



universidad  
de león



**FACULTAD DE CIENCIAS BIOLÓGICAS Y AMBIENTALES**

**BIG DATA APLICADO AL ÁMBITO  
BIOSANITARIO: ASPECTOS ÉTICOS Y  
LEGALES  
BIOMEDICAL BIG DATA: ETHICAL AND  
LEGAL ASPECTS**

**JUAN JOSÉ VIDAL NÚÑEZ  
GRADO EN BIOTECNOLOGÍA**

**JULIO, 2020**

# ÍNDICE

1. INTRODUCCIÓN .....	- 1 -
2. OBJETIVOS.....	- 2 -
3. MATERIAL Y MÉTODOS .....	- 2 -
4. BIG DATA .....	- 3 -
4.1. DEFINICIÓN .....	- 4 -
4.2. CARACTERÍSTICAS:.....	- 5 -
4.3. EL NUEVO PARADIGMA DE LA INFORMACIÓN.....	- 7 -
4.4. APLICACIONES Y BENEFICIOS .....	- 8 -
4.5. RIESGOS .....	- 11 -
5. LOS DATOS DE CARÁCTER PERSONAL.....	- 12 -
5.1. CONCEPTO.....	- 13 -
5.1.1. “Toda información”.....	- 13 -
5.1.2. “Sobre” .....	- 14 -
5.1.3. “Persona física” .....	- 15 -
5.1.4. “Identificada o identificable” .....	- 15 -
5.2. RELEVANCIA .....	- 17 -
5.3. REGULACIÓN EXISTENTE .....	- 18 -
6. BIG DATA EN EL ÁMBITO BIOSANITARIO .....	- 19 -
6.1. SITUACIÓN ACTUAL .....	- 19 -
6.2. DERECHOS DEL PACIENTE.....	- 20 -
6.3. TRANSFORMACIÓN DEL SISTEMA .....	- 22 -
6.4. PROBLEMAS QUE AFECTAN AL PACIENTE.....	- 24 -
6.4.1. <i>Transparencia para el empoderamiento del individuo</i> .....	- 24 -
6.4.2. <i>Los peligros de predecir el futuro</i> .....	- 26 -
6.5. POSIBLES SOLUCIONES.....	- 27 -
6.5.1. <i>Aumentar las exigencias como ciudadanos</i> .....	- 27 -
6.5.2. <i>Aplicar el principio de responsabilidad proactiva</i> .....	- 27 -
6.5.3. <i>Proteger desde el diseño o por defecto</i> .....	- 27 -
7. EPÍLOGO: CORONAVIRUS Y BIG DATA.....	- 28 -
8. CONCLUSIÓN .....	- 29 -

## RESUMEN

Los datos han estado siempre con nosotros, multiplicándose en los últimos años con la llegada de la Era de la Información. En este contexto, el *Big Data* se presenta como la opción más adecuada para procesarlos y obtener información de ellos. Esta tecnología puede resultar de gran utilidad en uno de los ámbitos donde la Biotecnología tiene más peso y en el que el análisis de datos es esencial: el ámbito biosanitario. Sin embargo, toda implantación conlleva riesgos y, en este caso concreto donde entran en juego los datos personales más íntimos, el peligro se vuelve más evidente. Mediante una revisión bibliográfica de la doctrina existente, podremos conocer el verdadero impacto y riesgo que acompaña a esta tecnología, definir posibles soluciones e inferir cuál debe ser nuestra actitud frente al *Big Data*.

## PALABRAS CLAVE

*Ámbito biosanitario, Big Data, datos personales, derechos, paciente, privacidad.*

## ABSTRACT

Data has always been with us, multiplying in recent years with the advent of the Information Age. In this context, *Big Data* is presented as the most appropriate option for processing and getting information from them. This technology can be particularly useful in one of the areas where Biotechnology plays an important role, the biosanitary field. However, any implementation carries its risks. In this case, where people's most intimate personal data comes into play, danger becomes more apparent. Through a bibliographic review of the existing doctrine, we will be able to know the true impact and risk that are brought with this technology. We will also define some possible solutions and conclude what attitude we shall have towards *Big Data*.

## KEYWORDS

*Big Data, biosanitary field, patient, privacy, rights, risks.*

# 1. INTRODUCCIÓN

El *Big Data* nace con la denominada “revolución digital” (Rivas, 2018) en la “Era de la Información”, siendo quizá uno de los elementos claves en su desarrollo y permitiendo comprender el mundo de una manera inimaginable, gracias a la interconexión de los individuos (Fernández de Lis, 2019; Cano, 2020). Allí, donde antes solo había una cabeza pensante dispuesta a hablar y entender el mundo, mientras los demás permanecían apresados en los dogmas, se enlazarán ahora todas esas ideas enclaustradas (Molero, 2014).

Cada vez más personas con sus dispositivos se convierten en usuarios que observan el mundo, lo fotografían, visitan y documentan. Plasman sus inquietudes allá dónde pueden a la espera de que otros simplemente las miren o, en el mejor de los casos, las compartan. Nuestra innata curiosidad, alimentada como nunca, a golpe de tecla y generando con cada búsqueda, cada comentario, cada movimiento, una cantidad ingente de datos (Cisco Systems Incorporation, 2020). Estos datos crecerán, duplicándose cada año, y son ya para muchas empresas un recurso con mucho valor (Huawei Technologies Company Limited y Oxford Economics Limited, 2017; Reinsel *et al.*, 2018; Roser *et al.*, 2020). Aquellas empresas que sepan invertir en *Big Data* se verán beneficiadas en el largo plazo, pues cada vez tienen más peso las tecnologías traídas por esta revolución (Huawei Technologies Company Limited y Oxford Economics Limited, 2017; Reinsel *et al.*, 2018).

Unas tecnologías que empezaron a dar sus primeros pasos con la invención de los transistores y procesadores (Roser y Ritchie, 2013). A partir de ese momento, el desarrollo ha sido a pasos agigantados. Los procesadores comenzaban a duplicar el número de transistores presentes en ellos, un escalado que fue descrito por Moore y que, según su teoría, duraría 10 años (Moore, 1965). Su teoría se cumplió, convirtiéndose en Ley, pero lo sorprendente no ha sido su precisión, sino que hoy en día, y gracias a esa mejora exponencial, otras muchas actividades han visto incrementada su capacidad en la misma medida: el almacenamiento, la velocidad de procesamiento, el abaratamiento de los dispositivos, etc (Roser y Ritchie, 2013). Estas mejoras pueden explicar el mayor acceso por parte de la población a los dispositivos, generando más datos que pueden ser estudiados y, a su vez, favoreciendo el desarrollo de tecnologías que los procesen (Roser y Ritchie, 2013).

Han pasado ya más de 10 años y es evidente que todos estos avances y el *Big Data* ya están implantados en nuestra Sociedad. Serán la base de aquello que este por venir, afectarán a todos

los sectores: sanidad, investigación, economía, cultura, etc. No debemos preguntarnos cuándo o cómo llegará el *Big Data*, sino qué uso le daremos.

## 2. OBJETIVOS

Los objetivos del presente trabajo son:

- Conocer qué es y qué caracteriza al *Big Data* para después reflexionar sobre su impacto, con sus límites y riesgos.
- Entender la importancia de proteger nuestros datos personales y porqué son la base de nuestra privacidad.
- Buscar la relación del *Big Data* y la Biotecnología a través del ámbito donde esta última tiene más presencia, el biosanitario.
- Reflejar someramente la situación del ámbito sanitario y del paciente.
- Saber qué beneficios y riesgos genera la implantación de tecnologías como el *Big Data* en el ámbito biosanitario.
- Discutir las posibles vulneraciones de los derechos del paciente al usar el *Big Data* en el ámbito biosanitario y aportar soluciones a estas.
- Inferir, tras haber cumplido el resto de los objetivos, cuál debe de ser nuestra actitud hacia el *Big Data*.

## 3. MATERIAL Y MÉTODOS

Para cumplir con los objetivos de este trabajo se ha realizado una revisión bibliográfica de la literatura aportada por distintas instituciones y organizaciones que operan a nivel internacional.

Algunas de las instituciones consultadas han sido:

- La Organización Mundial de la Salud (OMS).
- Distintas organizaciones dentro de la Unión Europea: Comisión, Parlamento y el Portal Europeo de Datos.
- La Organización para el Desarrollo y la Cooperación Económicas (OCDE).
- La base de datos de Our World in Data.
- La Organización de las Naciones Unidas (ONU).
- La Federación Europea de Industrias y Asociaciones Farmacéuticas (EFPIA por sus siglas en inglés).
- La página web del proyecto europeo Bigmedylitics, con sede en Valencia.

También se ha recurrido a los textos proporcionados por los propios estados, como es el caso del Boletín Oficial del Estado (BOE) español u otros documentos que marcan la doctrina actual en este tema. Así como otros textos de diversa índole: noticias o artículos periodísticos o científicos, libros de carácter divulgativo, etc.

## 4. BIG DATA

El término inglés *Big Data* hace referencia a las grandes bases de datos, siendo *Data* el plural de *Datum* (Cambridge University Press, 2020) que, en latín, quiere decir “lo dado” (Real Academia Española, 2020c). Esto es el *Big Data*: una colección de datos que se nos ha dado y que nosotros usamos como base para otras operaciones o para procesarlo.

No hicieron falta grandes ordenadores para que el concepto de *Big Data* comenzase a dar sus primeros pasos (Mayer-Schonberger y Cukier, 2013). El ser humano vive expuesto, como cualquier otro ser vivo, a un flujo constante de información producto de nuestras interacciones. Aunque durante siglos no hubo ningún registro de ellas y solo las manteníamos en nuestras mentes, con la llegada de las primeras sociedades fue necesario inventariar todo lo que sucedía (Mayer-Schonberger y Cukier, 2013). Podemos decir que el primer *Data* nace así, con tablillas de barro que recogían las cosechas, propiedades o ganado de los ciudadanos. Tablillas que los afanados escribas sumerios utilizaban no sólo para guardar la información, sino para realizar interpretaciones más complejas que permitieran a su rey recaudar bien los impuestos, redactar leyes o abastecer sus ejércitos.

Unos inicios modestos, pero que sentaron las bases para el nacimiento del *Big Data*. Los avances siguieron y el ser humano generaba cada vez más y más información para interpretarla. No solo eso, también fue capaz en algunos casos de descubrir relaciones ocultas en datos que aparentemente no tenían mayor relevancia (Lewis, 1927; Mayer-Schonberger y Cukier, 2013).

No obstante, fue con el nacimiento de Internet en 1989, cuando los empresarios, científicos y analistas empezaron a fijar realmente su atención en el almacenamiento de datos y como procesarlos (Roser *et al.*, 2020). Las bases de datos aumentaron de tamaño y los esfuerzos se centraron en buscar relaciones que antes no se podían ver. Es en un artículo de 1997, publicado por dos trabajadores de la NASA Michael Cox y David Ellsworth, cuando apareció por primera vez el término *Big Data* en un texto científico (Cox y Ellsworth, 1997; Vodafone España, 2018). No es solamente un conjunto de datos, sino todas las relaciones que podemos establecer y que pueden ser interpretadas con un sentido más o menos lógico para aplicarlas en algo muy

concreto. Desde aquí comenzaron a parecer elementos que se han quedado con nosotros: las 3 V del *Big Data*, Hadoop, plataformas seguras y reguladas para el almacenamiento (Cloudera o Hortonworks), las “Smart Cities”, los sistemas de recomendación de páginas web como Amazon o Netflix, etc (Vodafone España, 2018). Todo esto hasta llegar a campañas políticas donde no solo se ha hecho uso de esta tecnología, sino que ha sido parte de algunos de los programas electorales (Juste, 2019).

