Universidad Internacional de La Rioja

Facultad de Derecho

Máster Universitario en Protección de Datos

El impacto de la tecnología *big data* en el derecho fundamental de protección de datos

Trabajo fin de estudio presentado por:	Ana María Cruz Horrillo
Tipo de trabajo:	Trabajo Final de Máster
Director/a:	Santiago Bermell Girona
Fecha:	23 de julio del 2020

Resumen

La revolución tecnológica ha llegado y con ello la era del *big data*, el tratamiento de grandes volúmenes de datos ofrece muchas ventajas para la economía y el progreso humano. Sin embargo, también va acompañado de riesgos que afectan el derecho fundamental de protección de datos. El presente trabajo académico propone analizar la normativa vigente, en especial, el Reglamento General de Protección de Datos, a la luz de los servicios que ofrece el *big data*, con el fin de identificar sus riesgos desde la perspectiva jurídica y determinar si dicha normativa es capaz de dar soluciones a las problemáticas que se plantean.

Palabras clave: Protección de datos, *big data*, discriminación algorítmica, decisiones automatizadas.

Abstract

The technological revolution has arrived and with it the era of big data, the processing of large volumes of data offers many advantages for the economy and human progress. However, it also involves severe risks that adversely affect the data protection right. This paper looks to analyze the current laws, especially the General data protection regulation, in light of the services offered by big data, in order to identify its risks from a legal perspective and determine if the applicable regulation is capable to give solutions to the problems that arise.

Keywords: Data protection, *big data*, algorithmic discrimination, automated decisions.

Índice de contenidos

Índice de tablas

1. Introducción

De la misma manera que Henry Ford revolucionó la industria automovilística, el entendimiento del trabajo humano y la sociedad, la tecnología *big data* revolucionará el conocimiento, la estructura social y la comunidad humana (Boyd y Crawford, 2012). Como bien establece Latour (2009): "cambia los instrumentos y cambiarás toda la teoría social que los acompaña."

El *big data* ofrece no sólo a las disciplinas humanísticas una nueva forma de ciencia y método objetivo, sino también la posibilidad de cuantificar realidades sociales, nunca antes cuantificadas (Boyd *et al.* 2012), pero independientemente de esta capacidad astronómica de cuantificar, el *big data* no está lejos de limitaciones, riesgos, amenazas y discriminaciones hacia los interesados. El despliegue de la revolución tecnológica y las sociedades de la información requiere repensar la protección jurídica del derecho de protección de datos en los entornos de *big data*. ¿Es suficiente la actual regulación para proteger el derecho fundamental? ¿Cuáles son las implicaciones legales de utilizar datos masivos? Sin duda, este es el tema de debate en el siglo XXI.

Lo cierto es que, al siglo XXI le corresponde la religión del dataísmo[1]. Como bien expresa el historiador Harari (2017): "el *big data* será la nueva religión, la religión no trata tanto sobre Dios como de autoridad. Cuando tienes que tomar decisiones en tu vida ¿Hacia dónde te vuelcas a buscar la solución? En la edad media la fuente de autoridad suprema era Dios, así que la buscabas en la biblia. En el siglo XX la gente creía que la fuente suprema era el propio ser humano y su libertad. En el siglo XXI la autoridad son los algoritmos del *big data*. Cuando te enfrentas a un problema recurres a Google (…)."

[1] Término acuñado por el historiador Yuval Noah Harari en su libro *Homo deus*, quien establece que el liberalismo será reemplazado por la ideología o religión: dataísmo. Al igual que el capitalismo comenzó como una teoría científica neutral, pero ahora está mutando en una religión. Su valor supremo es el flujo de información. Véase: Harari, Y. *Homo Deus.* Barcelona : Penguin Random House Grupo Editorial, 2017. págs. 400-430.

En consecuencia, la humanidad se enfrenta a tiempos revolucionarios, donde el antropocentrismo es despojado por el datacentrismo (Harari, 2017). El *big data* es sólo el principio, de un largo camino de tecnología disruptiva, que podrá afectar los principios básicos liberales, entre ellos, el derecho fundamental de protección de datos. La pregunta es: ¿la sociedad se encamina a dictaduras digitales, o podrá marcar pautas éticas y legales, para propiciar el desarrollo tecnológico?

Precisamente este trabajo se centrará en analizar el *big data* desde la perspectiva jurídica, en especial a partir el Reglamento General de Protección de Datos, abordando las cuestiones claves como la transparencia, bases de legitimación, elaboración de perfiles, principios y anonimización.

Así, los primeros capítulos se centrarán en la definición de *big data*, sus principales características, beneficios y riesgos. En concreto, se hará énfasis en cómo la privacidad puede verse erosionada por la toma de decisiones automatizadas con efectos jurídicos hacia los interesados.

Análogamente cabe preguntarse si, dichos problemas encuentran soluciones en la actual normativa aplicable o si existen deficiencias. ¿El *big data* podrá llegar a discriminar los interesados? ¿Cuáles son las protecciones jurídicas en nuestro ordenamiento jurídico?

Este enfoque es de crucial importancia para la época en que vivimos, donde parece ser que el *big brother* de la literatura de Orwell (2013) cobró vida bajo el nombre de *big data*. Es más, la información se transformó en un arma letal. Después de todo la afirmación de Sun Tzu destruir al enemigo divulgando las comunicaciones internas (Tzu, 1999), es del todo pertinente, y más aún cuando todos los días somos testigos del mal uso de la información, de la invasión en los datos y la privacidad.

En consecuencia, el presente escrito académico se adentrará en describir el marco jurídico del derecho de protección de datos, intentando determinar si dicha regulación es suficiente para abordar los riesgos del *big data* en nuestra sociedad digital.

1.1.Justificación del tema elegido

El interés por esta temática radica en ser partícipe de la construcción de conocimiento, y de tejer diálogos para afrontar uno de principales problemas del siglo XXI: la tecnología disruptiva del *big data*. Así, en tiempos antiguos el bien más preciado era la tierra, luego en la época moderna las máquinas y las fábricas resultaron más importantes que la tierra. Ahora en la actualidad los datos eclipsarán la tierra y las máquinas (Harari, 2018). La carrera para poseer datos ha comenzado, y la maquinaria del derecho va a toda marcha para enmarcar esta nueva realidad social. Su más insigne producción: el Reglamento General de Protección de Datos.

Sin embargo, es inevitable el planteamiento de interrogantes como: ¿es suficiente dicha regulación? Conocer sus vacíos es verdaderamente apasionante, cuando de ello depende esbozar sociedades presentes y futuras, dado que, como bien sabemos el derecho es construcción social. Después de todo, la teoría *da mihi facta, dabo tibi ius* (dame los hechos y te daré el derecho) al fin y al cabo no es exacta. Más bien, el derecho contribuye también a moldear realidades sociales e imaginarios colectivos (Lemaitre, 2009).

En cualquier caso, no hay que perder de vista que el presente trabajo no pretende demonizar la tecnología o poner freno a la innovación tecnológica, sino estudiar el ciclo de vida de los datos en el ámbito de *big data*, los intervinientes y brindar una protección a los interesados en el desarrollo de una sociedad digital.

1.2. Problema y finalidad del trabajo

Las organizaciones, y la sociedad en general, están experimentando una transformación con la llegada de las tecnologías emergentes, en especial con la tecnología *big data*. Como bien expone Steven Pinker: “al sustituir los átomos por bits, la revolución digital está desmaterializando el mundo ante nuestros ojos.” (Pinker, 2019). Evidentemente, al tiempo que se propicia el progreso y la innovación tecnológica, se generan grandes riesgos a los derechos y libertades de los interesados. Nuestro legislador comunitario al percatarse de dicho suceso otorgó un mayor poderío a los interesados mediante la aprobación del Reglamento General de Protección de Datos.

Ahora bien, es cierto que el RGPD apuesta por mecanismos proactivos, bajo el manto del principio de responsabilidad proactiva y del principio de privacidad por diseño y por defecto (AEPD, 2017). Sin embargo, a medida que avanza la tecnología esta situación podría agravarse cada vez más.

No es sorprendente el crecimiento ágil de los macrodatos y sistemas algorítmicos que pueden resultar en vulneraciones de derechos fundamentales y en discriminaciones indirectas a grupos de personas o grupos minoritarios (Parlamento Europeo, 2017). De ahí la importancia del presente trabajo, el cual busca analizar si dicha normativa otorga las soluciones legales eficaces para disminuir el impacto en los derechos de los interesados, planteándose si las figuras jurídicas como la anonimización, elaboración de perfiles, evaluación de impacto etc., ¿son suficientes para contrarrestar los riesgos?

Dicho lo anterior, piense el lector por ejemplo en la desprotección del titular de los datos en el contexto tecnológico manifestando sesgos discriminatorios. (Cotino, 2017). ¿Es posible que el algoritmo tenga un cierto sesgo discriminatorio? ¿Es eficaz el poder de control de los interesados en el ámbito del *big data*?

En conclusión, este trabajo pretende analizar la protección del derecho fundamental de protección de datos a la luz de la normativa vigente, contrarrestando las disposiciones legales dentro del ámbito del *big data*, con el objetivo de enriquecer el debate al conciliar ambos mundos, cómo innovar con privacidad, y sin duda alguna cómo armonizar tecnología, avance y derecho.

1.3. Objetivos

Los objetivos generales más importantes de este trabajo son los siguientes:

- Observar, estudiar, analizar y desarrollar la problemática de estudio centrada en la vulneración del derecho de protección de datos en el ámbito del *big data*.
- Analizar el marco regulatorio de protección de datos y sus figuras jurídicas frente a las implicaciones tecnológicas del *big data*.
- Determinar si la nueva legislación es suficiente para contrarrestar las amenazas y riesgos de la tecnología del *big data*.

Los objetivos secundarios más importantes de este trabajo son:

- Estudiar el impacto del *big data*, sus beneficios y riesgos.
- Identificar las deficiencias de la nueva regulación de protección de datos en el ámbito del *big data*.

1. ¿Qué es el *big data*?

El *big data*[2] es un término que alude al enorme crecimiento en el acceso y uso de información automatizada. Se refiere a las gigantescas cantidades de información digital y que están sujetas a un análisis extenso basado en los algoritmos[3] (GT29, 2013).[4]

De este modo, presenciamos un fenómeno explosivo de información, en donde cada día se genera más de 2.5 exabytes de datos, equivalente a 1.000.000 de terabytes (Puyol, 2014). El volumen de datos proviene de incontables fuentes: sensores, teléfonos, motores de búsqueda, etc. El cual se acrecentará aún más con la consolidación del internet de las cosas y las ciudades inteligentes.

En concreto, si dispusiéramos de una pila de discos tan larga cubriría 1,5 veces la distancia hasta la luna, y lo más sorprendente es que dicha cifra se ha producido en apenas los últimos dos años, creciendo cada anualidad a un ritmo aproximado del 50 por ciento (Big data: la galaxia de los datos, 2020).

¿Qué es lo interesante de grandes volúmenes de datos? Según (Puyol, 2014) es posible utilizar esta herramienta para generar valor y servicios como marketing personalizado, ayudarnos a comprender el mundo, e incluso predecir lo que ocurrirá o resolver una crisis económica.

---

[2] Frente a la definición del *big data* encontramos múltiples acepciones:
***big data* como oportunidad.** Este es un concepto acuñado por Matt Aslett y define el *big data* como análisis de datos que fue ignorado previamente debido a las limitaciones de la tecnología.
***big data* como tecnología**. El término de *big data* nació y se impulsó por las nuevas tecnologías y en particular por el código abierto como HADDOP y NoSQL, formas de almacenar y manipular datos. (Elliot, 2013).

El *big data* también es emparentado con la minería de datos. Dado que, la minería de datos al igual que el *big data* utiliza los métodos de inteligencia artificial y la estadística para analizar los patrones. Sin embargo, de acuerdo con la guía de la (AEPD, 2017) el *big data* integra información de una mayor diversidad de fuentes (internas y externas), formatos (variedad) y en muchos casos el resultado se ha de obtener con mucha mayor celeridad (velocidad).

[3] De acuerdo con Harari (2017), un algoritmo es: "un conjunto metódico de pasos que pueden emplearse para hacer cálculos, resolver problemas, y alcanzar decisiones. Un algoritmo no es un cálculo concreto, sino el método que se sigue cuando se hace el cálculo."

