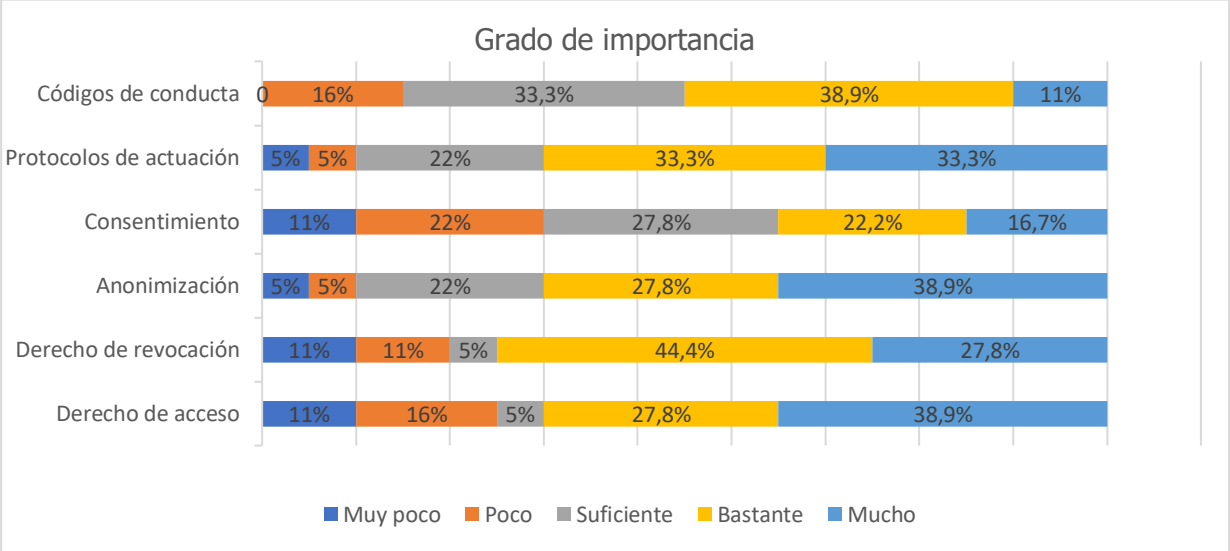


Figura 8: Opinión de los encuestados acerca del grado de importancia de diversas medidas para cumplir con los estándares éticos y legales en el uso del *big data* en España. Fuente: Elaboración propia.

En cuanto al grado de implantación que consideran los profesionales que existe en las organizaciones que usan inteligencia de datos en España, es notorio que los porcentajes decrecen, en el sentido de que no son semejantes las percepciones de importancia que tienen estas medidas respecto a las percepciones de instauración que tienen en las empresas españolas.

El consentimiento era la medida que menos importancia le otorgaban los profesionales consultados entre todas las planteadas. En cambio, el 50% considera que se aplica bastante o mucho en las empresas españolas y solamente un 10% considera que se aplica en muy pocas o pocas empresas.

Otro hecho destacable es el grado de implantación que consideran los profesionales que tienen los protocolos de actuación. Casi el 67% de los profesionales consideran se aplica en bastantes o muchas empresas mientras que tan sólo un 38% considera que los códigos de conducta se aplican en muchas o bastantes empresas.

Esta gran diferencia se puede explicar en base a lo estudiado en la primera parte del trabajo. Es más fácil para la empresa y los empleados, adoptar un protocolo de actuación en caso de vulneraciones de datos personales que adoptar toda una cultura en base a la protección de la privacidad.

Por último, también es remarcable que aproximadamente el 35% de los profesionales consideran que la anonimización y el derecho de revocación y de acceso a la información se aplica en las empresas españolas frente al 38,9% que considera que se aplica lo suficiente. Estas herramientas son de las más eficaces actualmente para garantizar la privacidad de las personas y evitar que se vulneren los derechos de las personas y según los profesionales son las medidas que menos se aplican en las empresas españolas que usan la inteligencia de datos. Lo que nos indica que hay mucho por hacer en cuanto al cumplimiento ético y legal del uso del *big data* en España.

La técnica de anonimizar datos es una medida que plantean los expertos consultados en la primera parte del trabajo, como solución para evitar vulneraciones a la privacidad. Es una medida obligatoria para empresas que comercien con datos o los cedan, pero para las empresas que recopilan datos para su uso exclusivo no es obligatorio y por eso, los profesionales consideran que esta medida se usa en menor medida.