En resumen, sabemos que el *Big Data* está presente en nuestras vidas, forma parte de muchas otras cosas e incluso, por lo expuesto hasta ahora, podríamos hablar de un concepto cuyas bases llevan siglos en nuestra sociedad. Sin embargo, ¿qué definición damos al *Big Data* y cuáles son sus características? ¿para qué lo usamos? o, simplemente, ¿está controlado algo que parece tan importante y que está tan implantado en nuestra era? Algo que está tan implantado en nuestro mundo, lo estará también en las ciencias como la Biotecnología y, por lo tanto, en el ámbito biosanitario. Es aquí donde la información, ya sea procedente de ensayos clínicos de una medicina o de un extenso estudio sobre algún genoma desconocido, es cada vez mayor. ¿Qué impacto tendrá una tecnología de este tipo? ¿deberán los científicos ver frenados sus avances por los peligros que suponga para el ciudadano? Son algunas de las preguntas que se plantean y que se responderán en los apartados siguientes.

#### 4.1. DEFINICIÓN

El *Big Data* puede definirse como el conjunto de recursos informativos o datos, de gran volumen, velocidad y variedad que demandan un procesamiento innovador y eficiente que mejore su interpretación, la toma de decisiones en torno a ellos y la automatización de procesos (Gartner Incorporation, 2020).

La primera definición nos habla de uno de los rasgos característicos del *Big Data* y son las “3- uves”: el volumen, la velocidad y la variedad (Laney, 2001; Richards y King, 2014; Gil González, 2016). Para que una base de datos pueda ser considerada dentro de la definición de *Big Data*, esta debe contener una gran cantidad de datos que no solo se generen, sino que también sean procesados por la organización poseedora de estos, a la mayor velocidad posible (Laney, 2001; Gil González, 2016). No todos los datos que se han recolectado serán de utilidad y la propia base que los colecta debe estar optimizada para clasificarlos o eliminarlos, lo que pone de manifiesto la tercera “uve”: la variedad (Laney, 2001; Gil González, 2016). Esta variedad no se debe sólo a los distintos datos que hay en la propia base, sino a la existente también entre las distintas bases y la recolección de estos (Laney, 2001).

El Diccionario Jurídico de la Real Academia Española, añade un componente humano a esta definición, el *Big Data* es el “conjunto de técnicas que permiten analizar, procesar y gestionar conjuntos de datos extremadamente grandes que pueden ser analizados informáticamente para revelar patrones, tendencias y asociaciones, especialmente en relación con la conducta humana y las interacciones de los usuarios” (Real Academia Española, 2020a).

Estas definiciones, aunque correctas, omiten cada una alguno de los elementos del *Big Data* que, sin embargo, sí figura en la homóloga. Entender la importancia de estos elementos y en qué consisten nos permitirá dar una definición más precisa al final de este apartado, que tenga en cuenta los planos técnico y social del *Big Data*.

Otro aspecto importante para la definición y que, sin embargo, no aparece en la que da el glosario de Gartner es: la interacción del usuario. Las bases de datos que se incluyen en el concepto de *Big Data* basan su existencia en los datos recogidos de las interacciones del ser humano (Mayer-Schonberger y Cukier, 2013; Gil González, 2016). Gracias a estas interacciones llega a ser posible analizar tendencias o incluso conductas del ser humano (Ginsberg *et al.*, 2009).

Por tanto, propongo una definición más completa y que es más acorde con las que se dan en los distintos ámbitos, teniendo en cuenta los planos social y técnico del *Big Data*: El *Big Data* es toda base de datos de gran volumen, cuya recolección y procesamiento se realiza a grandes velocidades y con una gran variedad de información, procedente de distintas interacciones humanas. Su uso y análisis tiene como objetivo generar predicciones y mejorar nuestra capacidad de decisión para la automatización de procesos relacionados con las conductas y tendencias de los individuos o grupos que aportaron los datos.

## 4.2. CARACTERÍSTICAS:

Partiendo de la definición anterior, vamos a explicar, en mayor profundidad, las características que hacen al *Big Data* tan distinto de las tecnologías anteriormente usadas en el procesamiento de datos. Estas características van más allá de las tres uves de volumen, velocidad y variedad, pero antes de explicarlas vamos a detenernos un momento en la importancia de la simultaneidad de estas tres letras (Gil González, 2016).

Las tres uves parecen algo trivial que hemos visto en otras bases de datos, pero ¿por qué cobran especial relevancia con el *Big Data*? Como ya se ha mencionado antes, se produce una



simultaneidad de las tres, algo que nunca había sido posible. Podemos tener grandes bases de datos y procesarlas velozmente con independencia de las diferencias existentes en los datos. Es por ello por lo que se dice que las tres uves son únicas del *Big Data*, pues en el antiguo tratamiento de las bases de datos había que renunciar a alguna de las tres para cumplir las otras dos (Gil González, 2016). Si queríamos procesar datos muy diversos a gran velocidad, no podíamos tener grandes conjuntos de información.

Estas “tres-uves” se convierten en las “seis uves” si tenemos en cuenta otras, que son: la veracidad, la visualización y el valor (Schroeck *et al.*, 2012; Gil González, 2016).

- Veracidad: haciendo referencia a la calidad de los datos y su fiabilidad. Es un campo en el que cada vez somos más precisos, pudiendo medir cosas que antes eran impensables. Sin embargo, esta tecnología es propensa a generar demasiados datos, sobre todo cuando la fuente de procedencia son entornos humanos como las redes sociales (Schroeck *et al.*, 2012). En estos casos, la calidad se pierde y aumenta la incertidumbre. Es por ello por lo que será necesario gestionarlos de manera adecuada y, en muchas ocasiones, recurrir a distintas fuentes que nos permitan contrastar para obtener como resultado una información más fiable.
- Visualización: la acumulación y utilización de los datos, así como su publicación, queda inutilizada si no somos capaces de darles una visualización adecuada (Gil González, 2016). Pongamos como ejemplo la aplicación de “Google Maps”. Nos puede mostrar en tiempo real el tráfico existente en las calles por las que nos desenvolvemos de manera habitual. Por desgracia, como apuntan algunos autores con otros ejemplos, un listado de calles con una palabra o símbolo al lado de cada una, indicando el nivel de tráfico, sería difícil de manejar (Gil González, 2016). Ante esta situación, en la aplicación de Google, podemos ver las distintas carreteras coloreadas, correspondiéndose la mayor congestión con el color rojo y los niveles más bajos con el verde (habiendo otros colores y tonalidades para los diferentes niveles de congestión).
- Valor: aquel que se obtiene por la investigación asociada al uso de la tecnología o el propio valor económico derivado del uso de los datos (Gil González, 2016). Como cualquier otra actividad humana, se dirige a generar algo que para nosotros tenga valor.

Algunos autores destacan también el uso intensivo de algoritmos por parte del *Big Data* o la utilización de los datos con otros fines como otras características añadidas (Al Nuaimi *et al.*, 2015; Gil González, 2016). Este uso de nuevos algoritmos, que cada vez son más potentes, no

es exclusivo de esta tecnología, pero si se le da un nuevo uso: el aprendizaje del propio algoritmo a partir de la información extraída de los primeros datos. Aquí es donde entra en juego la reutilización de los datos o el uso de los ya existentes para fines en los que en un inicio no se había pensado.

### 4.3. EL NUEVO PARADIGMA DE LA INFORMACIÓN

Todas las características mencionadas anteriormente son las que hacen que el *Big Data* sea diferente a cualquier tecnología desarrollada hasta ahora y que sea capaz de modificar el paradigma de la información. Este cambio se produce en tres direcciones (Mayer-Schonberger y Cukier, 2013; Gil González, 2016):

- El **cambio hacia el “todo”**: la estadística tradicional trataba los datos a base de muestreos representativos de la realidad. Esto se traduce en pequeñas cantidades de información, manejables y capaces de explicar cuestiones generales que fueran complejas. Sin embargo, estos muestreos no permiten estudiar subgrupos dentro de ellos sin perder fiabilidad. El *Big Data* no recurre a muestreos, sino que recoge todos los datos, para ser usados cuando nosotros queramos. Gracias a ellos podemos estudiar comportamientos particulares de los subgrupos que nos interesen, sin tener que renunciar a la fiabilidad.
- El **cambio hacia el “caos”**: algo inherente a recolectar semejante cantidad de datos y de tan diversa índole es, su desorden. Sin embargo, hay que convivir con él, rompiendo nuestra costumbre de estructurarlo todo y dándonos cuenta de que las inexactitudes (siempre que no sean datos completamente incorrectos), son compensadas por los beneficios a obtener (Mayer-Schonberger y Cukier, 2013; Gil González, 2016).
- El **cambio hacia la “correlación”**: no se trata de saber *por qué* ocurre, sino *qué* (Mayer-Schonberger y Cukier, 2013). No queremos buscar causas en los eventos, sino analizarlos en busca de patrones que hubieran pasado desapercibidos ante nuestros ojos. Esta metodología tiene un gran potencial, pero tiene una serie de riesgos asociados.

Puede parecer que este cambio es algo deliberado, como si la comunidad científica se guiara por un impulso ideológico, pero nada más lejos de la realidad. Estos cambios en los paradigmas de la Ciencia no son infundados, sino contrastados. La adopción del *Big Data* está en sus inicios, dando respuestas a interrogantes que la experimentación no ha podido, pero sembrando los suyos propios. No podemos asegurar que este cambio ya se haya producido, ni mucho menos relegar a un segundo plano el paradigma actual de las Ciencias experimentales.

#### 4.4. APLICACIONES Y BENEFICIOS

Un ejercicio útil para comprender cómo el *Big Data* está cambiando nuestra forma de pensar y las bases del conocimiento científico es ver sus aplicaciones. Atendiendo a aquellos campos donde se hace un uso de esta tecnología, podemos ver qué camino está tomando y el beneficio que supone. Esta herramienta tiene aplicaciones que van desde el diseño de nuevas ciudades inteligentes a otras como la gobernación o el ámbito biosanitario (Al Nuaimi *et al.*, 2015; Gesundheit Österreich Forschungs- und Planungs GmbH *et al.*, 2016; Coma *et al.*, 2020). En este último se centra este trabajo y más adelante se ahondará en los aspectos éticos y legales derivados de su implantación.

Cuando hayamos comentado algunos ejemplos prometedores de aplicación de esta tecnología, viendo el potencial beneficio, comentaremos los riesgos que supone su implementación para el ser humano. Tres campos en los que se ha hecho o hace uso del *Big Data* son:

- Las **redes sociales**: el primer ejemplo que viene a la mente de muchos cuando se habla de información y usuarios son las redes sociales. Todos los datos que definen a cada individuo dispuestos a unas pocas compañías. Facebook junta en sus bases de datos la información de un 10% de la población mundial y junto con otras compañías, ya está estudiando cómo sacarle provecho a la recolección y análisis de estos datos (Mayer-Schonberger y Cukier, 2013; Van Dijck, 2014).

