



HAL
open science

La potencialidad del Big Data en ámbito sanitario. Especial referencia al caso español

Jose Maria Biedma, Christian Bourret

► To cite this version:

Jose Maria Biedma, Christian Bourret. La potencialidad del Big Data en ámbito sanitario. Especial referencia al caso español. *Revista de Economía & Administración*, 2019, 16 (2), pp.93 - 109. hal-03996599

HAL Id: hal-03996599

<https://hal.science/hal-03996599>

Submitted on 7 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LA POTENCIALIDAD DEL BIG DATA EN EL ÁMBITO SANITARIO. ESPECIAL REFERENCIA AL CASO ESPAÑOL*

THE POTENTIAL OF BIG DATA IN THE HEALTH FIELD. SPECIAL REFERENCE TO THE SPANISH CASE

José María Biedma Ferrer[∞]

Christian Bourret^º

Resumen

Big Data (o inteligencia de datos) está adquiriendo una gran importancia en el ámbito sanitario. Los objetivos del trabajo son: reflexionar sobre el conocimiento y la aplicación de *Big Data* en el ámbito sanitario, con especial referencia al estudio de sus ventajas, riesgos e inconvenientes, así como los desafíos que presenta. La metodología se basa en la revisión sistemática de la literatura. Los resultados permiten concluir que, a pesar de la gran cantidad de información recogida en los procesos sanitarios, su explotación es baja. *Big Data* tendrá un impacto positivo sobre la mejora de los cuidados sanitarios, siendo muy útil para los estudiantes de medicina y el personal sanitario. No obstante, será necesario reforzar las medidas de protección de datos personales.

Palabras clave

Big Data, salud, riesgos, Diraya.

Abstract

Big Data (or data intelligence) is becoming very important in the health field. The objectives of the work are: to reflect on the knowledge and application of Big Data in the health field, with special reference to the study of its advantages, risks and disadvantages, as well as the challenges it presents. The methodology is based on the systematic review of the literature. The results allow us to conclude that, despite the large amount of information collected in health processes, its exploitation is low. Big Data will have a positive impact on the improvement of health care, being very useful for medical students and health personnel. However, it will be necessary to strengthen personal data protection measures.

Key words

Big Data, health, risk, Diraya.

Clasificación J.E.L.: M10; M15.

* Artículo recibido el 04-10-2019 y aprobado el 28-01-2020.

[∞] Doctor en Derecho. Profesor del Departamento de Organización de Empresas de la Universidad de Cádiz. Miembro del Instituto Universitario de Investigación para el Desarrollo Social Sostenible (INDESS), España. Correo electrónico: josemaria.biedma@uca.es.

^º Doctor en Historia y Civilización. Profesor de Ciencias de la Información y de la Comunicación en la Universidad de Paris Est Marne-la-Vallée- Institut Francilien d'Ingénierie des Services (IFIS) et IUT (Institut Universitaire de Technologie), Francia. Correo electrónico: christian.bourret@u-pem.fr.

Introducción

La importancia creciente de *Big Data* (o inteligencia de datos) es incontestable. En la literatura, se ha considerado que *Big Data* es una tendencia mundial, con un crecimiento del mercado y de los ámbitos de investigación asociados. En este sentido, Chen y Chun-Zhang (2014) han considerado que representa el principio de una era de competitividad, innovación, productividad y de revolución científica. *Big Data* se ha convertido en una herramienta muy utilizada en las organizaciones, para tratar de responder a los distintos problemas e incógnitas surgidos en su operativa diaria (Montoya & Gil, 2018).

Estudios recientes subrayan que nuestro mundo presenta una tendencia marcada por la generación de datos. Ello es consecuencia de la interacción de las personas con las nuevas tecnologías y de los flujos de datos personales. (Bollier, 2010). Así pues, las fuentes de datos se convierten en oportunidades para la investigación científica y la innovación de las empresas.

En el ámbito de la salud, se ha afirmado que, en el transcurso del próximo decenio, la medicina estará caracterizada por la utilización intensiva de datos (Sólo se explota entre un 10 y 15 por ciento de los datos en sanidad,

2017). En este sentido, la literatura señala que existen numerosos datos tratados en el ámbito de la medicina, como, por ejemplo, las imágenes médicas, las historias clínicas, los datos de ensayos clínicos y los datos de nivel de protección alto (Menasalvas, Gonzalo, & Rodríguez, 2017). Dentro de dicho contexto, es interesante estudiar el impacto que tiene *Big Data* en el ámbito sanitario, así como los riesgos que puede comportar para la confidencialidad de los pacientes, con especial atención a la protección de datos personales. En el caso de la región de Andalucía (España), la implantación de la Historia Clínica Electrónica (HCE, en adelante), dentro del sistema Diraya¹ ha implicado un gran volumen de información recogida y constituye un potencial interesante para *Big Data*.

Tras estas consideraciones generales, cabe señalar que los objetivos del presente trabajo son los siguientes: reflexionar sobre el conocimiento y la aplicación de *Big Data* en el ámbito sanitario, con especial referencia al estudio de sus ventajas, riesgos e inconvenientes, así como los desafíos que presenta. También se abordarán los límites legales derivados de las reglamentaciones aplicables. En lo concerniente a España y, particularmente, a Andalucía, se hará referencia a las posibilidades de *Big Data* en el ámbito sanitario.