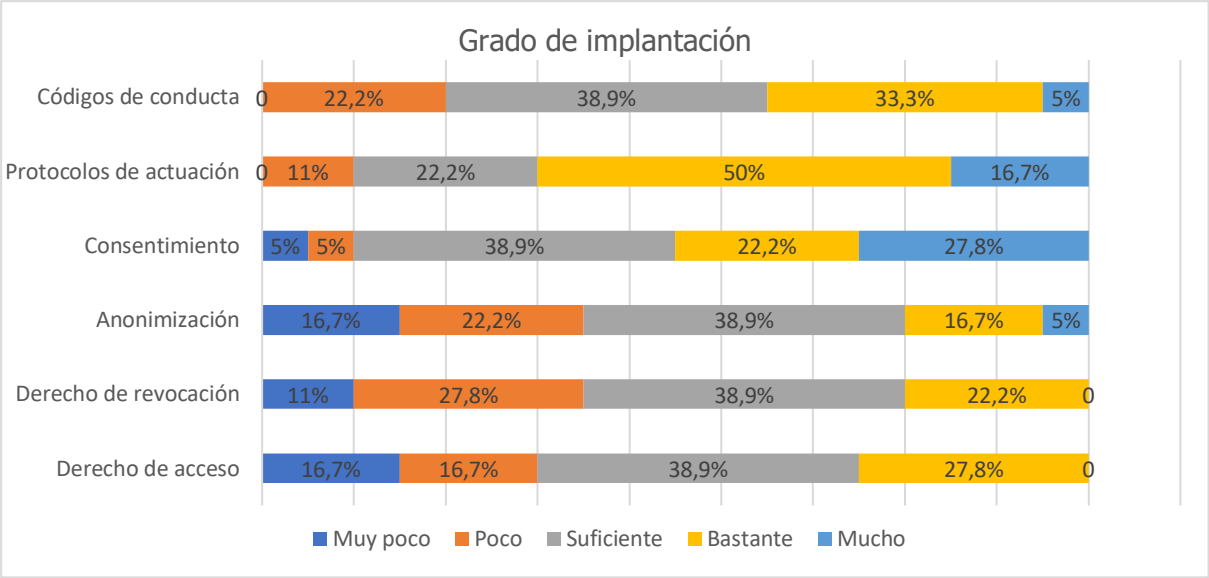


Figura 9: Criterio de los profesionales acerca del grado de implantación de diversas medidas en organizaciones españolas que usan *big data*. Fuente: Elaboración propia.

A continuación, los profesionales han valorado la necesidad de una serie de medidas que hemos estudiado en la primera parte del trabajo para proteger la privacidad de las personas en los casos de venta o cesión de datos personales a terceras personas.

Estas medidas son el consentimiento informado explícito, que significa que, en casos de cesión de datos a terceros, se debe volver a informar al titular de los datos de esta circunstancia y se debe explicar de una forma clara para que el consentimiento que otorga el titular sea eficaz. También hablamos de la transparencia en el proceso porque en muchos casos, se desconoce el lugar a donde van a parar los datos personales de las personas y las empresas ocultan a quienes ceden sus ficheros de datos. La última medida es la de técnica de anonimizar datos que sirve para eliminar toda característica que pueda vincular a una serie de datos con su titular cuando estos datos son cedidos a terceras personas.

En la figura 10 es destacable que existe casi unanimidad sobre la importancia y necesidad que le otorgan los profesionales al consentimiento informado explícito, la transparencia en el proceso y a la anonimización de datos. Tan sólo el 5% o el 10% de los encuestados consideran que estas tres medidas son muy poco o poco importantes.

Estas medidas son las que se han planteado en la primera parte para evitar que se vulneren los datos de las personas y que, según lo estudiado, no se cumplen porque continuamente ocurren escándalos como el de Cambridge Analitycs.

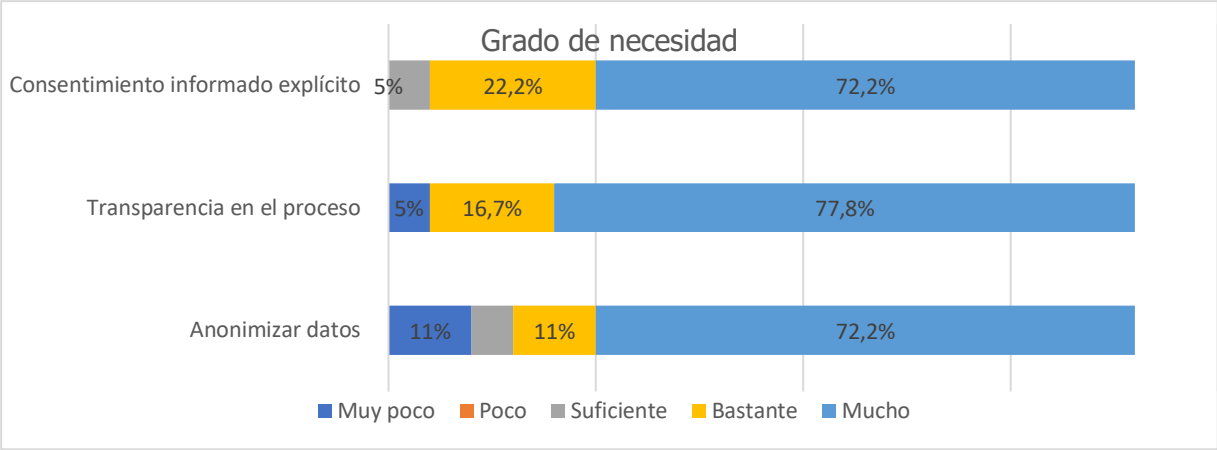


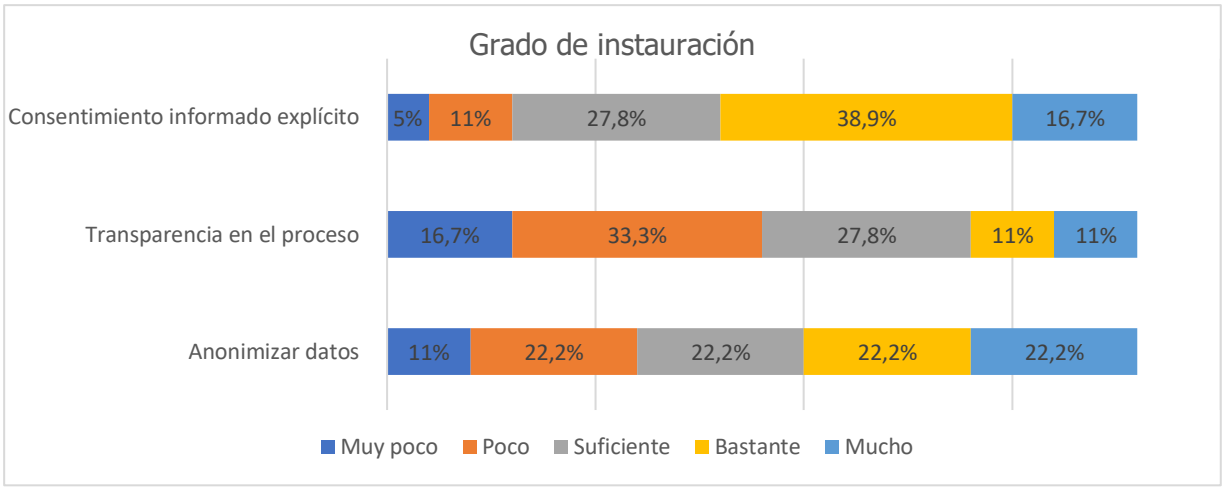
Figura 10: Opinión de los expertos en *big data* sobre medidas de protección de la privacidad en casos de cesión de datos a terceras personas. Fuente: Elaboración propia.

