Universidad de Alcalá

EL IMPACTO DEL BIG DATA EN LA PROTECCIÓN DE DATOS PERSONALES

THE IMPACT OF BIG DATA ON PERSONAL DATA PROTECTION

Máster Universitario en Acceso a la Profesión de Abogado

Presentado por:
Dª ANA LUCÍA MACHUCA GONZÁLEZ
Dirigido por:
Dª MÓNICA ARENAS RAMIRO

Alcalá de Henares, a 28 de octubre de 2020

ÍNDICE

1. INTRODUCCIÓN

Es bien sabido que las tecnologías son un tema de actualidad que avanzan a pasos agigantados, y que irrumpen en la mayoría de los aspectos de nuestra vida. Día a día, la tecnología de *Big Data* es un tema emergente que proporciona numerosos beneficios en nuestra sociedad, pero que crea un nuevo paradigma que representa un reto para la protección de los datos personales de los individuos y su privacidad. Por ello, he creído conveniente abarcar en este Trabajo de Fin de Máster (TFM) la aparente contradicción entre la protección de datos personales y el tratamiento masivo de datos.

Con las tecnologías y, principalmente, con Internet, la privacidad se ha visto amenazada en estos últimos años. No obstante, la entrada en vigor de la nueva normativa de protección de datos ha supuesto un gran avance legislativo, pues la antigua normativa se encontraba desfasada frente a la avanzada tecnología, y no abarcaba todas las finalidades y aplicaciones que hoy día tienen las nuevas tecnologías. Y es que día a día millones de personas se conectan a Internet para compartir y adquirir todo tipo de información de cualquier clase, haciendo uso de plataformas como redes sociales, blogs y otros. Sin embargo, existe la posibilidad de acceder a información de manera que anteriormente no controlábamos, como tu localización precisa exacta las 24 horas del día, la velocidad a la que caminas, con quién y cómo hablas, la identidad de tus amigos, tu estado de salud y mucha más información, todo ello gracias a los dispositivos electrónicos que llevamos día a día encima a todas partes.

Pese a que lo anterior puede poner en riesgo nuestra privacidad y datos personales, el tratamiento masivo de datos trae consigo un gran avance y muchos beneficios en nuestras vidas y en las empresas. Gracias al tratamiento masivo de datos podemos gestionar aspectos como el tráfico de las ciudades, convirtiéndose en “ciudades inteligentes”, también aspectos como control médico en tiempo real, mejoras en la salud pública al recopilar datos masivos que se utilizan en la codificación de material genético. Asimismo, para las empresas el tratamiento masivo de datos o *Big Data* se convierte en un instrumento muy útil que ayuda a la toma de decisiones estratégicas de las empresas, aumentando así sus beneficios.

Debemos encontrar el equilibro entre la necesidad y las ventajas de utilizar técnicas de *Big Data* y la obligación de cumplir con las normas de protección de datos personales, como el Reglamento (UE) 679/2016, General de Protección de Datos (RGPD) y la Ley Orgánica 3/2018, de Protección de Datos y Garantía de los Derechos Digitales (LOPDGDD).

A lo largo de este trabajo intentaremos explicar qué es el *Big Data* y cómo puede encuadrarse dentro de la normativa europea y española de protección de datos. Entenderemos la gran aplicabilidad del *Big Data* en diversos aspectos de nuestra vida, así como el peligro que supone su divulgación y su utilización de manera incorrecta. También analizaremos los pros y los contras respecto de la protección de datos personales. Consideraremos además situaciones mejorables para la privacidad de los usuarios, como es el consentimiento expreso para la anonimización de datos personales y la utilización de estos con fines estadísticos y de mejora, así como posibles medidas tecnológicas para la mejora de la privacidad, seguridad y confianza. El trabajo finaliza con conclusiones donde se recogen los resultados más relevantes obtenidos en el mismo y con una bibliografía donde aparecen las fuentes consultadas.

2. BIG DATA

2.1. LOS DATOS PERSONALES

Sin perjuicio de que luego nos detengamos en el análisis de lo que son los datos personales y los requisitos para su tratamiento, debemos señalar que un dato personal es "*toda información sobre una persona física identificada o identificable («el interesado»); se considerará persona física identificable toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un identificador, como por ejemplo un nombre, un número de identificación, datos de localización, un identificador en línea o uno o varios elementos propios de la identidad física, fisiológica, genética, psíquica, económica, cultural o social de dicha persona*"[1].

De manera general, cuando hablamos de datos hablamos de datos estructurados, y nos referimos a la información que se suele encontrar en la mayoría de bases de datos. Son archivos de tipo texto que se suelen mostrar en filas y columnas con títulos. Son datos que pueden ser ordenados y procesados fácilmente por todas las herramientas de minería de datos. Lo podríamos ver como si fuese un archivador perfectamente organizado donde todo está identificado, etiquetado y es de fácil acceso[2].

Un ejemplo de lo que se entiende por datos estructurados se produce cuando dejamos un mensaje telefónico a alguien. Por ejemplo, "María llamó a las 2 p.m. para preguntar sobre el pago de la factura". Esta oración normal se considera no estructurada porque los datos que contiene no están categorizados. Las máquinas no podrían entender esta oración. Si esta información fuera estructurada, tendríamos diferentes categorías como hora, nombre y mensaje:

Nombre	**Hora**	**Mensaje**
María	2 p.m.	Pago de la factura

[1] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos). Publicado en el DOUE el 4 de mayo de 2016. Art. 4.1. definición de datos personales: https://www.boe.es/doue/2016/119/L00001-00088.pdf

[2] Kyocera, definición de *datos estructurados*: https://www.kyoceradocumentsolutions.es/es/smarter-workspaces/insights-hub/articles/diferencia-entre-datos-estructurados-y-no-estructurados.html

Cuando los datos se estructuran de esta manera utilizando campos que son ampliamente reconocidos, las máquinas pueden comenzar a comprender de qué se trata el contenido[3].

Además, los datos estructurados pueden tener distintas fuentes[4]:

- Creados por la empresa: registros en tablas, ficheros XML, etc.
- Provocados: datos creados de manera indirecta a partir de una acción previa, como pueden ser valoraciones de restaurantes o de películas.
- Dirigidos por transacciones: datos que tienen lugar al finalizar una acción previa de manera correcta. Son este tipo de datos las facturas de compra o recibos de un cajero.
- Compilados: resúmenes de datos de empresa o servicios públicos de nivel grupal, como el censo electoral, vehículos matriculados, etc.
- Experimentales: datos generados como parte de un análisis.

Por otro lado, los datos no estructurados son aquéllos que no tienen una estructura específica, que no están procesados o tratados conforme a un criterio o sistema. Manipular este tipo de datos es algo más complejo, y no es posible su almacenamiento en una tabla como sí sucede en los estructurados. Son datos no estructurados los archivos multimedia, archivos PDF o Word, contenido de emails, comentarios en las redes sociales o interacciones con otros usuarios, pero esto no significa que no se traten de datos personales y deben ser también protegidos[5].

Tal y como ocurre en los datos estructurados, los no estructurados también proceden de distintas fuentes. Capturados, esto es, datos creados a partir del comportamiento de un usuario. Estos datos pueden ser extraídos a partir de aplicaciones de seguimiento de actividades (carrera, ciclismo, natación), o posición GPS.Y, por otro lado, los datos personales también pueden ser generados por los usuarios, esto es, datos que especifica un usuario, como son las publicaciones en redes sociales o vídeos reproducidos en YouTube.

[3] WooRank, ejemplo para entender qué es un dato estructurado: https://www.woorank.com/es/edu/seo-guides/que-son-los-datos-estructurados

[4] ENEB, fuentes de datos estructurados: https://claustroeneb.es/2019/01/09/tipos-y-calidad-de-datos-en-big-data/

[5] ENEB, definición de *datos no estructurados*: https://claustroeneb.es/2019/01/09/tipos-y-calidad-de-datos-en-big-data/

Finalmente, nos encontramos con los datos híbridos o semiestructurados, que podrían entenderse como una combinación de los anteriores. Estos datos no tienen una estructura fija como los datos estructurados, sin embargo, están organizados mediante metadatos ("datos sobre datos", información asociada) o mediante relaciones simples entre ellos. Debido a estas características, los datos semiestructurados son más fáciles de procesar que los datos no estructurados. Un ejemplo de datos semiestructurados son los datos almacenados en JSON o XML[6].

2.2. DEFINICIÓN DE *BIG DATA*

Con el nacimiento de Internet, sobre todo con la llegada de las redes sociales y, por consiguiente, la generación de grandes volúmenes de datos, surge el concepto de *Big Data.*

Big Data es una palabra de origen anglosajón, la cual el científico informático John Mashey utilizó por primera vez en el año 1998, en su artículo publicado en el New York Times denominado "*Big Data and the Next Wave of Infrastress*". En él, Mashey definió *Big Data* como "un término que se aplica a sets de datos cuyo tamaño está más allá de lo que las herramientas de software habitualmente utilizadas pueden capturar, administrar y procesar en un período de tiempo razonable". Mashey también predecía el estrés que iban a padecer las infraestructuras físicas y humanas *("infraestres*") de la informática ante el imparable crecimiento de datos que ya se avistaba.

Actualmente existen diferentes tipos de definiciones de *Big Data* de distintos autores, sin existir un consenso claro en cuanto a su definición. Sin embargo, de manera general, presentaremos dos definiciones de las más relevantes:

- Tal y como señala Gartner, *Big Data* se puede definir como información de gran volumen procesada a gran velocidad y muy variada que requiere sistemas de información innovadores y efectivos para poder facilitar la obtención de conocimiento y la toma de decisiones[7].

[6] Definición de *datos semiestructurados*: https://aprenderbigdata.com/que-es-el-big-data

[7] Definición de Garner sobre el *Big Data*: https://blogs.gartner.com/svetlana-sicular/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/

- Por otro lado, O'Reilly Media define el *Big Data* como los datos que sobrepasan la capacidad de procesamiento de las bases de datos tradicionales. Los datos se mueven demasiado rápido o no cuadran en la arquitectura tradicional. Para obtener valor de estos datos, deben buscarse sistemas tradicionales para procesarlos[8].

De las anteriores definiciones podemos extraer tres características comunes: las denominadas tres V del *Big Data*: volumen, velocidad y variedad.

La primera de las V es Volumen, donde usaré la definición de la Agencia Española de Protección de Datos (en adelante, AEPD), la cual denomina Volumen como "la característica más obvia y que recoge el propio nombre de *Big Data*. Se pasa de manejar magnitudes de megabytes, gigabytes, como mucho Terabytes, a manejar Petabytes (1.000.000.000.000.000 Bytes) de forma cada vez más frecuente"[9]. Recientemente se ha pasado a manejar Zettabytes (1.099.511.627.776 Gigabytes) e incluso Yottabytes, no siendo esta última siquiera la capacidad más grande de almacenamiento que existe. Según un estudio publicado en 2014 por EMC Coportarion (actual Dell EMC) a lo largo de este año 2020 se prevé alcanzar los 44 Zettabytes de información almacenada en todo el mundo[10]. Para poder hacernos una idea de la gran velocidad con la que crecen cada año, Jesús Morillo, columnista en El Nacional, nos hace una comparativa con el año 2008, en el cual sólo Google procesaba 24 Petabytes de datos al día[11].

La segunda de las V es Variedad, la cual se enmarca en la naturaleza de los datos y según la AEPD ha crecido de manera exponencial, tanto por la tipología de datos como por sus fuentes. Se ha pasado de manejar datos estructurados en bases de datos procedentes, en su mayoría, de fuentes internas, a tratar datos estructurados, semiestructurados y desestructurados, los cuales definiremos más adelante a lo largo de este trabajo. También se ha pasado de ser datos cuasi estáticos, sin sufrir ninguna variación, a datos

[8] Definición de O'Reilly Media sobre el *Big Data:* https://www.oreilly.com/radar/what-is-big-data/

[9] AEPD, definición de *Volumen*: https://www.aepd.es/sites/default/files/2019-09/guia-codigo-de-buenas-practicas-proyectos-de-big-data.pdf

[10] DELL EMC, data growth study: https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm

[11] El Nacional, artículo de Jesús Morillo: https://www.elnacional.com/opinion/columnista/explosion-generacion-datos_283236/

dinámicos o en continuo cambio; de originarse en un número de fuentes limitadas a proceder de personas, máquinas, sensores, etc. Utilizando también la definición de "Variedad" de la AEPD, esta variedad y volumen requieren un tratamiento diferente para poder convertirse en información[12]. De manera que con Variedad nos referimos a las distintas fuentes de datos y a los diferentes tipos de archivo o formato de estos. Así, cuantas más fuentes tengamos, siempre que podamos relacionar estos datos, conllevará tener una base de datos superior o más completa. Los datos pueden proceder de diversos sitios, tales como redes sociales, dispositivos de tecnología de Radio frecuencia, smartphones o encuestas. Estos datos permiten conocer información como hábitos de vida, dispositivos electrónicos conectados a la red, e-mails, páginas webs, blogs, etc. La innovación principal es tratar todos estos datos de diferentes fuentes, estructuradas, semiestructuradas o desestructuradas, es decir, tal y como hemos mencionado anteriormente, datos dinámicos en continuo cambio a como era anteriormente (datos estáticos y de fuentes internas).

La última de estas tres V es la Velocidad. Tal y como señala la AEPD, el tiempo es clave, y la captura, movimiento y proceso de los datos se hace a gran velocidad, llegando a ser en tiempo real en algunos casos[13]. Se requiere que los datos sean procesados en el mínimo tiempo posible. También se necesita que los datos se produzcan, procesen, analicen rápidamente para conocer el resultado en muy poco tiempo y así llevar a cabo la mejor acción para nosotros y nuestro negocio. Así, por ejemplo, si un negocio se encuentra en plena campaña, requerirá seguramente analizar los comentarios de sus seguidores a tiempo real, para ir modificando y mejorando su actuación promocional. También, el análisis en tiempo real puede ayudar a seguir la trayectoria e intensidad de un huracán, pudiendo realizar predicciones de dónde puede producir daños con horas o días de antelación[14].