¹ *Diraya es el sistema que se utiliza en el Servicio Andaluz de Salud como soporte de la historia clínica electrónica (Junta de Andalucía, 2004).*

El trabajo se estructura del siguiente modo: en primer lugar, se realiza una introducción y la descripción de la metodología utilizada. Posteriormente, se expone el marco teórico, dividido en los siguientes apartados: una aproximación al *Big Data*; ventajas e inconvenientes de aplicar el *Big Data*; *Big Data* y el ámbito sanitario; y límites legales derivados de las reglamentaciones aplicables. Luego, se estudia el caso de Andalucía: Diraya. Finalmente, se presentan las conclusiones.

Metodología

El presente trabajo, desde el punto de vista de la metodología, se ubica en los procesos de revisión documental del estado del arte, señalados por Carramolino (2009). Se realizó una revisión sistemática de la literatura (artículos científicos en revistas indexadas, comunicaciones publicadas, monografías de editoriales de prestigio internacional) como técnica exploratoria y analítica para la recolección de información y posterior análisis y tratamiento de aquella que presenta interés para el tópico del trabajo. La investigación se basó en la revisión de artículos publicados en algunas bases de datos como: *ABI INFORM*, *ScienceDirect*, *SciELO* y *Google Scholar*.

Para la obtención de publicaciones de interés sobre la temática de la investigación, se filtró la búsqueda de aquellos documentos cuyas palabras clave –tanto en inglés como en español–

guardaran relación con dicho tema, tales como *Big Data*, ámbito sanitario, historia clínica electrónica y Diraya. Después, se procedió al tratamiento de la información para la elaboración del presente trabajo. Su redacción se realizó a partir de la consolidación de los hallazgos o ideas más relevantes de cada artículo.

Una aproximación al concepto *Big Data*

En la actualidad, no existe una definición plenamente aceptada sobre *Big Data*, y, posiblemente, no exista jamás. A pesar de ello, este hecho se ha contemplado en la literatura de forma positiva, por cuanto se ha considerado útil contar con una concepción amplia y compartida sobre *Big Data* (Letouzé et al. 2017). A continuación, y sin ánimo de exhaustividad, se muestran algunas definiciones de *Big Data*.

La directiva UIT-T Y.3600 de la Unión Internacional de las Telecomunicaciones (*International Communication Union* [ITU], por su sigla en inglés), aprobada el 6 de noviembre de 2015, plantea que *Big Data* es un paradigma que permite la recogida, el almacenamiento, la gestión, el análisis y la visualización, potencialmente, en condiciones de tiempo real, de grandes conjuntos de datos con características heterogéneas. Para Ishwarappa y Anuradha (2015), una característica importante que define *Big Data* es

la gran cantidad de información que maneja. Así pues, la describen según las cinco V (5V): volumen, velocidad, variedad, veracidad y valor.

De acuerdo con la definición del tratamiento de datos que ha sido consolidada tanto en la reglamentación internacional como en las normas españolas, *Big Data* constituye un tratamiento de datos de carácter personal (Serrano, 2015), así como los conjuntos de datos cuyo tamaño excede la capacidad de las herramientas típicas de las bases de datos para capturar, almacenar y analizar (Henke, Libarikian, & Wiseman, 2016).

En la reputada sociedad de investigación Gartner (2018), se ha concebido al *Big Data* como los recursos de información caracterizados por su volumen, velocidad o variedad, los cuales requieren formas de tratamien-

to innovadoras y eficaces para mejorar los conocimientos y la toma de decisiones. Por otro lado, algunos autores subrayan que *Big Data* consiste en consolidar toda la información de una organización y ponerla al servicio de las empresas (Russom, 2011).

Tras la revisión de la literatura sobre la definición de *Big Data*, resultan de interés para los objetivos de este trabajo conocer las ventajas e inconvenientes que representa:

Ventajas e inconvenientes de la aplicación de *Big Data* en las organizaciones

La literatura constata que las ventajas y los inconvenientes de *Big Data*, a nivel general, son múltiples. En el Cuadro 1, se indican las principales ventajas que presenta para las organizaciones.

Cuadro 1. Principales ventajas del *Big Data* para las organizaciones

Mejoramiento de la estrategia de las empresas y su posicionamiento en el mercado.
Creación de nuevos productos y servicios.
Posibilidad de ofrecer mejores productos.
Mejoramiento de las relaciones con los clientes.
Las empresas pueden tomar decisiones más rápidamente y ser más competitivas.
Mejoramiento en la toma de decisiones.
Mejoramiento del análisis de riesgos con nuevos medios para controlar, medir y atenuar los riesgos comerciales y financieros.
Las empresas pueden plantear estrategias internas y externas para segmentar los clientes y poder actuar específicamente en cada segmento.

Fuentes: elaboración propia a partir de Velasco (2016); Álvarez (2013); Paniagua (2012); Duque-Jaramillo & Villa-Enciso (2016); Gil (2016); DataScope, 2018.

Sin embargo, *Big Data* no solo tiene ventajas. También presenta ciertos inconvenientes. Así, por ejemplo, el riesgo de caer en conclusiones erróneas que ningún ser humano ha revisado. Es importante analizar la verdadera relación entre las variables, con el fin de crear un modelo predictivo. En este sentido, es fundamental poder diferenciar la causalidad del azar.