Dentro de las plataformas podemos encontrarnos con sugerencias de amistad, mensajes o páginas más afines a nuestros gustos e incluso el impacto de eventos o sucesos recientes. Muchas veces los datos de estas redes no son procesados directamente por ellas, sino que son terceros los que usan los datos con un fin determinado (Gil González, 2016). La excesiva confianza en estos datos para determinados estudios puede ser un grave error (Gil González, 2016).

- Las **plataformas de venta**: todo el mundo ha comprado libros o es seguidor de alguna serie que seguramente esté disponible en alguna de las múltiples plataformas. Todas ellas, cada vez que adquirimos uno de sus productos, toman nuestros datos y con ellos algo más importante: los metadatos. Todos nuestros “favoritos”, compras pasadas, las etiquetas y opiniones de los artículos son tenidas en cuenta por los algoritmos de estas plataformas para conocer la mentalidad de su cliente (Mayer-Schonberger y Cukier, 2013; Richards y King, 2014).

Gracias a la imagen que generan estos algoritmos del cliente, las plataformas han logrado hacer recomendaciones cada vez más acertadas al usuario. En el caso concreto de Amazon esta evolución pasó de un comité de críticos de libros, conocidos como “la voz de Amazon”, a un algoritmo que hacía recomendaciones simples basadas en otros clientes (Mayer-Schonberger y Cukier, 2013; Gil González, 2016). El algoritmo cogía una pequeña muestra de clientes que hubieran adquirido el mismo artículo y se nos recomendaba en función de la similitud con el grupo (Mayer-Schonberger y Cukier, 2013). Este modelo era muy limitado, si comprábamos un libro de repostería el resultado eran miles de recomendaciones de libros de repostería, sin apenas variación.

Con el tiempo se dieron cuenta de que era mucho mejor hacer una relación, no entre clientes, sino entre productos (Mayer-Schonberger y Cukier, 2013). De esta forma la recomendación se establecía más rápido y era mucho más acorde a los gustos del cliente.

- Las grandes **firmas electrónicas**: aquí cabe destacar a uno de los gigantes en nuestra era de la información, Google. Sus bases de datos y sus algoritmos les han permitido desarrollar motores de búsqueda potentes, adquirir y mejorar redes sociales existentes e incluso traducir textos de todos los idiomas posibles. Sin embargo, vamos a destacar aquí uno de sus hitos que, dejando de lado la mala visión que pueda existir de la compañía (Suárez, 2019), supuso un gran avance en el estudio de datos.

Google fue capaz en el año 2009 de crear un modelo de predicción y monitoreo de la gripe (Ginsberg *et al.*, 2009), usando para ello potentes algoritmos y una base de datos inmensa: las búsquedas de sus usuarios. Repararon en que muchas de las búsquedas previas a una cita con el médico, ante sospecha de estar enfermo, eran coincidentes y que en muchos el resto de las búsquedas, no contempladas para la generación de su modelo, terminaban teniendo una relación indirecta. Gracias a todo ello pudieron hacer un modelo que, con datos casi en tiempo real, analizaba la evolución de la gripe con un día de retraso (Ginsberg *et al.*, 2009).

Mientras tanto, los servicios epidemiológicos siempre han llevado un retraso en sus análisis de hasta 2 semanas. La causa es clara: es necesaria una visita al médico, una analítica y su inclusión en la estadística. Un trabajo riguroso y preciso, pero muy lento.

Estos tres ejemplos son una simple muestra de los tantos usos que tiene el *Big Data* en el mundo. Sirven para hacerse una idea rápida de todo el potencial de esta tecnología, pero no representan todos los campos en los que se puede utilizar. Sin embargo, sí nos servirán para exponer los beneficios y riesgos que llevan asociados.

Los Beneficios del *Big Data* para las empresas podrían clasificarse en cinco niveles, según el aspecto que mejoren de la empresa (Wang *et al.*, 2018). Los niveles o categorías que se presentan a continuación suelen usarse también a la hora de valorar la implantación de cualquier otro sistema en una empresa (Gefen y Ragowsky, 2005; Esteves, 2009).

- **Beneficios de infraestructura:** son todas las mejoras que se realicen en las tecnologías de la información de la empresa u organización y que podrán ser aprovechadas para futuros eventos (Wang *et al.*, 2018). Google y su modelo de predicción de la epidemia de la gripe son el ejemplo perfecto. Un recurso informático que, si bien puede tener un uso inmediato, servirá de base para proyectos similares que quiera llevar a cabo la empresa.
- **Beneficios operacionales:** cualquier mejora en el sistema o los procesos (Wang *et al.*, 2018). Podríamos poner como ejemplo las recomendaciones basadas en el análisis de las enormes bases de datos de las redes sociales o las plataformas de compra.
- **Beneficios para la gestión:** referentes al monitoreo de nuestros recursos y la automatización o mejora en la toma de decisiones.
- **Beneficios estratégicos:** obtenidos para actividades o tareas muy concretas dentro de la empresa (Wang *et al.*, 2018). Suelen implicar decisiones muy importantes en el largo plazo. El ejemplo que puede ser más claro con las redes sociales, quienes buscan sacar provecho de su ingente cantidad de datos y crear lazos con empresas que, a través de alguna idea innovadora, den un valor añadido al análisis de sus datos.
- **Beneficios en la organización:** cambian por completo, o solo donde se requiera, aspectos como el organigrama, el enfoque de la empresa, etc (Wang *et al.*, 2018). Un claro caso en el que la empresa obtuvo un beneficio a este nivel es Amazon. Haciendo uso del *Big Data* no solo mejoró su sistema de recomendaciones, como comentábamos, sino que cambió su equipo al ser más eficiente y rentable centrar su plataforma en saber cuál será el próximo libro que recomendar al cliente y no en tener a un brillante comité de críticos dando su opinión de determinados títulos (Wang *et al.*, 2018).

Es importante hacer notar que este sistema, desarrollado por los autores para estudiar los beneficios, parte de una base muy difícil y es reflejar de manera objetiva el beneficio. Para hacerlo subdividieron cada uno de los niveles en categorías y dentro de cada proyecto encajaban las distintas declaraciones de posibles beneficios en una categoría u otra (Wang *et al.*, 2018). Después, con el compendio de todos los proyectos y declaraciones, podían dilucidar a qué nivel

se presenta un mayor beneficio aplicando el *Big Data* en un campo determinado (Wang *et al.*, 2018). En su caso, en el ámbito biosanitario.

Finalmente, hay que aclarar que los ejemplos propuestos para cada uno de los niveles pretenden reflejar cómo distintas aplicaciones del *Big Data*, en campos completamente diferentes, suelen tener mayor incidencia en determinados niveles de beneficio. Sin embargo, estos ejemplos no son resultado de aplicar de forma minuciosa el modelo propuesto por el autor en el artículo. Si quisiéramos conocer la influencia en los beneficios de Google causada por una herramienta relacionada con el *Big Data*, de acuerdo con la metodología descrita en el trabajo, tendríamos que buscar las distintas declaraciones en favor de estos beneficios y clasificarlas en las subcategorías creadas por el autor para comprobar a qué nivel hay mayor beneficio.

#### 4.5. RIESGOS

Aunque el *Big Data* tiene un enorme potencial, se enfrenta a ciertas limitaciones. Dejando de lado posibles inconvenientes a nivel técnico en el almacenamiento o el análisis, tenemos tres principales riesgos:

- **Riesgo de la causa errónea:** se pueden hacer predicciones con facilidad y buscar patrones, pero hay que ser cautelosos (Gil González, 2016). Encontrar una relación no tiene por qué implicar causa y puede ser una mera casualidad, en pocas palabras, es una relación espuria o falsa.

Este error no es algo nuevo (Neyman, 1952) (se puede mentir por educación, malicia o hacer una interpretación estadística falaz), pero si no es nuevo ¿por qué habríamos de mostrar una especial preocupación con el *Big Data*? La respuesta es porque es más fácil cometer el error. Cuando tenemos una mayor cantidad de datos es más probable que hagamos una relación que no procede y esta nueva tecnología basa gran parte de su impacto en detectar conexiones a gran velocidad y en gran cantidad desde grandes bases de datos. Solamente hay que imaginar el terrible desastre que podría desencadenar el uso de alguna de estas relaciones falsas en campos como el biosanitario y el efecto negativo que tendría en el paciente (Cohen *et al.*, 2014).

Resumiendo, debemos ser críticos con los resultados, intentar buscar experimentalmente la causa y exigir rigor científico en cualquier estudio. El *Big Data* puede arrojar luz sobre las pistas, pero se debe complementar con la experimentación durante años para comprobar que no estamos siguiendo pistas falsas.

- **Riesgo de automatización indeseada:** suponiendo que nuestras relaciones son correctas y no ha habido ningún error conoceremos *qué* hace falta y no será necesario decidir (Gil González, 2016). Esta automatización conlleva varios riesgos ¿y si el algoritmo falla? Podríamos tener a personas revisando los resultados arrojados, algo que cada vez es menos frecuente, pues es demasiada información para ser procesada. Además, esta idea no está plenamente apoyada por toda la comunidad científica, pues el ser humano tiene sesgos que la máquina no (Gil González, 2016). Otra solución al problema podría ser perfeccionar nuestros algoritmos *a priori*, evitando así que tengamos que comprobar *a posteriori* si la automatización está tomando las decisiones correctas y puede ser aplicable (Cohen *et al.*, 2014; Gil González, 2016).

Conviene subrayar que nuestro actual Reglamento General de Protección de Datos nos ampara y, salvo situaciones específicas, no permitirá la automatización de decisiones que afecten directamente a un individuo y, cuando lo haga, se deben garantizar plenamente los derechos del individuo y su no discriminación (Consejo de la Unión Europea y Parlamento Europeo, 2016).

- **Riesgo en la privacidad:** el más evidente de todos los riesgos que rodean al *Big Data* (Gil González, 2016). Todo ciudadano de la Unión Europea tiene derecho a una privacidad, siendo esta un concepto más amplio que la intimidad, pues incluye toda información referente al individuo (Real Academia Española, 2020b). Esta información son datos personales que pueden obtenerse directamente o ser inferidos de otros datos aportados por el individuo (Sweeney, 2000).

El *Big Data* amenaza, en su afán por recopilar datos e interpretarlos, la privacidad de las personas, poniendo en riesgo sus datos personales. Actualmente, nuestros reglamentos intentan salvaguardar nuestro derecho a una vida privada y a que nuestros datos sean anónimos, a que no se nos pueda vincular a una información o a un análisis derivado de esta. Sin embargo, esto no es fácil y veremos a lo largo de este trabajo que, perder esta privacidad y exponernos, es el mayor riesgo al que nos enfrentamos con el *Big Data*.

## 5. LOS DATOS DE CARÁCTER PERSONAL

Hemos comentado antes que el mayor riesgo que podemos correr al confiar nuestros datos al *Big Data* es perder nuestra privacidad, exponiendo nuestros datos personales. Veremos más adelante que no es el único de los problemas que surgirán de la relación entre esta herramienta y el ámbito biosanitario, pero sí es el de más peso. Para entender la situación de nuestros datos

en el ámbito sanitario primero tenemos que saber qué son, porqué son importantes y cómo se garantiza su protección.