[12] AEPD, definición de *Variedad*: https://www.aepd.es/sites/default/files/2019-09/guia-codigo-de-buenas-practicas-proyectos-de-big-data.pdf

[13]AEPD, definición de *Velocidad*: https://www.aepd.es/sites/default/files/2019-09/guia-codigo-de-buenas-practicas-proyectos-de-big-data.pdf

[14] GIL GONZÁLEZ, E., *Big Data, privacidad y protección de datos*, Boletín Oficial del Estado, Madrid, 2016.p.22

Además de ello, debemos considerar el hecho de que las empresas reciben a diario una gran cantidad de datos sobre sus clientes a una velocidad realmente alta, incluso, como hemos dicho, a tiempo real, implicando ello la necesidad inmediata de análisis.

La velocidad es de una importancia elevadísima en la transmisión y en el procesamiento y análisis de los datos, 5 minutos nos resulta poco tiempo, sin embargo, en este contexto pueden ser más que suficientes para detectar un fraude en una transacción. En otras ocasiones es necesario que los datos se analicen a tiempo real, Google sugiere palabras para nuestra búsqueda en la medida que vamos introduciéndolas, no pasando varios segundos después. No disponer de la información en el momento oportuno no sólo resta valor a la misma, sino que inutiliza su aplicación.

Estas tres V pueden ser, además, ampliadas con otras V que se han ido introduciendo conforme se investigaba en el uso y tratamiento de la información: Veracidad, Valor, Verificación, Variabilidad y Viabilidad.

La Veracidad hace referencia al nivel de fiabilidad o calidad de los datos. Obtener datos de calidad es indispensable para la correcta obtención de los datos. Si un dato no está estructurado puede crear incertidumbre. Así, la captura de datos de diversas fuentes puede ayudar a minimizar la incertidumbre, al igual que la verificación o limpieza, para poder sacar el máximo provecho de estos y ser lo más fiables posibles[15]. El concepto "*Garbage in, garbage out*" (*GIGO*) es muy utilizado en el campo de la información, y hace referencia a que la entrada de los datos sin sentido provoca la salida de información también sin sentido[16].

Valor no hace referencia únicamente al valor monetario de los datos, sino al valor informativo que poseen. La finalidad de hacer un tratamiento de datos masivos es obtener un valor añadido que reporte conocimiento. Para las empresas resulta, por lo tanto, de gran utilidad que la gestión y el análisis de los datos e informaciones ayuden a crear valor, que será percibido por los clientes gracias a las acciones que llevan a cabo[17].

[15] ENEB, definición de *Veracidad*: https://claustroeneb.es/2019/01/09/tipos-y-calidad-de-datos-en-big-data/

[16] Wikipedia, definición de *GIGO*: https://es.qwe.wiki/wiki/Garbage_in,_garbage_out

[17] ENEB, definición de *Valor*: https://claustroeneb.es/2019/01/09/tipos-y-calidad-de-datos-en-big-data/

La Visualización de los datos es fundamental para poder comprenderlos y tomar decisiones en consonancia. Las plataformas que gestionan los datos deben tener en cuenta la forma en la que se presentan los datos. Para que podamos extraer de manera sencilla la información de los datos masivos, es importante que visualmente se muestren de un modo práctico y que además vayan acompañados de un contexto para que el análisis no sea tan complejo[18]. Se conoce como *visual analytics* (VA) al campo de investigación que estudia y explora soluciones de visualización. Este se encarga de que la complejidad que reside en los datos masivos y el exceso de información que las empresas pueden encontrar, se transforme en una oportunidad para ellas y su negocio.

Verificación para asegurar y confirmar la integridad de los datos, especialmente la de aquellos que procedan de fuentes externas o los que proceden de la nube. La verificación puede darse mediante certificados o firmas digitales[19].

La Variabilidad es la característica que más se relaciona con otra de las V: la Velocidad. Esto es así debido a que los datos van cambiando, surgen nuevos y otros resultan por ello obsoletos[20].

Como última de las V tenemos la Viabilidad, y es que cuando una empresa quiere llevar a cabo un proyecto de *Big Data*, debe tener en cuenta con qué herramientas e infraestructuras cuenta, las que necesitan para llegar a su objetivo y calcular los costes de estos, ya que cada plataforma y software tiene unas características distintas, y por ello, su coste también es distinto. La empresa debe poder hacer frente a los gastos, y que estos estén justificados y sean los necesarios para lograr su objetivo y sacar beneficios para su negocio[21].

2.3. VENTAJAS DEL *BIG DATA*

El análisis de un gran volumen de datos a una velocidad antes inimaginable genera grandes oportunidades y beneficios, principalmente en la

[18] ENEB, definición de *Visualización*: https://claustroeneb.es/2019/01/09/tipos-y-calidad-de-datos-en-big-data/

[19] ENEB, definición de *Verificación*: https://claustroeneb.es/2019/01/09/tipos-y-calidad-de-datos-en-big-data/

[20] ENEB, definición de *Variabilidad*: https://claustroeneb.es/2019/01/09/tipos-y-calidad-de-datos-en-big-data/

[21] ENEB, definición de *Viabilidad*: https://claustroeneb.es/2019/01/09/tipos-y-calidad-de-datos-en-big-data/

economía. Resulta bien acertado citar la frase "información es poder", y es que son numerosas las ventajas que genera en las empresas, dependiendo fundamentalmente de las estrategias que estas se marquen[22].

La cantidad de datos generados en los últimos años supera a la previamente generada en toda la historia de la humanidad. Nos encontramos en la denominada "Era de los datos", los datos son considerados como el "nuevo petróleo", lo cual implica que el hecho de disponer de un gran volumen de datos estructurados que se puedan interpretar ayuda a las empresas a tomar mejores decisiones. Es fundamental abordar los datos de manera inteligente.

La tecnología del *Big Data* ha supuesto un análisis de comportamientos de clientes que antes no se encontraba al alcance, ayudando a las empresas a tomar decisiones acertadas, reduciendo los riesgos que conlleva a una empresa tomar decisiones dejándose llevar por la intuición. El hecho de que las empresas tengan acceso a mayor número de información aumenta también la velocidad en la que se toman estas decisiones, abordando de manera más eficiente el mundo globalizado. Las empresas pueden combinar distintos tipos de datos para realizar segmentación de clientes y campañas de marketing y publicidad centradas en clientes potenciales, convirtiéndose en las empresas más competitivas[23].

Además de tomar mejores decisiones gracias al análisis de datos, las empresas pueden llegar a soluciones diferentes que no podían plantearse en un principio. Ejemplo de esta situación es la Inteligencia Artificial, en concreto el aprendizaje automático, en la cual se enseña a un ordenador a resolver cuestiones por sí mismo a partir de los datos que se le transfiere, con lo cual es de gran ayuda a la hora de entender el origen de un problema o a hallar la solución óptima. Esto fue lo que hizo el informático Arthur Samuel de International Business Machines Corporation (IBM) en los años 50.

El Instituto de Ingeniería de la Universidad Autónoma de Madrid, en un post de 2016[24], nos cuenta cómo al informático le apasionaba jugar a las damas, por lo que diseñó un programa específico para que su máquina *Defense*

[22] VALLS GIMÉNEZ, J. F., *Big Data: atrapando al consumidor*, Profit Editorial, Barcelona, 2017.p.143.

[23] GIL GONZÁLEZ, E., *Big Data, privacidad y protección de datos*, cit.,p.29.

[24] Instituto de Ingeniería del Conocimiento: https://www.iic.uam.es/innovacion/5-ventajas-clave-big-data/

Calculator jugara con él. Después del entrenamiento de varias partidas, *Defense Calculator* recogía datos sobre los posibles movimientos y los estudiaba, por lo que Arthur continuó mejorando el programa y enseñándole estrategias para que eligiera el movimiento que le reportara más piezas de ventaja. El resultado fue un ordenador que superó la habilidad y fue capaz de ganarle las partidas.

Este aprendizaje automático es el mismo que se aplica hoy en día, por ejemplo, a los coches que conducen solos, los sistemas de traducción automática, de reconocimiento de voz, etc.

Por otro lado, una de las mayores ventajas del análisis de datos masivos son las nuevas oportunidades de negocio que ofrece. Como bien hemos mencionado anteriormente, que una empresa base sus decisiones de negocio en la información disponible le proporcionan sin duda una ventaja competitiva importante. Como ejemplo de ello nos encontramos con la información que se recoge de los datos vertidos en las redes sociales, útiles para identificar oportunidades para generar negocio y vender anuncios basados en los intereses de cada usuario. Esta información permite a las empresas redirigir sus estrategias a grupos determinados de usuarios que personifican el cliente ideal o potencial. Los sectores de turismo y la hostelería, donde el fenómeno *Big Data* ha tenido un gran impacto, deben actualizarse continuamente y revolucionar la experiencia del cliente cada día. Este tipo de inteligencia aplicada al negocio puede anticiparse a lo que va a ocurrir en el futuro, a ciertos cambios o resultados mediante la analítica predictiva para valorar, por ejemplo, si ampliar el negocio o diversificarlo.

El almacenamiento del gran volumen de datos que supone el *Big Data* puede conllevar un problema de infraestructura de almacenamiento, por eso es conveniente trabajarlos en un entorno que no ponga límites como la nube, que supone, además, un ahorro de costes de hardware. Además, es una mejora en la accesibilidad y la fluidez de la información para los propios empleados de la empresa, con lo cual se gana en eficacia, rapidez y eficiencia.

2.4. INCONVENIENTES DEL *BIG DATA*

Como era de prever, pese a las grandes ventajas que nos brinda el *Big Data*, la aplicación del *Big Data* también posee aspectos negativos. Existen,

principalmente, tres aspectos negativos. El primero de ellos es el riesgo de incurrir en conclusiones erróneas sin ninguna supervisión humana, ya que nos encontramos expuestos a encontrar relaciones entre información que no tienen ningún tipo de relación que puede ser debido a la casualidad o al puro azar[25].

El segundo es el riesgo de la toma de decisiones automatizadas. Estas decisiones automatizadas tampoco conllevan una supervisión humana. Las empresas utilizan algoritmos, con la consecuencia de que en muchas ocasiones toman decisiones sobre nosotros sin que podamos saber por qué las han tomado[26]. Un ejemplo de esto último que mencionamos sucede cuando acudimos a una tienda para realizar una compra a plazos, pues se lleva a cabo un estudio basado en introducir nuestros datos en un programa que éste decide, en base a unas puntuaciones y a la revisión de unas bases de datos, si nos conceden o no ese crédito para efectuar dicha compra. Todo esto, gracias al uso de nuestros perfiles que brindan cierta información y a la que nosotros tenemos acceso, puede determinar la toma de decisiones en diferentes aspectos de nuestra vida cotidiana y determinar, de cierta forma, la personalidad de las personas[27].

Continuamente generamos información debido a nuestra interacción con la red, de manera que se puede predeterminar que una persona tenga acceso o no a un determinado producto o servicio. El Reglamento de Protección de Datos, como veremos más adelante en este trabajo, ha tenido en cuenta este aspecto.

Por último, la aplicación del *Big Data* conlleva un importante riesgo para la privacidad de las personas y la violación de los datos personales[28]. El empleo de los macrodatos produce una intrusión en nuestro derecho fundamental a la protección de datos personales que deriva directamente de la Constitución, y que se encuentra regulado en la Ley Orgánica 3/2018, de Protección de Datos Personales y garantía de los derechos digitales (en adelante, LOPDGDD).

En los últimos años esta amenaza a estos derechos se ha visto aumentada debido a la gran expansión de Internet, a través de las denominadas *cookies* o programas de rastreo se posibilita el funcionamiento de las

[25] GIL GONZÁLEZ, E., *Big Data, privacidad y protección de datos*, cit.,pp.28-29.

[26] GIL GONZÁLEZ, E., *Big Data, privacidad y protección de datos*, cit.,p.42.

[27] GARRIGA DOMÍNGUEZ, A., *Nuevos retos para la protección de datos personales. En la era del Big Data y de la computación ubicua*, Dykinson, Madrid, 2016.p.69.

[28] GIL GONZÁLEZ, E., *Big Data, privacidad y protección de datos*, cit.,p.32.

denominadas “redes de seguimiento”[29]. Además, el uso del denominado Internet de las cosas (*Internet of Things*) donde encontramos objetos que monitorizan nuestro día a día, como los dispositivos de geolocalización o los controladores inteligentes, también ponen en peligro nuestros derechos de protección de datos y privacidad[30].

Por esto es importante garantizar la seguridad de los datos y la protección de la privacidad con entidades que recaban estos datos y controlan que su protección se cumpla. Son entidades como la Agencia Española de Protección de Datos (AEPD), encargada de hacer cumplir toda la normativa vigente y promover el desarrollo de normativas y guías a seguir tanto por consumidores, como por empresas. También la Oficina de Seguridad del Internauta (OSI) es otra entidad que ayuda a controlar que se respeten los derechos de los usuarios, proporcionando la información y el soporte necesarios para evitar y resolver los problemas de seguridad que pueden existir al navegar por Internet. Y no podemos olvidar el Instituto Nacional de Ciberseguridad (INCIBE), sociedad dependiente del Ministerio de Asuntos Económicos y Transformación Digital a través de la Secretaría de Estado de Digitalización e Inteligencia Artificial, siendo su misión reforzar la ciberseguridad, la confianza y la protección de la información y privacidad en los servicios de la Sociedad de la Información, aportando valor a ciudadanos, empresas, Administración, red académica y de investigación española, sector de las tecnologías de la información y las comunicaciones y sectores estratégicos en general[31].