Otro riesgo para las personas es la toma de decisiones automatizadas, sin intervención humana. La confianza ciega en los algoritmos lleva a menudo a las empresas a tomar decisiones que afectan a las personas, sin que se sepa la razón (Gil, 2016). De entre las soluciones señaladas por los analistas, se destaca aquella que consiste en programar o refinar los algoritmos, con el fin proteger a las personas y reducir el riesgo de tomar una decisión errónea a propósito de una persona en particular, así como la disminución del riesgo para la privacidad de las personas. Dicho tema está ligado al aspecto jurídico (Gil, 2016), el cual será tratado más adelante.

Finalmente, Lerman (2013) ha subrayado el riesgo de exclusión para aquellos cuya información no se recoge ni analiza. Ello es debido, según argumenta el citado autor, a que miles de millones de personas de todo el mundo permanecen en la periferia del *Big Data*. Ciertamente, la información sobre ellas no se recolecta ni analiza, porque no participan en actividades

diseñadas por *Big Data* para capturar información.

***Big Data* en el ámbito sanitario: potencialidad, inconvenientes y riesgos**

El estudio de *Big Data* en el ámbito sanitario presenta un gran interés. Al respecto, se ha afirmado que, a lo largo del próximo decenio, la medicina se caracterizará por la utilización intensiva de datos (Sólo se explota entre un 10 y 15 por ciento de los datos en sanidad, 2017). Así pues, la literatura ha subrayado que existe una cantidad significativa de datos tratados en el ámbito de la medicina. Se han señalado algunos ejemplos, tales como las imágenes médicas, las historias clínicas, los datos personales de nivel alto de protección, los datos genéticos que pueden ser puestos en evidencia, entre otros (Menasalvas et al., 2017). El análisis de estos datos tiene un papel fundamental en todo lo que, según los expertos, será la medicina del futuro o medicina "4P" (Hood & Galas, 2008): una medicina personalizada, predictiva, preventiva y participativa. Los ámbitos de la salud que presentan mayor potencial y más posibilidades de aplicación son el genómico, relacionado con el estudio de datos proporcionados por el material genético para la prevención y el tratamiento de enfermedades de origen genético; el "*Business Intelligence*", o el conjunto de estrategias y herramien-

tas enfocadas a la administración y la creación de conocimiento, a través del análisis de datos generados a partir de todo el proceso de atención médica); el de los modelos predictivos, enfocado en los diagnósticos y el cálculo de la probabilidad de morbilidad y mortalidad de la llamada minería de datos; el *crowdsourcing*, que se refiere a la investigación científica mediante las redes sociales, la financiación colectiva y el compartir datos de forma voluntaria. También se cuentan: la clínica operatoria, la autoasistencia; la mejora de la atención personalizada acordada con el paciente; la vigilancia a distancia de los pacientes; la medicina personalizada para todos; las autopsias virtuales; el seguimiento de los pacientes crónicos; y la mejora de los procedimientos médicos (Vincent, 2010; Rock Health Foundation citada por Joyanes & Poyatos, 2013).

Más concretamente, Fernández (2017) ha señalado el enorme potencial que tienen tanto el procesamiento como el análisis del *Big Data* Obstétrico para apoyar la toma de decisiones informadas y confiables por parte de la embarazada y los profesionales sanitarios. También se ha puesto de manifiesto el papel de *Big Data* en la prevención de enfermedades (Gong, Liu, & Chen, 2015). Más recientemente, se ha hecho hincapié en que, a través de *Big Data*, se favorece la duración del sistema sanitario y la aparición de nuevas formas de hacer en medicina (medicina personalizada,

preventiva, predictiva, participativa), así como luchar contra el fraude y los abusos (San Segundo & Soler, 2017).

Por último, se cree que el *Big Data* tendrá un impacto positivo sobre la mejora de los cuidados dispensados a los pacientes, siendo muy útil para los estudiantes de medicina y el personal sanitario, en interés de los pacientes (Serrano, 2015). No obstante, a pesar de la gran cantidad de datos recogidos en los procedimientos sanitarios, su explotación es baja (Cordellat, 2017).

Ahora bien, aunque *Big Data* cuenta con importantes ventajas y un gran potencial, como se recopiló anteriormente, la verdad es que también presenta una serie de inconveniente y riesgos. Estos de acuerdo con San Segundo y Soler (2017), han sido agrupados en cuatro bloques: barreras organizativas, obstáculos reglamentarios, barreras técnicas y riesgo éticos.

Acerca de las barreras organizativas, se puede afirmar que *Big Data* precisa un sistema de organización que permita compartir la información. Esta cuestión, en el caso de España y de otros países, adquiere una especial importancia por la particular organización del sistema nacional de salud pública, en el cual las Comunidades Autónomas (CC.AA.) son las que cuentan con la potestad en materia de salud pública. En este sentido, se ha afirmado cierta falta de coordinación, tanto entre las CC.AA. como entre el

sistema público y privado, así como entre diferentes centros sanitarios y departamentos (Mariscal, 2017). Por ello, la coordinación a nivel del consejo interterritorial supone un elemento esencial para conseguir el éxito (Biedma & Bourret, 2015; San Segundo & Soler, 2017).