## 5.1. CONCEPTO

Según el Reglamento General de Protección de Datos, los datos personales se definen como: “toda información sobre una persona física identificada o identificable («el interesado»); se considerará persona física identificable toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un identificador, como por ejemplo, un nombre, un número de identificación, datos de localización, un identificador en línea o uno o varios elementos propios de la identidad física, fisiológica, genética, psíquica, económica, cultural o social de dicha persona” (Consejo de la Unión Europea y Parlamento Europeo, 2016).

En la primera línea podemos ver los cuatro elementos que construyen la definición (Article 29 Data Protection Working Party, 2007). Todos ellos están fuertemente interconectados, pero los analizaremos por separado para entender su papel dentro del concepto de los datos personales.

### 5.1.1. “Toda información”

Este primer bloque pone de manifiesto la intención de que el concepto de dato personal sea lo más amplio posible y con una interpretación abierta (Article 29 Data Protection Working Party, 2007). De esta manera se incrementa la protección, pues alcanza cualquier dato objetivo/subjetivo acerca del individuo.

Incluir “cualquier dato” entraña un problema y es que el individuo puede introducir en la base datos incorrectos. El Reglamento ya lo prevé y garantiza el acceso del individuo a sus datos para poder corregirlos o suprimirlos (Consejo de la Unión Europea y Parlamento Europeo, 2016). Si la norma no tuviera ese alcance sólo lograríamos proteger una pequeña parte de los datos personales, otros como los gustos, las relaciones o la situación como empleado, quedarían excluidos de esta protección (Article 29 Data Protection Working Party, 2007). Estos datos no se recolectan de forma “objetiva” y directa, sino que pueden ser inferidos por otras vías. También debemos mencionar que esta información y el derecho a la protección de esta, gozan de artículo propio (Artículo 8) en la Carta de Derechos Fundamentales de la Unión Europea (Estados miembros, 2016), aunque estén relacionados con la *privacidad*, que tiene su propio espacio reservado en el artículo 7, (Estados miembros, 2016). Esto prueba la voluntad por parte de las instituciones de garantizar que el alcance de la normativa va más allá de la protección de la privacidad y la vida familiar.



Por último, hay que aclarar que “toda información” nos está avisando de que el formato tampoco es limitante para que algo sea considerado un dato personal (Article 29 Data Protection Working Party, 2007). En este trabajo nos centraremos en lo referente a *Big Data* y sus grandes bases de datos, pero una fotografía o un correo (siempre y cuando nos sirvan para identificar a un individuo) pueden ser considerados también como “datos personales”. Este tipo de datos están clasificados como *categorías especiales* dentro del Artículo 9 del Reglamento y su tratamiento es especial por todo lo que conllevan (datos fisiológicos, genéticos, sanitarios, etc.) (Consejo de la Unión Europea y Parlamento Europeo, 2016). Ocupan un lugar especial dentro del ámbito biosanitario y es que son a la vez *dato e identificador*, es decir, unas veces será la información privada de interés y otras un mero vínculo que nos permitirá deducirla (Article 29 Data Protection Working Party, 2007; Gil González, 2016).

### 5.1.2. “Sobre”

Decimos que los datos están relacionados, es información *acerca de* una determinada persona (Article 29 Data Protection Working Party, 2007). Aunque pueda parecer baladí detenerse a discutir sobre el significado y valor de esta preposición en la definición que damos de un dato personal, no lo es. Para demostrarlo nos puede bastar con un ejemplo:

Imaginemos que se va a subastar una vivienda y el precio de partida que nos marcan es bajo, pero es una buena casa. El atributo del precio es una información “sobre” el inmueble y no tiene por qué ser un dato personal. Sin embargo, si logramos vincular este dato a un nombre y descubrimos que el precio de partida es bajo debido a la apurada situación económica del vendedor, jugaremos con ventaja. Estamos estableciendo una relación y en este caso asociando el precio de una vivienda con un individuo.

El ejemplo anterior extrema la situación, pues da por hecho que tenemos fuentes complementarias para conocer quién era el anterior propietario y para desvelar su difícil situación. No obstante, permite hacernos una idea de la importancia de valorar qué información se puede asociar con nosotros.

De esta forma, para considerar que un dato o información es “sobre” un individuo tiene que haber un elemento de contenido, propósito o resultado en él (Article 29 Data Protection Working Party, 2007).

- **Contenido:** hablamos de este elemento cuando la información está directamente relacionada con el individuo y no requiere de ninguna deducción previa. Un análisis de sangre, una factura, etc.
- **Propósito:** puede considerarse que existe este elemento cuando hay un uso, o probable uso, de los datos con el fin de evaluar o influenciar el estado o comportamiento del individuo.
- **Resultado:** su existencia implica que previamente se han procesado datos que, no siendo referentes al individuo, el resultado de su análisis sí puede serlo.

Los tres elementos no son acumulativos, la sola presencia de uno de ellos es suficiente para considerar un dato como personal y relacionado con un individuo (Article 29 Data Protection Working Party, 2007).

### 5.1.3. “*Persona física*”

El Reglamento protegerá a todas las personas físicas, esto es, a todos los seres humanos dentro de la Unión sin ninguna distinción. El concepto de persona física lo marca la Declaración Universal de los Derechos Humanos en su Artículo 6 (Asamblea General de las Naciones Unidas, 1948). Sin embargo, podemos tener dudas cuando nos salimos de estos límites y nos preguntamos en qué situación estarán los fallecidos, los no nacidos y las personas jurídicas.

Es necesario saber que no se aplicaría el Reglamento sobre los datos de ninguna de ellas (Consejo de la Unión Europea y Parlamento Europeo, 2016). Esto no quiere decir que no haya una legislación específica para fallecidos, empresas o no nacidos y que vele por sus datos, pero está terminantemente prohibido que cualquier país europeo aplique, más allá del alcance con el que fue planteado, el Reglamento (Consejo de la Unión Europea y Parlamento Europeo, 2016). Anotar que, en el caso concreto de los fallecidos la normativa se aplicará siempre que, derivado de algún estudio histórico, el tratamiento de sus datos afecte a sus sucesores vivos (Consejo de la Unión Europea y Parlamento Europeo, 2016).

### 5.1.4. “*Identificada o identificable*”

Una persona identificada es aquella que hemos podido aislar en un determinado grupo, ha sido distinguida de entre el resto de los individuos (Consejo de la Unión Europea y Parlamento Europeo, 2016). Alguien que es posible que sea identificado es, por lo tanto, identificable. Parece algo sencillo, pero este segundo término marca el límite de la legislación. Siempre que

una persona no sea identificable, no habrá datos personales que proteger (Article 29 Data Protection Working Party, 2007; Consejo de la Unión Europea y Parlamento Europeo, 2016).

Esta identificación puede lograrse a través de unos datos que reciben el nombre de “identificadores” y que obviamente guardan una estrecha relación con el individuo. En la propia definición de los *datos personales* aparecen como el elemento necesario para llevar a cabo la identificación, directa o indirectamente (Consejo de la Unión Europea y Parlamento Europeo, 2016).

La identificación directa es sencilla y ya la hemos mencionado antes. El identificador directo más habitual es el nombre que, acompañado de otros datos, nos permite aislar por completo a la persona. Es por ello por lo que, comúnmente, reconocemos a alguien como identificado en el momento en el cuál sabemos cuál es su nombre (Article 29 Data Protection Working Party, 2007). Sin embargo, la mayor parte de las veces no tendremos acceso a estos identificadores que nos permiten reconocer inmediatamente al individuo, pero podemos seguir identificándolo.

La manera de identificar a cualquier individuo indirectamente es mediante la combinación de indicadores (Sweeney, 2000). Puede que incluso no nos haga falta su nombre, siempre que tengamos a nuestro alcance otros indicadores. El propio Reglamento, en sus consideraciones previas, nos indica que para saber si existe o no la probabilidad de identificar a un individuo debemos tener en cuenta todos los medios que se puedan usar para ello (Consejo de la Unión Europea y Parlamento Europeo, 2016). Esto es algo complejo porque estas “pruebas” de seguridad deben tener en cuenta el dinamismo del *Big Data* y los cambios a pasos agigantados que se producen en otras tecnologías que podrían acabar con esta seguridad. También se debe tener en cuenta otro elemento y es el propósito con el que se recogen estos datos. Si solo vamos a publicar información que no referencie a ningún individuo, no se considerarán datos personales y la legislación no actuará sobre ellos (Consejo de la Unión Europea y Parlamento Europeo, 2016).

Por último, debemos saber que cualquier dato que queramos publicar tiene que estar *anonimizado*, es decir, que la persona no sea identificable (Article 29 Data Protection Working Party, 2007, 2014). La única manera de que una persona no sea identificada es que sus datos sean anónimos y esto debe cumplirse siempre. Normalmente, por error, las organizaciones publican datos que, a primera vista, dan la impresión de estar publicados de una forma segura, pero no es así (Article 29 Data Protection Working Party, 2014). Estos datos se presentan, en muchos casos, como *pseudoanonimizados* y si conocemos el mecanismo por el que se han

pseudoanonimizado, podemos revertirlo. Un ejemplo de esta metodología es la asignación de claves para una determinada información del individuo (nombre, apellidos, edad, etc.). Estos datos son eliminados y sustituidos por una clave, quedando solo la información que después será procesada. Siempre que podamos descifrar la clave o asociar la información que procesamos de manera individual, habrá motivos para pensar que la persona es identificable.

En resumen, a la hora de tratar con datos personales, no podemos bajar la guardia y se debe suponer que la probabilidad de que un individuo sea identificado existe siempre (Narayanan y Felten, 2014). Cualquier tratamiento de los datos debe asegurar, por todos los medios posibles, que las personas no serán identificadas. No obstante, el testeado deberá ser continuo por el dinamismo del *Big Data*, pues lo que un año no es inidentificable, al siguiente sí puede serlo (Article 29 Data Protection Working Party, 2014; Narayanan y Felten, 2014).

## 5.2. RELEVANCIA

Cualquiera en su sano juicio puede llegar a preguntarse si su nombre, apellidos, correo o DNI son datos que entrañen un riesgo para su persona si se ponen en conocimiento de las manos equivocadas. Sin embargo, nuestros datos personales incluyen información más comprometedoras y de la que no somos conscientes.

Todo nuestro historial clínico, analíticas, cuentas bancarias, domicilio, ocupación e incluso patrones de comportamiento, podrían ser obtenidos y son datos personales. Algunos autores especulan con la idea de que, en el futuro, muchos de nuestros datos presentes en redes sociales y otras plataformas podrían ser extraídos para analizar comportamientos o tendencias y modelar la sociedad (Mayer-Schonberger y Cukier, 2013). Sin embargo, no hay evidencia contrastada de esto último, solo casos de fugas de datos o algoritmos que han logrado identificar a individuos o algunos patrones (Van Dijck, 2014). Por precaución nuestros datos personales deben ser protegidos, pues siempre habrá miedo a sociedades distópicas donde estos datos puedan ser usados en nuestra contra (Mayer-Schonberger y Cukier, 2013; Van Dijck, 2014; Rodríguez-Rata, 2020).

Además, si perdiéramos nuestros datos o se vieran expuestos, la confianza en las instituciones y empresas que forman los tejidos de nuestra nueva era podría venirse abajo (Ortiz-Ospina y Roser, 2020). Si esto llega a ocurrir, podríamos ver una caída de las propias tecnologías que tanto prometen beneficiarnos.