La cuestión cambia cuando se habla del grado de implantación de dichas medidas. Sólo el 22% de los profesionales consideran que hay muchas o bastantes empresas que sean transparentes en dichos procesos mientras que el consentimiento explícito, los expertos consideran que se aplica en una gran cantidad de empresas ya que el 27,8% considera que hay suficientes empresas que lo tengan implantando, mientras que el 38,9% opina que hay bastantes y un 16,7% muchas.

Respecto a la transparencia, se ha tratado en el trabajo que una mayor transparencia en los procesos supondría un aumento de la confianza de la sociedad en las empresas que realizan estas prácticas. El problema surge cuando estas empresas no muestran o tratan de esconder el tratamiento de dichos datos cuando se pide el consentimiento. Esta es una práctica muy poco ética que hemos estudiado ampliamente ya que las empresas usan lenguaje complejo y mucha información irrelevante para que los titulares de los datos muestren su consentimiento.

Tan sólo el 16% considera que hay pocas o muy pocas empresas que no soliciten el consentimiento en casos de cesiones de datos a terceras personas. Respecto a la anonimización de datos, hay divergencia de opiniones entre los expertos ya que casi se repiten los mismos porcentajes para las opiniones que consideran que hay muchas o bastantes empresas que anonimicen datos en casos de cesión de datos como que hay pocas o muy pocas empresas que lo hagan.

La técnica de anonimizar datos, como hemos visto, es una técnica compleja y que aun así no garantiza la protección de los datos, es por eso que los profesionales no están de acuerdo en base a esta técnica y su grado de instauración.



Fuente 11: Criterio de los profesionales del grado de instauración de medidas de protección de la privacidad en casos de cesión de datos a terceras personas. Fuente: Elaboración propia.

En la primera parte del trabajo, hemos dedicado un epígrafe a la gestión del cumplimiento de los estándares éticos y legales en el uso de la inteligencia de datos por parte de los equipos de *governance*.

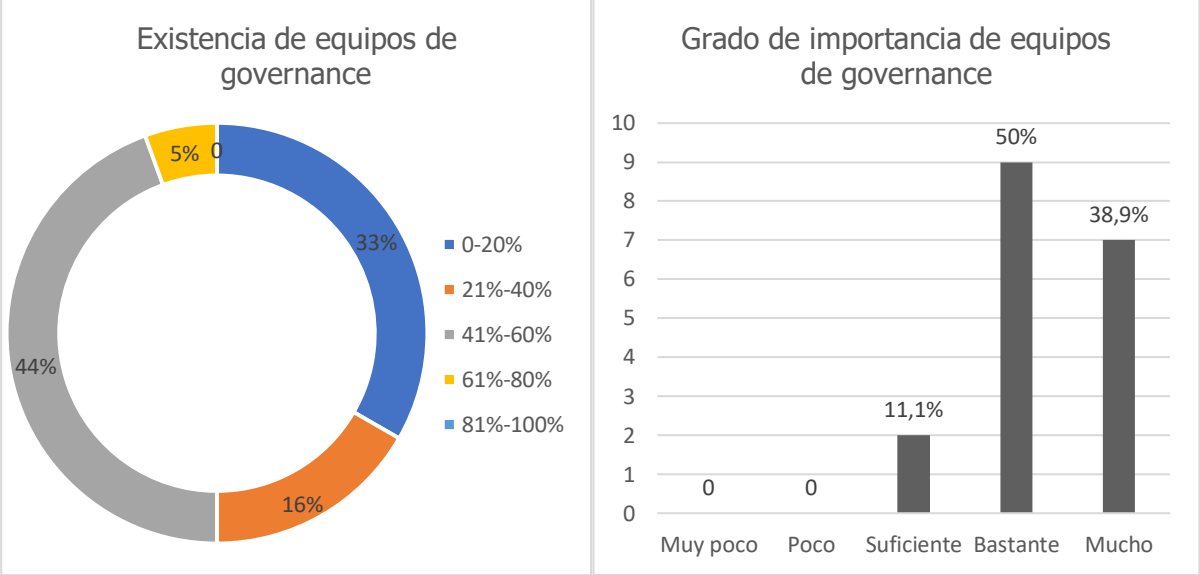
Por tanto, se les ha cuestionado a los profesionales de los macrodatos acerca del grado de importancia que, según consideren que tienen los equipos de *governance* y la existencia de dichos equipos en grandes empresas españolas.

En la figura 13 observamos que el 89% de los profesionales consideran que el gobierno de los datos es muy importante. El problema surge cuando volvemos a preguntar su opinión acerca del grado de implantación del gobierno de los datos en grandes empresas.