[4] Se debe aclarar que el GT29 se ha ocupado de cuestiones relacionadas con la protección de la privacidad y los datos personales. Creado por la Directiva 95/46/CE del Parlamento Europeo y del Consejo, de 24 de octubre de 1995. Sin embargo, fue sustituido por el Comité Europeo de Protección de Datos, regulado por los artículos 68 a 70 del RGPD. Ahora bien, desde el inicio de su actividad, el Comité ha refrendado los dictámenes emitidos por el GT29, aunque se está también a la espera de nuevos documentos del Comité Europeo de Protección de Datos.

La utilización de algoritmos es fundamental para poder hacer búsquedas rápidas e identificar datos relevantes asociados con el patrón. Es decir, con la herramienta de *big data* se puede correlacionar diversas fuentes de información, implicando en cada fase nuevos hallazgos y descubrimientos. Esta nueva información se puede volver a introducir para realizar búsquedas más exactas, mejorando el funcionamiento del algoritmo y así el sistema aprende, lo que se conoce como *machine learning* (Gil, 2016).

De esta manera, Amazon y Netflix nos hacen recomendaciones basadas en selecciones previas. Poco a poco el algoritmo cada vez tomará nuestras decisiones cotidianas. Por ejemplo, de acuerdo con Yuval Noah Harari este es sólo el principio. Si bien es cierto, dispositivos como Kindle pueden supervisar qué partes del libro leen deprisa y cuáles despacio, en qué página hicimos pausas, y en qué frase abandonamos el libro y no volvimos a abrirlo. Si Kindle mejora en reconocimiento facial y sensores biométricos, podrá saber cómo influyó cada frase que leímos en nuestro ritmo cardíaco. Pronto los libros terminarán leyéndonos mientras lo leemos, todo gracias al *big data* y bajo el lema de mejorar la calidad de servicios[5] (Harari, 2017).

Tal es, por lo demás como bien lo expresa Gil (2016), gracias a la tecnología del *big data* nuestra geolocalización ha sido datificada, del mismo modo nuestras palabras, amistades y gustos son transformados en datos constantemente.

1.1. Las cuatro V del *big data*

*A*l *big data* frecuentemente se le caracteriza mediante cuatro "V"[6] que se explican a continuación:

Volumen: el *big data* permite tratar grandes cantidades de datos y metadatos, se pueden manejar magnitudes de megabytes, gigabytes, y terabytes (AEPD, 2017).

[5] De hecho, una de las preocupaciones del autor respecto a estos algoritmos, es que puedan pasar de ser oráculos a soberanos. Señala, en ese sentido, que la riqueza podría quedar concentrada en manos de los algoritmos, de una clase algorítmica. "Si ya existen entidades intersubjetivas que pueden ser titulares de propiedad (como empresas y naciones), ¿por qué no podrían serlo los algoritmos?" (Harari, 2017).

[6] Ahora bien, usualmente se le suele agregar otra V de veracidad. Hace referencia a la fiabilidad o calidad de los datos. Lo anterior es relevante cuando se tratan datos no estructurados. Mediante la depuración y limpieza de datos, el *big data* puede reducir dicha incertidumbre y convertir la información más fiable. (NIST, 2015)

Variedad: se puede mencionar que muy de la mano con el volumen, pues de acuerdo con éste y con el desarrollo de la tecnología, existen muchas formas de representar los datos; es el caso de datos estructurados y no estructurados**.**

Con la plétora de sensores, dispositivos inteligentes, tecnologías de colaboración social, los datos que se generan se presentan en distintas formas, textos, datos webs, tuits audio, secuencias de clic y archivos de registros (Puyol, 2014).

Velocidad: la velocidad de transferencia de datos que llegan de manera constante que se integran con los datos existentes. El flujo de datos es masivo y continuo, por eso se requiere de sistemas que trabajen a toda máquina y sean capaces de procesar rápidamente los datos requeridos (NIST, 2015).

Valor: los análisis del *big data* pretenden dar valor a la organización, es decir, oportunidades económicas, reducción de costes o innovación (Gil, 2016).

1.2. Beneficios y riesgos del *big data*

1.2.1. ¿Para qué sirve el *big data*? Primacía del algoritmo.

El *big data* conlleva unos beneficios económicos y sociales, lo cierto es que son claras las oportunidades que genera el *big data* en nuestra sociedad digital. Precisamente, *The Economist* publicó el 26 de mayo de 2011 un artículo en donde se resalta que la revolución de los datos está cambiando la forma de hacer negocios, "las personas habían almacenado datos suficientes para rellenar 60.000 bibliotecas del Congreso de los Estados Unidos" (Schumpeter, 2011).

Una de las ventajas del uso de *big data* radica en el marketing digital y la generación de perfiles de consumidores. Al acumular los perfiles de datos de los clientes, estos sirven para tomar decisiones más eficientes. Por vía de ilustración, la empresa Tesco, una empresa de ventas al por menor, reúne 1500 millones de paquetes de datos de clientes cada mes, y los utiliza para ajustar precios y promociones. Esto combinado con la revolución de los móviles y la geolocalización, ha generado líneas de negocios como Starbucks, que ofrece descuentos a los clientes que se encuentran cerca, mejorando la prestación de servicios sin precedentes (Aguilar, 2013).

En efecto, el *big data* facilita la toma de decisiones en tiempo real, en particular para la gestión de marketing, estudios comparativos de precios o segmentación de mercado, permitiendo un mayor conocimiento de clientes y personalización del usuario (Aguilar, 2013).

Otro sector altamente beneficiado, es el sector público con el fin de proveer servicios eficientes y mejorar la calidad de vida de los ciudadanos. Por ejemplo, el transporte de Londres recoge los datos de 31 millones de viajes cada día incluyendo 20 millones de uso del sistema de billetes, información de ubicación y predicción para los 9,200 autobuses e información de flujo de tráfico. El *big data* ayuda para que esos datos revelen los patrones de viaje y crear beneficios para los viajeros, planificando los cierres y la construcción de nuevas entradas y salidas a fin de aumentar la capacidad en las estaciones del metro (ICO, 2017).

De este mismo modo, la tecnología del *big data* hace posible la consolidación de las *Smart Cities*, pero a una mayor escala, gracias a la gestión de toda la información de análisis que ofrece el *big data*, se permite obtener predicciones y recomendaciones para los ciudadanos. Esto facilita la optimización y reducción de costes de servicios públicos como el alumbrado, la recogida de basura, la predicción de fenómenos naturales o el control de los niveles de contaminación. Por vía de ilustración, el ayuntamiento de la ciudad de Nueva york ha utilizado esta tecnología con el fin de controlar el tráfico en la ciudad como una solución estable y una mejora continua a la movilidad de grandes ciudades (Gil, 2016).

Paralelamente es indudable el valor del uso del *big data* en el sector sanitario, el cual aporta beneficios significativos para reducir el tiempo de ingreso hospitalario o predecir futuras enfermedades y riesgos sanitarios (AEPD, 2017).

Así es como, la digitalización de los historiales médicos ya ha permitido a los médicos y a los pacientes tomar decisiones más rápidas, con el *big data* se puede ayudar a descubrir relaciones entre genes, las enfermedades y las respuestas a los tratamientos que revolucionará la medicina personalizada. Por ejemplo, un estudio holandés de 2015 demostró que el diagnóstico computarizado de cáncer de próstata utilizando imágenes de resonancia magnética era tan bueno como el de los radiólogos humanos (Tegmark, 2018).

Dentro de este contexto, la tecnología de *big data* se ha utilizado para analizar los datos provenientes de sensores para supervisar la salud y las actividades de las personas. Por

ejemplo, Google se alió con la gigantesca compañía farmacéutica Novartis que desarrolla un lente de contacto que comprueba cada poco segundo los niveles de glucosa en sangre analizando el contenido de las lágrimas. *Pixie Scientific* vende pañales inteligentes que analizan la caca del bebé en busca de indicios de la salud del niño (Scott, 2014).

De hecho, la herramienta de *big data* puede resultar muy útil para hacer seguimiento de enfermedades infecciosas. En el 2008, se puso en marcha el *Google Flu trends*, que rastreaba los brotes de gripe mediante el seguimiento de búsquedas de Google y podía alertar el comienzo de las pandemias o epidemias diez días antes que los sistemas tradicionales (Butler, 2013).

En concreto, el *big data* supone la confluencia de tendencias que han venido evolucionado a lo largo de esta última década: redes sociales, aplicaciones, internet de las cosas, y *cloud computing,* creando así nuevas oportunidades y modificando los modelos de negocios, la economía y la sociedad en sí misma.

En resumen, todo lo que nos fascina de la civilización es producto de la inteligencia humana, por lo que si podemos amplificar el conocimiento mediante el uso de algoritmos y *big data* se abrirá ante nosotros una puerta con considerables mejoras en ciencia, tecnología, enfermedades y bienestar común. Sin embargo, no se debe olvidar que las grandes oportunidades también vienen acompañadas de riesgos, como se explicará a continuación.

1.2.2. ¿Hacia dictaduras digitales?

En los últimos veinte años hemos experimentado más cambios tecnológicos que en los últimos dos milenios. Así es como, el ingeniero de Google, Ray Kurzweil, afirma que el mundo cambiará tal y como lo conocemos por el avance tecnológico exponencial (Duarte, 2020).

Por otra parte, el historiador Harari (2016) observa igualmente un crecimiento acelerado en la sociedad. Pasaron 600 años entre el descubrimiento de la pólvora y el momento en que fue utilizado para cañones, 40 años entre el momento en que Einstein determinó su ecuación E=MC2 y el suceso de Hiroshima, y 27 años entre el nacimiento del WWW (*world wide web*) y la consolidación del *big data* en nuestra sociedad digital (Monleón-Getino, 2015).

En tiempos recientes el mundo ha sido beneficiado con los avances extraordinarios de la ciencia y tecnología, hemos vencido la mayor parte de enfermedades infecciosas, mortalidad infantil, pobreza, hambre, en palabras del autor Pinker (2019): "el homo sapiens, el hombre sabio es la especie que utiliza la información para resistir la putrefacción de la entropía y el peso de la evolución."

En este orden de ideas, la tecnología multiplica de manera exponencial el conocimiento, y el *big data* es una de las herramientas principales para ello. No obstante, no hay que olvidar que conlleva una estela de riesgos de lo que debemos ser conscientes.

Como bien expone la AEPD (2017), la imposición del algoritmo en generación de perfiles de consumidores y sus predicciones podrían ser utilizados de forma discriminatoria a los interesados, incluso excluyendo a sectores minoritarios[7]. De manera que, si el proceso que genera los datos refleja sesgos a favor o en contra, los resultados generados podrían perpetuar dichos sesgos indefinidamente.

Confiar ciegamente en los algoritmos, y no entender las razones por las que se han tomado las decisiones, es una de las grandes cuestiones sobre el *big data* (Gil, 2016). En este sentido, cuando no se reconoce los sesgos, la mala calidad de los datos puede conducir a predicciones inexactas, lo que a su vez puede conducir a que las empresas nieguen erróneamente las ofertas o beneficios de los consumidores (FTC, 2016). Análogamente, cabría enfatizar que resulta cuestionable la solución de imponer intervención humana sobre dichos algoritmos, dado que, como bien lo delinea Lohr (2015): "el impulso de querer que una persona supervise los resultados que vomita el ordenador es muy humano."

De este ángulo, la *Federal Trade Commission* encontró evidencia de análisis masivo de datos que pueden conducir a tomar decisiones basadas en el comportamiento de otros consumidores que comparten algunas características con el sujeto de estudio. En particular, se evidenció que algunas entidades de crédito habían reducido el límite de crédito de un

[7] El análisis de *Big data* puede brindar a las empresas nuevas formas de intentar justificar su exclusión de ciertas poblaciones de oportunidades particulares. Por ejemplo, un estudio de análisis de *big data* mostró que los candidatos que completaban solicitudes de trabajo en línea utilizando navegadores que no venían con la computadora, pero tuvieron que instalarse deliberadamente (como Firefox o Google Chrome) funcionan mejor y cambian de trabajo con menos frecuencia. Si un empleador usara esta correlación para abstenerse de contratar a personas que usaron un navegador en particular, podrían estar excluyendo a solicitantes calificados por razones no relacionado con el trabajo en cuestión (FTC, 2016).

cliente, no en función de la historia crediticia del cliente, sino basado en el análisis de otros clientes que habían comprado en los mismos establecimientos (FTC, 2016).