Desde el punto de vista de los obstáculos reglamentarios, es necesario que la normativa garantice los derechos en el ámbito sanitario, asegurando la confidencialidad de las informaciones, sin que ello suponga un freno al desarrollo de *Big Data*. Así, por una parte, Serrano (2015) ha advertido la necesidad de reforzar las medidas de protección de las personas; mientras que, de otra parte, se ha subrayado también la necesidad de adaptar la normativa de protección de datos a la realidad de las nuevas tecnologías (Mariscal, 2017).

Con respecto a los riesgos éticos, los autores han señalado que este nuevo tipo de barrera queda definido por el desigual acceso a las TIC, sobre todo entre grupos vulnerables, así como por las nuevas situaciones de pérdida de libertad y autonomía, debido al aumento del control realizado por los diversos agentes del ámbito de la salud puesto sobre los ciudadanos. En este sentido, se ha alertado que, si bien la normativa europea regula la protección de los datos personales, entre ellos los datos de salud, existen países donde la regulación es inexistente o más laxa.

En tales países, la obtención de estos datos puede ser más fácil (Pérez, 2016). En una publicación del portal Correo Farmacéutico (Sólo se explota entre un 10 y 15 por ciento de los datos en sanidad, 2017), José Manuel Laperal ha llamado la atención sobre los riesgos que supone la ausencia de seguridad absoluta en lo referido a la protección de los datos.

Reflexión sobre el límite del respeto a la privacidad en el tratamiento de los datos y el Reglamento (UE) 2016/679. Estudio de posibles soluciones

El respeto a la vida privada es una obligación moral de la sociedad (Allen, 2016). En esta línea, dicho respeto es un principio esencial para permitir cualquier deliberación democrática y el ejercicio de la libertad de expresión (Krotoszynski & Ronald, 2015). Por ello, Rodrigo de Larrucea (2016) subraya la importancia de la privacidad de las personas. Además, añade que esta se debe preservar y tiene que prevalecer frente a cualquier ventaja del *Big Data* en el ámbito de la salud.

En el caso de Europa, la normativa sobre protección de datos personales, está basada en el consentimiento del propietario de los datos personales (Rubinstein, 2013). El problema, como señalan Becerra et al. (2018) y Gil (2016), es que el consentimiento se obtiene masivamente por defecto. Dichos autores consideran que no es

realista creer que existe un control efectivo de las informaciones personales por la vía del consentimiento. Además, una gran mayoría de personas no leen las políticas de confidencialidad antes de dar su consentimiento. Otras leen dichas políticas sin comprender su importancia. Así, dar el consentimiento es, en general, un ejercicio vacío de sentido.

Por otra parte, la sociedad no quiere renunciar a la utilización de las TIC, pero no tiene una cultura de la privacidad. Ello significa que la garantía del consentimiento es casi irreal o ineficaz (Oliver & Muñoz, 2014; Heeger 2015). Además, como apunta Gil (2016), el principio de “minimización de datos” no se respeta en la práctica. Este supone que los datos recogidos no deben ser excesivos. Es decir, solo se recabará la cantidad mínima necesaria para los fines requeridos. Como recuerda el autor, la realidad es que las autoridades de protección de datos rara vez solicitan a las empresas revisar sus procesos para minimizar los datos recopilados. La paradoja es que el principio de minimización de datos se opone a la lógica misma de *Big Data*. Los nuevos modelos analíticos descansan precisamente sobre el estudio de enormes cantidades de datos, sin los cuales sería imposible explotar el conocimiento que *Big Data* permite (Gil, 2016). Esta problemática invita a reflexionar acerca de posibles soluciones.

Una posible alternativa se encuentra en la anonimización de los datos. Sin embargo, como señala Gil (2016), se han detectado numerosos casos de reidentificación en bases de datos anonimizadas. Esta misma autora señala que la puesta en marcha de medidas de transparencia es un medio de contribuir a la protección de la confidencialidad. Por ejemplo, las informaciones concernientes al lugar donde han sido obtenidos los datos, las autorizaciones que comportan, en qué medida las personas son identificadas o no, así como el tiempo en el cual los datos estarán almacenados. En esta línea, el artículo 35.1 del Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 abril 2016 estipula que:

Cuando sea probable que un tipo de tratamiento, en particular si utiliza nuevas tecnologías, por su naturaleza, alcance, contexto o fines, entrañe un alto riesgo para los derechos y libertades de las personas físicas, el responsable del tratamiento realizará, antes del tratamiento, una evaluación del impacto de las operaciones de tratamiento en la protección de datos personales. Una única evaluación podrá abordar una serie de operaciones de tratamiento similares que entrañen altos riesgos similares. (Reglamento (UE) 2016/679 L119, p. 53)

La Agencia Española de Protección de Datos (AEPD en adelante) ha considerado positivamente la implementación de evaluaciones de impacto en la privacidad o protección de datos, lo cual van más allá de la simple verificación del cumplimiento de la ley y permiten, de acuerdo con un proceso sistemático y reproducible, anticipar respuestas a cada riesgo, involucrar a terceros en su solución, educar y motivar a los empleados y aumentar la credibilidad de la empresa. Además, el informe de la *Agència Valenciana del Turisme Invat. tur* (2015) ha señalado que tales evaluaciones son particularmente útiles en casos como *Big Data*, Internet de las cosas o el desarrollo y construcción de ciudades inteligentes.