En la figura 12 observamos, que a pesar de otorgarle mucha importancia a los equipos de *governance*, el 33% de los profesionales consideran que existen equis de gobierno de datos en menos del 20% de las grandes empresas españolas. Mientras que el 44% de los expertos en *big data* consideran que hay equipos de *governance* entre el 40% y 60% de las empresas. Por último, tan sólo el 5% de los encuestados considera que la gran mayoría de grandes empresas españolas tienen equipos de gobierno de datos.

Aquí nos encontramos con otra gran diferencia entre lo estudiado en la primera parte del trabajo y con la opinión de los expertos. En el epígrafe quinto del trabajo se ha investigado acerca de la importancia de que tienen los equipos de *governance* para el cumplimiento ético-legal en el uso de los macrodatos. Esta divergencia entre la teoría y la práctica se debe a que el gobierno de los datos debe implantarse en grandes empresas ya que no tiene mucho sentido

en pequeñas y medianas empresas. Segundo, la implantación del *governance* es una tarea difícil y costosa para las empresas y aunque tenga muchos beneficios, como hemos estudiado, sigue siendo más sencillo y económico, asegurarse de cumplir la normativa que instaurar una cultura en la empresa en torno a la protección de la privacidad y los datos.



Figuras 12 y 13: Opinión de los profesionales en macrodatos acerca de la existencia de equipos de *governance* en grandes empresas españolas y la importancia que le atribuyen a los equipos de *governance.* Fuente: Elaboración propia.

Para finalizar, los profesionales que se dedican a la inteligencia de datos en organizaciones españolas opinaron sobre la importancia que tienen una serie de retos o desafíos de los que mencionamos en el epígrafe 2.3 de este trabajo en la adopción de proyectos de inteligencia de datos.

Los desafíos sobre los que se han tenido que pronunciar los profesionales son la adquisición y mantenimiento de conocimientos complejos para el uso de *big data,* la implantación de un sistema de seguridad efectivo, la dificultad en la recopilación y análisis de datos, la continua aparición y adaptación a nuevas tecnologías, el cumplimiento de la normativa española actual sobre protección de datos y el cumplimiento de los estándares éticos en materia de datos.

Como se observa en la figura 14, según la opinión de los profesionales, la implantación de un sistema de seguridad efectivo y el cumplimiento de la normativa son los mayores desafíos a los que se puede enfrentar una empresa cuando quiera adoptar la inteligencia de datos.

Como estudiamos en el epígrafe 2.3 del trabajo, es muy necesario implantar un sistema de seguridad efectivo ya que sino el proyecto fracasaría y tendrían pérdidas de reputación. Al

igual pasaría si se incumple con la normativa de protección de datos ya que se podrían ver envueltos en un escándalo, sufrir sanciones y pérdida de beneficios y de reputación. Por eso, el 95% de los encuestados considera que es muy importante o bastante importante con ambos casos para una correcta implantación de los macrodatos.

Como hemos visto en la primera parte del trabajo, el análisis de los datos no estructurados es una tarea de gran complejidad y es por eso que el 82% de los profesionales consideran que la complejidad en el análisis de los datos es el segundo gran desafío al que se debe hacer frente.

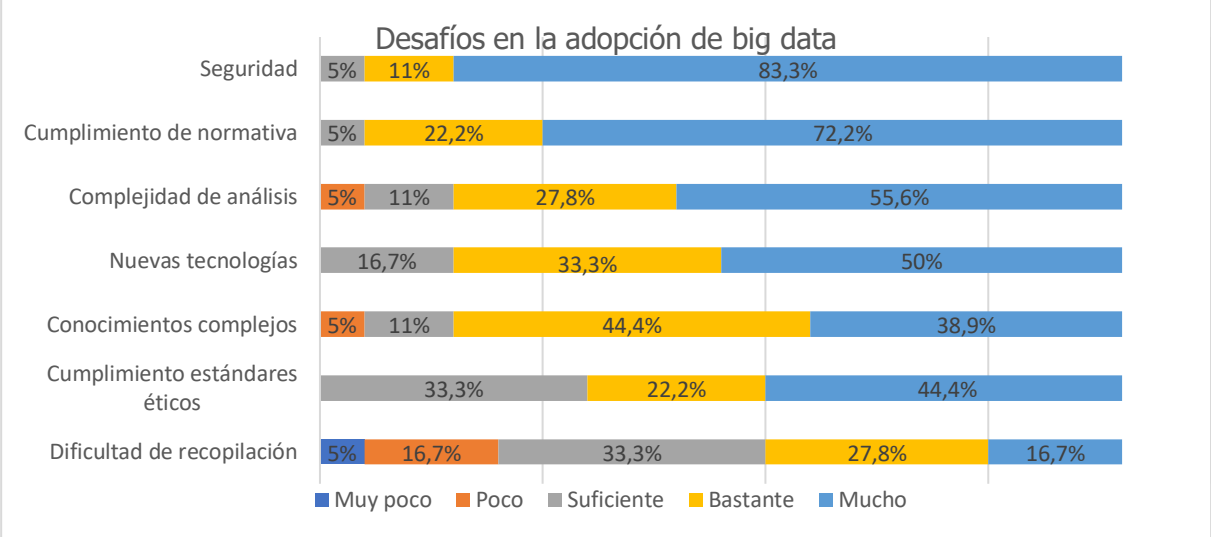


Figura 14: Criterio de los profesionales en inteligencia de datos acerca de los desafíos que supone adoptar un proyecto de *big data*. Fuente: Elaboración propia.