[29] GARRIGA DOMÍNGUEZ, A., *Nuevos retos para la protección de datos personales. En la era del Big Data y de la computación ubicua*, cit.p.23.

[30] GARRIGA DOMÍNGUEZ, A., *Nuevos retos para la protección de datos personales. En la era del Big Data y de la computación ubicua*, cit.pp.25-26.

[31] Incibe.es: https://www.incibe.es/que-es-incibe

3. LOS DATOS DE CARÁCTER PERSONAL Y SU TRATAMIENTO NORMATIVO

3.1. DEFINICIÓN

La definición de un dato personal, descrita anteriormente, la encontramos en el artículo 4.1 del Reglamento General de Protección de Datos (RGPD): *"«datos personales»: toda información sobre una persona física identificada o identificable («el interesado»); se considerará persona física identificable toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un identificador, como por ejemplo un nombre, un número de identificación, datos de localización, un identificador en línea o uno o varios elementos propios de la identidad física, fisiológica, genética, psíquica, económica, cultural o social de dicha persona;".*

Un gran número de datos que son utilizados en *Big Data* pueden ser datos de carácter personal, siendo de aplicación de esta manera la legislación vigente en protección de datos con las obligaciones que impone. Sin embargo, no todos los datos deben cumplir con las normas de protección de datos. Todo dato que no identifique a una persona no será considerado un dato personal, lo cual puede ser la clave para cumplir con la normativa de protección de datos personales en el caso del *Big Data.*

3.2. TRATAMIENTO NORMATIVO DE LOS DATOS DE CARÁCTER PERSONAL

Con tratamiento normativo hacemos referencia a la legislación general que se debe conocer a la hora de hablar de protección de datos y privacidad. Si bien es cierto que existen otros tipos de normas más concretas en función del ámbito tecnológico, únicamente haremos referencia a algunas normas generales y que constituyen la base de este ámbito, limitándonos exclusivamente a las normas españolas, comunitarias e internacionales.

A pesar de que Internet es global, cada Estado regula de una forma diferente la privacidad, por lo que las leyes de privacidad, las medidas y

organismos reguladores difieren de un país a otro[32]. En Europa, por ejemplo, cada Estado traspone las Directivas de una manera concreta y con ciertas diferencias.

Fuera de la Unión Europea nos encontramos aún con más desigualdad en las normas, por lo que el enfoque legislativo de los distintos países no es el mismo. Así, la percepción de privacidad y vida privada también cambia de un país a otro, por lo que es más que evidente que el enfoque legislativo también lo hará. En Estados Unidos, por ejemplo, en el debate entre la seguridad y la privacidad, la seguridad en muchas ocasiones posee mayor relevancia que la privacidad, y más aún, a raíz del atentado del 11 de septiembre de 2001[33].

Las leyes de protección de datos están diseñadas para proteger nuestra información personal tanto en línea como sin conexión, sin embargo, las leyes de retención de datos determinan cuánto tiempo los datos, incluidos los personales, deben ser retenidos por una entidad para fines legales o comerciales. Ambos aspectos pueden tener un gran impacto en la privacidad de las comunicaciones, el comportamiento y la persona en diferentes maneras[34]. Tal y como señalan Craig y Ludloff, Europa ha adoptado un modelo regulatorio de leyes integrales, en el cual las leyes generales rigen la recopilación y el uso de la información personal por parte de los sectores públicos y privados, y estas leyes suelen ir acompañadas de un órgano supervisor para garantizar su cumplimiento, como es el caso en España de la AEPD, la cual se encarga de que la legislación en materia de protección de datos sea aplicada de manera correcta y se lleven a cabo las acciones correspondientes para indemnizar los derechos de los titulares de los datos cuando han sido lesionados.

Las leyes actuales en materia de protección de datos son consideradas leyes de tercera generación, pues buscan el equilibro entre la protección de datos personales y el derecho a la información, además de dar respuesta a los riesgos que la tecnología nos presenta y su incidencia en los derechos fundamentales[35].

---

[32] CRAIG, T. y LUDLOFF, M. E. *Privacy and Big Data,* O'Reilly Media, 2011.p.2.

[33] CRAIG, T. y LUDLOFF, M. E. *Privacy and Big Data,* cit., p.9.

[34] CRAIG, T. y LUDLOFF, M. E. *Privacy and Big Data,* cit., p.15.

[35] HERNÁNDEZ LÓPEZ, J. M. *El derecho a la protección de datos personales en la doctrina del Tribunal Constitucional*, Aranzadi, 2013.p.26.

3.2.1. NORMATIVA INTERNACIONAL

Aunque el derecho a la protección de datos es considerado reciente, las raíces del mismo se encuentran en Estados Unidos a finales del siglo XIX, con la protección a la esfera privada desarrollada por Warren y Brandeis, quienes en su ensayo "*The Right to Privacy*"[36] dieron forma a la clásica definición de la privacidad (*privacy*), entendida genéricamente como el derecho a ser dejado solo o a no ser molestado, "*the right to be let alone"*[37]*.* El enfoque de dicho ensayo se debe tener en cuenta en la actualidad para comprender desde una perspectiva más completa el derecho a la protección de datos, ya que defendía la facultad de establecer unos límites entre la vida privada de las personas y la intromisión en ella por parte de terceros[38]. Esta obra estableció unas bases para que el derecho a no ser molestado fuera evolucionando de manera paulatina en la actual "privacy", siendo una de las primeras ocasiones en la historia en que se aludía a la posibilidad de que la intimidad alcanzada protección jurisdiccional.

Tanto el derecho a la intimidad (que es el concepto utilizado en nuestro país) como el derecho a la "privacy" evolucionaron y se desarrollaron, siendo prueba de ello la Declaración Universal de Derechos Humanos (DUDH), la cual, en su artículo 12[39] ofrece una protección que abarca una importante parte de los derechos de la persona frente a intromisiones de terceros. Este derecho a la intimidad siguió evolucionando hasta desarrollar un nuevo derecho a partir del año 1970 referente a la protección de datos personales en una esfera más similar a la que existe hoy en día[40].

---

[36]The Right to Privacy, December 15, 1890: https://www.jstor.org/stable/1321160?seq=1#metadata_info_tab_contents

[37] SALDAÑA DÍAZ, M. N. "The Right to Privacy. La génesis de la protección de la privacidad en el sistema constitucional norteamericano: el centenario legado de Warren y Brandeis", *Revista de derecho político*, núm.85, 2012.p.198.

[38] SALDAÑA DÍAZ, M. N. "El derecho a la privacidad en los Estados Unidos: aproximación diacrónica a los intereses constitucionales en juego", *Teoría y realidad constitucional*, núm.28, 2011.p.280.

[39] Artículo 12 DUDH: "*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. Toda persona tiene derecho a la protección de la ley contra tales injerencias o ataques".*

[40] BANISAR, D. y SIMON, D. "Global Trends in Privacy Protection: An International Survey of Privacy, Data Protection, and Surveillance Laws and Developments", *John Marshall Journal of Computer & Information Law, Vol. XVIII,* núm.1, 2012.p.10.

Resulta fundamental mencionar en este apartado el Convenio Europeo de Derechos Humanos (CEDH)[41] firmado en el año 1950 por el Consejo de Europa, el cual se formó después de la Segunda Guerra Mundial con el fin de reunir a los Estados de Europa para promover el Estado de derecho, la democracia, los derechos humanos y el desarrollo social. El CEDH es un tratado internacional, sin embargo, también posee un gran carácter protector de los derechos humanos y las libertades fundamentales en Europa[42].

Con el auge de la tecnología de la información en la década de 1960 se generó una creciente necesidad de contar con normas más detalladas para salvaguardar a las personas físicas protegiendo sus datos personales. A mediados de la década de 1970, el Consejo de Europa adoptó diversas resoluciones en materia de protección de datos personales referidas al artículo 8 del CEDH. En 1981 quedó abierto para su firma el Convenio para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal[43] (Convenio 108)[44].

El Convenio 108 fue y sigue siendo el único instrumento internacional jurídicamente vinculante en el ámbito de la protección de datos. El Convenio 108 se aplica a todo tratamiento de datos realizado por los sectores público y privado, incluidas las autoridades judiciales y los cuerpos de seguridad. Protege a las personas físicas contra los abusos que pueden llevarse a cabo en el tratamiento de datos personales, y busca, al mismo tiempo, regular los flujos transfronterizos de datos personales. En lo que respecta al tratamiento de datos personales, los principios establecidos en el Convenio se refieren, en particular, a la recopilación y el tratamiento automático de datos de manera lícita y leal, con fines legítimos especificados.

---

[41] Instrumento de Ratificación del Convenio para la Protección de los Derechos Humanos y de las Libertades Fundamentales, hecho en Roma el 4 de noviembre de 1950, y enmendado por los Protocolos adicionales números 3 y 5, de 6 de mayo de 1963 y 20 de enero de 1966, respectivamente. Publicado en BOE el 10 de octubre de 1979: https://www.boe.es/buscar/doc.php?id=BOE-A-1979-24010

[42] El CEDH, en su artículo 8 reconoce el derecho a la vida privada, derivando posteriormente de él el derecho a la protección de datos. Artículo 8 CEDH: "*Toda persona tiene derecho al respeto de su vida privada y familiar, de su domicilio y de su correspondencia*".

[43] AGENCIA DE LOS DERECHOS FUNDAMENTALES DE LA UNIÓN EUROPEA Y CONSEJO DE EUROPA, *Manual de legislación europea en materia de la protección*, 2018.p.27.

[44] Convenio para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, hecho en Estrasburgo el 28 de enero de 1981. Publicado en BOE el 15 de diciembre de 1985: https://www.boe.es/buscar/doc.php?id=BOE-A-1985-23447

Esto significa que los datos no deben utilizarse con propósitos incompatibles con estos fines y que no deben conservarse más tiempo del necesario. También se refieren a la calidad de los datos, que concretamente deben ser adecuados, pertinentes y no excesivos (proporcionalidad), además de exactos. No solo establece garantías en relación con el tratamiento de datos personales y obligaciones relativas a la seguridad de los datos, sino que prohíbe, a falta de garantías jurídicas adecuadas, el tratamiento de los datos «sensibles» de una persona, como la raza, las opiniones políticas, la salud, la religión, la vida sexual o los antecedentes penales[45].

El Convenio 108 es vinculante para los Estados que lo han ratificado y han sido todos los Estados miembros de la UE los que lo han hecho. No está sujeto al control judicial del TEDH, pero se ha tomado en consideración en la jurisprudencia del TEDH en el contexto del artículo 8 del CEDH. A lo largo de los años, el Tribunal ha determinado que la protección de los datos personales es parte importante del derecho al respeto de la vida privada (artículo 8) y se ha regido por os principios del Convenio 108 para determinar si se ha producido o no injerencia en este derecho fundamental[46]. Este Convenio ha sido adaptado a los cambios tecnológicos y, en la actualidad, podemos hablar del Convenio 108+[47].

3.2.2. NORMATIVA EUROPEA

Centrándonos en la normativa vigente aplicable a la protección de datos en la Unión Europea (UE), resulta conveniente citar la entrada en vigor del Tratado de Lisboa el 1 de diciembre de 2009, el cual supuso el establecimiento de una novedosa base jurídica para el derecho a la protección de datos, al

[45] AGENCIA DE LOS DERECHOS FUNDAMENTALES DE LA UNIÓN EUROPEA Y CONSEJO DE EUROPA, *Manual de legislación europea en materia de la protección*, cit.p.28.

[46] AGENCIA DE LOS DERECHOS FUNDAMENTALES DE LA UNIÓN EUROPEA Y CONSEJO DE EUROPA, *Manual de legislación europea en materia de la protección*, cit.p.29.

[47] Este Convenio pasó a firma el 10 de octubre de 2018 en Estrasburgo con gran acogida entre los Estados miembros de la UE, siendo España uno de los firmantes. Se trata del único instrumento internacional que confiere a las personas el derecho a la protección de sus datos personales e incluye garantías suplementarias para hacer frente a los retos de las nuevas tecnologías en armonía con el Reglamento general sobre la protección de datos de la UE. Refuerza los principios de proporcionalidad, minimización de datos y legalidad del tratamiento, amplia el catálogo de datos sensibles, con la inclusión de los genéticos y biométricos, y la pertenencia a un sindicato.

introducir en su artículo 16[48] del nuevo Tratado de Funcionamiento de la Unión Europea. Además, convirtió este derecho en fundamental al atribuir a la Carta de los Derechos Fundamentales de la UE el carácter de jurídicamente vinculante, reconociéndolo en su artículo 8[49].

Desde 1995 hasta mayo de 2018, el principal instrumento jurídico de la UE en materia de protección de datos fue la Directiva 95/46/CE del Parlamento Europeo y del Consejo, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. Se adoptó en 1995, en un momento en el que varios Estados miembros habían adoptado ya leyes nacionales de protección de datos y surgió la necesidad de armonizar dichas leyes para garantizar un elevado nivel de protección y la libre circulación de datos personales entre los Estados miembros[50].

Sin embargo, las Directivas no son de aplicación directa, sino que deben ser transpuestas a las legislaciones nacionales de los Estados miembros. Inevitablemente, los Estados miembros tienen cierta discrecionalidad en la transposición de las disposiciones de la Directiva, adoptando diversas normas de protección de datos en el conjunto de la UE, con normas y definiciones interpretadas de manera diferente en las legislaciones nacionales. Además, las tecnologías de la información experimentaron cambios significativos desde que se redactó la Directiva a mediados de la década de 1990, por lo que se resultó necesario reformarla, dando lugar a la adopción del Reglamento general de protección de datos en abril de 2016 (RGPD)[51].