Otra posible solución es la adopción de sistemas de confidencialidad desde el diseño (Gil, 2016). Ello se realizaría para tener en cuenta la confidencialidad desde el primer momento. Al respecto, Cavoukian (2009) avanzaba hace unos años que el concepto *privacidad desde el diseño* debe basar el modelo de operar por defecto de cualquier organización. Este tipo de sistemas de confidencialidad implica que las tecnologías se construyen, teniendo en cuenta la necesidad de proteger la confidencialidad. Se ha considerado que, en este caso, además de cumplir con los principios de protección de datos, es preciso tener en cuenta la confidencialidad desde el mismo momento cuando se utiliza la

tecnología. En este sentido, la AEPD (Agencia Española de Protección de Datos [AEPD], 2014) subraya que el concepto de confidencialidad, desde el diseño, facilita una mayor efectividad en la protección de los derechos de los afectados y evita los costes posteriores de las empresas para la recuperación y adaptación legal de la tecnología, así como para el daño potencial a su reputación.

Por otra parte, con fines de protección, el nuevo Reglamento Europeo estableció el denominado "derecho al olvido". Tal derecho es la respuesta a la amenaza que supone, para el desarrollo libre de la personalidad, el almacenamiento permanente en Internet de información personal, cuya difusión, con el tiempo, puede dañar a la persona, cuando existe una discrepancia entre los datos publicados y la realidad actual (Mieres, 2014).

Desde un punto de vista ético, también se debe tener en cuenta la integridad de los investigadores que reutilizarán los datos. Esta integridad se presume, debiendo llevarse a cabo con honestidad, independencia y respeto por su deber de secreto, según las buenas prácticas de la investigación. Antes de volver a usar los datos, los investigadores deben comprometerse a no usar los datos para fines distintos a los autorizados, a no transferirlos a otros, ni a identificar de nuevo los temas (Sánchez, 2019).

Después de revisar las limitaciones legales y los grandes datos, podemos decir que la protección de datos y la confidencialidad son fundamentales. A pesar del progreso, es necesario continuar buscando más soluciones que protejan al individuo y, al mismo tiempo, sirvan como una ventaja competitiva para las organizaciones. No debemos olvidar que el éxito de los grandes datos en salud se conseguirá mediante la integración conjunta de tecnología, políticas de privacidad y confidencialidad apropiadas, infraestructura y una cultura de intercambio de datos (Menasalvas et al. 2017).

El caso de Andalucía: Diraya

En el caso de Andalucía, con la implementación progresiva de la Historia Clínica Electrónica dentro del marco del proyecto Diraya, el volumen de información recopilada es muy importante y constituye un potencial interesante para *Big Data*. Cabe recordar que Diraya es un sistema de información de soporte para el Servicio Andaluz de Salud (SAS, en adelante) de la Junta de Andalucía, el cual consta de diferentes módulos vinculados entre sí y que comparten información, de modo que cada información se registra solo una vez en el sistema (Servicio Andaluz de Salud, 2004). El sistema informático Diraya registra más de 383 millones de registros de salud y 163 millones de recibos electrónicos cada año, entre otros datos (Sáez, 2017).

En España, el impacto de los grandes datos en la salud es más una promesa que una realidad. Entre otras razones, se pueden citar las siguientes: datos de salud fragmentados entre diferentes comunidades autónomas y entre sistemas públicos y privados; heterogeneidad de políticas y gestión de la salud; los profesionales actuales probablemente no estén preparados para aplicar *Big Data* a la práctica clínica (Martos, 2017; Puyol, 2014). No obstante, se ha defendido la necesidad de implementar una estrategia nacional de *Big Data*, por considerar que es una oportunidad para unir voluntades políticas y tecnologías (Parra, 2016).

Actualmente, hay pocos estudios sobre *Big Data* y salud. Sin embargo, un estudio en Sevilla (España) sobre pacientes crónicos que utiliza técnicas de *Big Data* concluye que los modelos predictivos que analizan la base de datos del historial médico facilitan el acercamiento de los pacientes crónicos, después del alta hospitalaria. Así, el envío de alertas a los profesionales de referencia de un paciente facilita la priorización de las intervenciones para pacientes crónicos complejos. Esta herramienta permite intervenciones ambulatorias inmediatas que evalúan la gestión del régimen de tratamiento y la adherencia al tratamiento, siendo necesaria para tratar de evitar las causas que conducen a los reingresos al hospital (López, 2017).

Por otro lado, Andalucía ha firmado un acuerdo entre el SAS y Red.es, entidad pública adscrita al Ministerio de Energía, Turismo y Agenda Digital de España. Dicho acuerdo permitirá, especialmente, trabajar en la aplicación de tecnologías emergentes como *Big Data*, *Analytics*, *Machine Learning* o *Internet of Things* (Junta de Andalucía, 2018).

Fuera de la Comunidad Autónoma de Andalucía, cabe resaltar que Indra, empresa líder en tecnología y consultoría, lidera el proyecto P-Ictus. Como señala Gutierrez (2018) en un comunicado de prensa, el objetivo del proyecto es prevenir el accidente cerebrovascular basándose en el conocimiento obtenido de las tecnologías *Big Data* de la historia clínica electrónica y otras fuentes de datos no clínicos. Para desarrollar el citado proyecto, según indica la empresa, utilizan los datos suministrados por la Fundación para la Formación e Investigación de Profesionales de la Salud (Fundesa-lud) Extremadura.