Las nuevas tecnologías y la adquisición y mantenimiento de conocimientos complejos han obtenido resultados similares debido a que ambos desafíos están relacionados. Ya que la continua aparición de nuevas tecnologías conlleva a que los profesionales de los macrodatos deban estar formándose y actualizando sus conocimientos constantemente. Por eso, en ambos casos, más del 80% de los profesionales considera que es muy importante o bastante importante este reto.

La dificultad de recopilación de datos es el menor de los desafíos planteados a los que tienen que hacer frente los profesionales de la inteligencia de datos y eso se debe a la inmensa cantidad de bases de datos que existen en la actualidad. También se debe a la aparición del internet de las cosas y los múltiples dispositivos que generan millones de datos. Por eso, tan sólo el 43% de los profesionales lo considera un gran reto.

Por último, debemos comentar el desafío que supone el motivo de la realización de este trabajo. El 44,4 % de los profesionales considera que es un reto muy importante cumplir con los estándares éticos mientras que el 33% considera que es suficiente pero no supone un gran reto.

La figura 14 confirma el motivo de la realización de este trabajo porque se ha investigado acerca de las aplicaciones de la inteligencia de datos, los beneficios que reporta a la sociedad y el aumento de su uso en los últimos años. Todo esto implica que los retos que hemos planteado en esta pregunta se van superando, pero es destacable que tenga más importancia, para los profesionales cumplir con los desafíos técnicos que los éticos.

En este estudio se ha intentado tratar sobre todos los aspectos éticos y legales que rodean a la inteligencia de datos porque es una innovación tecnológica muy potente pero que en esos aspectos no está tan avanzada como en las aplicaciones que tiene y eso se debe en parte a la importancia que le otorgan los profesionales y los responsables de su uso.

7. CONCLUSIONES

7.1 Implicaciones teóricas

En este trabajo se han estudiado las implicaciones éticas y legales que se derivan del uso del *big data*. Además, se ha analizado la gestión de los datos para cumplir con estas implicaciones.

El fenómeno *big data* ha transformado la sociedad en muchos sentidos. La inteligencia de datos tiene múltiples aplicaciones como ha intentado demostrar este trabajo. Por todo ello, ha supuesto una auténtica revolución para el ámbito empresarial ya que los beneficios que suponen una buena implementación de los macrodatos son enormes.

A pesar de todas las aplicaciones que se han estudiado en este trabajo, se ha manifestado las carencias existentes en materias legales y éticas. En el trabajo se han expuestos diversos casos de vulneraciones a la privacidad para tratar de demostrar esas carencias. Además, se han mostrado diversas medidas para evitar tales transgresiones como la obligatoriedad del consentimiento del titular de los datos o la anonimización.

Hemos tratado de informar sobre los aspectos en los que se centra la normativa actual para proteger la privacidad de las personas y aunque se cumpla con la normativa eso no significa

que se cumpla con los valores éticos ya que en este trabajo hemos tratado de probar que el cumplimiento legal no es suficiente para garantizar el cumplimiento ético.

Como hemos intentado mostrar en este trabajo, se intenta proteger una serie de derechos, pero la normativa actual no es efectiva porque siguen produciéndose transgresiones de la privacidad.

Es por todo esto, que es indispensable regular adecuadamente todos los aspectos legales sobre la inteligencia de datos para evitar que sigan produciéndose vulneraciones de la privacidad y que la sociedad no se sienta vigilada por los dispositivos que la rodean.

7.2 Implicaciones prácticas

Es necesario que se afronte un debate ético que sirva para mejorar la normativa actual ya que como hemos podido comprobar en la segunda parte del trabajo, todas las medidas y herramientas existentes que se han tratado en este estudio sirven para paliar estos defectos de ley y transgresiones a la ética, pero no para solucionarlos.

La mayoría de los profesionales de la inteligencia de datos, corroboran este problema que plantean los expertos en ética y *big data* que hemos analizado en la primera parte de este estudio.

En la práctica, teóricos y profesionales consideran que el sistema de consentimiento de los titulares de datos tiene múltiples defectos, la autorregulación es insuficiente, al igual que la técnica de anonimizar datos. Todas estas medidas son insuficientes para proteger los datos personales de la inteligencia de datos. Igualmente, las responsabilidades por incumplimiento ético o legal no están correctamente especificadas porque siguen produciéndose vulneraciones a la privacidad en la actualidad.

Por consiguiente, es imprescindible que expertos en materia de datos, ingenieros, expertos en ética, en derecho, empresarios, políticos asuman la responsabilidad para debatir y mejorar las consecuencias éticas y legales.

En el mundo empresarial, generalmente, las principales preocupaciones de las empresas son cumplir con la ley y obtener ingresos. Por tanto, les preocupa más aplicar correctamente todo aquello que les beneficie para conseguir esos objetivos y dejan en un segundo plano aquello que les afecta en menor medida, como es la ética.

Es por eso que en la práctica no existan las suficientes personas encargadas de la dirección y gestión de los datos o equipos de *governance* que garanticen el cumplimiento ético y legal de los macrodatos.

7.3 Limitaciones

Las principales limitaciones del cumplimiento ético y legal en la inteligencia de datos se deben a los tres actores principales que se han tratado en este trabajo y son los principales responsables para realizar un debate acerca de las implicaciones éticas y legales.

En primer lugar, las empresas, son los sujetos activos. Es decir, son las encargadas de aplicar los macrodatos y por tanto principales responsables del cumplimiento de la normativa y estándares éticos al usar la inteligencia de datos.

En segundo lugar, la sociedad, son los sujetos pasivos. Las personas son titulares de los datos que se usan para la inteligencia de datos. La sociedad también se beneficia de muchas de las aplicaciones que tienen los macrodatos.

El último protagonista son las administraciones y gobiernos encargados de realizar un marco normativo completo ineludible y que garantice plenamente los derechos de las personas.