Pero, sin lugar a dudas, el RGPD es la norma con mayor importancia en el régimen jurídico actual de la protección de datos.

---

[48] Artículo 16: "*Toda persona tiene derecho a la protección de los datos de carácter personal que le conciernan".*

[49] Carta de los Derechos Fundamentales de la Unión Europea. Publicado en DOUE el 30 de marzo de 2010: https://www.boe.es/buscar/doc.php?id=DOUE-Z-2010-70003 Artículo 8: "*Protección de datos de carácter personal 1. Toda persona tiene derecho a la protección de los datos de carácter personal que le conciernan. 2. Estos datos se tratarán de modo leal, para fines concretos y sobre la base del consentimiento de la persona afectada o en virtud de otro fundamento legítimo previsto por la ley. Toda persona tiene derecho a acceder a los datos recogidos que le conciernan y a obtener su rectificación. 3. El respeto de estas normas estará sujeto al control de una autoridad independiente*".

[50] AGENCIA DE LOS DERECHOS FUNDAMENTALES DE LA UNIÓN EUROPEA Y CONSEJO DE EUROPA, *Manual de legislación europea en materia de la protección*, cit.p.33.

[51] AGENCIA DE LOS DERECHOS FUNDAMENTALES DE LA UNIÓN EUROPEA Y CONSEJO DE EUROPA, *Manual de legislación europea en materia de la protección*, cit.p.34.

Fue en el año 2009 cuando el Parlamento Europeo, en su Resolución sobre el Programa de Estocolmo[52], concedió una especial importancia a las necesidades actuales de protección de datos frente al avanzado desarrollo tecnológico, tal y como se establece en su apartado 83[53].

En el año 2011, tras un periodo de consultas y reuniones con Gobiernos nacionales, ONG's, expertos y autoridades en protección de datos, el Parlamento Europeo emitió el 6 de julio de 2011 la Resolución sobre un enfoque global de la protección de datos personales en la UE[54], la cual señalaba que los desarrollos tecnológicos habían conllevado peligros que la anterior normativa, la Directiva 95/46/CE, no era capaz de proteger. Resultaba necesario una nueva norma que fuese capaz de garantizar una aplicación uniforme sobre la protección de datos en Europa. El RGPD no fue una norma sencilla de aprobar, podemos decir que fue una de las más complicadas que tuvo que afrontar la UE[55].

Finalmente, el Parlamento y el Consejo llegaron a un acuerdo sobre el texto final de la norma y sería publicado en el Diario Oficial de la UE el 4 de mayo de 2016, entrando en vigor el día 24 del mismo mes. Su aplicación, sin embargo, no sería preceptiva hasta el 25 de mayo de 2018, otorgando tiempo tanto a los Estados Miembros como a las empresas para adaptarse a las nuevas exigencias de la norma.

El objeto del RGPD, tal y como señala su artículo 1, es doble: defender los derechos fundamentales comunitarios y en especial la protección de datos, comprendiendo la normativa relativa a la protección de los mismos. Además, el RGPD regula la libre circulación de los datos, conteniendo un aspecto más

[52]Resolución del Parlamento Europeo, de 25 de noviembre de 2009, sobre la Comunicación de la Comisión al Parlamento Europeo y al Consejo titulada «Un espacio de libertad, seguridad y justicia al servicio de los ciudadanos − Programa de Estocolmo»:https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+TA+P7-TA-2009-0090+0+DOC+PDF+V0//ES

[53] Apartado 83 de la Resolución sobre el Programa de Estocolmo: "*Insiste en que en todas las políticas de la Unión se garantice el respeto de la dimensión de los derechos fundamentales que tienen la protección de datos y el derecho a la vida privada*".

[54]Resolución del Parlamento Europeo, de 6 de julio de 2011, sobre un enfoque global de la protección de los datos personales en la Unión Europea: https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2011-0323+0+DOC+XML+V0//ES

[55] Esta dificultad se hizo evidente por el hecho de ser la propuesta legislativa que ha recibido el mayor número de enmiendas, más de 3.000, en toda la historia del Parlamento; y no fue hasta diciembre de 2015, ya con las nuevas instituciones surgidas tras las elecciones del 2014, cuando se alcanzó el acuerdo (Gómez Barroso, Feijoo y Martínez, 2017: 116).

relacionado con el libre mercado, algo que preocupaba a la UE. Así, las actividades económicas de la UE pueden realizarse de una manera más ágil en comparación con la anterior normativa, al ser una norma de aplicación directa en cada Estado miembro[56].

Por otro lado, en el ámbito de aplicación material, el RGPD en su artículo 2 indica que es de aplicación respecto al tratamiento de datos personales, ya sea total o parcial e independientemente de si se tratan de datos automatizados o manuales. Resulta necesario que los datos puedan ser incluidos en un fichero, pues de lo contrario quedarían fuera del alcance del RGPD (Martín, 2018: 8).

En lo que concierne a la aplicación territorial, ha sido extendida respecto de las normativas anteriores. A partir del RGDP se protege el tratamiento de datos sin tener en cuenta de dónde se encuentre el establecimiento de un responsable o de un encargado de la UE. Así, lo relevante no es si el tratamiento se lleva a cabo dentro de la UE, sino que el elemento que va a conllevar la aplicación del RGPD es que los datos tratados pertenezcan a ciudadanos de la UE. Así, el RGPD se aplica frente a sociedades encargadas de tratar datos personales de europeos desde un país externo[57].

3.2.3. NORMATIVA NACIONAL

En lo que respecta a nuestro país, cabe mencionar en primer lugar nuestra Constitución Española de 1978 (CE), la cual es una de las primeras en introducir la protección de datos frente al uso de la informática en su artículo 18. Asimismo, encontramos que la dignidad de la persona se muestra como el contenido esencial de nuestra Constitución[58].

En la sección primera de la CE, relativa a los derechos fundamentas y las libertades públicas, se recogen los diferentes conceptos que hemos ido mencionando a lo largo de este TFM, clasificándolos como derechos fundamentales. De esta manera, nos interesa su artículo 18.1, donde hace

[56] HERRÁN ORTIZ, A. I., "Aproximación al derecho a la protección de datos personales en Europa. El reglamento general de protección de datos personales a debate*", Revista de Derecho, Empresa y Sociedad (REDS)*, núm.8, 2016.p.4.

[57] DURÁN ARROYO, A., "El nuevo reglamento de protección de datos personales. Análisis de su eficacia en la determinación de su ámbito territorial y los remedios en caso de tratamiento ilícito", *Revista Jurídica de la Universidad Autónoma de Madrid*, núm.37, 2018.p.425.

[58] REBOLLO DELGADO, L. *El derecho fundamental a la intimidad*, Dykinson, Madrid, 2005.p.110.

referencia al "*derecho al honor, a la intimidad personal y familiar y a la propia imagen*", consagrando así el derecho a la intimidad como derecho fundamental. Además, en el artículo 18.4 vemos como se establece que "*la ley limitará el uso de la informática para garantizar el honor y la intimidad personal y familiar de los ciudadanos y el pleno ejercicio de sus derechos*".

Si bien nos centraremos en la actual y vigente Ley Orgánica 3/2018, de protección de datos personales y garantía de los derechos digitales (LOPDGDD), debemos citar las primeras normas sobre la materia, estas son la Ley Orgánica 5/1992, de 29 de octubre, de regulación del tratamiento automatizado de los datos de carácter personal (LORTAD) y la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal (LOPD). La LOPDGDD entró en vigor el 6 de diciembre de 2018 sustituyendo a la anterior Ley Orgánica 15/1999 de Protección de Datos de Carácter Personal. La LOPDGDD se desarrolló con el objetivo de adaptar la legislación española a la normativa europea, definida por el RGPD, que se encontraba vigente desde el 25 de mayo de 2018.

La LOPDGDD no es una norma de transposición del RGPD, cuyas disposiciones eran desde el 25 de mayo de 2018 directamente aplicables en España, sino que tiene como fin armonizar la legislación española con las disposiciones ya vigentes del RGPD y detallar la regulación de protección de datos en diferentes materias que, o bien no están expresamente recogidas en el RGPD, o bien se abordan en el RGPD con la intención de que pudieran ser reguladas con más detalle por parte de los Estados Miembros. Asimismo, la nueva LOPDGDD incorpora a nuestro ordenamiento jurídico un elenco de "derechos digitales" de nuevo cuño. La LOPDGDD sigue la senda de la derogada LOPD, y proyecta su ámbito de aplicación sobre el conjunto de los datos personales, si bien con ciertas modificaciones en los requisitos para obtener información, guardarla o compartirla, y establece cambios en relación al tratamiento de datos de usuarios en Internet[59].

Al margen de la adaptación de la normativa interna al RGPD, a través de la LOPDGDD, se incorporan al ordenamiento jurídico español diecisiete nuevos

[59] Uría Menéndez: https://www.uria.com/documentos/circulares/1030/documento/8327/Aprobacion_LOPD-Garantia_Derechos_Digitales_.pdf?id=8327

"derechos digitales", que pretenden dar respuesta a cuestiones derivadas de la incorporación de las nuevas tecnologías en el día a día de las personas. Como ejemplo de estos nuevos "derechos digitales" podemos citar el derecho al testamento digital y el derecho a la rectificación en Internet.

Además, se concretan algunos principios o requisitos necesarios a la hora de tratar los datos personales, siendo el consentimiento, por ejemplo (del cual hablaremos en relación con el *Big Data*), el principio más relevante en este aspecto. Este consentimiento no puede ser automático ni por omisión, sino que el usuario debe ser consciente de que está aceptando que sus datos sean registrados.

3.3. LOS PRINCIPIOS DEL TRATAMIENTO

El RGPD establece que el tratamiento de los datos personales debe de llevarse a cabo conforme a los principios definidos en su artículo 5. Todo tratamiento debe estar basado en un fin legítimo. El RGPD enumera seis principios legítimos, los cuales veremos a continuación, y el tratamiento de datos personales debe estar vinculado a uno de ellos.

En primer lugar, nos encontramos con el principio de licitud, lealtad y transparencia. Según este principio, los datos deben ser "*tratados de manera lícita, leal y transparente en relación con el interesado*". Queda vinculado al principio de transparencia, el cual queda igualmente vinculado con la información, ya que la misma debe facilitarse de forma comprensible y accesible. Por tanto, el tratamiento no será leal y lícito si la información no está accesible o no es comprensible[60]. Así lo establece el considerando 39 del RGPD[61].

En segundo lugar, el RGPD establece el principio de limitación de la finalidad. Los datos deben ser "*recogidos con fines determinados, explícitos y*

[60] DPO&it law: http://www.dpoitlaw.com/reglamento-general-de-proteccion-de-datos-rgpd/unidad-i-5-principios-relativas-al-tratamiento-de-datos-personales/

[61] Este Considerando 39 establece que "*toda información y comunicación relativa al tratamiento de dichos datos sea fácilmente accesible y fácil de entender, y que se utilice un lenguaje sencillo y claro. Dicho principio se refiere en particular a la información de los interesados sobre la identidad del responsable del tratamiento y los fines del mismo y a la información añadida para garantizar un tratamiento leal y transparente con respecto a las personas físicas afectadas y a su derecho a obtener confirmación y comunicación de los datos personales que les conciernan que sean objeto de tratamiento. Las personas físicas deben tener conocimiento de los riesgos, las normas, las salvaguardias y los derechos relativos al tratamiento de datos personales, así como del modo de hacer valer sus derechos en relación con el tratamiento*".

legítimos, y no serán tratados ulteriormente de manera incompatible con dichos fines". El RGPD aclara la posibilidad de realizar tratamientos de datos con finalidades distintas de las recogidas siempre y cuando se dé una serie de presupuestos. Así, en el artículo 6.4 se establece que el responsable de tratamiento tendrá en cuenta una serie de cuestiones con objeto de determinar si dicho fin es compatible[62].

En tercer lugar, los datos deberán ser tratados conforme al principio de minimización de datos. Así, los datos deberán ser "*adecuados, pertinentes y limitados a lo necesario en relación con los fines para los que son tratados*". Es decir, los datos personales serán adecuados, pertinentes y limitados a la necesidad para la que fueron recabados. Cobra especial valor el sentido de la "necesidad", entendiéndola de tal manera que, si el objetivo puede alcanzarse sin realizar un tratamiento de datos, los mismos no deberían ser tratados. Por otro lado, dicha limitación a lo necesario debe ser evaluada desde un punto de vista cuantitativo (volumen de datos) como cualitativo (categoría de datos). Así se establece en el considerando 39 del RGPD[63].

En cuarto lugar, nos encontramos con el principio de exactitud. Así, los datos personales serán *"exactos y, si fuera necesario, actualizados; se adoptarán todas las medidas razonables para que se supriman o rectifiquen sin dilación los datos personales que sean inexactos con respecto a los fines para los que se tratan*". A este respecto señala también el considerando 39 que "*deben tomarse todas las medidas razonables para garantizar que se rectifiquen o supriman los datos personales que sean inexactos."*

En quinto lugar, los datos personales deberán ser tratados en virtud del principio de limitación del plazo de conservación. Los datos deberán ser

[62] Tendrá en cuenta, entre otras (art. 6.4 RGPD):

a. *cualquier relación entre los fines para los cuales se hayan recogido los datos personales y los fines del tratamiento ulterior previsto;*
b. *el contexto en que se hayan recogido los datos personales, en particular por lo que respecta a la relación entre los interesados y el responsable del tratamiento;*
c. *la naturaleza de los datos personales, en concreto cuando se traten categorías especiales de datos personales, de conformidad con el artículo 9, (Categorías especiales de datos personales) o datos personales relativos a condenas e infracciones penales, de conformidad con el artículo 10 (Datos relativos a condenas e infracciones);*
d. *las posibles consecuencias para los interesados del tratamiento ulterior previsto;*
e. *la existencia de garantías adecuadas, que podrán incluir el cifrado o la seudonimización.*

[63] Considerando 39 RGPD: "*Los datos personales solo deben tratarse si la finalidad del tratamiento no pudiera lograrse razonablemente por otros medios*."