A continuación, tal como señala el mencionado comunicado de prensa, procederán a la extracción de los datos de pacientes que hayan sufrido un accidente cerebrovascular previamente identificados, junto con información de otras áreas de su historia clínica, como recaídas, hospitalizaciones, pruebas de diagnóstico o medicamentos. Para conseguir una mejora del proceso de estratificación individual posterior,

los datos anteriormente indicados se complementarán con otros datos como los económicos, sociodemográficos, de contaminación, climatológicos y sociales. Toda la información será procesada a través de técnicas de análisis estadístico de *Big Data*. Ello producirá conocimiento sobre la prevención y el diagnóstico de distintos tipos de accidente cerebrovascular.

Gutierrez (2018) concluye señalando que los resultados favorecerán el desarrollo de reglas y modelos que tienen su fundamento en la ciencia, incorporando una herramienta de apoyo a la decisión de fácil acceso para los profesionales. El coordinador del proyecto ha señalado que, si bien en todas las áreas se habla del potencial de las tecnologías de *Big Data* para la salud, por el momento no se pueden aplicar “en datos clínicos relevantes validados por profesionales sanitarios como diagnósticos, antecedentes o alergias” (Gutiérrez, 2018, párraf. 5).

Conclusiones

En general, *Big Data* representa el inicio de una era de innovación, competitividad, productividad y revolución científica (Chen & Chun-Zhang, 2014). Las diversas fuentes de datos suponen oportunidades para la investigación científica y la innovación empresarial, las cuales deben ser aprovechadas.

La literatura constata la inexistencia de una definición plenamente aceptada,

pero, a la vez, se afirma la utilidad de una concepción amplia y común de aquello que se entiende por *Big Data* (Letouzé et al., 2017). En este sentido, resulta de interés la definición aportada por la sociedad de investigación Gartner (2018), cuando considera que *Big Data* como los recursos de información caracterizados por su volumen, velocidad o variedad que requieren formas de tratamiento innovadoras y eficaces para mejorar los conocimientos y la toma de decisiones.

De la revisión realizada, se concluye que las ventajas del uso de *Big data* por parte de las organizaciones son numerosas e importantes. No obstante, se han detectado ciertos inconvenientes que deben ser superados. En el caso concreto del ámbito sanitario, cabe señalar que la medicina está caracterizada por la utilización intensiva de datos (Cordellat, 2017). El análisis de estos datos va a jugar un papel fundamental en lo que los autores denominan *medicina del futuro o medicina 4P*: personalizada, predictiva, preventiva y participativa (Hood & Galas, 2008). A pesar de la gran cantidad de información recogida, su explotación en el ámbito de la salud es baja (Cordellat, 2017).

Se estima que el *Big Data* va a tener un impacto positivo en la mejora de los cuidados de salud de los pacientes, así como en la prevención, el diagnóstico de tratamiento de enfermedades y en la reducción de gastos sanitarios (Me-

nasalvas et al, 2017; Serrano, 2015). Sin embargo, se han reconocido los riesgos que comporta para la confidencialidad de los pacientes (Laperal, 2017). La solución del consentimiento, prevista en las normas, no es suficiente, porque el consentimiento se presta de forma masiva por defecto y no existe un control efectivo de los datos personales (Gil, 2016; Becerra et al., 2018), y, en ese sentido, es necesario trabajar en soluciones para proteger dicha confidencialidad.

Frente a ello, se han propuesto diversas soluciones, tales como la anonimización de los datos y la puesta en marcha de medidas de transparencia, como medio que ayuda a la protección de la confidencialidad. El Reglamento Europeo ha establecido la evaluación del impacto de las operaciones de tratamiento en la protección de datos personales. Otra posible solución apunta a la adopción de sistemas de confidencialidad desde el diseño (Gil, 2016). Para ello, se requiere construir los sistemas de confidencialidad desde el primer momento, implicando que las tecnologías se forjan teniendo en cuenta la necesidad de proteger la confidencialidad.

Adicionalmente, la revisión de literatura señala que, en la actualidad, no existen muchos estudios sobre el *Big Data* en el ámbito sanitario. En España, y particularmente en Andalucía de forma tímida, el *Big Data* comienza a ser explotado en el ámbito sanitario.

Tras la revisión realizada, se puede afirmar que el *Big Data* constituye una gran oportunidad para los científicos, los pacientes y los profesionales. No obstante, es necesario reducir los riesgos descritos y se requiere incrementar la investigación en torno a su utilidad en campos como la salud. Es un desafío importante para los próximos años.

Referencias

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS (2014). *Guía para una Evaluación de Impacto en la de Protección Datos Personales*. Recuperado de <https://www.aepd.es/sites/default/files/2019-09/guia-analisis-de-riesgos-rgpd.pdf>

AGÈNCIA VALENCIANA DEL TURISME [INVAT. TUR]. (2015). *BIG DATA: retos y oportunidades para el turismo*. Recuperado de <http://www.thinktur.org/media/Big-Data.-Retos-y-oportunidades-para-el-turismo.pdf>

ALLEN, A. (2016). Protecting one's own privacy in a big data Economy. *Harvard Law Review Forum*, 130(71), 71-78.

ÁLVAREZ, J. (2013). El Big Data: Una Gran Oportunidad. *La Nueva España*. Recuperado de <http://www.lne.es/opinion/2013/11/10/big-data-una-gran-oportunidad/1497242.html>

BECERRA, J., COTINO, L., LEÓN, I. P., SÁNCHEZ-ACEVEDO, M. E., TORRES, J., & VELANDIA, J. (2018). Derecho y Big Data. *Jus Público*, 25.