La predictibilidad preventiva del algoritmo, como lo denomina (Martínez, 2014) puede llegar a ser riesgosa, dado que, existe una transferencia de la autoridad de los humanos a los algoritmos. Si el lector vio el episodio "Descolgar el DJ" de *Black Mirror*, sabe que es peligroso que el algoritmo buceé en preferencias y características para encontrar la pareja perfecta, o como bien recuerda *Minority Report*, un sistema de derecho penal preventivo que categoriza a las personas y castiga aquellas que no han cometido un delito aún, puede ser sumamente perjudicial.

Precisa advertir que más allá de los escenarios hiperbólicos, el *big data* ofrece una herramienta crucial para predictibilidad en cualquier campo y para la toma de decisiones, pero una consecuencia nefasta será la tiranía del algoritmo, encaminándonos a una pesadilla Orweliana (Martínez, 2017). Como bien lo expone Harari (2017): "el algoritmo se convertirá en soberano. Al tener tanto poder en sus manos, y a saber mucho más que nosotros, podría empezar a moldearnos nuestros deseos y tomar decisiones por nosotros[8]." Desde luego, esta situación nos podría conducir a una dictadura de errores masivos, o peor aún, de estupidez artificial (Cotino, 2017).

Un segundo riesgo, que expone el Grupo de Trabajo del Artículo 29, es la falta de transparencia en la tecnología de *big data*: regularmente los interesados no entienden cómo se van a utilizar sus datos personales y por lo tanto no pueden ejercer control sobre ellos (GT29, 2013).

Adicionalmente el sistema al volverse más sofisticado y complejo, aumenta la dificultad para entender cómo funciona el algoritmo y sus decisiones. A esto se le suman, que muchas ocasiones hay un cambio de finalidad, dado que, el algoritmo arroja un hallazgo inesperado.

Lo anterior son cuestiones difíciles de solventar y armonizar entre privacidad e innovación, cierto es que el Reglamento General de Protección de Datos ha establecido unas previsiones

[8] Esto se relaciona con el problema del *big data* denominado por (Richards Neil y Jonathan King, 2013) la paradoja de la identidad. El cual se ilustra detalladamente con la siguiente cita textual: "El poder de *big data* es, por lo tanto, el poder de utilizar la información para empujar, persuadir, influir e incluso restringir nuestras identidades. Tal influencia sobre nuestras identidades individuales y colectivas corre el riesgo de erosionar el vigor y la calidad de nuestra democracia. Si no tenemos el poder de decir individualmente quién soy, si los filtros y los empujones y las recomendaciones personalizadas socavan nuestras elecciones intelectuales, nos habremos identificado, pero perderemos nuestras identidades como las hemos definido y apreciado en el pasado".

y parámetros para contrarrestar estas dificultades, las cuales serán analizadas a lo largo de este trabajo.

En consecuencia, resulta evidente cómo estos escenarios atraviesan los principios fundamentales de protección de datos. El resto del trabajo se centrará en dilucidar dichos conflictos desde el punto de vista jurídico.

2. Derecho de protección de datos en una cáscara de nuez

El derecho de protección de datos, tal y como se reconoce hoy en día, ha tenido una evolución constante, adaptándose a las nuevas realidades sociales. En efecto, la historia del derecho de protección de datos comienza en el año 1890.

Warren y Brandeis[9] marcaron un hito en la doctrina norteamericana cuando escribieron un artículo titulado *The right to privacy* invocando su alcance y protección, por las preocupaciones que tenían respecto a las publicaciones en los periódicos de su vida privada. Por lo tanto, dicho derecho se ha entendido por el derecho anglosajón como el *right to be left alone*, a no ser molestado y estar solo. Fue concebido como un derecho amplio, fundamental y limitado. Cimentado en el principio de la inviolabilidad de la persona, a partir de las exposiciones iniciales que había realizado el Juez Cooley, quien identificó la base de este derecho relacionado con la propiedad (Frosini, 1988).

Dicho artículo académico tuvo una repercusión en la doctrina norteamericana. Por todo, no es sorprendente que el *Right to privacy*, fuese objeto de revisión por el jurista William Prosser y categorizará acciones civiles para su protección denominado agravio de privacidad. Posteriormente, la concepción formulada por los dos jóvenes juristas influyó en el sistema constitucional norteamericano. Así, aunque ni las enmiendas ni la constitución de 1787 mencionaban expresamente el derecho, el Tribunal Supremo a lo largo de una amplia trayectoria, lo ha considerado implícito a partir de la Cuarta Enmienda, frente a registros y requisas arbitrarias que limita la intrusión del gobierno en las personas, domicilios, documentos y efectos personales [10] (Saldaña, 2012).

[9] Comenta Frosini (1988) que: "en el siglo XIX Warren y Brandeis, dos jóvenes abogados de Boston inventaron esta expresión. Warren, después de haberse casado con la hija de un senador, había principado a llevar una vida de lujo y rumbosa. Este hecho atrajo la curiosidad y chismografía de los periódicos en sus crónicas amarillas, hasta el punto de suscitar escándalo. Warren, irritado por esta invasión en su vida privada, se asoció con su antiguo compañero de estudios de la Universidad de Harvard, Louis Brandeis, y juntos escribieron un ensayo titulado *The right to privacy*, que fue publicado en la *Harvard Law Review.* Según estos dos autores, todo individuo tiene derecho a ser dejado en paz, a proteger a su soledad, a su vida íntima, del mismo modo que tiene derecho a proteger su vida privada.

[10] Ha sido la labor del Tribunal Supremo de Justicia en afirmar que la protección de la privacidad emana de su articulado, principalmente de la Cuarta Enmienda, como se estipula en la sentencia de *Griswold vs. Connecticut.*

Con el paso del tiempo, anudado con la convergencia del desarrollo de nuevas tecnologías de información, el derecho de *right to privacy* tomó una nueva dimensión, como el poder de controlar el flujo de información personal. En esta línea, el Congreso norteamericano aprobó la Ley de Privacidad de 1974, la primera ley general de protección de la información personal en poder de las Agencias Federales de los Estados Unidos (Saldaña, 2011).

Ahora bien, al otro lado del Atlántico, el continente europeo influenciado por los acontecimientos de Estados Unidos empezó a preocuparse por regular el tratamiento de datos. La primera generación de leyes de protección de datos, ocurre en el *Land* alemán de Hesse donde se promulga la primera norma vinculante el 7 de octubre de 1970. Dicha iniciativa normativa fue seguida por el parlamento sueco, al establecer una ley dirigida a la protección de bases de datos tanto públicas como privadas, a diferencia de la Ley de Privacidad estadunidense de 1974, que sólo regulaba la protección de bases de datos públicas (Quesada, 2015).

En cuanto a España, la constitución de 1978 prevé implícitamente el derecho de protección de datos en su artículo 18.4 CE, donde se establece que el legislador limitará el uso de la informática para proteger los derechos fundamentales de los ciudadanos.

Posteriormente el Tribunal Constitucional en su jurisprudencia (STC 292/2000)[11] establece que el derecho a la protección de datos debe ser entendido como un derecho fundamental

[11] En la sentencia analizada el TC establece con claridad las notas diferenciales del derecho fundamental a la protección de datos respecto del derecho fundamental a la intimidad, poniendo de manifiesto que la diferencia principal radica en su función diferencial, lo que a su vez implica que tengan objeto y contenido diferentes. En cuanto a su función y finalidad: Mientras el derecho fundamental a la protección de datos persigue garantizar el control y el poder de disposición sobre el uso y destino de los datos personales propios, el derecho fundamental a la intimidad persigue proteger frente a cualquier invasión de la vida personal y familiar que la persona desee excluir del conocimiento ajeno y de las intromisiones de terceros; es decir, resguardar la vida privada de una publicidad no querida. En cuanto al objeto: el derecho a la protección de datos tiene una singularidad propia, con un objeto más amplio que el derecho a la intimidad, alcanzando la protección de la dignidad personal y el derecho al honor; es decir, está vinculado al pleno ejercicio de los derechos de la persona. De este modo, como ya se ha anticipado, el objeto de protección del derecho fundamental a la protección de datos no se limita a los datos privados o íntimos de la persona, sino a cualquier tipo de dato personal, sea o no íntimo, cuyo conocimiento por terceros pueda afectar a su titular en modo no deseado. Es decir, el objeto del derecho a la protección de datos no es la protección de la intimidad individual, sino la protección de los datos personales, sean o no íntimos en tanto se trate de datos que permitan la identificación del individuo. En cuanto al contenido: el derecho a la intimidad implica para los terceros el deber de abstención de toda intromisión en la esfera íntima ajena, mientras que el derecho a la protección de datos atribuye a su titular un haz de facultades y, al hacerlo, impone a los terceros una serie de obligaciones de hacer (que no deberes de mera abstención); obligaciones sin las cuales se

autónomo, cuyo contenido está integrado por los principios y derechos que se contemplaban en la antigua normativa. En virtud de este derecho fundamental, el ciudadano, goza del derecho a la libertad informática frente a potenciales agresiones a la dignidad y a la libertad proveniente de un uso ilegítimo del tratamiento mecanizado de datos. En palabras del Tribunal Constitucional (STC 292/2000): "persigue garantizar a esa persona un poder de control sobre sus datos personales, sobre su uso y destino, con el propósito de impedir su tráfico ilícito y lesivo para la dignidad y derecho del afectado."

De este modo, el derecho de protección de datos se consolida como un derecho autónomo y confiere facultades y herramientas necesarias para que el individuo conozca la información, su finalidad, y su tratamiento. Desde el punto de vista legal, la protección de datos personales ha sido definida como: "la protección jurídica de las personas en lo que concierna al tratamiento de sus datos personales, o de otro forma, el amparo debido a los ciudadanos contra la posible utilización por terceros, en forma no autorizada, de sus datos personales susceptibles al tratamiento, para confeccionar una información que, identificable con él, afecta su entorno personal, social, o profesional en los límites de su intimidad, incide directamente en un derecho fundamental de elevado contenido" (Ortiz, 2005).

A continuación, se explicará en los siguientes apartados qué son los datos personales y cuál es el marco jurídico de regulación aplicable para el tratamiento de dichos datos.

2.1.¿Qué son los datos personales?

Los datos personales se encuentran definidos en el artículo 4 del RGPD:

> *Datos personales: toda información sobre una persona física identificada o identificable («el interesado»); se considerará persona física identificable toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular*

vacía de contenido el poder de disposición y control sobre los datos personales propios. Este haz de facultades, en palabras del TC, queda conformado por: (i) el derecho a que se requiera el previo consentimiento para la recogida y uso de los datos personales, (ii) el derecho a saber y ser informado sobre el destino y uso de esos datos; y, (iii) el derecho a acceder, rectificar y cancelar dichos datos.

> *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* (art. 4 RGPD).

En concreto, los datos personales no se circunscriben únicamente a nombres y apellidos, sino que contiene una gama amplia de datos que incluye la voz, datos de navegación, geolocalización, preferencias o *likes* en redes sociales, nuestra forma de andar, etc. (Gil, 2016). De igual modo, para determinar si una persona es identificable, se debe utilizar los medios razonables y sin esfuerzos desproporcionados. Por lo contrario, cuando no sea posible dicha identificación, o requiera tales esfuerzos desproporcionados no se aplicará la normativa de protección de datos.

Lo anterior, es crucial, dado que, en los análisis de *big data* es común la utilización de datos anonimizados. En otras palabras, no se le aplicaría la normativa de protección de datos, dado que, la anonimización supone que no sería posible identificar a la persona con los datos, teniendo en cuenta todos los medios que puedan ser razonablemente utilizados como lo dispone el Grupo de trabajo del artículo 29 (2007) en su Dictamen sobre el concepto de datos personales.

Justo es decir que, cuando los datos no hacen identificable a una persona, no se les aplica la regulación relativa a datos personales. De esta manera, si se tratan los datos anónimos, se convierten en datos no personales, y su privacidad por lo tanto se encuentra protegida siendo irrelevante a la aplicación normativa. No obstante, el *big data* puede desafiar esta cuestión, dado que, muchas veces facilita la re-identificación de los sujetos[12], al tratar grandes volúmenes de datos y permitir la correlación entre distintas fuentes, lo que se considera un riesgo potencial jurídico, el cual será analizado más adelante en este trabajo (Gil, 2016).

Una vez comprendida la verdadera extensión de aquellos datos considerados personales, se debe realizar una breve mención sobre el marco jurídico aplicable.

[12] Por ejemplo, el caso de Netflix en el 2007 identificó a los usuarios cuando se comparó con un conjunto de datos proveniente de las fuentes de calificaciones de películas en internet (ENISA, 2015).