"mantenidos de forma que se permita la identificación de los interesados durante no más tiempo del necesario para los fines del tratamiento de los datos personales". Si bien en nuestra normativa ya se establece que deberán ser cancelados cuando los datos dejen de ser útiles para la finalidad en la que fueron recabados (art. 17 RGPD derecho de supresión o "derecho al olvido"), el RGPD además de limitar el plazo de conservación establece la obligación al responsable de incluir plazos para la supresión o revisión periódica. Considerando 39: "*Para garantizar que los datos personales no se conservan más tiempo del necesario, el responsable del tratamiento ha de establecer plazos para su supresión o revisión periódica*."

En último lugar, los datos personales serán tratados conforme al principio de integridad y confidencialidad. Así, los datos serán "*tratados de tal manera que se garantice una seguridad adecuada de los datos personales, incluida la protección contra el tratamiento no autorizado o ilícito y contra su pérdida, destrucción o daño accidental, mediante la aplicación de medidas técnicas u organizativas apropiadas*".

Los encargados del tratamiento de datos personales serán responsables del cumplimiento de estos principios y deben ser capaces de demostrarlos, así lo recoge el RGPD en su artículo 5.2, denominándolo principio de "responsabilidad proactiva". A través de este principio el responsable y encargado de tratamiento estarán obligados a demostrar que sus actividades de tratamiento de datos cumplen con los principios relativos al tratamiento de datos. Para ello, deberán implantar unas medidas técnicas y organizativas apropiadas a fin de demostrar que los tratamientos que realizan son conformes con el RGPD. Estas medidas deberán ser actualizadas y revisadas periódicamente a través de procedimientos internos o externos de auditoría, o con la adhesión a códigos de conducta o procesos de certificación[64].

[64] DPO&it law: http://www.dpoitlaw.com/reglamento-general-de-proteccion-de-datos-rgpd/unidad-iii-accountability-o-principio-de-responsabilidad-proactiva/

4. EL IMPACTO DEL *BIG DATA* EN LA NORMATIVA DE PROTECCIÓN DE DATOS

Como bien hemos mencionado anteriormente, la analítica de datos a partir del *Big Data* es capaz de brindarnos numerosas ventajas para empresas, particulares y Administraciones, siendo capaz de interpretar una cantidad ingente de datos. No obstante, como particulares, también hemos visto como inconveniente del *Big Data* la exposición que sufren nuestros datos, sin saber el uso ni el fin que se le va a dar, convirtiéndose en un evidente riesgo.

Como consecuencia de sus peligros y del avance de las tecnologías y con el objetivo de proteger a los ciudadanos, surgen en el seno de la UE las normas para la protección de datos personales que hemos ido mencionando en el anterior epígrafe, materializados en el actual y vigente RGDP.

Cuando hablamos del *Big Data*, del mundo tecnológico y de Internet, hay que tener en cuenta que los datos son capaces de seguir en Internet durante un tiempo indeterminado aun después de haber sido solicitado su borrado u "olvido", ya que esta información puede permanecer en los buscadores (como Google) y en la memoria denominada caché, que permite que los contenidos puedan permanecer en la web sin ningún tipo de control. De esta manera, resulta confuso saber la cantidad de datos personales e información que circula por Internet[65], aunque tengamos la posibilidad de solicitar ayuda a la AEPD para su retirada, o ejercer lo que se conoce como el derecho al olvido[66].

La normativa de protección de datos se aplica únicamente cuando la información de las personas físicas hace que éstas sean identificados directamente o identificables. Sin embargo, cuando los datos no hacen identificable a una persona, no se aplica esta regulación. Es decir, cuando los datos se hacen anónimos a través de técnicas de anonimización, se convierten en datos no personales, y la privacidad de los individuos queda protegida, de modo que no resulta necesario aplicar ninguna norma sobre protección de datos[67]. Como bien hemos mencionado anteriormente, un dato personal es toda

[65] GARRIGA DOMÍNGUEZ, A., *Nuevos retos para la protección de datos personales. En la era del Big Data y de la computación ubicua*, cit.p120.

[66] La AEPD define el d*erecho al olvido* como el derecho de solicitar que los datos personales se supriman de las búsquedas en Internet.

[67] GIL GONZÁLEZ, E., *Big Data, privacidad y protección de datos*, cit.,pp.51-52.

información sobre una persona física identificada o identificable, tal y como señala el artículo 4.1 RGPD, y que únicamente a los mismos se les aplicará la normativa de protección de datos.

El *Big Data* desafía las normas de protección de datos al facilitar la reidentificación de los sujetos[68], ya no solo a partir de los datos seudónimos[69], sino también a partir de datos que considerábamos anónimos y que no permitían identificar a los sujetos. Es decir, las técnicas de anonimización ya no siempre son suficientes con la llegada del *Big Data* y las posibilidades que ofrece la tecnología.

En conclusión, el *Big Data* amenaza la normativa de protección de datos, debido a diversos motivos.

En primer lugar, la anonimización posee limitaciones en entornos del *Big Data*. La anonimización se presentaba como la solución óptima para tratar los datos protegiendo la privacidad de los individuos. No obstante, en los últimos años se han dado numerosos casos de reidentificación de bases de datos que habían sido anonimizadas. Así, cada vez resulta más sencillo reidentificar a los sujetos, gracias a que el *Big Data* permite cruzar datos procedentes de fuentes muy diversas, que pueden contener datos personales parciales sobre una persona, o incluso identificarnos a través de datos que antes eran considerados no personales, como las puntuaciones que un usuario otorga a una película en portales de Internet[70].

En segundo lugar, encontramos que los principios de "minimización de datos"[71] y de "limitación de los fines"[72] no se cumplen en la práctica. El principio

[68] Entendemos por reidentificación de los sujetos en el ámbito de protección de datos cuando los datos personales consiguen relacionarse de nuevo con los datos identificativos de la persona. Grupo Ático 34: https://protecciondatos-lopd.com/empresas/seudonimizacion-anonimizacion/

[69] Datos que han sufrido un proceso de seudonimización, definido por el RGPD en su artículo 4.5): "«se*udonimización»: el tratamiento de datos personales de manera tal que ya no puedan atribuirse a un interesado sin utilizar información adicional, siempre que dicha información adicional figure por separado y esté sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos personales no se atribuyan a una persona física identificada o identificable*"

[70] Legal Today: https://www.legaltoday.com/practica-juridica/derecho-civil/nuevas-tecnologias-civil/que-es-el-big-data-y-por-que-debe-interesarme-si-soy-abogado-2016-10-18/

[71] Principio de "minimización de datos" recogido en el art. 5.1.c) RGPD: "*Los datos personales será (…) c) adecuados, pertinentes y limitados a lo necesario en relación con los fines para los que son tratados («minimización de datos»)*".

[72] Principio de "limitación de los fines" recogido en el art. 5.1.b) RGPD: "*Los datos personales serán (…) b) recogidos con fines determinados, explícitos y legítimos, y no serán tratados ulteriormente de manera incompatible con dichos fines; de acuerdo con el artículo 89, apartado 1, el tratamiento ulterior*

de minimización de datos, conforme al cual se utilizarán el mínimo posible de datos personales, se contrapone contra la lógica del *Big Data*. Los nuevos modelos analíticos se basan precisamente en el estudio de cantidades masivas de datos, sin los cuales no podría extraerse el conocimiento que nos permite el *Big Data*. Además, el *Big Data* se basa precisamente en reutilizar datos que fueron obtenidos para una primera finalidad, otorgándole una nueva finalidad, lo que contradice el principio de limitación de la finalidad. Es de hecho, en este aspecto donde reside la mayor fuente de beneficios del *Big Data*[73].

En tercer lugar, la normativa confía demasiado en el consentimiento informado del individuo para recopilar y tratar sus datos de carácter personal[74]. De hecho, esta cuestión no es nueva, aunque es cierto que el consentimiento se encuentra reforzado desde el 2016 con la aplicación del RGPD al exigir que sea expreso. Esto supone un grave problema, pues la mayoría de los individuos no leen las políticas de privacidad antes de prestar su consentimiento, y aquéllos que lo hacen no las comprenden. De esta manera, prestar nuestro consentimiento en el entorno digital en el que surgen la mayor parte de los datos utilizados en los sistemas *Big Data* es, en muchas ocasiones, un ejercicio vacío[75].

Por último, como hemos podido ver, el *Big Data* aumenta el riesgo relacionado con la toma de decisiones de forma automática, produciendo que decisiones trascendentales para nuestra vida, tales como calcular nuestro perfil sanitario utilizado por una empresa de seguros médicos queden sujetas a algoritmos ejecutados de forma automática y sin intervención humana[76]. A este

de los datos personales con fines de archivo en interés público, fines de investigación científica e histórica o fines estadísticos no se considerará incompatible con los fines iniciales («limitación de la finalidad»)".

[73] GIL GONZÁLEZ, E., *Big Data, privacidad y protección de datos*, cit.,pp.52-53.

[74] "Consentimiento informado". El consentimiento debe ser libre, específico, informado e inequívoco, según el art. 4.11 RGPD: "*11) «consentimiento del interesado»: toda manifestación de voluntad libre, específica, informada e inequívoca por la que el interesado acepta, ya sea mediante una declaración o una clara acción afirmativa, el tratamiento de datos personales que le conciernen*". Así, en el "consentimiento informado" se trata de informar al usuario de la finalidad del tratamiento, el nombre del responsable del tratamiento, cómo van a ser tratados los datos y los derechos de los que es titular la persona. Iberley: https://www.iberley.es/revista/funciona-consentimiento-rgpd-184

[75] Legal Today: https://www.legaltoday.com/practica-juridica/derecho-civil/nuevas-tecnologias-civil/que-es-el-big-data-y-por-que-debe-interesarme-si-soy-abogado-2016-10-18/

[76] GIL GONZÁLEZ, E., *Big Data, privacidad y protección de datos*, cit.,p.53.

aspecto, la normativa de protección de datos regula la elaboración de perfiles, dando la posibilidad de oponerse a ellos (artículo 21.2 RGPD)[77].

Todos estos motivos suponen un gran reto para el *Big Data*, unidos, además, al hecho de que la tecnología avanza a pasos agigantados, mientras que la normativa no es capaz de adaptarse tan rápido al nuevo entorno tecnológico.

4.1. CONSENTIMIENTO *VS* BIG DATA

Centrándonos en el RGPD, existen varias bases legitimadoras para realizar un tratamiento de datos personales. Sin embargo, existe la obligación de cumplir al menos una de ellas, sea el consentimiento u otra base legítima establecida conforme a Derecho (Considerando 40 RGPD).

En cuanto al entorno del *Big Data*, analizaremos algunas de estas bases legitimadoras principales para el posible tratamiento de datos, centrándonos en el consentimiento.

La definición de consentimiento se encuentra recogida en el artículo 4.11 del RGPD, que dicta como sigue: "*toda manifestación de voluntad libre, específica, informada e inequívoca por la que el interesado acepta, ya sea mediante una declaración o una clara acción afirmativa, el tratamiento de datos personales que le conciernen*". Este consentimiento debe ser expreso, tal y como señala el artículo 7 del RGPD. Este consentimiento expreso constituyó, en su día, la mayor novedad en el Reglamento y en la LOPDGDD, reforzándolo y dejando sin efecto el consentimiento tácito (artículo 6 LOPDGDD). Así, este consentimiento conlleva la aprobación por parte del afectado de la inclusión de sus datos personales en un fichero de su tratamiento[78].

El RGPD pretende ampliar el control que los titulares tengan de sus propios datos y para eso el RGPD utiliza el requisito del consentimiento como instrumento principal. El consentimiento se convierte por tanto en "la llave de todo tratamiento de datos personales. Salvo las excepciones legalmente

[77] El RGPD, en su artículo 21.2 nos dice que, en el caso de que una Entidad utilice como base de legitimación el interés legítimo para efectuar el perfil de un cliente, en el mismo momento de la recogida de los datos de éste, se debe ofrecer al mismo la posibilidad de oponerse a que sus datos sean utilizados para elaborar perfiles.

[78] HERNÁNDEZ LÓPEZ, J. M. *El derecho a la protección de datos personales en la doctrina del Tribunal Constitucional*, cit.,pp.70-71.

previstas, el consentimiento da acceso al tratamiento de nuestros datos personales, lo legitima, y permite hablar de un mayor o menor control de los mismos, esto es, haber sido conscientes o no de que nuestros datos están siendo tratados"[79].

El Considerando 32 del RGPD, además, recoge que "*el consentimiento debe darse mediante un acto afirmativo claro que refleje una manifestación de voluntad libre, específica, informada e inequívoca…*". De esta manera, resulta obligatorio que el consentimiento sea, por tanto, claro, libre, específico, informado e inequívoco, siendo esto último especialmente relevante en el tema que nos concierne. El consentimiento resulta clave en la protección de datos personales, pues proporciona autonomía a los interesados y una mayor transparencia en el tratamiento de los datos. Este tratamiento posee una conexión directa con el consentimiento informado e inequívoco.

Este consentimiento debe ser inequívoco para cada tratamiento de datos, sin embargo, teniendo en cuenta la gran cantidad de los tratamientos del *Big Data* y la correlación de datos personales, resulta imposible recabar el consentimiento de esa gran cantidad de personas afectadas, por lo que la utilización de esta nueva tecnología conlleva este riesgo que resulta inherente[80].