BIEDMA, J. M., & BOURRET, C. (2015). Les enjeux des projets d'interopérabilité en santé pour la mobilité des citoyens. Le cas de l'Espagne, *Revue Projectique*, (13/1), 23-36.

BOLLIER, D. (2010). *The Promise and Peril of Big Data*. Washington, DC: The Aspen Institute.

CARRAMOLINO, B. (2009). *La revisión del estado del arte. Metodología de la investigación*. Valladolid: Universidad de Valladolid.

CAVOUKIAN A. (2009). *Privacy by Design. The 7 Foundational Principles*. Recuperado de <https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf>

CHEN, C. L. P., & CHUN-ZHANG, Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, 275, 314–347.

DATASCOPE. (22 de marzo de 2018). ¿Cómo tu empresa puede beneficiarse del Big Data?. *My Data Scope*. Recuperado de <https://www.mydatascope.com/blog/es/2018/03/22/como-tu-empresa-puede-beneficiarse-del-big-data/>

CORDELLAT, A. (5 de junio de 2017). Retos en la transformación digital de la sanidad española. *Diario Médico*. Recuperado de: <https://www.diario-medico.com/profesion/retos-en-la-transformacion-digital-de-la-sanidad-espanola.html>

DUQUE-JARAMILLO, J. C., & VILLA-ENCISO, E. M. (2016). Big Data: desarrollo, avance y aplicación en las organizaciones de la era de la información. *Revista CEA*, 2(4), 27-45.

FERNÁNDEZ, M. I. (2017). Resultados del Big Data aplicado a la gestión Obstétrica, una revisión narrativa. *Revista Tesela*, (21). Recuperado de <http://www.index-f.com/tesela/ts21/ts11595.php>

GARTNER. (2018). Big Data. *Gartner Glossary*. Recuperado de <http://www.gartner.com/it-glossary/big-data/>

GIL, E. (2016). *Big data, privacidad y protección de datos*. Recuperado de <https://www.aepd.es/sites/default/files/2019-10/big-data.pdf>

GONG, J., LIU, S., & CHEN, D. (2015). Use of big data - targeted prevention and control of regional and seasonal diseases in obstetrics. *CJPR*, 35(10), 1033-1036.

GUTIÉRREZ, C. (2018). Indra lidera un proyecto de I+D+i para prevenir el ictus en los pacientes de riesgo a través de tecnología Big Data. *Indra*. Recupe-

rado de "<https://www.indracompany.com/es/noticia/indra-lidera-proyecto-idi-prevenir-ictus-pacientes-riesgo-traves-tecnologia-big-data%20%20>"

HEEGER, E. (August, 2015). Controlling your online profile: reality or an illusion? A research into informed consent as a mechanism to regulate commercial profiling. *SSRN*. Recuperado de <http://dx.doi.org/10.2139/ssrn.2658651>.

HENKE, N., LIBARIKIAN, A., & WISEMAN, B. (October, 2016). Straight talk about big data. *McKinsey Digital*. Recuperado de <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/straight-talk-about-big-data#>

HOOD, L. E., & GALAS, D. J. (2008). P4 Medicine: Personalized, predictive, preventive, participatory. A Change of view that changes everything. *Computing Research Association*. Recuperado de http://www.cra.org/ccc/files/docs/init/P4_Medicine.pdf

INTERNATIONAL COMMUNICATION UNION [ITU]. (2015). Big data – Cloud computing based requirements and capabilities. Series Y: Global Information Infrastructure, Internet Protocol Aspects and Next-Generation Networks Cloud Computing (Y.3600). Recuperado de <https://dig.watch/sites/default/files/T-REC-Y.3600-201511-1%21%21PDF-E.pdf>

- ISHWARAPPA, K., & ANURADHA, J. (2015). A brief introduction on Big Data 5vs characteristics and Hadoop technology. *Procedia Computer Science*, 48, 319-324.
- JOYANES, L., & POYATOS, J. M. (2013). *Big Data y el sector de la salud: el futuro de la sanidad*. Recuperado de <http://poyatosdiaz.com/index.php/big-data-y-el-sector-de-la-salud-el-futuro-de-la-sanidad>
- KROTOSZYNSKI, J., & RONALD, J. (2015). Reconciling privacy and speech in the era of Big Data: A comparative legal analysis. *William & Mary Law Review*, 56(4), 1279-1338.
- LERMAN, J. (2013). Big Data and Its Exclusions. 66 *Stanford Law Review Online*, 55. Recuperado de https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2293765
- LETOUZÉ, E., CAÑÓN, V., CLAVIJO, A., GODOY, L., PESTRE, G., & RICARD, J. (2017). *Definición de la estrategia de Big data para el estado colombiano y para el desarrollo de la industria de Big Data en Colombia estado del arte y análisis comparativo de estrategias nacionales de Big data*. Recuperado de http://datapopalliance.org/wp-content/uploads/2018/09/Documento1_VersionFinal_DNP.pdf
- LÓPEZ, B. (2017). La utilización de un modelo predictivo para gestionar los reintegros hospitalarios en pacientes crónicos complejos. *Enfermería Clínica*, 27 (Espec Congr 2), 18, 1-2.
- MARISCAL, R. A. (27 de octubre de 2017). ¿Cuáles son los principales obstáculos para implantar el Big Data en salud? *iSanidad*. Recuperado de <http://isanidad.com/99065/cuales-son-los-principales-obstaculos-para-implantar-el-big-data-en-salud/>
- MARTOS, A. (2017). *Big data, un reto para las organizaciones sanitarias*. Ponencia presentada en el XI Foro de Sanidad. Universidad de Alcalá, Alcalá, España.
- MENASALVAS, E., GONZALO, C., & RODRÍGUEZ, A. (2017). Big Data en salud: retos y oportunidades. *Economía industrial*, (405), 87-97.
- MIERES, L. J. (2014). El derecho al olvido digital. Documento de trabajo, (186). Recuperado de https://www.fundacionalternativas.org/public/storage/laboratorio_documentos_archivos/e0d97e985163d78a27d6d-7c23366767a.pdf
- MONTOYA, M.C., & GIL, G. A. (enero-junio, 2018). Actualidad e importancia de la implementación de Big Data utilizando las herramientas Hadoop y Spark. *Lámpsakos*, (19), 67-72.
- OLIVER, D., & MUÑOZ, J. F. (2014). El mito del consentimiento, o por qué un sistema individualista de protección de datos (ya) no sirve para (casi) nada. En