### 5.3. REGULACIÓN EXISTENTE

Cabe preguntarse si nuestras leyes han sabido adaptarse a los nuevos tiempos, como ya ha ido ocurriendo en otras ocasiones, desde que apreció el Derecho moderno (Álvarez Robles, 2019). Es evidente que el desarrollo trae consigo un cambio en la sociedad y en los derechos que esta demanda. En esa línea, la nueva Ley Orgánica 3/2018 de Protección de Datos Personales y *Garantía de los Derechos Digitales*, va un paso más allá del Reglamento (España, 2018; Álvarez Robles, 2019). Esta nueva Ley reconoce, en su Título X, unos nuevos Derechos Digitales y muestra que nuestras instituciones son conscientes de la necesidad de crear un nuevo entorno, ya no es válido trasponer los Derechos que habíamos logrado hasta ahora, propios de un mundo analógico, al ámbito digital (España, 2018; Álvarez Robles, 2019).

Es dentro de este entorno de nuevos Derechos Digitales, donde cobra especial importancia la protección de nuestros Datos Personales. Actualmente, nuestros datos personales están protegidos por el Reglamento General de Protección de Datos (RGPD). Gracias a él, los ciudadanos tienen más control sobre sus datos, pudiendo ejercer los siguientes derechos (Consejo de la Unión Europea y Parlamento Europeo, 2016; España, 2018) para garantizar ese control:

- El derecho a acceder a nuestros datos.
- El derecho a rectificar o corregir alguno de los datos aportados, previa solicitud.
- El derecho de supresión de nuestros datos. De manera coloquial se le conocer como el “derecho al olvido”, es decir, a que se eliminen los datos del tratamiento.
- El derecho a la limitación del tratamiento. Los responsables deberán limitar el tratamiento de los datos si así lo desea el interesado.
- El derecho a la portabilidad de los datos.
- El derecho de oposición del interesado a que sus datos sean tratados. Esto tiene especial importancia más adelante, cuando hablemos de la automatización y la creación de perfiles.

Estos son los principales puntos sobre los que se apoya el RGPD (Comisión Europea, 2015; Consejo de la Unión Europea y Parlamento Europeo, 2016) para lograr su objetivo y mejorar el uso de los datos y la confianza de las personas en su tratamiento. Algo que es acorde al trabajo realizado sobre la legislación, con el fin de lograr una adaptación completa a la Era Digital.

## 6. BIG DATA EN EL ÁMBITO BIOSANITARIO

En este ámbito se incluyen todas las instituciones u organizaciones y recursos que se destinan con el fin de garantizar la salud, entendiendo la salud como el estado de completo bienestar físico, mental y social, y no solamente la ausencia de afecciones o enfermedades (Organización Mundial de la Salud (OMS), 2014). En resumidas cuentas, nos referimos en este trabajo a todo aquello referente a las ciencias de la salud y a sus aspectos éticos y legales actuales. Introduciremos brevemente todas estas ideas para poder conocer después la influencia del *Big Data* en este ámbito.

### 6.1. SITUACIÓN ACTUAL

Dentro de todo el ámbito biosanitario no sólo debemos incluir al personal médico y los farmacéuticos que brindan la asistencia al paciente, también hay que incorporar en la ecuación a todos aquellos que garantizan el correcto abastecimiento del sistema sanitario, es decir, las industrias dedicadas a la producción de los materiales utilizados, así como de los medicamentos o tratamientos de los que dispondrán los pacientes.

Hasta hace unos años el valor de este amplio sector se centraba en la atención y producción mencionadas antes. El paciente no participaba activamente en este ámbito y solo se consideraba su papel de forma pasiva, recibiendo la asistencia que necesitaba. Esto no quiere decir que el paciente no fuese importante o no se garantizara su tratamiento, pero si provocó que la comunidad científica centrara su atención en la eficiencia de la asistencia o de la industria farmacéutica. Sin embargo, con la evolución de nuestra sociedad hemos visto un cambio en el centro de nuestra mentalidad, enfocada ahora en el individuo (Organización para la Cooperación y el Desarrollo Económicos (OCDE) y Unión Europea, 2018).

En paralelo con esta evolución, todo el sector se ha ido desarrollando (European Federation of Pharmaceutical Industries and Associations (EFPIA), 2019). No sólo vemos que las industrias farmacéuticas generan un gran valor por sí solas, sino que cada vez buscan nuevas formas de aumentar este valor, aprovechando las nuevas tendencias del mercado (Comisión Europea *et al.*, 2017; International Data Corporation (IDC) y Open Evidence, 2017). En la Unión Europea, en línea con lo que hemos ido comentando a lo largo del trabajo, cada vez más empresas recurren al mercado de datos y la propia unión tienen entre sus planes lograr una mayor digitalización de nuestro entorno (Comisión Europea, 2020).

La Comisión Europea es consciente de esta sociedad digitalizada a la que nos dirigimos, donde términos como “eHealth” o “eGovernance” se harán cada vez más frecuentes y los datos de ámbitos como el biosanitario, tendrán cada vez más importancia (Comisión Europea, 2019, 2020). Un futuro en el que los propios sistemas de salud, que actualmente basan su existencia en la relación profesional-paciente, podrían sufrir cambios (Comisión Europea, 2019). El *Big Data* tendrá, por lo tanto, mucho que decir en este hipotético futuro en el que los avances digitales y los datos jugarán un papel fundamental en el tratamiento de los pacientes.

Ahora, se debe ser consciente de los posibles cambios, pero también será necesario conocer los riesgos que puede implicar cualquiera de estos cambios. Si una tecnología puede vulnerar alguno de nuestros derechos, tendrá que ser revisada con el fin de establecer alguna posible solución. Para comprender esta tensión existente entre los avances y los derechos del paciente y poder plantear soluciones, tendremos que conocer por separado cada una de las partes involucradas.

## 6.2. DERECHOS DEL PACIENTE

La tensión existente es el resultado de enfocar la figura del paciente hacia un papel más activo dentro del ámbito biosanitario. Esta tensión podría dañar los derechos del paciente, pero ¿qué derechos han alcanzado los individuos en este ámbito?

La proclamación en 1948 de la Declaración Universal de Derechos Humanos marca el comienzo de una época en la que las leyes, con independencia de su ámbito, deberán respetar la dignidad y libertad de las personas. Esta idea se introduce en el preámbulo de la ley básica de sanidad del año 2002 (España, 2002) y que nos muestra también, en este mismo preámbulo, cómo la Constitución Española considera en su artículo 43 la protección de la salud de un individuo como uno de los derechos de sus ciudadanos (España, 1978).

También, dentro de este contexto, los Estados miembros del Consejo de Europa, los demás Estados y la Comunidad Europea firman en España el Convenio de Oviedo en lo relativo a los derechos humanos y la biomedicina, afirmando como primer objetivo el garantizar la dignidad del individuo por encima de cualquier interés social, económico o científico (Consejo de Europa, 1997). En esta misma línea, la Unión Europea con su Carta de Derechos fundamentales y su artículo 35 proclama que: “toda persona tiene derecho a acceder a la prevención sanitaria y a beneficiarse de la atención sanitaria en las condiciones establecidas por las legislaciones y prácticas nacionales” (Estados miembros, 2016).

Por último, la ley que en España garantiza los derechos del paciente es la ya citada antes, *Ley 41/2002 básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica*. Los principios sobre los que se apoya (España, 2002; Sánchez-Caro y Abellán, 2003) esta ley son los siguientes:

- **Dignidad de la persona:** como ya hemos comentado anteriormente, cualquier actividad dentro del sector debe respetar la dignidad humana, su autonomía y, ante cualquier dato extraído de tal actividad, la privacidad del individuo.
- **Consentimiento informado:** en relación con lo comentado antes sobre el Convenio de Oviedo y teniendo en cuenta lo que está previsto por este, cualquier procedimiento deberá contar con el consentimiento del paciente después de haber sido correctamente informado de todos los riesgos.
- **Autonomía:** el respeto a la autonomía del paciente queda garantizado en esta ley y el máximo exponente de esto es el consentimiento informado (Sánchez-Caro y Abellán, 2003).
- **Obligación de colaboración:** la ley manifiesta a través del punto 5 de su artículo 2 el deber del paciente para con el sistema sanitario. Esta obligación por parte del ciudadano a aportar información sobre su salud responde al interés público o al motivo de la asistencia sanitaria (Sánchez-Caro y Abellán, 2003).
- **Información:** enlazado con el consentimiento informado y con el derecho a saber que poseen todos los individuos, el punto 6 del artículo 2 establece que el profesional deberá informar al paciente, para que este elija libremente.
- **Reserva o secreto profesional:** garante, no sólo de la profesionalidad, sino del tan mencionado derecho a la privacidad. Gracias a la aplicación de este secreto se protege al paciente de que alguien se haga con una información tan sensible y también algo mayor como es la privacidad, evitando cualquier otro daño a la esfera de la intimidad (familia, amigos, etc.) (Sánchez-Caro y Abellán, 2003).

En resumen, los pacientes tienen garantizada la asistencia médica y sus derechos sobre la salud, pero también la normativa vigente, en concordancia con lo dispuesto por el Reglamento General de Protección de Datos, vela por el derecho del paciente a permanecer anónimo y a que sus datos personales, y con ello su privacidad, sean respetados. Si bien es lógico preguntarse si tecnologías como el *Big Data* podrán terminar con estos derechos o hacer que nos veamos forzados a cambiar alguna de las normas vigentes. En los siguientes apartados veremos como



el *Big Data* está influenciando al ámbito biosanitario y si nuestras modificadas leyes y el reciente Reglamento serán suficientes para aprovechar sus beneficios, sin que tengamos que vernos expuestos a los riesgos.

### 6.3. TRANSFORMACIÓN DEL SISTEMA

Hasta ahora nuestras instituciones y gobiernos han sido conscientes, en mayor o menor medida, de la importancia de la salud de sus ciudadanos. Sin embargo, esta nueva era requiere de un mayor dinamismo por parte de los países de nuestro entorno para adaptar sus sistemas y puede que el *Big Data* presente algunos beneficios frente a los problemas que enfrentamos (Gesundheit Österreich Forschungs- und Planungs GmbH *et al.*, 2016).

Algunos de estos remarcados problemas de nuestra sociedad son las enfermedades causadas por el alcohol, el tabaquismo o la obesidad, junto a otras como las enfermedades raras o los contratiempos que devienen por tener una sociedad más envejecida (Ramos Cordero y Pinto Fontanillo, 2015; Dirección General de Comunicación de la Comisión Europea, 2018). También empiezan a aparecer viejos fantasmas del pasado con crecientes movimientos en contra de las vacunas o las resistencias de algunos patógenos a los antibióticos (Dirección General de Comunicación de la Comisión Europea, 2018).

Debemos considerar a la digitalización y al *Big Data* como posibles herramientas de combate que optimizan nuestros sistemas y que nos permiten mantener todo aquello que hemos conseguido hasta ahora. Gracias a este tipo de tecnologías, podremos ver una mejora en las comunicaciones entre países, científicos y empresas que aumentarán la capacidad de asistencia, investigación y desarrollo de productos (BigMedilytics, 2020). Sin embargo, podría resultar en una trampa, pues los estados miembros tienen profundas diferencias a nivel socioeconómico y de funcionamiento, tanto de puertas para afuera como en el propio país (Dirección General de Comunicación de la Comisión Europea, 2018). Si no se ejecutan de manera correcta las reformas, la implantación de las tecnologías no será efectiva y no contará con el apoyo de los ciudadanos, algo que se desvela fundamental para cambiar el sistema.