Además, el RGPD sigue reforzando la figura del consentimiento al señalar en su Considerando 42 que "*cuando el tratamiento se lleva a cabo con el consentimiento del interesado, el responsable del tratamiento debe ser capaz de demostrar que aquel ha dado su consentimiento a la operación de tratamiento…*" y en su Considerando 43 que: "*…Se presume que el consentimiento no se ha dado libremente cuando no permita autorizar por separado las distintas operaciones de tratamiento de datos personales pese a ser adecuado en el caso concreto, o cuando el cumplimiento de un contrato, incluida la prestación de un servicio, sea dependiente del consentimiento, aun cuando este no sea necesario para dicho cumplimiento*". De esta manera, el consentimiento debe poder ser demostrado, tal y como recoge el RGPD, por parte del responsable del tratamiento de los datos personales y ser, como hemos mencionado, específico

[79] ARENAS RAMIRO, M., "Reforzando el ejercicio del derecho a la protección de datos personales: viejas y nuevas facultades", en GARCÍA MAHAMUT, R. y RALLO LOMBARTE, A. (coord.), *Hacia un nuevo derecho europeo de protección de datos*, Tirant lo Blanch, 2015.p.329.

[80] GIL GONZÁLEZ, E., *Big Data, privacidad y protección de datos*, cit.,p.68.

e inequívoco para cada uno de los tratamientos da manera individual, por lo que, en el caso de tratamiento masivo de datos, como es el *Big Data,* resulta casi inviable, pues el gran volumen de datos y la demostración de la procedencia de ellos es un trabajo que roza lo imposible.

Antes de la entrada en vigor del RGPD y de la LOPDGDD este consentimiento se presuponía dado al navegar por páginas web, es decir, ésta entendía que habías aceptado que se trataran los datos personales al responsable del tratamiento de dicha página web. Sin embargo, con la entrada en vigor de ambas normas y como hemos visto, el consentimiento ha sufrido un gran cambio, pues el responsable de la página web ahora tiene que demostrar que ha obtenido este consentimiento. Esta obligación por parte del responsable del tratamiento de demostrar este consentimiento aparece recogida en el artículo 7 del RGPD, el cual nos expresa dicha obligación del responsable de demostrar que obtuvo el consentimiento del interesado[81], y que éste conoce su derecho a retirar su consentimiento en cualquier momento[82]. Este consentimiento, que como hemos visto ya no es tácito, en el ámbito *online* se cede con la aceptación a través de un clic en una política de privacidad, siempre que las casillas no se encuentren premarcadas y exista una casilla para cada consentimiento[83].

El RGPD es claro al diferenciar cada consentimiento con su finalidad dada, siendo éste un punto sustancial para el tratamiento de datos. Además, el RGPD señala que la facilidad de revocación debe ser igual de sencilla, chocando con la esencia del *Big Data*, pues busca la reutilización de los datos personales en una variedad de campos inimaginables, utilizando los datos para usos secundarios, la rigidez de la legislación en esta materia es claramente una barrera. Los responsables del tratamiento de datos personales deben conseguir un consentimiento informado, inequívoco y revocable cada vez que traten datos

[81] Art. 7.1 RGPD: "*1. Cuando el tratamiento se base en el consentimiento del interesado, el responsable deberá ser capaz de demostrar que aquel consintió el tratamiento de sus datos personales*".

[82] Art. 7.2 RGPD: "*3. El interesado tendrá derecho a retirar su consentimiento en cualquier momento. La retirada del consentimiento no afectará a la licitud del tratamiento basada en el consentimiento previo a su retirada. Antes de dar su consentimiento, el interesado será informado de ello. Será tan fácil retirar el consentimiento como darlo*"

[83] Ángel Benito Rodero: https://www.angelbenitorodero.es/proteccion-de-datos/consentimiento-rgpd/

con una finalidad distinta, pues cada finalidad requiere el consentimiento expreso del interesado[84]. Así lo recoge la LOPDGDD en su artículo 6.2[85].

Si bien es cierto que la LOPDGDD desarrolla el RGPD, también lo es que no aporta apenas nada nuevo respecto de este último, pues el consentimiento es una base legitimadora para utilizar datos personales de manera eficaz y cristalina. Pero, sin embargo, deja patente la obligación de cumplir con varias premisas. Entre estas premisas, un consentimiento para cada tratamiento de datos, lo que hace que en el ámbito del *Big Data* esta base legitimadora sea residual, ya que la ventaja del *Big Data* es el tratamiento de datos con diferentes finalidades. Esto supone una desventaja en el caso, por ejemplo, de las empresas, pues recabar o solicitar este consentimiento del interesado en función de las finalidades a que se va a destinar esa información de nuevo supone un mayor coste para ella[86].

Respecto al tratamiento de datos de menores de edad, debemos tener en cuenta el Considerando 38 del RGPD[87], el cual origina la idea de que los menores precisan una protección específica de sus datos personales, debido a sus vulnerabilidades evidentes a la hora de tomar decisiones, siendo especialmente necesaria dicha protección para fines de mercadotecnia o para la elaboración de perfiles[88]. El *Big Data* es utilizado con fines de mercadotécnica, lo que supone un esfuerzo mayor en la protección y en la captura de estos datos personales para evitar el tratamiento de menores por error o sin el consentimiento de sus padres/tutores legales.

---

[84] DPO&it law: http://www.dpoitlaw.com/reglamento-general-de-proteccion-de-datos-rgpd/unidad-i-6-el-consentimiento/

[85] Artículo 6.2. LOPDGDD: "*Tratamiento basado en el consentimiento del afectado. 2. Cuando se pretenda fundar el tratamiento de los datos en el consentimiento del afectado para una pluralidad de finalidades será preciso que conste de manera específica e inequívoca que dicho consentimiento se otorga para toda de ellas.*"

[86] GIL GONZÁLEZ, E., *Big Data, privacidad y protección de datos*, cit.,p.67.

[87] Considerando 38 RGPD: "*(38) Los niños merecen una protección específica de sus datos personales, ya que pueden ser menos conscientes de los riesgos, consecuencias, garantías y derechos concernientes al tratamiento de datos personales. Dicha protección específica debe aplicarse en particular, a la utilización de datos personales de niños con fines de mercadotecnia o elaboración de perfiles de personalidad o de usuario, y a la obtención de datos personales relativos a niños cuando se utilicen servicios ofrecidos directamente a un niño. El consentimiento del titular de la patria potestad o tutela no debe ser necesario en el contexto de los servicios preventivos o de asesoramiento ofrecidos directamente a los niños*"

[88] Legal Today: https://www.legaltoday.com/practica-juridica/derecho-publico/proteccion-datos/el-tratamiento-de-datos-de-menores-de-edad-que-dice-el-rgpd-al-respecto-2018-06-28/

En la LOPDGDD encontramos el tratamiento de datos en su artículo 7. El artículo señala que para menores de edad, mayores de catorce años, el tratamiento es permitido con el consentimiento del menor cuando éste sea expreso e informado. Sin embargo, para el caso de menores de edad, con menos de 14 años, el tratamiento requiere el consentimiento del titular de la patria potestad o tutela[89].

A la hora de utilizar el *Big Data* se debe prestar atención en el tratamiento de los datos de menores de edad, pues puede suponer un riesgo para ellos al ser una herramienta peligrosa en casos como la elaboración de perfiles y la publicidad personalizada, pudiendo influir en sus ideas perjudicando su desarrollo, máxime cuando se encuentran en la etapa de formación de su personalidad.

A pesar de todo lo mencionado, el problema principal con el que nos encontramos en la actualidad es que la mayoría de las personas, mayores o menores de edad, no prestan atención a la hora de prestar su consentimiento y leer las políticas de privacidad, debido a la gran complejidad, de manera que se podría considerar que las relaciones, en estos casos, son desequilibradas.

Por todo lo visto, podemos concluir que una forma de cumplir con la normativa y respetar el requisito del consentimiento en las Técnicas de *Big Data* puede ser la anonimización de los datos personales, pues se reduce el riesgo de una posible reidentificación de los datos anonimizados, siendo la única solución viable en el tratamiento de grandes volúmenes de datos.

4.2. ANONIMIZACIÓN Y SEUDONIMIZACIÓN DE LOS DATOS

4.2.1. SEUDONIMIZACIÓN

En el ámbito de la protección de datos, la seudonimización es una posibilidad más para poder protegerlos y hacerlos más seguros. El tratamiento y uso de seudónimos en protección de datos es una salvaguarda de la privacidad, siendo capaces de paliar los riesgos para el tratamiento de datos.

[89] Iberley: https://www.iberley.es/temas/consentimiento-menores-materia-proteccion-datos-62818

Esta técnica no convierte los datos personales en anónimos, siendo reversible la técnica de no identificación y pudiendo descubrir la identidad de las personas.

La definición que nos brinda el RGPD de la seudonimización en su artículo 4.5) es la siguiente: *"5) «seudonimización»: el tratamiento de datos personales de manera tal que ya no puedan atribuirse a un interesado sin utilizar información adicional, siempre que dicha información adicional figure por separado y esté sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos personales no se atribuyan a una persona física identificada o identificable".* Esto es, consiste en tratar los datos personales sin los datos identificativos del interesado, pero sin suprimir la vinculación entre los datos que consigan determinar la persona titular de los mismos. Un ejemplo sería la sustitución de los nombres de clientes por un código o por un identificador numéricos, es decir, cambiar los datos personales por seudónimos[90].

El RGPD, en su artículo 25, considera una serie de medidas desde el diseño y por defecto. Esto es, que de entre todas las medidas se utilice la que resulte menos invasiva para la protección de datos. Es evidente que estas medidas deben adoptarse de manera previa al tratamiento. Es por ello que la seudonimización de datos favorece los principios de protección y minimización de los mismos, y hacen, por defecto, que únicamente sean objeto de tratamiento los datos personales estrictamente necesarios para cada uno de los fines específicos del tratamiento.

Debemos tener claros los objetivos de la seudonimización de datos, los cuales podríamos resumirlos en[91]:

1. Garantizar la confidencialidad, integridad, disponibilidad y resiliencia permanentes de los sistemas y servicios de tratamiento.
2. Restaurar la disponibilidad y el acceso a los datos personales de forma rápida en caso de incidente físico o técnicos (uno de los significados de resiliencia).

[90] Grupo Ático34: https://protecciondatos-lopd.com/empresas/seudonimizacion-anonimizacion/

[91] Confilegal: https://confilegal.com/20170129-la-importancia-del-seudonimizacion-en-el-nuevo-reglamento-de-proteccion-de-datos/

3. Implantar un proceso de verificación, evaluación y valoración regulares de la eficacia de las medidas técnicas y organizativas para garantizar la seguridad del tratamiento[92].

No debemos olvidar que un dato seudonimizado está sujeto a la normativa de protección de datos, ya que se trata de un dato personal, aunque no permita la identificación directa de la persona, pero sí hacerlo de forma indirecta.

Un buen ejemplo de seudonimización lo tenemos en la Ley 41/2002, de 14 de noviembre, básica reguladora de la Autonomía del paciente y de derecho y obligaciones en materia de información y documentación clínica, la cual, en su artículo 16.3 señala: "*3. El acceso a la historia clínica con fines judiciales, epidemiológicos, de salud pública, de investigación o de docencia, se rige por lo dispuesto en la legislación vigente en materia de protección de datos personales, y en la Ley 14/1986, de 25 de abril, General de Sanidad, y demás normas de aplicación en cada caso. El acceso a la historia clínica con estos fines obliga a preservar los datos de identificación personal del paciente, separados de los de carácter clinicoasistencial, de manera que, como regla general, quede asegurado el anonimato, salvo que el propio paciente haya dado su consentimiento para no separarlos*"[93].

El Dictamen 05/2014 del Grupo de Trabajo del Artículo 29, de 10 de abril sobre técnicas de anonimización, recoge algunas de las técnicas más relevantes de seudonimización, como cifrado con clave secreta o con clave de borrado de claves; función hash; función con clave almacenada o descomposición en tokens, entre otras. Para poder comprender estos tipos de seudonimización pondremos de ejemplo la siguiente tabla:

DNI	Nombre y apellido	Edad	Profesión	Salario anual	Patología
54376156S	Enrique Calvo	28	Abogado	23.000	Diabetes

[92] Protecciónde datos.org: https://www.protecciondatos.org/seudonimizacion-de-datos-segun-rgpd/

[93] Gahazas: https://gahazas.com/2017/02/27/analisis-de-los-conceptos-de-anonimizacion-seudonimizacion-y-disociacion-en-el-ambito-de-proteccion-de-datos/

70387128P	María Villanueva	37	Ingeniera	35.000	Enfermedad de Crohn

En la anterior tabla podemos diferencias tres tipos de identificadores:

- Identificadores directos: se tratan de características de una persona que son capaces de identificarla (nombre, DNI o dirección).
- Identificadores indirectos: se tratan de atributos que pueden ser compartidos por varias personas y cuya relación puede conducir a la reidentificación de alguna de ellas (edad o profesión).
- Atributos sensibles: se tratan de datos de carácter especial, se encuentran recogidos en el artículo 9 del RGPD (tratamiento de categorías especiales de datos personales), y en el artículo 9 de la LOPDGDD (categorías especiales de datos). Alguno de estos datos son la ideología, salud, vida sexual, religión, etc. En la tabla anterior sería la patología[94].