2.2. Marco jurídico de protección de datos

En el ámbito europeo, la protección de datos se encuentra contemplada como derecho fundamental en el artículo 16 del Tratado de Funcionamiento de la Unión Europea y en el artículo 8 de la Carta de los Derechos Fundamentales de la Unión Europea. [13]

Asimismo, el Consejo de Europa adoptó en 1981 el Convenio del Consejo de Europa para la protección de las personas, conocido habitualmente como el Convenio 108, abierto a firma el 28 de enero de 1981 en Estrasburgo, es el único instrumento internacional jurídicamente vinculante de protección de datos. (Quesada, 2015). No obstante, actualmente existe una nueva versión, Convenio 108 plus, que abrió firma el 10 de octubre del 2018, con el fin de hacer frente a los retos originados por las nuevas tecnologías. Dicho Convenio refuerza los principios de proporcionalidad, minimización y legalidad del tratamiento (Esteban, 2018).

Conscientes a la transición de un mercado único digital junto con la evolución tecnológica, la aparición de los servicios de la sociedad de información y el Tratado de Lisboa que otorgó fuerza vinculante a la CFDUE, fue necesario un cambio legislativo de la antigua Directiva 95/46/CE del Parlamento Europeo y del Consejo de 24 de octubre de 1995, por el nuevo Reglamento Europeo 2016/679. El cual introdujo nuevos cambios y derechos respondiendo a dichas necesidades sociales informáticas.

Este reglamento por su naturaleza jurídica (de aplicación directa y obligatorio cumplimiento), pone fin a las disparidades que aparecieron como resultado de la transposición de la Directiva a las leyes nacionales de los Estados miembros, e instaura en todos ellos el mismo alto nivel de protección de los datos personales.

En este sentido, el Reglamento, a diferencia de la Directiva es, como ya se describió, de ámbito obligatorio en cada uno de sus aspectos y aplicable de un modo directo en todos los Estados miembros. No necesitando de transposición alguna, pero sí del desarrollo en determinados

[13] Artículo 16 del Tratado de funcionamiento de la Unión Europea: toda persona tiene derecho a la protección de los datos de carácter personal que le conciernan.
Artículo 8 de la Carta de Derechos Fundamentales de la Unión Europea: Toda persona tiene derecho a la protección de los datos de carácter personal que la conciernan.
2. Estos datos se tratará de modo leal, para fines concretos y sobre la base del consentimiento de la persona afectada o en virtud de otro fundamento legítimo previsto por la ley. Toda persona tiene derecho a acceder a los datos recogidos que la conciernan y a su rectificación.
3. El respeto de estas normas quedará sujeto al control de una autoridad independiente.

elementos que se adoptan por una norma interna que, en nuestro caso, ha venido determinada por la Ley Orgánica 3 del 2018, la cual tiene como objeto: adaptar el ordenamiento jurídico español al Reglamento y garantizar los derechos digitales de la ciudadanía confirme al mandato constitucional (art. 1 LOPDGDD).

De igual modo, el RGPD refuerza el papel de los interesados y el control sobre sus datos. En esta misma línea, se establecen nuevos elementos de la protección de datos, como el principio de transparencia, la creación del principio de responsabilidad proactiva, el principio de privacidad desde el diseño y por defecto, y la realización de evaluación de impacto cuando existe un alto riesgo para los interesados (AEPD, 2018). Precisamente estas figuras tienen una especial importancia para la tecnología del *big data*, que imponen la protección de los datos desde el diseño y a lo largo del ciclo de vida del tratamiento de los datos personales.

En consecuencia, el resultado no sólo es la afirmación de nuevos derechos para los individuos (derecho de portabilidad, limitación del tratamiento y derecho de supresión o derecho de olvido). Sino también reforzar el deber de información del Responsable del tratamiento, con el fin de que el interesado dé su consentimiento libre e informado, el cual debe ser expreso (*opt-in*), a diferencia del consentimiento implícito (*opt-out*), que se venía regulando en la normativa anterior (AEPD, 2018).

Por otro lado, el Reglamento protege el desarrollo económico y la libre circulación de los datos, como así lo afirma su Considerando 13 RGPD: "el buen funcionamiento del mercado interior exige que la libre circulación de los datos personales en la Unión no sea restringida ni prohibida por motivos relacionados con la protección de las personas físicas en lo que respecta el tratamiento de datos personales." Por lo tanto, el objetivo de la normativa es armonizar los derechos de los interesados y el desarrollo económico, permitiendo el avance tecnológico (AEPD, 2017).

En consecuencia, en los siguientes capítulos, se analizará los fundamentos relevantes de la normativa para el tratamiento de los datos mediante el *big data,* como bases legitimadoras y elaboración de perfiles. Desde luego, conviene preguntarnos si la normativa aplicable es suficiente para afrontar el advenimiento de una tecnología sin precedentes como el *big data*.

3. *Big data* a través del Reglamento General de Protección de Datos

3.1. Deber de información y bases legitimadoras

El Reglamento General de Protección de Datos introduce un cambio esencial, en donde prima el deber de transparencia e información, es decir, al titular de los datos se le otorga un poder de disposición mayor sobre sus datos personales. En este sentido, el interesado tiene derecho a saber sobre los fines del tratamiento, quién es el Responsable del tratamiento, y sus derechos de conformidad con los artículos 12, 13, y 14 del RGPD.

No obstante, la complejidad de la tecnología del *big data* puede traer problemas en el terreno de la transparencia. Los interesados podrían no comprender bien para qué fines se están utilizando los datos o cómo se están procesando los datos, debido a la funcionalidad del algoritmo al ser opaco y complejo (ICO, 2017).

En este orden de ideas, el deber de transparencia plantea muchos problemas, entre ellos, primero porque los escasos interesados que logran leer las políticas de privacidad no las comprenden, y segundo porque si las políticas de privacidad se redactan en un lenguaje sencillo se corre el riesgo de que no se pueda elaborar un entendimiento total del tratamiento: lo anterior se denomina la paradoja de la transparencia (Nissebaum, 2014).

Así la AEPD (2018), ha propuesto una solución que mitiga la conspicua paradoja: la información basada en capas. El enfoque de información multinivel ofrece una mayor compatibilidad entre la exigencia de información detallada y concisa. Por lo tanto, al presentarse una primera capa[14] de información básica de forma resumida y la remisión a una información adicional en un segundo nivel, se garantiza una mayor comprensión, y a la vez se facilita la lectura al interesado.

[14] En la primera capa se recomienda introducir: el responsable del tratamiento, la finalidad, legitimación, destinatarios, derechos y procedencias (cuando los datos no procedan del propio interesado) (AEPD,2018).

Ahora bien, el responsable o encargado del tratamiento también deberá legitimar el tratamiento con fundamento a las bases legales del artículo 6 del RGPD cuando procesan datos con la tecnología de *big data.*

La primera base legitimadora que se encuentra en dicho articulado es el consentimiento. El consentimiento debe cumplir con las características del (considerando 32 del RGPD), es decir, una manifestación de voluntad libre, específica, informada e inequívoca del interesado. No obstante, esta base legitimadora no es práctica en el contexto del *big dat*a, por la naturaleza opaca de las técnicas algorítmicas (ICO, 2017), lo cual será analizado con profundidad en el capítulo siguiente.

Adicionalmente, el requerimiento del consentimiento es incompatible con los usos secundarios cuando se utiliza el *big data*. Esta forma de concebir el consentimiento obligaría a que cada vez que se descubra un nuevo uso para los datos, el responsable debería volver a pedir el consentimiento a cada uno de los individuos cuyos datos estén siendo tratados por segunda vez. Esto, en muchas ocasiones, puede ser técnicamente inviable, por no decir que las empresas no podrían asumir los costes. (Gil, 2016). Y es que con el sistema de decisiones automatizadas es muy difícil otorgar el consentimiento para fines que ni siquiera el responsable del tratamiento prevé desde sus inicios[15] (Cotino, 2017).

Por otro lado, se contempla la base legal del interés legítimo, la cual aplica cuando es necesario para la satisfacción de intereses legítimos perseguidos por el responsable siempre que dichos intereses no prevalezcan sobre los derechos del interesado. De esta manera, una organización puede utilizar el interés legítimo para prevención de fraude o marketing directo, sin embargo, las autoridades de control recomiendan realizar un análisis de proporcionalidad y necesidad (AEPD, 2017). De este modo, se sopesa la naturaleza y fuente del interés legítimo, y, por otra parte, las repercusiones para los derechos de los interesados. Así, tras analizar los dos aspectos en la balanza, se puede llegar a una conclusión preliminar de prevalencia sobre los

[15] No obstante, La autoridad de control del Reino Unido recomienda un nuevo acercamiento para el consentimiento que supere el sistema binario del sí o no. Es decir, presentar de manera gradual el consentimiento a medida que la relación con el proveedor de servicios evoluciona, esto puede realizarse con notificaciones o alertas. Por ejemplo, si la aplicación necesita analizar los datos de geolocalización para nuevos fines no incompatibles, recabar el consentimiento (ICO, 2017).

derechos de los interesados o el refuerzo de estos mediante garantías adicionales,[16] como lo expone el GT29 (2014) en su Dictamen sobre el concepto de interés legítimo.

En efecto, el interés legítimo puede ser una condición alternativa a recabar el consentimiento de los interesados, siempre y cuando se informe sobre dicho tratamiento. Lo cierto es que, la condición de intereses legítimos no es una opción flexible para la organización, en cambio impone la responsabilidad de llevar a cabo una evaluación de necesidad, respetando los derechos e intereses de las personas (ICO, 2017).

La siguiente base legitimadora que contempla el artículo 6 del RGPD consiste en el cumplimiento de una relación contractual, dicha base legal puede resultar controversial en el ámbito de *big data*. En vista que, considerando la naturaleza del big data, y su constante reutilización de datos personales para otros fines, resulta extremadamente difícil argumentar la realización de un análisis de *big data* como rigurosamente necesario para ejecutar un contrato (ICO, 2017).

No obstante, respecto a este punto se puede encontrar un caso excepcional en donde las elaboraciones de perfiles mediante tecnología de *big data* se encuentre legitimado en el marco de un contrato, siendo estrictamente esencial. Este es el caso del contrato de seguros, donde la elaboración de perfiles con fines estadístico-actuariales son necesarios para la determinación del riesgo y de la prima del contrato de seguro. En efecto, se constituye como base jurídica el cumplimiento de una obligación legal (art. 99.7 LOSSEAR) y ejecución de un contrato (art. 6.1 literal b del RGPD). Dado que, en el ámbito del sector asegurador, el tratamiento automatizado de los datos personales de los clientes y potenciales clientes es un procedimiento inherente y absolutamente imprescindible para el desenvolvimiento de la actividad de cualquier compañía aseguradora.

Es preciso aclarar, que el Reglamento permite el uso de perfiles cuando existe consentimiento, pero también cuando sean necesario para la celebración o la ejecución de un contrato entre el interesado y un responsable de tratamiento como ocurre cuando se pretende celebrar un contrato, además de permitirse cuando esté autorizado por una norma Derecho de la UE o

[16]Se deben seguir los siguientes pasos: 1). Considerar la naturaleza de los datos personales (datos de categorías especiales, infracciones penales, etc.) 2). Considerar las expectativas razonables de los interesados en el uso de los datos personales 3). Considerar el potencial impacto de los individuos y los daños derivados del tratamiento (GT29, 2014).

nacional, de acuerdo con el artículo 22 literal 2 del RGPD, como sucede en el caso de contrato de seguro.

Por último, queda por mencionar el uso del *big data* por parte de las administraciones públicas, legitimado por una obligación legal o en el marco de interés público. En este sentido, el uso de *big data* en el ámbito de la sanidad pública puede ofrecer grandes oportunidades y reducir costes en el tratamiento de la información clínica, o el seguimiento de epidemias o pandemias y, especialmente, en el ámbito de la investigación genómica, en la que se genera y se trabaja con grandes cantidades de información y datos (Parra, 2017).

Así es como, la AEPD en tiempo de COVID-19, ha establecido que el uso de tecnologías como la geolocalización de los móviles a partir de redes sociales o por lo operadores de telecomunicaciones que resultan eficaces para conocer los patrones de movilidad de la población, y evaluar el progreso de contagio entre personas (AEPD, 2020).

Tras haberse expuesto la paradoja de la transparencia y las bases legitimadoras, se ha analizado cómo éstas inciden en el ámbito de *big data*. Precisamente, se encuentra una deficiencia en la forma de informar al interesado, la quiebra de la utilidad del consentimiento en el uso del *big data*, y el paso a otras bases legitimadoras que se ajustan más a las necesidades de esta tecnología en particular.