En resumen, para la transformación del sistema necesitamos un fuerte impulso por parte de la Unión Europea, sino nunca lograremos una evolución conjunta. Las tres prioridades que ha marcado la Comisión Europea (Dirección General de Comunicación de la Comisión Europea, 2018) para lograr este impulso son:

- **Garantizar el acceso seguro a los datos sanitarios por parte del ciudadano** (Dirección General de Comunicación de la Comisión Europea, 2018) dentro y fuera de las fronteras de cada país miembro, empezando por compartir el historial o las prescripciones de cada paciente, aumentando también el alcance de los registros electrónicos y compartiendo los formatos existentes para que haya un mayor entendimiento entre países dentro de la Unión.
- **Generar una infraestructura que permita compartir los datos y administrar tratamientos personalizados** (Dirección General de Comunicación de la Comisión Europea, 2018), permitiendo así a los investigadores y otros profesionales obtener recursos a lo largo de Europa. La creación de mecanismos de colaboración en las investigaciones, el establecimiento de especificaciones para el acceso e intercambio de datos de una forma segura y la gestión de proyectos piloto en enfermedades raras, infecciosas o de otro tipo son algunas de las tareas más concretas que podrían ayudar a alcanzar este objetivo.
- **Centralizar el sistema en el paciente y lograr un mayor empoderamiento de este** (Dirección General de Comunicación de la Comisión Europea, 2018), poniendo a su disposición todas las herramientas tecnológicas posibles para que ellos también puedan participar en su salud. Para lograr este objetivo se deberá facilitar el acceso a las nuevas tecnologías, garantizar la oferta a través de las autoridades y movilizar recursos hacia la digitalización.

Podemos ver que estos ambiciosos objetivos tienen en su planteamiento acciones más concretas que, a ojos de la Comisión Europea, garantizarán que sean alcanzados a lo largo de los próximos años. También podemos deducir que en esta digitalización el *Big Data* jugará un papel esencial.

La capacidad del *Big Data* para tratar grandes bases de datos que poseen una gran variedad de información puede presentarse como un apoyo relevante, si se tiene en cuenta la diversa cantidad de datos que implica la digitalización. Además, la propia Comisión Europea ve en el *Big Data* la posibilidad de (Gesundheit Österreich Forschungs- und Planungs GmbH *et al.*, 2016):

- Mejorar la calidad y efectividad de nuestros tratamientos.
- Aumentar las posibilidades de prevenir enfermedades.
- Garantizar la seguridad del paciente, gracias a la transmisión directa de la información.

- Obtener unos resultados de nuestras predicciones más precisos, pues comprenderemos mejor los cambios demográficos, las tendencias, las vías de transmisión, etc.
- Transmitir el conocimiento entre académicos.
- Reducir la ineficiencia y el gasto.

Estas son las posibilidades que explora la Comisión Europea en su *Estudio sobre el Big Data en la sanidad pública, telemedicina y el sistema sanitario* (Gesundheit Österreich Forschungs- und Planungs GmbH *et al.*, 2016). En este mismo estudio la Comisión presenta veinte iniciativas que ya han hecho uso de la tecnología del *Big Data* en el ámbito sanitario. No las comentaremos en este trabajo porque deberíamos analizar en profundidad la metodología usada durante el estudio con el fin de establecer el impacto y características de cada una de ellas. Sin embargo, si focalizaremos sobre una de las recomendaciones redactada por los autores del estudio (y reafirmada de nuevo en su conclusión): hay que “establecer una estrategia de comunicación que encamine a la opinión pública hacia la necesidad del *Big Data* en el ámbito biosanitario” (Gesundheit Österreich Forschungs- und Planungs GmbH *et al.*, 2016).

#### 6.4. PROBLEMAS QUE AFECTAN AL PACIENTE

Los riesgos del *Big Data* ya han sido comentados en apartados anteriores (4.5), pero aquí vamos a hablar de dos problemas concretos que pueden llegar a surgir si aplicamos el *Big Data* en el ámbito sanitario y son: la falta de transparencia, relacionada indirectamente con el consentimiento del paciente, y los problemas derivados de las predicciones analíticas.

##### 6.4.1. *Transparencia para el empoderamiento del individuo*

La transparencia es uno de los valores más importantes en las democracias occidentales. El ciudadano debe conocer toda la información que le pueda afectar para poder tener un papel activo en la sociedad. Solo de esta forma garantizamos que el ciudadano vea respetado su derecho a saber.

Si pretendemos usar tecnologías como el *Big Data*, debemos ser conscientes de todo lo que implica. El ciudadano debe conocer cada detalle de las operaciones que se hagan con sus datos. No hacerlo podría suponer un descenso aún mayor en la confianza que depositamos en nuestras instituciones y que, en el caso concreto de España, es ya bastante baja (Ortiz-Ospina y Roser, 2020). Sin la debida transparencia, nunca se obtendrá el consentimiento válido del ciudadano, es por ello por lo que relacionamos en este punto ambos conceptos.

El Reglamento General de Protección de Datos garantiza esta transparencia y la necesidad de que haya un consentimiento por parte del individuo del que se van a tratar sus datos. Es necesario anotar que este consentimiento específico puede no tener por qué existir si hay una obligación legal o de satisfacción de unos intereses con un tercero (Consejo de la Unión Europea y Parlamento Europeo, 2016). En ese caso y siempre que los intereses no prevalezcan sobre los derechos del interesado, no tiene por qué haber consentimiento.

El tratamiento de datos y la posible implantación de las tecnologías analíticas están sujetas, por tanto, al consentimiento del individuo. Este consentimiento es la base de todo y, aunque el Reglamento garantice su aplicación y deje claro en sus consideraciones previas (39) que toda comunicación con el fin de obtener un consentimiento debe ser clara (Consejo de la Unión Europea y Parlamento Europeo, 2016), no sabemos si realmente es así. Es cierto que el Eurobarómetro nos muestra que, en el año 2019, casi la mitad de los encuestados estaban dispuestos a ceder sus datos para mejorar el sistema sanitario, colaborar con la investigación, etc (Dirección General de Comunicación de la Comisión Europea, 2020). Hay una predisposición del ciudadano a que se recurra a sus datos para mejorar el estado del ámbito biosanitario, pero en esta misma encuesta vemos que solamente un 46% de los ciudadanos quiere tener un mayor control sobre estos datos (Dirección General de Comunicación de la Comisión Europea, 2020).

Tenemos por tanto una población dispuesta a cooperar con la implantación de estas tecnologías analíticas. Sin embargo, a simple vista, podemos apreciar dos problemas que hacen tambalearse al consentimiento como base de esta cooperación hacia la digitalización:

- La **pasividad del ciudadano**, que prefiere no complicar su existencia con la gestión de sus datos o sencillamente no pretende otorgarlos.
- La incertidumbre existente cuando hablamos de la **claridad** que debería existir a la hora de solicitar el consentimiento. Los ciudadanos no suelen prestar atención a la información que se les brinda y en muchas ocasiones, aunque sea clara a nivel “legal”, puede no serlo para el “ciudadano de a pie”.

En pocas palabras, ¿podrá ser el consentimiento informado el camino hacia el empoderamiento de los individuos y la base para la introducción del *Big Data* en el tratamiento de datos biosanitarios? Si es así, hay que tener en cuenta que este consentimiento se fundamenta en la percepción y comprensión del ciudadano. Por lo tanto, hay que preguntarse si los países en la situación de España estarán preparados para delegar la responsabilidad en el ciudadano

(Ministerio de Educación Cultura y Deporte: Secretaría de Estado de Educación Formación Profesional y Universidades *et al.*, 2017). Posiblemente, el consentimiento siga formando parte de muchos procedimientos, pero viendo el peso que adquiere el propio *tratamiento* de los datos, no tiene sentido hacer recaer la responsabilidad solamente en el usuario (Gil González, 2016).

#### 6.4.2. *Los peligros de predecir el futuro*

La toma de decisiones automatizadas es uno de los puntos del *Big Data*, pero esto está discutido en el ámbito biosanitario (Cohen *et al.*, 2014) por dos motivos: el riesgo de una mala decisión y el riesgo a la discriminación del individuo.

- **El riesgo de una mala decisión:** Las decisiones automáticas se basan en modelos matemáticos, pero afinar un modelo para que sus predicciones no entrañen riesgo es muy difícil. Es cierto que el *Big Data* nos puede ayudar a conseguir una optimización más rápida, pero deberemos tener suficientes datos, estandarizar cualquier metodología y, por último, probar el modelo con unas garantías de seguridad muy estrictas (Cohen *et al.*, 2014). Cumplir todos estos requisitos es una tarea difícil y, aunque se logre, debemos tener en cuenta otras ideas. El modelo debe tener capacidad de recibir retroalimentación por parte de todos los usuarios, tampoco sus decisiones pueden girar en torno a cuestiones que requieran de una profunda deliberación y donde el sesgo humano podría mostrarse útil para evitar daños en el paciente. Esto restringiría los modelos a predicciones en la ocupación, la hospitalización, el avance de una infección, etc.
- **El riesgo a la discriminación del individuo:** La agrupación de los individuos es algo inherente al *Big Data* y al propio ser humano quién, en su proceso de entender el mundo, categoriza. Por desgracia, el *Big Data* en su afán por buscar patrones y tendencias, como ya hemos visto anteriormente, crea infinidad de categorías. Si hablamos de movimientos de personas o del tráfico en una zona, esto puede no ser muy importante, pero en el ámbito biosanitario tiene mucho peso.

La prevalencia de ciertas enfermedades, el consumo de medicinas y otros parámetros que se estudian en el campo de la salud podrían tener más presencia en unas categorías que en otras y, con ello, ser la semilla para una posible discriminación de estos grupos (Cohen *et al.*, 2014; Gil González, 2016).

Estos dos problemas, en principio, son tenidos en cuenta por la regulación existente. El Reglamento General de Protección de Datos en su artículo 22 nos protege de que podamos ser

objeto de cualquier decisión automatizada y tomada a partir de perfiles o categorías, salvo que el individuo sea plenamente consciente de ello y lo acepte (Consejo de la Unión Europea y Parlamento Europeo, 2016). Sin embargo, es inevitable pensar que un mal uso de estas agrupaciones pueda conducir, no solo a errores, sino a una discriminación de cualquier tipo.

## 6.5. POSIBLES SOLUCIONES

Los problemas planteados en el *Big Data* aplicado al ámbito sanitario pueden ser resueltos, siempre que no aparezcan nuevos riesgos con esta tecnología. Tres posibles soluciones a estos problemas son:

### 6.5.1. *Aumentar las exigencias como ciudadanos*

La existencia de un consentimiento informado claro y honesto no es, ni puede ser, la única exigencia que deba hacer la ciudadanía hacia los responsables del tratamiento de datos. Aunque este consentimiento exista, se debe obligar al responsable a tener una mayor consideración con nuestros datos y a que cumplan los fines con los que se ideó el tratamiento. De esta forma, si alguno de los que trata los datos no responde a los objetivos ni requisitos establecidos, tendrá su parte de responsabilidad y no será culpa del ciudadano que firmó un consentimiento.