J. Valero Torrijos (Ed.), *La protección de los datos personales en Internet ante la innovación tecnológica* (pp. 153-196). Pamplona: Thomson Reuters Aranzadi.

PANIAGUA, S. (23 de julio de 2012). Josep Curto, de IDC: Big Data es un cambio transgresor y disruptivo. *Soraya Paniagua Blog*. Recuperado de <https://www.sorayapaniagua.com/2012/07/23/josep-curto-de-idc-big-data-es-un-cambio-transgresor-y-disruptivo/>

PARRA, C. L. (2016). Big Data en sanidad en España: la oportunidad de una estrategia nacional. *Gaceta Sanitaria*, 30(1). Recuperado de http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0213-91112016000100013

PÉREZ, G. (enero-febrero, 2016). Peligros del uso de los Big Data en la investigación en salud pública y en epidemiología. *Gaceta Sanitaria*, 30(1). Recuperado de http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0213-91112016000100014

PUYOL, J. (2014). Una aproximación a Big Data. *Revista de Derecho UNED*, (14), 471-505.

REGLAMENTO (UE) 2016/679 DEL PARLAMENTO EUROPEO Y DEL CONSEJO (27 de abril), *relativo a la protección de las personas físicas en lo que respecta al*

tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos). Recuperado de <https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX%3A32016R0679>

RODRIGO DE LARRUCEA, C. (2016). Mhealth y Bigdata en sanidad. *Derecho y salud no van siempre de la mano*. Recuperado de <https://carmenrodrigodelarrucea.wordpress.com/2016/04/14/mhealth-y-bigdata-en-sanidad/#more-793>

RUBINSTEIN, I. (2013). Big Data: The end of privacy or a new beginning? *International Data Privacy Law*, 3(2), 74-87.

RUSSOM, P. (2011). Big Data Analytics, *TDWI Best Practices Report*. Recuperado de <https://vivomente.com/wp-content/uploads/2016/04/big-data-analytics-white-paper.pdf>

SÁEZ, D. (2017). Noticias del Sector. *Revista de la Sociedad Española de Informática y Salud*, (123). Recuperado de <https://seis.es/wp-content/uploads/2018/02/Revista-123.pdf>

SÁNCHEZ, J. M. (2019). La investigación con Big Data desde la perspectiva de la ética. *El Médico Interactivo*. Recuperado de: <http://elmedicointeractivo.com/la-investigacion-con-big-data-desde-la-perspectiva-de-la-etica/>

SAN SEGUNDO, J. M., & SOLER, I. (2017). *Big Data en Salud Digital. Informe de resultados 2017*. Recuperado de <https://www.ontsi.red.es/ontsi/sites/ontsi/files/Informe%20Big%20Data%20en%20Salud%20Digital.pdf>

SERRANO, M. (junio, 2015). Big data o la acumulación masiva de datos sanitarios: derechos en riesgo en el marco de la sociedad digital. Ponencia presentada en el XXIV Congreso sobre Derecho y Salud, Granada, España.

SERVICIO ANDALUZ DE SALUD [SAS]. (2004). Qué es Diraya. *SAS Información*, (24). Recuperado de http://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/contenidos/publicaciones/datos/176/html/DATOS/bol_sas24.pdf

_____. (s. f.). *Servicio Andaluz de Salud, Consejería de Salud* [web]. (2018). Recuperado de <https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/>

SÓLO SE EXPLOTA ENTRE UN 10 Y 15 POR CIENTO DE LOS DATOS EN SANIDAD. (5 de junio de 2017). *Correo Farmacéutico*. Recuperado de <https://www.correo-farmacaceutico.com/politica-sanitaria/solo-se-explota-entre-un-10-y-15-porciento-de-los-datos-en-sanidad.html>

VELASCO, L. (2016). Big Data aplicado al proceso de toma de decisiones. Informe Big Data. Recuperado de <http://www.computing.es/siteresources/files/744/39.pdf>

VINCENT J. L. (2010). We should abandon randomized controlled trials in the intensive care unit. *Crit Care Med*, 38(10), 534–538.