En las próximas páginas, se adentrará el lector a un análisis sobre los riesgos cruciales que acarrea el *big data*, particularmente: los problemas de preservar la privacidad con las técnicas de anonimización y el sesgo discriminatorio del algoritmo en el marco de decisiones automatizadas.

3.2.Anonimización

La anonimización es una técnica que se aplica a los datos personales eliminado todos los elementos suficientes para que no se pueda identificar a los interesados (GT29, 2014).

En este sentido, si a los datos personales se le aplican técnicas de anonimización, significa que no es posible identificar a una persona, y es irrelevante la aplicación de las disposiciones jurídicas relativas a protección de datos como bien lo establece el Reglamento General de Protección de datos:

> *(…) Por lo tanto los principios de protección de datos no deben aplicarse a la información anónima, es decir información que no guarda relación con una persona física identificada o identificable, ni a los datos convertidos en anónimos de forma que el interesado no sea identificable, o deje de serlo. En consecuencia, el presente Reglamento no afecta al tratamiento de dicha información anónima, inclusive con fines estadísticos o de investigación*. (Considerando 26 del RGPD).

Cabe recalcar, que como bien lo propone el GT29 (2014) en su Dictamen sobre técnicas de anonimización: la obtención de datos personales a anonimizar es propiamente un tratamiento posterior, y por lo tanto debe cumplir con las obligaciones impuestas por el RGPD. Especialmente, el tratamiento deberá estar legitimado por las bases legales del artículo 6 del RGPD y se deberá informar debidamente a los interesados de acuerdo con los artículos 13 y 14 del RGPD.

Desde este ángulo, muchas organizaciones optan por este mecanismo para los análisis de datos en entornos de *big data*. Este fue el célebre caso del Instituto Nacional de Estadísticas cuando estudió los datos de movilidad de los españoles a través de sus teléfonos móviles, utilizando datos completamente anónimos, con el fin de conocer cuántos ciudadanos se movían de un municipio a otro, o cómo fluctuaba la población de su residencia al trabajo (Maqueda, 2019).

De cualquier modo, el GT29 reconoce que existe un riesgo de reidentificación residual incluso después de aplicar las técnicas de anonimización, y más aún en el universo del *big data* donde las fuentes son variadas y múltiples permitiendo la correlación con distintas bases de datos, lo cual posibilita de manera exponencial la reidentificación de los sujetos[17](ICO, 2017).

En concreto, el Grupo de Trabajo del artículo 29 propone dos criterios para determinar si una base de datos es anónima: verificar que el fichero no tenga las propiedades de singularización,

[17] Uno de los casos más famosos sobre la reidentificación de sujetos fue el caso de NETFLIX, los datos identificativos habían sido anonimizados, excepto por las calificaciones de las películas y fechas. Sin embargo, se añadió la técnica de ruido a las calificaciones de los usuarios. Estos datos fueron cruzados con la fuente externa pública de *INTERNET MOVIE DATABASE.*

asociación e inferencia[18], y realizar un análisis sobre el riesgo de reidentificación de los datos (GT29, 2014).

En términos generales, existen dos aproximaciones a las técnicas de anonimización: la aleatorización y la generalización. La aleatorización incluye técnicas que alteran la veracidad de los datos. De esta forma, no reduce la singularización pero si se protege contra ataques de inferencias. Las técnicas más utilizadas de aleatorización son la adición de ruido y la permutación (Gil, 2016).

La adición de ruido consiste en cambiar atributos para que sean menos exactos, pero conservando la distribución general. Así, por ejemplo, si una base de datos recoge las alturas de individuos, es decir, 1,80m, la adición de ruido podría consistir en modificar los datos para que se englobaran en rangos de valores de ±10 cms, es decir, de 1,70m a 1,90m (GT29, 2014).

Por su parte, la permutación consiste en intercambiar los atributos de los individuos, de forma que quedarían ligados artificialmente a otros sujetos, por ejemplo, intercambiar los salarios devengados entre los diferentes puestos de trabajo, como lo muestran las siguientes tablas a continuación.

[18] 1). "Singularización: la posibilidad de extraer de un conjunto de datos algunos registros (o todos los registros) que identifican a una persona; 2). Vinculabilidad: la capacidad de vincular como mínimo dos registros de un único interesado o de un grupo de interesados, ya sea en la misma base de datos o en dos bases de datos distintas. Si el atacante puede determinar (p. ej., mediante un análisis de correlación) que dos registros están asignados al mismo grupo de personas, pero no puede singularizar a las personas en este grupo, entonces la técnica es resistente a la singularización, pero no a la vinculabilidad; 3). Inferencia: la posibilidad de deducir con una probabilidad significativa el valor de un atributo a partir de los valores de un conjunto de otros atributos" (GT29, 2014).

Tabla 1. "Tablas" de técnicas de permutación

Individuo	Puesto de trabajo	Día de nacimiento	Tipo de membresía
A	Decano	3 de enero 1970	Plata
B	Vendedor	5 de febrero 1972	Platino
C	Abogado	7 mayo1985	Oro
D	Ingeniero de sistema	10 abril 1990	Plata
E	Enfermero	13 mayo 1995	Plata

Tabla 1. "Tablas" de técnicas de permutación (Elaboración propia)

Tabla 2. "Tablas" de técnicas de permutación/ luego de anonimizar

Individuo	Puesto de trabajo	Día de nacimiento	Tipo de membresía
A	Abogado	10 de abril 1990	Plata
B	Enfermero	7 de marzo 1985	Plata
C	Vendedor	13 de mayo 1995	platino
D	Ingeniero de sistemas	3 enero 1970	Plata
E	Decano	5 de febrero 1972	Oro

Tabla 2. "Tablas" de técnicas de permutación/luego de anonimizar (Elaboración propia)

La segunda familia de técnicas de anonimización es la generalización. Consiste en generalizar o diluir los atributos de los sujetos, modificando su escala (por ejemplo, haciendo referencia a un país en vez de una ciudad; o a datos mensuales en vez de semanales) (Gil, 2016).

Como ya se ha señalado, el riesgo de reidentificación siempre estará latente en el uso de tecnologías del *big data* cuando se analizan datos masivos. Sin embargo, como contempla el ICO (2017) la cuestión no radica en eliminar completamente el riesgo, pero de mitigarlo lo más posible, aplicando distintas técnicas de anonimización dependiendo de cada caso en particular.

En definitiva, las recomendaciones de las autoridades de control de protección de datos se encaminan a realizar un análisis de riesgos del proceso de anonimización para adoptar medidas de seguridad adecuadas de la mano del principio de responsabilidad proactiva, la realización de una evaluación de impacto, y seguimientos de los procesos de anonimización (AEPD, 2016).

Recapitulando, ninguna técnica de anonimización podrá garantizar la imposibilidad de la reidentificación, ya que existirá siempre un índice de probabilidad de reidentificación que se deberá intentar atenuar mediante la correspondiente gestión de riesgos. Por ello, el responsable del tratamiento deberá tener en cuenta el progreso de los riesgos a lo largo del tiempo, la reevaluación periódica, con el fin de impulsar la innovación, la investigación y el desarrollo en la cuantificación y análisis de datos a través de la herramienta de *big data* (AEPD, 2016).

Como expone Gil (2016), la implementación de estas medidas, podrán garantizar los derechos y libertades de los interesados, de forma que se puedan obtener innovación, tecnología y progreso, salvaguardando la privacidad, en últimas es una situación de *winwin* (aquella en la que todos ganan), de acuerdo con la teoría de juegos.

4. ¿*Big data* pone en jaque el derecho de protección de datos? Retos actuales

4.1. Elaboración de perfiles y decisiones automatizadas. ¿Hacia la estupidez artificial y discriminación?

La elaboración de perfiles y el tratamiento automatizado de datos cobra especial importancia en el ámbito de la tecnología de *big data.* En efecto, el RGPD define en su artículo 4 el concepto de elaboración de perfiles: "toda forma de tratamiento automatizado de datos personales consistente en utilizar datos personales para evaluar determinados aspectos personales de una persona física, en particular para analizar o predecir aspectos relativos al rendimiento profesional, situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, ubicación o movimientos de dicha persona física." (art. 4 RGPD).

De esta manera, la elaboración de perfiles permite encasillar a los interesados sobre la base de determinados aspectos con el fin de evaluar comportamientos, preferencias, etc. A su vez, el RGPD regula las decisiones automatizadas, las cuales se aplican una vez se han categorizado los individuos mediante la elaboración de perfiles. Existiendo el riesgo de que se creen estereotipos y cause discriminación hacia los individuos, afectando sus derechos y libertades (Gil, 2016).

De esta manera, el *big data* depende de los datos recogidos en diversas fuentes creadas por la sociedad, y la sociedad en sí misma contiene inequidad, exclusión, y rastros de discriminación (Goodman, B., y Flaxman, S. , 2017). Así, incontables veces el algoritmo replica el perjuicio humano una y otra vez. Precisamente, este fue el caso cuando se automatizaron los datos de descripción de imágenes hecho por humanos, quienes no describieron las imágenes neutralmente. Los bebés de raza blanca fueron descritos como "bebés", pero los bebés de raza negra fueron descritos como "bebés de color" (FRA, 2018).

La situación anterior también ocurrió en una facultad de medicina en 1980, donde se desarrolló un programa para ordenar las solicitudes de admisión de los estudiantes, recopilándose los datos de solicitudes anteriores. El sistema reprodujo dichos criterios de selección, discriminando a las mujeres y a las personas inmigrantes (CdE, 2018). Desde este

ángulo, cabe preguntarse, como bien reflexiona Nissenbaum (2014): ¿pueden los datos representar la verdad objetiva o cualquier interpretación necesariamente está sesgada por algún filtro subjetivo o en la forma en que se limpian los datos?"

Así, el RGPD es consciente de este escenario y evoca la discriminación algorítmica en su considerando 71 RGPD[19], advirtiendo a los responsables y encargados del tratamiento sobre la importancia de aplicar medidas técnicas y organizativas para corregir cualquier error o factor que propicie la discriminación. A su vez, el art. 13 literal f RGPD, contempla la obligación de informar a los interesados sobre la lógica aplicada a las decisiones automatizadas.

Mención especial merece esta disposición, dado que, este derecho de transparencia en la información es una de las primeras herramientas que ofrece la normativa para combatir la discriminación algorítmica. Aunque es cierto que, muchos doctrinantes cuestionan este instrumento, al resultar paradójico con el mandato de legibilidad y sencillez que aboga el Reglamento. En otros términos: ¿cómo debe explicarse dicha técnica? ¿Y qué se requiere para explicar las decisiones de los algoritmos? (FRA, 2018). Después de todo, una de las ventajas de los algoritmos, frente a los enfoques tradicionales, es que son modelos altamente complejos, sofisticados y capaces de representar límites de decisiones no lineales[20] (Goodman, 2016).

Así es como, (Burrell, 2016) distingue tres clases de barreras de la transparencia:

- Ocultación intencional por parte de las organizaciones, dado que, quieren proteger sus intereses y derechos de propiedad intelectual.

[19] "Aplicar medidas técnicas y organizativas apropiadas para garantizar, en particular, que se corrigen los factores que introducen inexactitudes en los datos personales y se reduce al máximo el riesgo de error, asegurar los datos personales de forma que se tengan en cuenta los posibles riesgos para los intereses y derechos del interesado y se impidan, entre otras cosas, efectos discriminatorios en las personas físicas por motivos de raza u origen étnico, opiniones políticas, religión o creencias, afiliación sindical, condición genética o estado de salud u orientación sexual, o que den lugar a medidas que produzcan tal efecto. Las decisiones automatizadas y la elaboración de perfiles sobre la base de categorías particulares de datos personales únicamente deben permitirse en condiciones específicas." (Considerando 71 RGPD).

[20] Las redes neuronales artificiales reciben este nombre porque se asemejan a su homólogo biológico. Pero claramente como el cerebro éste también contiene cierta opacidad. De acuerdo Goodman (2016): "las redes neuronales están compuestas por múltiples nodos, cada uno representando un particular conjunto de transformaciones funcionales y organizadas en capas visibles y ocultas, donde se encuentran cientos de nodos virtuales."

- Brechas de alfabetización técnica.
- Un desajuste entre el aprendizaje automático y las exigencias del razonamiento a escala humana.