### 6.5.2. *Aplicar el principio de responsabilidad proactiva*

En línea con las exigencias que debe hacer el ciudadano, debemos ser conscientes de este principio y luchar porque se cumpla. El nuevo Reglamento traslada la responsabilidad de la seguridad del tratamiento a la organización que tratará los datos (Agencia Española de Protección de Datos, 2018). El responsable, según el Artículo 5, deberá aplicar las medidas técnicas y organizativas necesarias para garantizar y poder demostrar que el tratamiento es conforme con el Reglamento (Agencia Española de Protección de Datos, 2018).

Por lo tanto, la entidad supervisora de cada territorio ya no es un organismo que deba garantizar la protección de todo fichero existente de forma directa, sino que se apoya en quien tratará esos datos. El principio exige así a las organizaciones un papel más proactivo frente a todos los tratamientos a realizar, pues serán ellos quienes deban garantizar su seguridad (Agencia Española de Protección de Datos, 2018).

### 6.5.3. *Proteger desde el diseño o por defecto*

Considerar la privacidad desde los inicios de cualquier proyecto, para que no haya brechas en el modelo (Gil González, 2016). Si se exige que todos los proyectos, basados en el tratamiento de datos, presenten una planificación de los riesgos en la privacidad y como atajarlos,

evitaríamos que salgan a la luz modelos que pongan en riesgo los datos personales de los individuos.

Este asunto puede ser abordado desde el diseño, es decir, la tecnología se construye teniendo en cuenta el nivel de privacidad necesario o por defecto, dónde se parte de las opciones más protectoras y luego el usuario puede reducir, si desea obtener otras utilidades, este nivel de seguridad (Gil González, 2016).

## 7. EPÍLOGO: CORONAVIRUS Y BIG DATA

Es pronto para hacernos una idea del impacto real que tendrá la pandemia causada por el SARS-CoV-2, además de ser un tema especialmente sensible. Por lo tanto, no se valorará aquí el impacto de la epidemia y simplemente se usará como ejemplo de la importancia y utilidad de los datos y las técnicas que nos permiten trabajar con ellos.

Manifiesta el Portal Europeo de Datos la necesidad de conocer los datos y de analizarlos, sosteniendo que tecnologías como el *Big Data* se han vuelto imprescindibles en el tratamiento de los datos (Portal Europeo de Datos, 2020b, 2020c). Gracias a todas las tecnologías disponibles estamos combatiendo la enfermedad, no sólo eso, sino que estamos conociendo mejor las pandemias que nos han azotado en el pasado y desarrollando mecanismos para estar preparados cuando vengan en el futuro (Coma *et al.*, 2020).

La comunidad científica está desarrollando ahora mismo modelos basados en el *Big Data* y sólo mediante la contrastación de estos podremos saber si funcionan (Portal Europeo de Datos, 2020a). La propia Unión Europea está habilitando, en las plataformas destinadas al estudio de los datos de la Covid-19, espacios para apoyar iniciativas que puedan ayudar a combatir la enfermedad (Connecting Europe Facility, 2020). También están poniendo a disposición de las comunidades sus plataformas para la implantación del *Big Data* (Mayer-Schonberger y Cukier, 2013).

Esto es un ejemplo de cómo el *Big Data* puede sernos de gran utilidad y, aunque esto nos beneficie, no debemos dejar de lado los riesgos que traen consigo estas tecnologías. Siendo muy importante el riesgo que corremos de tomar las decisiones equivocadas por interpretar mal estos datos.

## 8. CONCLUSIÓN

Hemos podido ver a lo largo de este trabajo, tal y como se advertía al principio, que hemos llegado a la era de la información. Cada vez son más las bases de Datos que, como *Our World in Data*, se encargan de extraer, procesar y publicar datos de diversa índole para que todo el mundo pueda verlos. En este contexto, se ha incidido en una nueva herramienta que podría sernos de gran ayuda para arrojar un poco de luz sobre la marea de información en la que vivimos. No solo eso, sino que la herramienta podría mejorar la situación de nuestros sistemas de salud, eficientes, pero que cada vez nos cuestan más.

Sin embargo, las preocupaciones han empezado a emerger. Muchos ciudadanos, cegados por el potencial de la herramienta, temen que su privacidad se pierda y no es para menos. A este pánico se suma ahora el paradigma de las redes sociales y plataformas de venta que, con su tecnología, pueden establecer los patrones de la población. Algo único de cada persona o grupo y que parecía impredecible.

Los datos se han convertido en un recurso que parece contener lo que nunca hemos podido ver. Una idea que se ha asentado con firmeza en muchas mentes, pero que tiene más de religioso o ideológico que de científico. No son un recurso, son “lo dado”. El recurso es el mundo en el que vivimos. Ese mundo que la Ciencia intenta comprender buscando el *por qué* y aplicando el método científico, pero ahora con la ayuda de una nueva herramienta. Antes del *Big Data* ya había datos, puede que menos, pero cometíamos y cometemos errores por no ser críticos. El *Big Data* podría cometer errores más graves si seguimos instaurando la creencia de que es nuestro único futuro. Hemos visto que algunos auguran que se predecirán patrones humanos desde las redes sociales, a lo que debemos preguntarnos ¿todos los perfiles son una representación fidedigna del usuario? ¿está toda la población representada en ellos? El *Big Data* no es algo venido del espacio exterior. Es una tecnología fruto de nuestros avances en la digitalización y cómo tal, si la usamos bien, nos aportará un gran beneficio. Si no somos críticos antes de usarla, nos perjudicará.

No obstante, ¿qué pasa con la información sensible, como la procedente del ámbito sanitario? No debemos escudarnos en la sensibilidad de nuestra información para no implantar algo. Debemos ser conscientes de los riesgos y exigir medidas efectivas a las instituciones que garanticen nuestra privacidad. El consentimiento informado ha sido un gran paso, pero no puede ser la base del sistema. Debemos lograr que quienes tratan los datos se responsabilicen y no debemos cargar todo a hombros del ciudadano. Tendríamos que garantizar que el ciudadano



pueda denunciar un uso no apropiado de los datos. No es prudente confiar en que será capaz de ver todos los peligros a la hora de firmar la cesión de sus datos. Esta idea solo se puede alcanzar con el apoyo de las instituciones y con individuos dispuestos a exigir responsabilidades a quienes se hacen cargo de sus datos. Eso es el auténtico empoderamiento del ciudadano.

La implantación de las nuevas tecnologías y la digitalización de nuestra sociedad, incluyendo el uso del *Big Data*, supondrán un reto. A pesar de ello, traerán un gran beneficio al que no podemos renunciar. La privacidad no morirá por ello, sino que su debate pasará a estar más vivo que nunca. En resumen, el *Big Data* es nuestra nueva tecnología y marca una nueva era, como el descubrimiento del fuego. En nuestras manos está si queremos mirar más allá de la penumbra con su ayuda o si por el contrario nos quedaremos mirando fijamente, a la espera de respuestas, ciegos.

## REFERENCIAS

Agencia Española de Protección de Datos (2018) *Guía del Reglamento General de Protección de Datos para responsables de tratamiento*. Madrid: Agencia Española de Protección de Datos.

Álvarez Robles, T. (2019) "Primer año de vigencia de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos y garantía de derechos digitales: los derechos en minúscula", *Revista de la Facultad de Ciencias Jurídicas*, 5.

Article 29 Data Protection Working Party (2007) "Opinion 4/2007 on the concept of personal data", *Working Party opinions and recommendations*, (WP 136), pp. 1-26.

Article 29 Data Protection Working Party (2014) "Opinion 05/2014 on Anonymisation Techniques", *Working Party opinions and recommendations*, (WP 216), pp. 1-37.

Asamblea General de las Naciones Unidas (1948) *La Declaración Universal de Derechos Humanos*, *Sitio web oficial de la Organización de las Naciones Unidas*. Disponible en: <https://www.un.org/es/universal-declaration-human-rights/> (Accedido: 24 de mayo de 2020).

BigMedilytics (2020) *BigMedilytics | Improving healthcare in Europe through big data*, *Sitio web oficial del proyecto BigMedilytics*. Disponible en: <https://www.bigmedilytics.eu/> (Accedido: 24 de mayo de 2020).

Cambridge University Press (2020) *Significado de «Datum»*, *Página web del Diccionario Cambridge*. Disponible en: <https://dictionary.cambridge.org/es/diccionario/ingles/datum> (Accedido: 24 de mayo de 2020).

Cano, F. (2020) *De las «Chicas del cable» al 5G: 96 años en los que Telefónica no deja de conectar personas*, *El Español*. Disponible en: [https://www.elespanol.com/invertia/empresas/tecnologia/20200426/chicas-cable-anos-telefonica-no-conectar-personas/484952446\\_0.html](https://www.elespanol.com/invertia/empresas/tecnologia/20200426/chicas-cable-anos-telefonica-no-conectar-personas/484952446_0.html) (Accedido: 24 de mayo de 2020).

Cisco Systems Incorporation (2020) *Cisco annual Internet report (2018–2023) White paper*.

Cohen, I. G., Amarasingham, R., Shah, A., Xie, B. y Lo, B. (2014) "The legal and ethical concerns that arise from using complex predictive analytics in health care", *Health affairs*, 33(7), pp. 1139-1147.

Coma, E., Mora, N., Prats-Urbe, A., Fina, F., Prieto-Alhambra, D. y Medina-Peralta, M. (2020) "Excess cases of influenza suggest an earlier start to the coronavirus epidemic in Spain than official figures tell us: an analysis of primary care electronic medical records from over 6 million people from Catalonia", *COVID-19 SARS-CoV-2 preprints from medRxiv and bioRxiv*. doi:<https://doi.org/10.1101/2020.04.09.20056259>.

Comisión Europea (2015) *Questions and answers - Data protection reform*, *Zona de prensa de la Comisión Europea*. Disponible en: [https://ec.europa.eu/commission/presscorner/detail/en/MEMO\\_15\\_6385](https://ec.europa.eu/commission/presscorner/detail/en/MEMO_15_6385) (Accedido: 24 de mayo de 2020).

Comisión Europea (2019) *Transformation of health and care, Shaping Europe's digital future*. Disponible en: <https://ec.europa.eu/digital-single-market/en/politicas/ehealth> (Accedido: 24 de mayo de 2020).

Comisión Europea (2020) *Creating a digital society, Shaping Europe's digital future*. Disponible en: <https://ec.europa.eu/digital-single-market/en/politicas/creating-digital-society> (Accedido: 24 de mayo de 2020).

Comisión Europea, International Data Corporation (IDC) y Open Evidence (2017) *Final results of the European Data Market study measuring the size and trends of the EU data economy, Shaping Europe's digital future*. Disponible en: <https://ec.europa.eu/digital-single-market/en/news/final-results-european-data-market-study-measuring-size-and-trends-eu-data-economy> (Accedido: 24 de mayo de 2020).

Connecting Europe Facility (2020) *Big Data test infrastructure*, *Sitio web oficial de Connecting Europe Facility*. Disponible en: <https://ec.europa.eu/cefdigital/bdti> (Accedido: 24 de mayo de 2020).