Cada parte es responsable en la actualidad de que se produzcan vulneraciones a la privacidad, ya sea porque las empresas abusan de su situación para obtener más beneficios, la sociedad no se muestra inquieta ante el destino de sus datos personales o los gobiernos no actúan ante las continuas innovaciones tecnológicas provocando lagunas de legislación.

Hasta hace unos años, el fenómeno *big data* era algo que se asociaba al futuro, que sus aplicaciones iban a cambiar el mundo, pero ya no es algo venidero, sino que es el presente. La inteligencia de datos está transformando la sociedad en casi todos sus aspectos. Todos tenemos la obligación de trabajar para mejorar en los aspectos éticos y legales ya que siempre van a la cola de las innovaciones, aunque sea uno de los aspectos más importante de toda innovación porque es aquello que sirve de garantía para el futuro de cualquier avance en la sociedad.

BIBLIOGRAFÍA

Al-Badi, A., Tarhini, A., y Khan, A. I. (2018). Exploring big data governance frameworks. *Procedia Computer Science*, *141*, 271–277.

Asadi, I., Breidbach, C. F., Davern, M. J., y Shanks, G., (2016). Ethical implications of big data analytics. *Proceedings of the 24th European Conference on Information System (ECIS), Estambul*, Turquía. Recuperado de http://aisel.aisnet.org/ecis2016_rip/24

Barocas, S., y Nissenbaum, H. (2014). Big data's end run around anonymity and consent. En J. Lane, V. Stodden, S. Bender, y H. Nissenbaum (Eds.), Privacy, big data, and the public good frameworks for engagement (pp. 44-75). Cambridge: Cambridge University Press.

Buytendijk, F., y Heiser, J. (2013). Privacy and Ethical Concerns Can Make Big Data Analytics a Big Risk Too. *Gartner* (G00296953). Recuperado de https://www.gartner.com/en/documents/2358315/privacy-and-ethical-concerns-can-make-big-data-analytics

Buytendijk, F., y Laney, D. (2014). Information 2020: Beyond Big Data. *Gartner* (G00261907). Recuperado de https://www.gartner.com/en/documents/2681316/information-2020-beyond-big-data

Buytendijk, F., y Oestreich, T. (2015). Organizing for Big Data Through Better Process and Governance. *Gartner* (G00274498). Recuperado de https://www.gartner.com/en/documents/3002918/organizing-for-big-data-through-better-process-and-gover

Calvo, P., y Osal, C. (2018). Whistleblowing y datos masivos: Monitorización y cumplimiento de la ética y la responsabilidad social. *El profesional de la información, 27 (1)*, 173-184.

Chen, W., y Quan-Haase, A. (2018). Big Data Ethics and Politics: Toward New Understandings. *Social Science Computer Review*, 1–7. https://doi.org/10.1177/0894439318810734

Colmenarejo, R., (2018). Ética aplicada a la gestión de datos masivos. *Anales de La Cátedra Francisco Suárez*, *52*, 113–129.

Cotino, L. (2017). Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales. *Dilemata*, *24*, 131–150.

Cotino, L. (2019). Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y *big data* confiables y su utilidad desde el derecho. *Revista Catalana de Dret Públic, 58,* 29-48.

Dinh, L., Karmakar, G., Kamruzzaman, J., y Stranieri, A. (2018). Significance Level of a Big Data Query by Exploiting Business Processes and Strategies. *Baltic DB&IS 2018.* Conferencia llevada a cabo en el 13th Joint Proceedings of the Conference Forum and Doctoral Consortium, Trakai, Lituania.

Dorasamy, N., y Pomazalová, N., (2016). Social Impact and Social Media Analysis Relating to Big Data. En Z. Mahmood (Ed.) *Data Science and Big Data Computing: Frameworks and methodologies.* (pp. 293-313). Suiza: Springer International Publishing.

Duncan, A. D., Buytendijk, F., y Logan, V. (2016). How Chief Data Officers Show Leadership in Influencing the Data-Driven Culture. *Gartner* (G00304771). Recuperado de https://www.gartner.com/en/documents/3398017/how-chief-data-officers-show-leadership-in-influencing-t

Favaretto, M., De Clercq, E., y Elger, B. S. (2019). Big Data and discrimination: perils, promises and solutions. A systematic review. *Journal of Big Data, 6 (12),* 1-27.

Galimany, A., (2014). *La creación de valor en las empresas a través del Big Data* (trabajo fin de grado). Universidad de Barcelona, España.

Gómez-Barroso, J. L., Feijóo, C., y Martínez, D., (2017). Política antes que regulación: la protección de la información personal en la era del Big Data. *Economía industrial, 405,* 113-119.

González, P., (2017). Responsabilidad proactiva en los tratamientos masivos de datos. *Dilemata, 24,* 115-129.

Goodman, E., (2014). Design and Ethics in the Era of Big Data. *Interactions, 21* (3), 22–24.

Hammer, M. (2017). Research Ethics in Big Data. *Oncology Nursing Forum*, *44* (3), 293–295.

Hoffman, D. (Agosto de 2018). Privacy with Big Data: A Framework. *Digital Disruption, AMCIS 2018.* Conferencia llevada a cabo en el 24th Americas Conference on Information Systems, Nueva Orleans, Estados Unidos. Recuperado de: https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1095&context=amcis2018

IBM, (2012). Analytics: el uso de big data en el mundo real. Cómo las empresas más innovadoras extraen valor de datos inciertos.

Jin, X., Wah, B., Cheng, X., y Wang, Y. (2015). Significance and Challenges of Big Data Research. *Big Data Research*, *2* (2), 59–64.