Justo es decir que, explicar la lógica aplicada tal vez no sea una tarea sencilla, dado que las redes neuronales y el lenguaje binario no siempre tienen una interpretación en el lenguaje natural. Especialmente con el advenimiento de tecnologías como *machine learning*, que desafía la comprensión del algoritmo, este proceso es tan complejo que a menudo se crea el efecto *black-box* (ICO, 2017). De esta manera, la información al igual que el cerebro, se encuentra codificada en conexiones múltiples en lugar de almacenarse en ubicaciones específicas. Como bien expone Castelvecchi (2016): "a pesar de que hacemos estas redes no estamos más cerca de entenderlas que un cerebro humano."

Por lo tanto, ¿qué comprensibilidad se puede esperar en una red neuronal multicapa con una arquitectura compleja? [21] (Goodman, B., & Flaxman, S. , 2017). No cabe duda de que, "los enfoques de aprendizaje automático están solos en el espectro por su falta de interpretabilidad." (Lisboa 2013).

Paralelamente, es preciso traer a colación la segunda herramienta contemplada en el artículo 22 del RGPD, la cual reconoce el derecho de las personas a no verse sometidas a las decisiones automatizadas con efectos jurídicos sobre ella. En este sentido, para que dicha prohibición opere deben concurrir las siguientes circunstancias: (i) debe existir una decisión, (ii) basada únicamente en un tratamiento automatizado, (iii) dicha decisión debe tener un efecto jurídico sobre los interesados. Un ejemplo de dichas decisiones puede ser una decisión judicial, o la denegación de un crédito (art. 22 RGPD).

Existen excepciones a esta regla general, cuando se basa en el consentimiento explícito del interesado, cuando se encuentra legitimado por la ejecución de un contrato o cuando esté autorizado por el derecho de la Unión o del Estado miembro. No obstante, si el responsable del tratamiento legitima el tratamiento mediante el consentimiento o la ejecución del

[21] Por ejemplo, Alphago de Google, el cual utiliza el sistema de *deep learning*, fue diseñado para jugar juegos de mesas, siendo campeón mundial. Muchos de sus movimientos fueron descritos como inhumanos. Dado que, sus decisiones y razonamiento se encuentran fuera de la comprensión humana (ICO, 2017).

contrato debe implementar las respectivas medidas de seguridad, como mínimo el derecho a obtener intervención humana (GT29, 2017).

Sin embargo, cabe plantearse que esta protección es bastante limitativa y a la larga poco efectiva. Dado que, muchas de las decisiones automatizadas quedan por fuera del ámbito de protección, porque sólo aplican exclusivamente a los procesos automatizados. En consecuencia, si un empleado de un banco deniega un crédito en base a la recomendación del sistema, siempre que no sea, una intervención humana superficial, la disposición no aplicaría.[22]

Para una mayor ilustración, se trae a colación el caso de Propublica, donde se analizaron los datos de 7000 acusados en algunos estados de Estados Unidos con el fin de predecir la vida criminal futura de estas personas (Angwin, 2016), muy al estilo de *Minority report*. Los hallazgos encontraron que la mayoría de los acusados de color eran clasificados como el doble de peligrosos para la sociedad. Indudablemente en este caso, la disposición del artículo 22 del RGPD aplicaría conforme a los criterios expuestos. Pero, suponga ahora el lector que intercediera un humano, como el procurador. ¿Automáticamente no se aplicaría la prohibición del artículo 22, dejando expuesto a los interesados la aplicación de decisiones con un alto sesgo discriminatorio? O ¿cuál es el papel que debe tener un ser humano para que se considere que la decisión no sea únicamente automatizada? (Gil, 2016).

Por otro lado, cabe analizar si en realidad la intervención humana es un medio efectivo para garantizar la protección del derecho de protección de datos o es simplemente una medida paliativa. Dado que, muchas veces no se detecta el sesgo discriminatorio del patrón, y la lógica del algoritmo frecuentemente es enigma para los humanos o como bien lo contempla (Burrell, 2016): "cuando un computador aprende y toma decisiones, lo hacen sin la comprensión humana." En consecuencia, como expone Gil (2016): "el objetivo no es necesariamente que un humano supervise el resultado a posteriori, sino mejorar la calidad de la clasificación de los individuos a priori."

[22] De acuerdo con el Grupo de Trabajo del artículo 29, la participación humana debe garantizar que sea una supervisión de la decisión significativa, en lugar de un gesto simbólico (GT29, 2017). Sin embargo, cabe preguntarse: ¿cuáles son las condiciones para que las organizaciones puedan garantizar una intervención humana significativa y no superficial?

Al llegar a este punto, es necesario recalcar que es prematuro concluir que las medidas que dispone el RGPD no tienen mérito, pero si es necesario prestar atención a los problemas de la opacidad del algoritmo en el ámbito del *big data*. En consecuencia, se hace un llamado para redireccionar el debate, visibilizar el problema, enfatizando su complejidad. Naturalmente, si ya los sesgos discriminatorios son considerados falencias en la estructura societaria, es imperativo evitar extrapolar dichos sesgos a la dimensión de los bits. Así pues, el siguiente subcapítulo abordará las posibles soluciones para los conflictos anteriormente expuestos.

4.2. Las soluciones que ofrece el RGPD ¿son suficientes?

Los algoritmos utilizados en el análisis de *big data*, no sólo deben ser predictivos y eficientes pero justos y transparentes. En efecto, el RGPD ofrece otros métodos para combatir el sesgo discriminatorio conocido como las auditorías del algoritmo, y apuesta por mecanismos preventivos y proactivos (ICO, 2017).

El primero de ellos, las evaluaciones de impacto que se requieren realizar cuando concurren los supuestos generales en el apartado 3 del artículo 35 RGPD —a saber: (i) la realización de evaluaciones sistemáticas y exhaustivas de aspectos personales cuando se basa en un tratamiento automatizado, como la elaboración de perfiles y sobre cuya base se tomen decisiones que produzcan efectos jurídicos; (ii) el tratamiento a gran escala de categorías especiales de datos o de datos personales relativos a condenas e infracciones penales ; y (iii) la observación sistemática a gran escala de una zona de acceso público. En este sentido, en el caso que se realicen tratamiento automatizados con intervención humana, como garantía adicional se deberá realizar una evaluación de impacto y abordar una serie de medidas adecuadas para la protección de los interesados como: designación de delegados de protección de datos, transparencia, u ofrecer a los interesados medios para oponerse, etc. (GT29, 2017).

La segunda herramienta para auditar el algoritmo consiste en la aplicación del principio de privacidad de diseño y por defecto[23] contemplado en el artículo 25 del RGPD. Es decir,

[23] De acuerdo con la AEPD (2017), la privacidad desde el diseño y por defecto contiene siete principios fundamentales: 1). Proactivo no reactivo; 2). La privacidad como configuración por defecto o privacidad por

incorporar desde la primera fases del proyecto, medidas de privacidad adecuadas, las cuales no sólo consisten en la anonimización, como se explicó anteriormente, sino también la minimización de datos o limitación de la finalidad, y por qué no incluir un *discrimination detection*[24] en los sistemas con el fin de prevenir en primer lugar que se tomen decisiones discriminatorias (ICO, 2017).

En este orden, algunos doctrinantes sostienen que dichos mecanismos no son suficientes e incluso advierten la necesidad de fortalecer los derechos de los interesados y el control que tienen sobre sus datos personales. En concreto, doctrinantes como (Goodman, B*. et al*. , 2017) establece que se debería acompañar la normativa con disposiciones de antidiscriminación aplicables en el ámbito digital. De manera similar, doctrinantes abogan por la posibilidad de introducir otras técnicas específicas como *discrimination impact assessments*[25]*,* es decir, una evaluación impacto concentrada en los sesgos discriminatorios de los algoritmos (Selbst, 2017).

En esta misma dirección, autores como Martínez (2017) establecen que se deben aplicar bases éticas desde el diseño y desde la aplicación de la tecnología de *big data*. En palabras del autor: "todo proyecto de *big data* debería incorporar principios éticos fundamentales. El primero de ellos, particularmente conocido, podría expresarse con el bien conocido lema corporativo de no hacer el mal."

De igual modo, doctrinantes como Watcher y Mittlelstadt (2019) defienden un derecho a inferencias razonables. En este sentido, el *big data* puede basar sus decisiones en datos inferidos mediante la correlación de bases de datos, prediciendo esos comportamientos y características de los sujetos. Los autores proponen un derecho denominado: *right to reasonable inferences,* el cual es aplicable a las predicciones basadas de los algoritmos no

defecto;3). La privacidad embebida en el diseño; 3). Funcionalidad completa- suma positiva; 4). Seguridad punto a punto; 5). Visibilidad y transparencia; 6). El respeto a la privacidad del usuario.

[24] En este sentido, se puede incorporar un sistema de anti- clasificación, que detectan proxy de características protegidas (por ejemplo, en razón del sexo o raza). Adicionalmente se recomienda que las organizaciones documenten el proceso de detección del sesgo discriminatorio y su posterior corrección mediante medidas de seguridad adecuadas en el diseño de la tecnología (Binns, 2019).

[25] Si bien el autor Selbst (2017), analiza esta cuestión en los ficheros policiales en Estados Unidos, advierte que este mismo mecanismo se puede extrapolar a cualquier otro tratamiento. El modelo consiste en abordar un análisis de potenciales efectos discriminatorios en el uso de los algoritmos, con el fin de que el Responsable del tratamiento considere las cuestiones claves y determinar si se precisan acciones o la intervención de la Autoridad de control o de los interesados.

verificables que afecten el derecho fundamental de protección de datos. Así, le proporciona al interesado un mayor control sobre las percepciones y los procesos de toma de decisiones del Responsable del tratamiento, y de convencerlo potencialmente de que uno o ambos están equivocados.[26]

Este derecho requeriría una justificación ex ante por parte del Responsable del tratamiento para establecer si una inferencia es razonable. Por ende, el Responsable del tratamiento tendría que divulgar: (1) por qué ciertos datos son una base relevante para hacer inferencias; (2) por qué estas inferencias son relevantes para la finalidad del tratamiento; y (3) si los datos y métodos utilizados para extraer las inferencias son precisos y estadísticamente confiables. Un mecanismo ex post permite a los interesados cuestionar inferencias irracionales, lo que puede respaldar los desafíos contra las decisiones automatizadas ejercidas en virtud del artículo 22 (3) del RGPD (Watcher *et al.* , 2019).

Valga la verdad, que la supremacía del *big data* es un tema que transciende cada vez más en nuestra sociedad, requiriendo una especial atención de sus posibles consecuencias y, en todo caso una exhaustiva supervisión. Es por ello, que se propone de la mano de la FRA (2018), la creación de un organismo similar a las autoridades de control de protección de datos, con el propósito de supervisar exclusivamente las tecnologías de *big data*, desarrollando informes o bases éticas para el uso de la tecnología, nutriéndose de diferentes expertos en diversas áreas, como abogados, sociólogos, ingenieros, matemáticos o filósofos con el fin de brindar más protección a los interesados.

En resumen, en este capítulo se ha analizado el articulado sobre decisiones automatizadas y elaboración de perfiles contemplados en el RGPD. De nuestro análisis se concluye, que a nivel práctico algunas de las disposiciones quedan en letra muerta, no ofreciendo una herramienta efectiva para solucionar los retos provocados por las tecnologías disruptivas, en especial, el *big data.*

[26] De acuerdo con los autores este tipo de derecho no es nada nuevo, ya que, los interesados tienen ya el derecho de rectificar sus datos (artículo 16 del RGPD). Esta propuesta solo sugiere ampliar el alcance del Art. 16 de meramente de los datos de entrada a datos de salida.

5. Conclusiones

Ciertamente, en el marco del ordenamiento jurídico el derecho de protección de datos se ha consolidado como un derecho fundamental y autónomo, cuyo resguardo no puede pasar desapercibido en el ámbito del *big data*. De esta manera, se ha analizado a lo largo de este trabajo la normativa con el fin de determinar si efectivamente se da una solución a los retos y riesgos que plantea dicha tecnología disruptiva, a continuación, se enumeran las conclusiones derivadas del análisis exhaustivo:

1. **Confianza ciega algorítmica.** Con el paso del tiempo, la tecnología del *big data* tomará más decisiones que afecten la vida de los ciudadanos, decisiones automatizadas para la publicidad comportamental, o para puestos de trabajo. El riesgo latente se evidencia cuando el algoritmo utilice un sesgo discriminatorio y lo reproduzca indefinidamente. La tecnología del siglo XXI podrá permitir que el algoritmo termine conociendo a los sujetos más de lo que se conocen a sí mismo, adicionalmente poco a poco la soberanía de las decisiones trascendentales de la vida diaria recaerá cada vez más sobre los algoritmos informáticos. La dictadura del algoritmo no es un escenario exagerado o de ciencia ficción, por el contrario, puede ser un futuro cercano que amenazará la sociedad del siglo XXI. La caja de pandora se ha abierto antes de entender las consecuencias. El *big data* puede dar muchas ventajas en el progreso humano, no obstante, no se debe ignorar el lado oscuro de sus riesgos expuestos, la curiosidad humana puede llevar a la sociedad a tierras peligrosas. Con la tecnología, la búsqueda infinita de conocimiento podría llevar a la humanidad a un paso más lejos. Por ello, es crucial visibilizar el problema con el fin de repensar el marco normativo, haciendo un llamado para incitar el debate.