Volviendo a las técnicas de seudonimización existentes, y en concreto al método de cifrado con clave secreta[95], la aplicaremos a la anterior tabla de manera que:

~~**54376156S**~~	~~**Enrique Calvo**~~	28	Abogado	23.000	Diabetes
~~**70387128P**~~	~~**María Villanueva**~~	37	Ingeniera	35.000	Enfermedad de Crohn

RM#1	**R#45**	28	Abogado	23.000	Diabetes
LP#2	**T#79**	37	Ingeniera	35.000	Enfermedad de Crohn

[94] AEPD: https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf

[95] En esta técnica, el poseedor de la clave puede reidentificar al interesado con suma facilidad. Para ello, le basta con descifrar el conjunto de datos, ya que este contiene los datos personales, aunque sea en forma cifrada. Si se aplican los sistemas de cifrado más avanzados, tan solo es posible descifrar los datos si se conoce la clave. Dictamen 05/2014 sobre técnicas de anonimización: https://www.aepd.es/sites/default/files/2019-12/wp216-es.pdf

En este ejemplo, en el cual hemos aplicado la técnica de cifrado en clave secreta, se ha sustituido el nombre y el DNI de la persona por un código cifrado, de esta manera únicamente la persona con la información que permite vincular el código cifrado al dato personal podrá identificarlo. Habría una clave que únicamente tendría el responsable del tratamiento de datos, por la cual sería capaz de identificar que R#45 es Enrique Calvo y lo mismo para las otras variables. Esta clave del cifrado debe guardarse de forma segura, pues es fundamental para poder reidentificar de manera inmediata. La dificultad respecto a reidentificar dependerá únicamente del proceso o método utilizado, y los valores o atributos sustituidos, teniendo en cuenta que este método no es infalible y que, con las nuevas tecnologías es imprescindible adoptar medidas extras de protección.

Además de este método con clave secreta, existen muchos otros muy utilizados también en la seudonimización, como la función hash, función con clave almacenada, cifrado determinista o función hash con clave con borrado de clave y la descomposición en tokens[96]

4.2.2. ANONIMIZACIÓN

La anonimización ha sido entendida por la AEPD como "*la ruptura de la cadena de identificación de las personas.*"[97] La anonimización es una forma de eliminar posibilidades de identificación de las personas en la sociedad actual, en la cual la información es un recurso fundamental para la toma de decisiones en todos los ámbitos de la misma. Además de la importancia de eliminar o reducir al mínimo los riesgos de identificación de los datos anonimizados, es de vital importancia garantizar la veracidad de los resultados de tratamientos de los mismos. En el diseño del proceso, hay que prever las consecuencias de una eventual reidentificación de las personas que pudiera generar un perjuicio para sus derechos[98]. De nuevo, vemos la importancia de cumplir con el citado artículo 25 RGPD de privacidad desde el diseño y por defecto.

[96] Dictamen 05/2014 sobre técnicas de anonimización: https://www.aepd.es/sites/default/files/2019-12/wp216-es.pdf

[97] AEPD: https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf

[98] Protecciónde datos.org: https://www.protecciondatos.org/anonimizacion-de-datos-personales/

El tratamiento de grandes volúmenes de datos, como es el caso del *Big Data*, ofrece numerosos beneficios, pero también es un riesgo para la privacidad y para la protección de datos de carácter personal. Así, la anonimización es la única solución viable para el *Big Data* cumpliendo con la normativa.

Cuando la AEPD indica que la finalidad del procedimiento de anonimización es "*la ruptura de la cadena de identificación de las personas*" resulta fundamental diferenciar entre ruptura directa y ruptura indirecta, entendiéndose esta última, tal y como señala la AEPD, como aquélla que pueda tener lugar como consecuencia de información de una o varias fuentes que por sí misma o en combinación de otros factores puede permitir la reidentificación de las personas cuando sus datos hubieran sido anonimizados. Por ejemplo, la combinación de sexo, edad, lugar de nacimiento y padecimiento de una determinada enfermedad pueden permitir la identificación indirecta de una persona concreta[99]. Además, en este proceso de anonimización resulta indispensable prever las consecuencias de una eventual reidentificación de las personas que pudiera generar un perjuicio o merma de sus derechos. La AEPD también señala como necesario prever una hipotética pérdida de información por negligencia del personal implicado, por falta de una política de anonimización adecuada o por una revelación de secreto intencionada que diera lugar a la pérdida de las variables de identificación o claves de identificación de las personas[100].

Tal y como hemos mencionado, la anonimización resulta la única solución viable para el *Big Data* debido a que la regulación, como bien sabemos, avanza con menos rapidez que la tecnología, y esto es lo que ocurre con el *Big Data*. Con la anonimización conseguimos romper la cadena de identificación de las personas al convertir sus datos personales en anónimos. Este proceso de anonimización consigue mitigar el miedo de la población ante las posibles invasiones de su privacidad al tratar sus datos personales de manera continua con dispositivos como smartphones, tablets u ordenadores, de manera que la

[99] AEPD: https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf

[100] AEPD: https://datos.gob.es/sites/default/files/doc/file/orientaciones_y_garantias_anonimizacion_0.pdf

anonimización consigue encontrar un equilibrio entre las nuevas tecnologías y el tratamiento de datos y la privacidad de las personas.

No obstante, el procedimiento de anonimización no es sencillo, debiendo comenzar, tal y como señala el artículo 25 RGPD y el Considerando 78, por la privacidad desde el diseño y por defecto, esto es, pensar en la privacidad desde el comienzo[101]. Antes de que se vayan a utilizar datos, si se prevé utilizar técnicas de *Big Data*, deberíamos aplicar dicho artículo.

Resulta fundamental destacar que, según la normativa[102], un dato anonimizado no es un dato personal, de manera que se evitaría la normativa de protección de datos.

En el proceso de anonimización resulta fundamental, en primer lugar, una evaluación de impacto en la protección de datos personales (en adelante EIPD)[103]. La EIPD, tal y como señala la AEPD[104], es una herramienta que permite evaluar de manera anticipada cuáles son los potenciales riesgos a los que están expuestos los datos personales en función de las actividades de tratamiento que se llevan a cabo en los mismos. El análisis de riesgos para un determinado tratamiento permite identificar los riesgos que se ciernen sobre los datos de los interesados y establecer una respuesta adoptando las salvaguardias necesarias para reducirlos hasta un nivel de riesgo aceptable.

El RGPD prevé que estas evaluaciones de impacto se lleven a cabo antes del tratamiento en los casos que sea probable que exista un alto riesgo para los derechos y libertades de los afectados.

Hoy en día, con la avanzada tecnología, existe una cantidad de datos públicos en plataformas como en redes sociales que permiten relacionar datos anonimizados con personas identificables. Así, la realización de la EIPD conllevaría un análisis de los pros y contras del proceso de anonimización para ver la viabilidad del proceso. La anonimización entraría en juego al diferenciar

[101] DPO&it law: http://www.dpoitlaw.com/reglamento-general-de-proteccion-de-datos-rgpd/unidad-iii-1-2-2-accountability-privacidad-por-defecto-y-privacidad-por-diseno/

[102] Dcitamen 04/2007, de 20 de junio, sobre el concepto de datos personales: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_es.pdf

[103] El RGPD recoge la EIPD en su artículo 35.

[104] AEPD: https://www.aepd.es/sites/default/files/2019-09/guia-evaluaciones-de-impacto-rgpd.pdf

cuáles son los identificadores potenciales y modificar estas variables para reducir el riesgo de reidentificación[105].

El *Big Data* correlaciona una cantidad masiva de datos que, junto con las últimas técnicas, puede llegar a desanonimizar datos personales, de manera que el proceso debe realizarse de manera segura y responsable.

Existen numerosos procesos de anonimización, sin embargo, nos centraremos a lo largo de los siguientes apartados en ejemplos sencillos que nos permitan entender la finalidad de los mismos. Podemos utilizar dos métodos, el básico, basado en utilizar el mínimo número de datos o el basado en introducción de variables, que dificulten la identificación personal.

Así, en primer lugar, tenemos el método basado en la reducción de datos, basado en la reducción o eliminación de variables con la finalidad de minimizar las posibilidades de identificación, a menos número de datos personales menor será la posibilidad de identificar a los individuos. Dentro de este método podemos encontrarnos con las siguientes posibilidades, las cuales definiremos brevemente[106]:

1. Eliminación de variables: Consiste en eliminar datos especialmente sensibles que pueden ser identificadores directos. Esta técnica afecta a la utilidad de los datos, pero, por el contrario, si estos datos no tuviesen relevancia a la hora del análisis podría ser una manera óptima:

564234567P	Blanca ~~**Sanz**~~	Traductora	Artrosis

**********	Blanca ****	Traductora	Artrosis

En este ejemplo, el DNI de la persona es el dato más significativo a nivel identificativo, de manera que eliminarlo resulta fundamental para llevar a cabo la anonimización. Si el

[105] AEPD: https://www.aepd.es/sites/default/files/2019-09/guia-evaluaciones-de-impacto-rgpd.pdf

[106] AEPD: https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf

DNI fuese un dato imprescindible resulta evidente que este método sería imposible de utilizar.

2. Reducción de registros: Se utiliza cuando tras aplicar otra medida los sujetos sigan siendo identificables.
3. Recodificación: Consiste en agrupar determinadas categorías de datos en una nueva categoría, reduciendo las posibilidades de reidentificación.
4. Supresión de registros: Consiste en la eliminación de registros de datos que contienen datos que permiten la identificación de sujetos. Esta medida se utilizará cuando sea imposible anonimizar un determinado sujeto y se hará indicación expresa de los registros eliminados y el motivo por el que se excluyen del resultado final de la anonimización.

Estos métodos consistentes en la reducción de los datos pueden tener alguna desventaja, tales como que el atacante que quiere reidentificar los datos personales tenga información previa o si al llevar a cabo reducción de los datos personales se han perdido datos.

En segundo lugar, tenemos el método de introducción de perturbaciones que, al contrario que el método de reducción de datos mencionado anteriormente, consiste en introducir nuevos datos para cambiar los datos personales, dificultando de esta manera la reidentificación de los mismos, aunque un sujeto accediese a los datos no podría estar a ciencia cierta seguro de cuáles son los datos originales. Dentro de este método nos encontramos con técnicas como el intercambio aleatorio[107], el redondeo[108] o permutación temporal[109], entre otras[110].

Además del método basado en la reducción de datos y el método basado en la introducción de perturbaciones, en la anonimización se utilizan otros como

[107] Según la AEPD, se trata de una técnica basada en la introducción de una distorsión aleatoria en un conjunto de microdatos manteniendo el detalle y estructura de la información original.

[108] Según la AEPD, se trata de la sustitución de variables por valores redondeados de forma aleatoria.

[109] Según la AEPD, se trata de un movimiento aleatorio de rangos temporales que no genera distorsión sobre los resultados medios finales

[110] AEPD: https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf

el algoritmo de Hash, algoritmo de cifrado, sello de tiempo o capas de anonimización.

En conclusión, el proceso de anonimización no puede asegurar la imposibilidad de reidentificación de las personas en términos absolutos. Las nuevas tecnologías hacen difícil que sea posible garantizar el anonimato absoluto cuando se tratan grandes conjuntos de datos personales. Habitualmente existirá el riesgo de que se pueda revertir la anonimización, haciendo posible esta reidentificación de personas.

4.2.3. K-ANONIMIZACIÓN

La AEPD quiere que los sujetos que deben recopilar datos personales garanticen su anonimato. Esto es, que no sea posible identificar a personas a través de su relación. Para ello, en 2019 se publicó una ficha técnica[111] sobre anonimato con una serie de pautas para empresas e instituciones que utilicen procesos de *Big Data* e inteligencia artificial para el tratamiento de datos.

Se busca así comprobar la efectividad de los procesos de anonimización a través del valor K-anonimización, que mide la efectividad de los procesos de anonimización llevados a cabo por un responsable de tratamiento respecto a un conjunto de datos supuestamente anónimos.

La K-anonimización mide la vulnerabilidad de datos que ya han sido anonimizados. Se trata de una propiedad de los conjuntos de datos anonimizados que permite medir cuán de anónimos son los sujetos relacionados con estos datos en los que previamente se han realizados procedimientos de desidentificación. Así, se analiza la probabilidad de que un tercero externo consiga relacionar datos que ya han sido tratados consiguiendo un perfil al que le sean atribuibles los datos[112].

La K-anonimización se puede conseguir a través de diferentes métodos, como son la generalización (haciendo que los valores sean menos precisos) y la eliminación (eliminando ciertos datos para que sea menos probable la identificación).

[111] AEPD, sobre la K-Anonimización: https://www.aepd.es/sites/default/files/2019-09/nota-tecnica-kanonimidad.pdf

[112] Noticias jurídicas: http://noticias.juridicas.com/actualidad/el-sector-legal/14094-k-anonimidad:-la-aepd-publica-una-guia-para-aprender-a-anonimizar-datos/

Este método de medición de riesgo o vulnerabilidad se centra en los datos cuasi-identificadores, es decir, en los que por sí mismos no identifican a una persona, pero en conjunto con otros podrían hacerlo. Así, para que un individuo sea considerado K-anónimo dentro del conjunto de datos en el que se encuentra incluido, es necesario que para cualquier combinación de los atributos cuasi-identificadores asociados, existan al menos otros individuos que comparten con él los mismos valores para esos mismos atributos. El mejor valor es el que resulte ser más elevado, es decir, el que contiene un menor riesgo de desanonimización[113]. Por ello, interesa un valor de K-anonimización alto para garantizar una buena anonimización, ya que mayores valores se corresponden con requisitos de privacidad más exigentes, dado que será necesaria la existencia de más sujetos dentro de un grupo, que satisfagan idéntica combinación de rasgos identificativos[114].

[113] Labeconsultores: https://labeconsultores.com/ciberseguridad-proteccion-datos/k-anonimidad-que-es/

[114] Prodat: https://www.prodat.es/blog/a-vueltas-con-la-anonimizacion-hablamos-de-la-k-anonimidad.html

5. CONCLUSIONES

Como hemos ido viendo a lo largo de este TFM, es posible que la privacidad y la protección de datos no sean compatibles con el *Big Data,* y es necesario buscar un equilibro. En este análisis, hemos podido llegar a las siguientes conclusiones. Todos los servicios que utilizamos a diario conllevan una intromisión en nuestra privacidad en la que actúa el *Big Data,* esos servicios captan nuestra información a cada instante, y, en este mundo digital en el que vivimos, pocas personas estarían dispuestas a dejar de hacer uso de estas aplicaciones que nos facilitan nuestro día a día por el hecho de ser anónimos. En plataformas tan simples y utilizadas como lo son Amazon o Netflix, el *Big Data* y el cruce de datos se produce al predecir nuestros gustos y recomendarnos ciertas películas, documentales o series, produciendo una intromisión en nuestra privacidad.