Keffer, S. (2019). Too big to surveil : the fourth amendment illuminated by ' modern lights ' and shadowed by obsta principiis in a post- Carpenter world concerned with privacy. *Information & Communications Technology Law, 28 (2),* 161-185.

Jones, G., y George, J., (2014). *Administración contemporánea. Octava edición.* Madrid, España: McGraw-Hill.

Kemp, R. (2014). Legal aspects of managing Big Data. *Computer Law & Security Review*, *30*, 482–491.

Martínez, A., (2019), La inteligencia artificial, el big data y la era digital: ¿una nueva amenaza para los datos personales?. *La propiedad Inmaterial, 27,* 5-23.

Martínez, R., (2017). Cuestiones de ética jurídica al abordar proyectos de Big Data. El contexto del Reglamento general de protección de datos. *Dilemata*, *24*, 151–164.

Mayer-Schönberger, V., y Cukier, K., (2013). *Big Data. La revolución de los datos masivos.* Madrid, España: Turner Publicaciones S.L.

Monleon-Getino, A., (2015). El impacto del Big-data en la Sociedad de la Información. Significado y utilidad. *Historia y Comunicación Social*, *20* (2), 427–445.

Morte, R., (2017). ¿Protección de datos/privacidad en la época del Big Data, IoT, wearables…? Sí, más que nunca. *Dilemata, 24,* 219-233.

Nersessian, D., (2018). The law and ethics of big data analytics: A new role for international human rights in the search for global standards. *Business Horizons*, *61* (6), 845–854.

Nielsen, T., Buytendijk, F., McMullen, L., Lopez, J., Hunter, R., y Casper, C., (2015). What to Do if Your Digital Business Strategy Violates Culture, Ethics or the Law. *Gartner* (G00276440). Recuperado de https://www.gartner.com/en/documents/3095617/what-to-do-if-your-digital-business-strategy-violates-cu

O'Leary, D. E. (2016). Ethics for Big Data and Analytics. *IEEE Intelligent Systems*, *31* (4), 81–84.

Olesen, T. (2019). The Politics of Whistleblowing in Digitalized Societies. *Politics and Society, 47* (2), 277-297.

Ortiz, P. (2018). La protección de datos, un asunto profundamente humano. *Telos. Cuadernos de comunicación e innovación, 109,* 132–135.

Pérez, C., (2016). Aspectos legales del Big Data. *Índice, 68,* 18-21.

Royakkers, L., Timmer, J., Kool, L., y van Est, R., (2018). Societal and ethical issues of digitization. *Ethics and Information Technology, 20,* 2, 127-142.

Saltz, J., y Dewar, N. (2019). Data science ethical considerations : a systematic literature review and proposed project framework. *Ethics and Information Technology*, *21* (3), 197-208.

Scotti, V. (2017). Big data or big (privacy) problem? *IEEE Instrumentation and Measurement Magazine*, *20* (5), 23–26.

Someh, I., Davern, M., Breidbach, C., y Shanks, G., (2019). Ethical Issues in Big Data Analytics : A Stakeholder Perspective. *Communications of the Association for Information Systems*, *44,* 718-747.

Soto, Y., (2017). Datos masivos con privacidad y no contra privacidad. *Revista de Bioética y Derecho*, *40*, 101–114.

Spiekerman, S., Acquisti, A., Böhme, R. y Hui, K-L., (2015). The challenges of personal data markets and privacy. *Electronic Markets*, 25 (2), 161-167.

Sun, Z., Strang, K., y Li, R., (Octubre de 2018). Big Data with Ten Big Characteristics. *ICBDR 2018.* Conferencia llevada a cabo en el Association for Computing Machinery, Weihai, China.

Trom, L., y Cronje, J. (2019). Analysis of Data Governance Implications on Big Data. En Arai, K., & Bhatia, R., (Eds.) *Advances in Information and Communication* (Vol. 69), 645-654. Springer International Publishing.

Watson, H., (2014). Tutorial: Big Data Analytics: Concepts, Technologies, and Applications. *Communications of the Association for Information Systems, 34,* 1247-1268.

ANEXO: ENCUESTA

Estimado/a Sr/Sra:

Soy Alejandro Durán y estudio el doble grado en Administración de empresas y Derecho en la Universidad de Sevilla. Me dirijo a usted porque estoy realizando el Trabajo de Fin de Grado sobre **los aspectos éticos y legales del uso del *big data*** y me gustaría conocer la percepción que usted tiene como profesional, del uso del *big data* en España y sus aspectos éticos y legales. Estaría muy agradecido si pudiese realizar este breve cuestionario. No le supondrá más de 5 minutos en realizarlo.

La encuesta es totalmente anónima y la información obtenida será tratada de forma confidencial, únicamente con fines académicos.

Gracias de antemano por su colaboración.

Alejandro Durán Domínguez

1. Sexo:

	Masculino
	Femenino
	No procede

2. Sector en el que trabaja

	Finanzas
	Turismo
	Tecnológico
	Médico
	Seguros
	Bancario
	Administrativo
	Comercio
	Otro

3. Años de experiencia: (Cumplimentar)

4. Según su experiencia como profesional, ¿cómo calificaría su nivel de conocimientos en el análisis y aplicación de técnicas de big data? (Valore en una escala de 1 a 7)

1	Pésimo
7	Excelente

A continuación, se van a presentar diferentes afirmaciones sobre las implicaciones ético-legales del uso del big data. Por favor, valore el grado de conformidad que en su opinión tiene las siguientes cuestiones éticas y legales en el uso del *big data* en organizaciones españolas donde:

1=Totalmente en desacuerdo

2=En desacuerdo

3=Ni de acuerdo ni en desacuerdo

4=De acuerdo

5=Totalmente de acuerdo

5. Los datos recopilados se destinan únicamente para los fines que fueron diseñados.

1	Totalmente en desacuerdo
2	En desacuerdo
3	Ni de acuerdo ni en desacuerdo
4	De acuerdo
5	Totalmente de acuerdo

6. Las personas no tienen la suficiente preocupación por la protección de los datos personales que facilitan a las empresas.