2. **Transparencia paradójica.** La transparencia en la información es una de las herramientas que ofrece la normativa para combatir la discriminación algorítmica. Como se analizó, se han encontrado varias dificultades para que esta herramienta sea verdaderamente efectiva en la práctica. La primera de ellas consiste en que la tecnología del *big data* es altamente compleja, y los interesados no podrían

comprender bien cómo se están procesando los datos o cómo funciona la técnica que aplica el patrón del algoritmo cuando se toman decisiones automatizadas, por su efecto *black-box*. La segunda de ellas consiste en que, al explicarse en un lenguaje sencillo y comprensible para los interesados, se corre el riesgo que resulte en un entendimiento empobrecido, aunque dicho supuesto se ha tratado de mitigar mediante la incorporación de informaciones multinivel o por capas siguiendo las recomendaciones de las autoridades de control. No obstante, cabe recalcar que la brecha paradójica no desvanece, al considerar que revelar la lógica aplicada del algoritmo es una hazaña casi imposible, debido al desajuste entre el aprendizaje automático y las exigencias del razonamiento a escala humana.

3. **El problema de la intervención humana.** La intervención humana como medida de protección para garantizar el derecho de protección de datos en el ámbito de *big data* encuentra muchos obstáculos para lograr ser realmente efectiva. Como se examinó, la habilidad de los humanos en intervenir las decisiones de los algoritmos es precaria, teniendo en cuenta la complejidad y opacidad del algoritmo, siendo un verdadero enigma para los humanos. De esta forma, se concluye que en realidad es una utopía afirmar que existe una intervención humana significativa, dado que, en últimas todas se convierten en un mero gesto simbólico. Por lo tanto, se propone más bien que el ser humano no supervise *a posteriori* en el *output* pero un intervención *a priori* en el *input*, siendo la privacidad desde el diseño la clave para combatir los efectos negativos en la privacidad de los interesados a causa de las decisiones automatizadas.

4. **Consentimiento insuficiente.** Como se ha expuesto a lo largo del trabajo, el consentimiento no es la base legitimadora más adecuada en el ámbito del *big data*. Debido a que el consentimiento es binario y ofrece sólo dos opciones a los interesados, lo cual es incompatible con los usos secundarios de la información personal que define el potencial del *big data*. Así es como, el responsable del tratamiento queda obligado a volver a recabar el consentimiento, incurriendo en altos costos o en acciones desproporcionadas. El RGPD abre las puertas a diferentes bases legitimadoras que pueden contrarrestar los usos de los datos personales en el ámbito de *big data*, como

el interés público o el interés legítimo siempre que se realice un análisis de proporcionalidad y necesidad.

5. **La anonimización no reduce el riesgo totalmente.** Como se expuso con anterioridad, las limitaciones que adolecen las técnicas de anonimización en el ámbito del *big data* es un riesgo inminente, las grandes cantidades de datos disponibles deja una alta posibilidad de reidentificación. Si bien el riesgo no es posible erradicarlo de todo, está en manos de los responsables del tratamiento mitigarlo, diseñado minuciosamente técnicas de anonimización en consideración con la naturaleza de los datos y de sus distintos usos. Adicionalmente, es fundamental que se hagan revaluaciones periódicas con el fin de robustecer su eficacia. En últimas, se debe conciliar los riesgos inherentes con la responsabilidad proactiva para fomentar el avance, progreso y tecnología. La anonimización no debe ser entendido como un proceso ocasional, sino un proceso continuado y crucial para potenciar el desarrollo y la privacidad. Adicionalmente, no es de olvidar que la anonimización es un tratamiento de datos, es decir, si el Responsable del tratamiento, requiere utilizar los datos anonimizados para otras finalidades, necesita legitimarlo en una base legal, y en consecuencia informar al interesado sobre el respectivo tratamiento de acuerdo al artículo 13 del RGPD.

6. **Introducción de nuevos derechos o fortalecimiento de garantías de antidiscriminación.** El Reglamento apuesta por mecanismos proactivos y preventivos en vez de reactivos, que precisamente tienen una especial importancia para el *big data*, impone la protección de datos desde el diseño y por defecto y el ciclo de vida del tratamiento de datos. Y lo mismo debería postularse respecto a la inclusión de medidas que integren la no discriminación. Así el RGPD no hace el intento de proveer una definición de la discriminación algorítmica. En consecuencia, se propone la introducción de un cuerpo normativo específico para esta materia y la creación de un organismo similar a las autoridades de control de protección de datos, con el propósito de supervisar exclusivamente las tecnologías de *big data*, y discriminación algorítmica. A su vez, es importante resaltar la proposición de los doctrinantes sobre el derecho a inferencias razonable, que pueden complementar o respaldar el derecho contra decisiones automatizadas, en miras a promover una justificación a las predicciones

(output data) no verificables y contradictorias que invaden la privacidad de una persona o dañan la reputación.

En todo caso, la opacidad del algoritmo no se aliviará con una sola herramienta o procedimiento, pero la combinación de diferentes técnicas de auditoría del algoritmo, y más importante aún, el uso de alternativas como los principios éticos, las evaluaciones de impacto, la educación general del público y la sensibilización por las decisiones automatizadas.

Referencias bibliográficas

Bibliografía básica

Revista

Burrell, J. How the machine thinks: Understanding opacity in machine learning algorithms. *Big data & society*. 2016.

Cotino, L. Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales. *Universidad de Valencia*, 2017, págs. 131-150.

Latour, B. Tarde`s idea of quantification. The *Social After Gabriel Tarde: Debates and Assessments*. 2009, págs. 145-162

Nissenbaum, H. Privacy, big data and the public good. *Cambridge University press*. 2014.

Parra, C. Big data en sanidad en España: la oportunidad de una estrategia nacional. *Gaceta Sanitaria*. 2017 págs. 63.

Puyol, J. Una aproximación a Big data. *Revista de derecho UNED*, 2014.

Wachter S y Mittelstadt B. A right to reasonable inferences: Re-thinking data protection law in the age of big data. *Columbia business law review*. 2019. Págs. 1-130

Revista electrónica

Binns, R. Human bias and discrimination in AI systems. *ICO news* [En línea], 2019. [Consulta: junio 2020] Disponible en: https://ico.org.uk/about-the-ico/news-and-events/ai-blog-human-bias-and-discrimination-in-ai-systems/

Elliot, T. 7 Definitions of Big data you should know about . [En línea] 2013. [Consulta: 19 de 04 de 2020.] Disponible en: https://timoelliott.com/blog/2013/07/7-definitions-of-big-data-you-should-know-about.html.

Esteban, A. El Consejo de Europa moderniza el Convenio 108 sobre Protección de Datos: nace el Convenio 108 +. *Cuatrecasas.* [En línea] 2018. [Consulta: abril de 2020.] Disponible en: https://blog.cuatrecasas.com/propiedad-intelectual/convenio-108-datos/

Goodman, B. A step towards accountable algorithms? Algorithm discrimination and European Union General data protection. *Oxford University* [En línea], 2016. [Consulta: abril de 2020.] Disponible en: http://www.mlandthelaw.org/papers/goodman1.pdf

Goodman, B., & Flaxman, S. European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation". *AI Magazine* [En línea], 2017, núm. 3, pp. 50-57 [consulta: mayo de 2020] Disponible en: https://www.aaai.org/ojs/index.php/aimagazine/article/view/2741

Martínez, R. Cuestiones de ética jurídica al abordar proyecto de big data El contexto del Reglamento general de protección de datos. *Universidad de valencia* [En línea], 2017. [Consulta: 15 de junio 2020] Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=6066833

Monleón-Getino, A. El impacto del Big-data en la Sociedad de la Información. *Universidad de Barcelona* [En línea], 2015. [Consulta: 15 de junio 2020] Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=5301499

Richards N. y Jonathan K. Three paradoxes of big data. *Stanford Law Review.* [En línea] 2013. [Consulta: abril de 2020.] Disponible en: https://www.stanfordlawreview.org/online/privacy-and-big-data-three-paradoxes-of-big-data/.

Saldaña, M. The right to privacy: la génesis de la protección de la privacidad en el sistema constitucional norteamericano: el centenario legado de Warren y Brandeis. *UNED revista de derecho político*. [En línea] 2012. [Consulta: abril de 2020.] Disponible en: http://revistas.uned.es/index.php/derechopolitico/article/view/10723

Schumpeter, B. Building with big data, the data revolution is changing the lanscape of business. *The economist.* [En línea] 2011. [consulta: mayo de 2020]. Disponible en: https://www.economist.com/business/2011/05/26/building-with-big-data

Scott, M. Novartis joins with Google to develop contact lens that monitors blood sugar. *The new york times.* [En línea] 2014. [Consulta: abril de 2020.] Disponible en: https://www.nytimes.com/2014/07/16/business/international/novartis-joins-with-google-to-develop-contact-lens-to-monitor-blood-sugar.html.

Selbst, A. Disparate impact in big data policing. *Georgia law review* [En línea], 2017. [Consulta: mayo 2020] Disponible en: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2819182

Libro

Aguilar, L. *Big data, análisis de grandes volúmenes de datos en organizaciones.* 1ª ed. México: Alfaomega Grupo Editor, 2013.

Boyd, D. & Crawford, K. *CRITICAL QUESTIONS FOR BIG DATA, information, communication & society.* 1ª ed. Cambridge: Routledge, 2012.

Frosini, V. *Información y derecho .* 2ª ed. Bogotá: Temis, 1988.

Gil, E. 2016. *Big data, privacidad y protección de datos.* 1ª ed. Madrid : Agencia española de protección de datos, 2016.

Harari, Y. *Homo Deus.* 1ª ed. Barcelona : Penguin Random House Grupo Editorial, 2017. págs. 400-430.

—. *Sapiens.* 1ª ed. Madrid : Penguin Random House Grupo Editorial, 2016.

—. *21 lecciones para el siglo XXI.* 1ª ed. Madrid: Penguin Random House Grupo editorial, 2018.

Lemaitre, J. *El derecho como conjuro.* 2ª ed. Bogotá : Universidad de los Andes, 2009.

Ortiz, C. *La protección de datos personales.* 1ª ed. Madrid: Paidós, *2005*.

Orwell, G. *1984.* 3ª ed. Madrid : Debolsillo, 2013.

Pinker, S. *En defensa de la ilustración.* 1ª ed. Barcelona : Paidós, 2019. págs. 176-177.

Quesada, A. *Protección de datos y telecomunicaciones convergentes.* 1ª ed. Madrid : AEPD, 2015.

Remolina, N. *Recolección de datos internacionales.* 1ª ed. Madrid : Agencia Nacional de Protección de datos , 2015.

Tegmark, M. *Vida 3.0.* 1ª ed. Barcelona : Penguin Random House grupo editorial, 2018.

Tzu, S. *El arte de la guerra.* 4ª ed. Madrid : EDAF S.A., 1999.

Periódico en línea

Angwin, J. Make Algorithms Accountable. *The New York Times*, 1 Agosto 2016. Disponible en: http://www.nytimes.com/2016/08/01/opinion/make-algorithms-accountable.html?_r=1

Butler, D. When google got flu wrong. *News in focus*. 13 febrero 2013. Disponible en: https://www.nature.com/news/when-google-got-flu-wrong-1.12413.

Castelvecchi, D. Can we open the black box of AI? *Nature,* 5 octubre 2016. Disponible en: https://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731

Harari, Y. Google eligirá a tu pareja; te conocerá mejor que tú. *El país.* 6 abril 2017. Disponible en: https://retina.elpais.com/retina/2017/04/05/talento/1491388233_697594.html

Lohr, S. If algorithms know all, how much should humans help? *The New York Times.* 6 abril 2015. Disponible en: https://www.nytimes.com/2015/04/07/upshot/if-algorithms-know-all-how-much-should-humans-help.html.