1. Es más que evidente que la normativa en protección de datos evoluciona de manera más lenta que la tecnología, y el *Big Data* ha sufrido un gran auge y su utilización es cada vez mayor, pero resulta necesario equilibrar las ventajas que nos ofrecen las tecnologías con su intromisión en nuestra privacidad, mediante, por ejemplo, la privacidad por diseño y por defecto recogido en el RGPD.

2. Resulta de gran ayuda tener una regulación actualizada y unificada en materia de protección de datos, tanto nacional, europea como internacional, ya que Internet no conoce de fronteras y la legislación debe tener en cuenta este aspecto e intentar ir por delante para ser uniforme en todos los Estados.

3. Si bien es cierto que la normativa en protección de datos ha ido reforzándose frente al uso de la informática o, lo que es lo mismo, en nuestro caso el *Big Data*, introduciendo, por ejemplo, la figura del consentimiento expreso o el derecho de informar, también lo es que la mayoría de las personas no leen con atención las cláusulas de protección de datos, pues se busca simpleza, rapidez y servicio inmediato.

4. Considerando las implicaciones en la privacidad que supone un tratamiento masivo de datos en el *Big Data*, donde se consigue un beneficio al tratar grandes cantidades de datos, esto conlleva un riesgo más que evidente para la privacidad de las personas, incluso con datos seudonimizados, al existir

una correlación de datos personales, pudiendo generar una reidentificación de las personas. Así, este riesgo va a seguir existiendo con la legislación actual siempre y cuando quieran tratar datos personales basándose en el consentimiento.

5. La única solución que parece viable a este crecimiento imparable del *Big Data* es la anonimización, en la cual, los datos anonimizados quedan fuera de la normativa, pudiendo utilizarse en el entorno del *Big Data* sin preocupaciones, utilizando siempre métodos seguros que no permitan la reidentificación.

6. No obstante, que los ciudadanos utilicen aplicaciones, redes sociales u otro tipo de plataformas o aplicaciones en las que ceden datos o exponen su vida pública no debería suponer un peligro para nuestro derecho a la privacidad. Esto se conseguirá únicamente si la tecnología y la norma avanzan a la par, algo que está claro que no es posible que suceda. Así, la única alternativa para impedir esta intromisión en nuestra privacidad ante el fenómeno del *Big Data*, dada la dificultad para la compatibilizar tecnología y privacidad, sería la utilización de técnicas como la anonimización.

7. Estas técnicas sólo funcionarán siempre y cuando, como hemos mencionado, sean efectivas y se actualicen sus métodos con las nuevas tecnologías para poder evitar que los datos sufran un proceso de reidentificación de datos, evaluando la cantidad de anonimización que contienen los datos con técnicas como la K-anonimización. Esta anonimización debería ir unida, por supuesto, a nuestra responsabilidad ante todo, siendo conscientes de qué compartimos, cómo y cuándo, y nuestras implicaciones de nuestros actos, debiendo abogar, así, por una mayor privacidad desde la responsabilidad.

6. BIBLIOGRAFÍA

AEPD (2019). Código de buenas prácticas en protección de datos para proyectos Big Data. https://www.aepd.es/sites/default/files/2019-09/guia-codigo-de-buenas-practicas-proyectos-de-big-data.pdf

AEPD (2019). *Guía práctica para las evaluaciones de impacto en la protección de datos sujetas al RGPD.* https://www.aepd.es/sites/default/files/2019-09/guia-evaluaciones-de-impacto-rgpd.pdf

AEPD (2019). *La K-anonimidad como medida de privacidad.* https://www.aepd.es/sites/default/files/2019-09/nota-tecnica-kanonimidad.pdf

AEPD (2019). *Orientaciones y garantías en los procedimientos de anonimización de datos personales*. https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf

Agencia de los Derechos Fundamentales de la Unión Europea y Consejo de Europa (2018). *Manual de legislación europea en materia de la protección de datos.*

Álvarez Hazas, Gonzalo (2017). Análisis de los conceptos de anonimización, seudonimización y disociación en el ámbito de protección de datos. Recuperado el 25 de septiembre de 2020 de https://gahazas.com/2017/02/27/analisis-de-los-conceptos-de-anonimizacion-seudonimizacion-y-disociacion-en-el-ambito-de-proteccion-de-datos/

Álvarez Sieiro, Claudia (2019). A vueltas con la anonimización. Hablamos de la K-Anonimidad. Recuperado el 25 de octubre de 2020 de https://www.prodat.es/blog/a-vueltas-con-la-anonimizacion-hablamos-de-la-k-anonimidad.html

APRENDERBIGDATA.COM (2020). ¿Qué es el BIG DATA? – Aprender BIG DATA desde cero. (2020). Recuperado el 25 de septiembre de 2020 de https://aprenderbigdata.com/que-es-el-big-data

Arenas Ramiro, Mónica (2015). Reforzando el ejercicio del derecho a la protección de datos personales: viejas y nuevas facultades. En *Hacia un nuevo derecho europeo de protección de datos (311-372).* Tirant lo Blanch.

Banisar, David y Davies, Simon G. (2012). *Global Trends in Privacy Protection: An International Survey of Privacy, Data Protection, and Surveillance Laws and Developments.* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2138799

Benito Rodero, Ángel (2019). El consentimiento en el RGPD. Recuperado el 25 de septiembre de 2020 de https://www.angelbenitorodero.es/proteccion-de-datos/consentimiento-rgpd

Castrillo de la Fuente, Marta (2018). Recuperado el 25 de septiembre de 2020 de https://www.legaltoday.com/practica-juridica/derecho-publico/proteccion-datos/el-tratamiento-de-datos-de-menores-de-edad-que-dice-el-rgpd-al-respecto-2018-06-28/

Craig, Terence y Ludloff, Mary E. (2011). *Privacy and Big Data.* O'Reilly Media.

Dell Technologies (2014). The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things. Recuperado el 25 de septiembre de 2020 de https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm

Díaz, Cristina (2020). K-anonimidad, ¿qué es?. Recuperado el 25 de septiembre de 2020 de https://labeconsultores.com/ciberseguridad-proteccion-datos/k-anonimidad-que-es/

DPO & it law (2017). RGPD - Unidad I: 6. El Consentimiento. Recuperado el 25 de septiembre de 2020 de http://www.dpoitlaw.com/reglamento-general-de-proteccion-de-datos-rgpd/unidad-i-6-el-consentimiento/

DPO & it law (2017). RGPD - Unidad III: 1.2.2 Accountability – Privacidad por defecto y Privacidad por Diseño. Recuperado el 25 de septiembre de 2020 de http://www.dpoitlaw.com/reglamento-general-de-proteccion-de-datos-rgpd/unidad-iii-1-2-2-accountability-privacidad-por-defecto-y-privacidad-por-diseno/

DPO & it law (2017). RGPD – Unidad I: 5. Principios Relativos al Tratamiento de Datos Personales. Recuperado el 25 de octubre de 2020 de http://www.dpoitlaw.com/reglamento-general-de-proteccion-de-datos-rgpd/unidad-i-5-principios-relativas-al-tratamiento-de-datos-personales/

Durán Arroyo, Alicia (2018): "El nuevo reglamento de protección de datos personales. Análisis de su eficacia en la determinación de su ámbito territorial y los remedios en caso de tratamiento ilícito", *Revista Jurídica de la Universidad Autónoma de Madrid,* núm. 37.

Fernández, Genaro (2018). Cómo funciona el consentimiento en el RGPD. Recuperado el 25 de septiembre de 2020 de https://www.iberley.es/revista/funciona-consentimiento-rgpd-184

Garriga Domínguez, Ana (2016). *Nuevos retos para la protección de datos personales. En la era del Big Data y de la computación ubicua.* Dykinson.

Gil González, Elena (2016). ¿Qué es el big data y por qué debe interesarme si soy abogado?. Recuperado el 25 de septiembre de 2020 de https://www.legaltoday.com/practica-juridica/derecho-civil/nuevas-tecnologias-civil/que-es-el-big-data-y-por-que-debe-interesarme-si-soy-abogado-2016-10-18/

Gil González, Elena (2016). *Big Data, privacidad y protección de datos.* Boletín Oficial del Estado.

Gómez Barroso, José L., Feijóo, Claudio y Martínez, Dolores F. (2017): "Política antes que regulación: la protección de la información personal en la era del Big Data", *Economía Industrial*, núm. 405.

González, Yolanda (2018). La seudonimización y anonimización de datos personales - ¿Qué es?. Recuperado el 25 de septiembre de 2020 de https://protecciondatos-lopd.com/empresas/seudonimizacion-anonimizacion/

Grupo de Trabajo sobre Protección de Datos del artículo 29 (2007). *Dictamen 4/2007 sobre el concepto de datos personales*. https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_es.pdf

Grupo de Trabajo sobre Protección de Datos del artículo 29 (2014). *Dictamen 05/2014 sobre técnicas de anonimización.* https://www.aepd.es/sites/default/files/2019-12/wp216-es.pdf

Hernández López, José M. (2013). *El derecho a la protección de datos personales en la doctrina del Tribunal Constitucional.* Aranzadi.

Herrán Ortiz, Ana I. (2016): "Aproximación al derecho a la protección de datos personales en Europa. El reglamento general de protección de datos personales a debate", *Revista de Derecho, Empresa y Sociedad (REDS),* núm. 8.

Iberley (2019). Consentimiento de menores en materia de protección de datos. Recuperado el 25 de septiembre de 2020 de https://www.iberley.es/temas/consentimiento-menores-materia-proteccion-datos-62818

Instituto de Ingeniería del Conocimiento, UAM (2016). 5 ventajas clave del Big Data. Recuperado el 25 de septiembre de 2020 de https://www.iic.uam.es/innovacion/5-ventajas-clave-big-data/

Kyocera Document Solutions España (2020) Diferencia entre datos estructurados y no estructurados. Recuperado el 25 de septiembre de 2020 de https://www.kyoceradocumentsolutions.es/es/smarter-workspaces/insights-hub/articles/diferencia-entre-datos-estructurados-y-no-estructurados.html

Martín, Bartolomé (2018): "Sobre el ámbito de aplicación material del Reglamento General de Protección de Datos", *Actualidad jurídica Aranzadi*, núm. 943.

Morillo, Jesús (24 de mayo, 2019). Explosión de generación de datos. *El Nacional.*

Noticias Jurídicas (2019). K-anonimidad: la AEPD publica una guía para aprender a anonimizar datos. Recuperado el 25 de septiembre de 2020 de http://noticias.juridicas.com/actualidad/el-sector-legal/14094-k-anonimidad:-la-aepd-publica-una-guia-para-aprender-a-anonimizar-datos/

Pérez, C. (2019). Tipos y Calidad de Datos en Big Data - Claustro ENEB. Recuperado el 25 de septiembre de 2020 de https://claustroeneb.es/2019/01/09/tipos-y-calidad-de-datos-en-big-data/

PROTECCIONDATOS.ORG (2017). Anonimización de datos personales. Recuperado el 25 de septiembre de 2020 de https://www.protecciondatos.org/anonimizacion-de-datos-personales/

PROTECCIONDATOS.ORG (2017). Seudonimización de datos según RGPD - Proteccion de datos. Recuperado el 25 de septiembre de 2020 de https://www.protecciondatos.org/seudonimizacion-de-datos-segun-rgpd/

Rebollo Delgado, Lucrecio (2005). *El derecho fundamental a la intimidad.* Dykinson.

Saldaña Díaz, María N. (2011): "El derecho a la privacidad en los Estados Unidos: aproximación diacrónica a los intereses constitucionales en juego", *Teoría y realidad constitucional*, núm. 28.

Saldaña Díaz, María N. (2012): "The Right to Privacy. La génesis de la protección de la privacidad en el sistema constitucional norteamericano: el centenario legado de Warren y Brandeis", *Revista de derecho político*, núm. 85.

Svetlana Sicular. (2013). Gartner's Big Data Definition Consists of Three Parts, Not to Be Confused with Three "V"s Recuperado el 25 de septiembre de 2020 de https://blogs.gartner.com/svetlana-sicular/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/

Uría Menéndez (2018). *Principales novedades de la nueva Ley Orgánica de Protección de Datos Personales y garantía de los derechos digitales.* https://www.uria.com/documentos/circulares/1030/documento/8327/Aprobacion_LOPD-Garantia_Derechos_Digitales_.pdf?id=8327

Valls Giménez, Josep F. (2017). *Big data: atrapando al consumidor*. Profit Editorial.

Vázquez, Sonia, y De Miguel, Javier (2017). La importancia del seudonimización en el nuevo Reglamento de Protección de Datos. Recuperado el 25 de septiembre de 2020 de https://confilegal.com/20170129-la-importancia-del-seudonimizacion-en-el-nuevo-reglamento-de-proteccion-de-datos/

Wikipedia. Basura dentro basura fuera - Garbage in, garbage out (s.f.). Recuperado el 25 de septiembre de 2020 de https://es.qwe.wiki/wiki/Garbage_in,_garbage_out

Wilder-James, Edd (2012). What is big data?. Recuperado el 25 de septiembre de 2020 de https://www.oreilly.com/radar/what-is-big-data/

WooRank (2020). ¿Qué son los datos estructurados?. Recuperado el 25 de septiembre de 2020 de https://www.woorank.com/es/edu/seo-guides/que-son-los-datos-estructurados