1	Totalmente en desacuerdo
2	En desacuerdo
3	Ni de acuerdo ni en desacuerdo
4	De acuerdo
5	Totalmente de acuerdo

7. El Reglamento General de Protección de Datos (RGPD) es eficaz para evitar vulneraciones de datos mediante el uso del *big data.*

1	Totalmente en desacuerdo
2	En desacuerdo
3	Ni de acuerdo ni en desacuerdo
4	De acuerdo
5	Totalmente de acuerdo

8. El RGPD y la Ley Orgánica de Protección de Datos (LOPD) garantizan de forma eficaz la privacidad de las personas.

1	Totalmente en desacuerdo
2	En desacuerdo
3	Ni de acuerdo ni en desacuerdo
4	De acuerdo
5	Totalmente de acuerdo

9. Debe plantearse un debate en el campo del uso de *big data* para mejorar el sistema de consentimiento para el tratamiento de los datos personales.

1	Totalmente en desacuerdo
2	En desacuerdo
3	Ni de acuerdo ni en desacuerdo
4	De acuerdo
5	Totalmente de acuerdo

Por favor, valore según su opinión como profesional sobre las siguientes cuestiones ético-legales en el uso del *big data* en organizaciones en España utilizando la siguiente escala enumerada donde:

1=Muy poco

2=Poco

3=Suficiente

4=Bastante

5=Mucho

10. Según el RGPD, los datos deben usarse estrictamente para el fin por el que fueron recopilados. Valore, según su opinión como profesional, el grado de cumplimiento que tiene esta medida en las empresas situadas en España.

1	Muy poco
5	Mucho

11. Valore el grado de importancia que en su opinión tienen, las siguientes medidas para controlar que se cumple con los aspectos éticos-legales en el uso del *big data* en las organizaciones españolas:

- Códigos de conducta realizados por las empresas

1	Muy poco
5	Mucho

- Protocolos de actuación en materia de protección de datos

1	Muy poco
5	Mucho

- El consentimiento del usuario o consumidor al aceptar las condiciones de uso y políticas de privacidad

1	Muy poco
5	Mucho

- Anonimizar datos

1	Muy poco
5	Mucho

- Derecho de rectificación o revocación por parte del titular de los datos

1	Muy poco
5	Mucho

- Derecho de acceso a la información por parte del usuario o consumidor

1	Muy poco
5	Mucho

12. Valore el grado de implantación en empresas españolas que en su opinión tienen las siguientes medidas en el uso del *big data* en España.

- Códigos de conducta realizados por las empresas

1	Muy poco
5	Mucho

- Protocolos de actuación en materia de protección de datos

1	Muy poco
5	Mucho

- El consentimiento del usuario o consumidor al aceptar las condiciones de uso y políticas de privacidad

1	Muy poco
5	Mucho

- Anonimizar datos

1	Muy poco
5	Mucho

- Derecho de rectificación o revocación por parte del titular de los datos

1	Muy poco
5	Mucho

- Derecho de acceso a la información por parte del usuario o consumidor

1	Muy poco
5	Mucho

13. En casos de cesión de datos personales a terceras personas, ¿considera necesarias las siguientes medidas para proteger la privacidad?

- Consentimiento informado explícito

1	Muy poco
5	Mucho

- Transparencia en el proceso

1	Muy poco
5	Mucho

- Anonimizar datos

1	Muy poco
5	Mucho

14. ¿En qué medida considera, como profesional, que se aplican las siguientes medidas en los casos de cesiones de datos a terceras personas?

- Consentimiento informado explícito

1	Muy poco
5	Mucho

- Transparencia en el proceso

1	Muy poco
5	Mucho

- Anonimizar datos

1	Muy poco
5	Mucho

15. ¿En qué porcentaje de grandes empresas españolas considera usted que existen equipos de *governance?*

	0%-20%
	21%-40%
	41%-60%
	61%-80%
	81%-100%

16. Valore el grado de importancia que en su opinión tiene un equipo de governance en las organizaciones españolas para la gestión de los datos.

1	Muy poca

2	Poca
3	Suficiente
4	Bastante
5	Mucha

17. Por último, a continuación, se van a plantear un conjunto de retos o desafíos que se deben adoptar para el uso del *big data* en España. Valore qué grado de importancia tienen estos desafíos según su opinión como profesional para un correcto uso del *big data* en las organizaciones situadas en España.

- Adquisición y mantenimiento de conocimientos complejos para el uso de big data

1	Muy poca
5	Mucha

- Implantación de un sistema de seguridad efectivo

1	Muy poca
5	Mucha

- Dificultad en la recopilación de datos

1	Muy poca
5	Mucha

- Complejidad en el análisis de datos

1	Muy poca
5	Mucha

- Aparición y adaptación a nuevas tecnologías

1	Muy poca
5	Mucha

- Cumplimiento de la LOPD y RGPD

1	Muy poca
5	Mucha

- Cumplimiento de los estándares éticos sobre los datos personales de las personas

1	Muy poca
5	Mucha

Si desea recibir los resultados del estudio, puede dejar una dirección de correo electrónico a efectos de enviar los resultados. (Cumplimentar).

Muchas gracias por su colaboración.