Consejo de Europa (1997) "Convenio para la protección de los derechos humanos y la dignidad del ser humano con respecto a las aplicaciones de la Biología y la Medicina: Convenio relativo a los derechos humanos y la biomedicina", *European Treaty Series*, 9 de abril de 1997, (164), pp. 1-10.

Consejo de la Unión Europea y Parlamento Europeo (2016) "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", *Diario Oficial de la Unión Europea, Serie L*, 27 de abril de 2016, (L119), pp. 1-88.

Cox, M. y Ellsworth, D. (1997) "Application-controlled demand paging for out-of-core visualization", en *Proceedings. Visualization '97 (Cat. No. 97CB36155)*, pp. 235-244.

- Van Dijck, J. (2014) "Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology", *Surveillance & Society*, 12(2), pp. 197-208.
- Dirección General de Comunicación de la Comisión Europea (2018) *Comunicación de la Comisión Europea relativa a la consecución de la transformación digital de la sanidad y los servicios asistenciales en el Mercado Único Digital, la capacitación de los ciudadanos y la creación de una sociedad más saludable*. Luxemburgo: Oficina de Publicaciones de la Unión Europea.
- Dirección General de Comunicación de la Comisión Europea (2020) *Attitudes towards the impact of digitalisation on daily lives*. Luxemburgo: Oficina de Publicaciones de la Unión Europea.
- España (1978) "Constitución Española", *Boletín Oficial del Estado*, 29 de diciembre de 1978, (311), pp. 29313-29424.
- España (2002) "Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica.", *Boletín Oficial del Estado*, 15 de noviembre de 2002, (274), pp. 40126-40132.
- España (2018) "Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.", *Boletín Oficial del Estado*, 6 de diciembre de 2018, (294), pp. 119788-119857.
- Estados miembros (2016) "Carta de los Derechos Fundamentales de la Unión Europea", *Diario Oficial de la Unión Europea, Serie C*, 7 de junio de 2016, (202), pp. 389-405.
- Esteves, J. (2009) "A benefits realisation road-map framework for ERP usage in small and medium-sized enterprises", *Journal of enterprise information management*, 22(1/2), pp. 25-35.
- European Federation of Pharmaceutical Industries and Associations (EFPIA) (2019) *The Pharmaceutical Industry in figures 2019*. Bruselas: European Federation of Pharmaceutical Industries and Associations (EFPIA) (The Pharmaceutical Industry in Figures).
- Fernández de Lis, P. (2019) "La revolución tecnológica necesita ética o nos esclavizará", *El País - Tecnología*. Disponible en: [https://elpais.com/tecnologia/2019/10/07/actualidad/1570442594\\_979114.html](https://elpais.com/tecnologia/2019/10/07/actualidad/1570442594_979114.html) (Accedido: 24 de mayo de 2020).
- Gartner Incorporation (2020) *Big Data concept, Gartner Glossary*. Disponible en: <https://www.gartner.com/en/information-technology/glossary/big-data> (Accedido: 24 de mayo de 2020).
- Gefen, D. y Ragowsky, A. (2005) "A multi-level approach to measuring the benefits of an ERP system in manufacturing firms", *Information systems management*, 22(1), pp. 18-25.
- Gesundheit Österreich Forschungs- und Planungs GmbH, Dirección General de Salud y Seguridad Alimentaria de la Comisión Europea, Habl, C., Renner, A.-T., Bobek, J. y Lashchkolning, A. (2016) *Study on Big Data in public health, telemedicine and healthcare*. Luxemburgo: Oficina de Publicaciones de la Unión Europea.
- Gil González, E. (2016) *Big data, privacidad y protección de datos*. Madrid: Agencia Española de Protección de Datos y Agencia Estatal del Boletín Oficial del Estado.
- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S. y Brilliant, L. (2009) "Detecting influenza epidemics using search engine query data", *Nature*, 457, pp. 1012-1014.
- Huawei Technologies Company Limited y Oxford Economics Limited (2017) *Digital spillover - Measuring the true impact of the digital economy*.
- International Data Corporation (IDC) y Open Evidence (2017) *European Data Market SMART 2013/0063*. Bruselas: Dirección General de Comunicación de la Comisión Europea.
- Juste, M. (2019) *Del «blockchain» al «Big Data»: La digitalización como arma electoral*, *Expansión*. Disponible en: <https://www.expansion.com/economia-digital/innovacion/2019/04/16/5cb4b0ba468aeb3b448b4574.html> (Accedido: 24 de mayo de 2020).
- Laney, D. (2001) *3D Data management: Controlling data volume, velocity, and variety*. Stamford: METAGroup.
- Lewis, C. L. (1927) *Matthew Fontaine Maury, the pathfinder of the seas*. Annapolis, Maryland: The United States Naval Institute.
- Mayer-Schonberger, V. y Cukier, K. (2013) *Big Data: A revolution that will transform how we live, work and think*. London: John Murray.
- Ministerio de Educación Cultura y Deporte: Secretaría de Estado de Educación Formación Profesional y

Universidades, Dirección General de Evaluación y Cooperación Territorial y Instituto Nacional de Evaluación Educativa (2017) *PIRLS 2016 Estudio internacional de progreso en comprensión lectora. Informe español*. Madrid: Ministerio de Educación, Cultura y Deporte: Secretaría General Técnica: Subdirección General de Documentación y Publicaciones.

Molero, V. M. (2014) *La Revolución Digital*. Madrid: Departamento de Estudios e Imagen Corporativa de la Universidad Complutense de Madrid.

Moore, G. E. (1965) "Cramming more components onto integrated circuits", *Electronics*, 38(8), pp. 114-117.

Narayanan, A. y Felten, E. W. (2014) *No silver bullet: De-identification still doesn't work*. Princeton, Nueva Jersey: Department of Computer Science at Princeton University.

Neyman, J. (1952) *Lectures and conferences on mathematical statistics and probability*. 2.<sup>a</sup> ed. Washington: Graduate School U.S. Department of Agriculture.

Al Nuaimi, E., Al Neyadi, H., Mohamed, N. y Al-Jaroodi, J. (2015) "Applications of big data to smart cities", *Journal of Internet services and applications*, 6(1), p. 25.

Organización Mundial de la Salud (OMS) (2014) "Constitución de la Organización Mundial de la Salud", en *Documentos Básicos*. 48.<sup>a</sup> ed. Ginebra: Ediciones de la OMS, Organización Mundial de la Salud, pp. 1-21.

Organización para la Cooperación y el Desarrollo Económicos (OCDE) y Unión Europea (2018) *Health at a glance: Europe 2018: State of health in the EU cycle*. París: OECD Publishing. doi:10.1787/health\_glance\_eur-2018-en.

Ortiz-Ospina, E. y Roser, M. (2020) *Trust, Our World in Data*. Disponible en: <https://ourworldindata.org/trust> (Accedido: 24 de mayo de 2020).

Portal Europeo de Datos (2020a) *Data related initiatives, Sitio web oficial del Portal Europeo de Datos*. Disponible en: <https://www.europeandataportal.eu/en/covid-19/initiatives> (Accedido: 24 de mayo de 2020).

Portal Europeo de Datos (2020b) *The power of Data in the everlasting battle: Humanity vs. Virus (1/2), Sitio web oficial del Portal Europeo de Datos*. Disponible en: <https://www.europeandataportal.eu/en/covid-19/stories/power-data-everlasting-battle-humanity-vs-virus-12> (Accedido: 26 de mayo de 2020).

Portal Europeo de Datos (2020c) *The power of data in the everlasting battle: Humanity vs. Virus (2/2), Sitio web oficial del Portal Europeo de Datos*. Disponible en: <https://www.europeandataportal.eu/en/covid-19/stories/power-data-everlasting-battle-humanity-vs-virus-22> (Accedido: 24 de mayo de 2020).

Ramos Cordero, P. y Pinto Fontanillo, J. A. (2015) "Las personas mayores y su salud: Situación actual", *Avances en Odontostomatología*, 31(3), pp. 107-116.

Real Academia Española (2020a) *Definición de «Big Data», Página web del Diccionario del español jurídico*. Disponible en: <https://dej.rae.es/lema/big-data> (Accedido: 24 de mayo de 2020).

Real Academia Española (2020b) *Definición de «Privacidad», Página web del Diccionario del español jurídico*. Disponible en: <https://dej.rae.es/lema/privacidad> (Accedido: 24 de mayo de 2020).

Real Academia Española (2020c) *Significado de «Dato», Página web del Diccionario de la lengua española*. Disponible en: <https://dle.rae.es/dato?m=form> (Accedido: 24 de mayo de 2020).

Reinsel, D., Gantz, J., Rydning, J., International Data Corporation (IDC) y SEAGATE (2018) *The Digitization of the World - From edge to core*. Framingham: International Data Corporation (IDC).

Richards, N. M. y King, J. H. (2014) "Big Data Ethics", *Wake Forest Law review*, 49, pp. 393-432.

Rivas, G. (2018) *La Revolución Digital: el potencial de estar en las nubes, El País - Opinión*. Disponible en: [https://elpais.com/elpais/2018/04/06/planeta\\_futuro/1523023364\\_337768.html](https://elpais.com/elpais/2018/04/06/planeta_futuro/1523023364_337768.html) (Accedido: 24 de mayo de 2020).

Rodríguez-Rata, A. (2020) *China receta «Big Data» para controlar a sus ciudadanos y luchar contra el coronavirus, La Vanguardia*. Disponible en: <https://www.lavanguardia.com/vida/20200302/473825002349/china-receta-big-data-control-ciudadano-lucha-coronavirus-inteligencia-artificial-app.html> (Accedido: 24 de mayo de 2020).

Roser, M. y Ritchie, H. (2013) *Technological progress, Our World in Data*. Disponible en: <https://ourworldindata.org/technological-progress> (Accedido: 24 de mayo de 2020).

Roser, M., Ritchie, H. y Ortiz-Ospina, E. (2020) *Internet, Our World in Data*. Disponible en: <https://ourworldindata.org/internet> (Accedido: 24 de mayo de 2020).

- Sánchez-Caro, J. y Abellán, F. (2003) *Derechos y deberes de los pacientes*. Granada: Editorial Comares.
- Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., Tufano, P., IBM Institute for Business Value y Escuela de Negocios Saïd en la Universidad de Oxford (2012) *Analytics: el uso de big data en el mundo real. Cómo las empresas más innovadoras extraen valor de datos inciertos*. Madrid: IBM Institute for Business Value.
- Suárez, G. (2019) *Tú también mientes: Cómo las búsquedas de Google revelan tus secretos más oscuros*, *El Mundo Papel Futuro*. Disponible en: <https://www.elmundo.es/papel/futuro/2019/03/26/5c991493fdddffba428b45ea.html> (Accedido: 25 de mayo de 2020).
- Sweeney, L. (2000) "Simple demographics often identify people uniquely", *Carnegie Mellon University, Data Privacy working paper*, 3.
- Vodafone España (2018) *El gran auge del Big Data: repaso desde los inicios hasta hoy, Ideas para tu empresa*. Disponible en: <https://ideasparatuempresa.vodafone.es/big-data-desde-los-inicios-hoy/> (Accedido: 24 de mayo de 2020).
- Wang, Y., Kung, L. y Byrd, T. A. (2018) "Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations", *Technological forecasting and social change*, 126, pp. 3-13.