Maqueda, A. El INE seguirá la pista de los móviles de toda España durante ocho días. *El país*, 29 de octubre 2019. Disponible en:

https://elpais.com/economia/2019/10/28/actualidad/1572295148_688318.html

Web

Big data: la galaxia de los datos. *Good rebels.* 20 mayo 2020, 12:30. Disponible en: https://www.goodrebels.com/es/big-data-la-galaxia-de-los-datos/

La ley de rendimientos acelerados o por qué todo está cambiando*. Diario de la universidad Loyola Andalucía.* 19 mayo 2020, 11:20. Disponible en: http://www.loyolaandnews.es/ley-de-rendimientos-acelerados/

Conferencia

Martínez, R. *Ética y privacidad de los datos.* Fundación Ramón Areces, conferencia. 2014.

Película

Hang the DJ. Dirigida por Tim Van Patten. EE. UU.: Netflix, 2017.

Minority report. Dirigida por Steven Spielberg. EE.UU.: Dream works, 2002.

Dictamen

Dictamen 04/2007 del GT29 sobre el concepto de datos personales (01248/07/ES). *Diario oficial de la Unión Europea.* 20 de junio de 2007, núm. 136, pp. 1-29. Disponible en: https://ec.europa.eu/justice/article29/documentation/opinionrecommendation/files/2007/wp136_es.pdf

Dictamen 05/2014 del GT29 sobre técnicas de anonimización (0829/14/ES). *Diario oficial de la Unión Europea.* 10 de abril de 2014. Disponible en: https://www.aepd.es/sites/default/files/2019-12/wp216-es.pdf

Dictamen 06/2014 sobre el concepto de interés legítimo del responsable del tratamiento de los datos en virtud del artículo 7 de la Directiva 95/46/CE. *Diario oficial de la Unión Europea*. 9 de abril de 2014. Disponible en: https://www.aepd.es/sites/default/files/2019-12/wp217_es_interes_legitimo.pdf.

Resolución

Resolución de 14 de marzo de 2017, sobre las implicaciones de los macrodatos en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley del Parlamento Europeo. Diario oficial de la Unión Europea, de 14 de marzo de 2017. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52017IP0076

Resolución de 15 de enero de 2018, de la Secretaría de Estado de Servicios sociales e Igualdad, por la que se publica el Acuerdo de Consejo Territorial de Servicios Sociales y del Sistema para la Autonomía y Atención a la Dependencia, libre determinación del contenido del servicio de teleasistencia básica y avanzada. Diario Oficial del Estado, 15 de febrero de 2018. Disponible en: https://www.boe.es/eli/es/res/2018/01/15/(1)

Informe

Agencia Española de Protección de Datos

Orientaciones y garantías en los procedimientos de anonimización de datos personales. Agencia Española de Protección de Datos, 2016. Disponible en: https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf

Código de buenas prácticas en protección de datos para proyecto de big data. Agencia Española de Protección de Datos e ISMS FORUM, 2017. Disponible en: https://www.aepd.es/sites/default/files/2019-09/guia-codigo-de-buenas-practicas-proyectos-de-big-data.pdf

Guía de protección de datos para el ciudadano, Agencia Española de Protección de datos, 2017. Disponible en: https://www.aepd.es/sites/default/files/2020-05/guia-ciudadano.pdf

Guía del Reglamento General de Protección de datos para responsables del tratamiento. Agencia Española de Protección de Datos, 2018. Disponible en: https://www.aepd.es/sites/default/files/2019-09/guia-rgpd-para-responsables-de-tratamiento.pdf

Guía para el cumplimiento del deber de informar. Agencia Española de Protección de Datos, 2018. Disponible en: https://www.aepd.es/sites/default/files/2019-09/guia-modelo-clausula-informativa.pdf.

El uso de tecnologías en la lucha contra el COVID19. Un análisis de costes y beneficios. Agencia Española de Protección de Datos, 2020. Disponible en: https://www.aepd.es/sites/default/files/2020-05/analisis-tecnologias-COVID19.pdf

Agencia de los Derechos Fundamentales de la Unión Europea

Big data: discrimination in data-supported decision making. FRA, 2018. Disponible en: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-focus-big-data_en.pdf

Agencia Europea de Seguridad de las Redes y de la Información

Privacy by design in big data: An overview of privacy enhancing technologies in the era of big data analytics. ENISA, 2015. Disponible en: https://www.enisa.europa.eu/topics/data-protection/privacy-by-design

Consejo de Europa

Discrimination, artificial intelligence, and algorithmic decision-making. CdE, 2018. Disponible en: https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73

Comisión Federal de Comercio

Big data: a tool for inclusion or exclusion. Federal Trade Commission**,** 2016. Disponible en: https://www.ftc.gov/reports/big-data-tool-inclusion-or-exclusion-understanding-issues-ftc-report

Grupo de Trabajo del Artículo 29

Opinion 03/2013 on purpose limitation. GT29, 2013. Disponible en: https://ec.europa.eu/justice/article29/documentation/opinionrecommendation/files/2013/wp203_en.pdf.

Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679. GT29, 2016. Disponible en: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf.

Oficina del Comisionado de Información

Big data, artificial intelligence, machine learning and data protection. ICO, 2017. Disponible en: https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf

Instituto Nacional de Estándares y Tecnología

NIST Big Data Public Working Group Security and Privacy Subgroup. NIST, 2015. Disponible en: https://bigdatawg.nist.gov/V2_output_docs.php

Legislación citada

Carta de los Derechos Fundamentales de la Unión Europea. *Diario oficial de las Comunidades Europeas*, 11 de diciembre de 2000. Disponible en: https://www.europarl.europa.eu/charter/pdf/text_es.pdf

Tratado de la Unión Europea, firmado en Maastricht el 7 de febrero de 1992. Diario Oficial de las Comunidades Europeas, 29 de julio de 1992, núm. 191, pp. 1-112- Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=LEGISSUM%3Axy0026

Convenio 108 (núm. 274) para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, adoptado en Estrasburgo el 28 de enero de 1981. Disponible en: https://www.boe.es/buscar/doc.php?id=BOE-A-1985-23447

Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE. *Diario Oficial de las Comunidades Europeas*, 4 de mayo de 2016, núm.119. Disponible en: https://www.boe.es/doue/2016/119/L00001-00088.pdf

Constitución Española. *Boletín Oficial del Estado*, 29 de diciembre de 1978 Disponible en: https://www.boe.es/eli/es/c/1978/12/27/(1)/con

Ley Orgánica 3/2018, de 5 de diciembre, de protección de datos personales y garantías de los derechos digitales. *Boletín Oficial del Estado*, 6 de diciembre de 2018, núm. 294, p. 119788 a 119857. Disponible en: https://www.boe.es/eli/es/lo/2018/12/05/3

Ley Orgánica 39/2006 de 14 de diciembre, de Promoción de la Autonomía Personal y Atención a las Personas en situación de dependencia. Boletín Oficial del Estado, 15 de diciembre 2006. Disponible en: https://www.boe.es/eli/es/l/2006/12/14/39/con

Ley 20/2015, de 14 de julio, de ordenación, supervisión y solvencia de las entidades aseguradoras y reaseguradoras. *Boletín Oficial del Estado,* 15 de julio de 2015, núm. 168, p. 58455 a 58611. Disponible en: https://www.boe.es/eli/es/l/2015/07/14/20

Jurisprudencia referenciada

Sentencia de Tribunal Constitucional de 30 de noviembre de 2000, C-292/2000, BOE: T:2001:332, apartado 6.

Sentencia de Tribunal Supremo de Estados Unidos Griswold V. Connecticut, 381 U.S. 479, 1965. Disponible en: https://supreme.justia.com/cases/federal/us/381/479/

Listado de abreviaturas

AEPD- Agencia Española Protección de Datos

Art – Artículo

CdE- Consejo de Europa

CE- Constitución Española

CFDUE- Carta de los Derechos Fundamentales de la Unión Europea

EEE- Espacio Económico Europeo

EIPD- Evaluación de Impacto en Protección de Datos

ENISA- Por sus siglas en inglés*: European Union Agency for Cybersecurity*

FRA- por sus siglas en inglés: *Fundamental Right Agency*

FTC- Federal Trade Commission

GT29- Grupo de trabajo del Artículo 29

LOPDGDD- Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos y Garantía de los Derechos Digitales

LOSSEAR- Ley de ordenación, supervisión y solvencia de las entidades aseguradoras y reaseguradoras

LSSI- Ley 34/2002 de 11 julio, de servicios de la sociedad de la información y de comercio electrónico

NIST- National Institute of Standards and Technology

PPE- Parlamento Europeo

RGPD- Reglamento General de Protección de Datos

TC- Tribunal constitucional

UE- Unión europea

Anexo A. Caso práctico de *big data* en telemedicina

Descripción:

La telemonitorización de hábitos es una de las claves de la evolución de una prestación social que comenzó funcionado como una simple telealarma. La información obtenida a través de una red de dispositivos de tecnología sensorial es analizada mediante una serie de algoritmos que permitirá extraer patrones de comportamiento habituales del usuario dentro de su hogar, de tal forma que podamos identificar, de forma proactiva, situaciones potencialmente sospechosas de constituir un riesgo cuando las rutinas habituales no se cumplen.

En efecto, la herramienta de teleasistencia desarrollada por una empresa tecnológica española busca monitorizar a las personas de tercera edad mediante un sistema de pulseras y sensores la actividad diaria del usuario y el estado de diversos parámetros (temperatura, calidad del aire, ruidos y silencios, etc.) dentro de la vivienda. Posteriormente los datos serán analizados en un procesador "al vuelo" por los procesos y modelos de IA, y almacenados en una plataforma de *big data*. Con el fin de ir generando notificaciones y alertas enviadas a la App móvil y a APIs de terceros, y a la vez pueden ser consultados por sus familiares en tiempo real en un dashboard que permite analizar visualmente los datos de los usuarios. Permitiendo, además videos conferencias con los familiares y la persona mayor.

Componentes tecnológicos:

- Smartwatch
- Video conferencias y Hub
- Termómetro
- Sensor de humedad
- Sensor de movimiento
- Sensor de calidad de aire

Plataforma de *big data*:

- Almacenamiento de los datos procedentes de dispositivos y sensores
- Visualización de datos en Dashboards en tiempo real

- Detección de patrones, tendencias y anomalías
- Creación de modelos predictivos

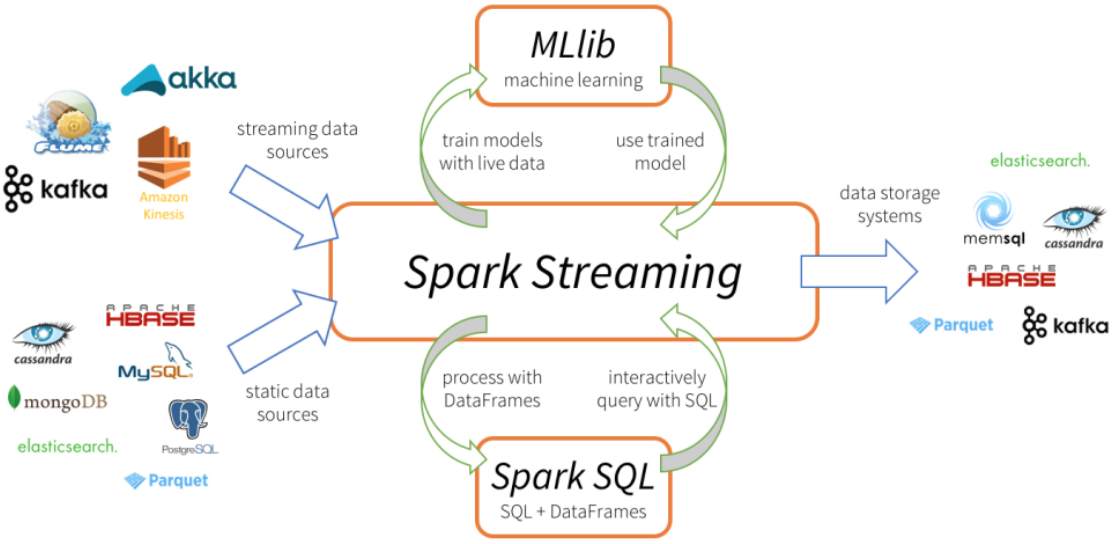

Imagen 1. Flujo de datos y componentes tecnológicos

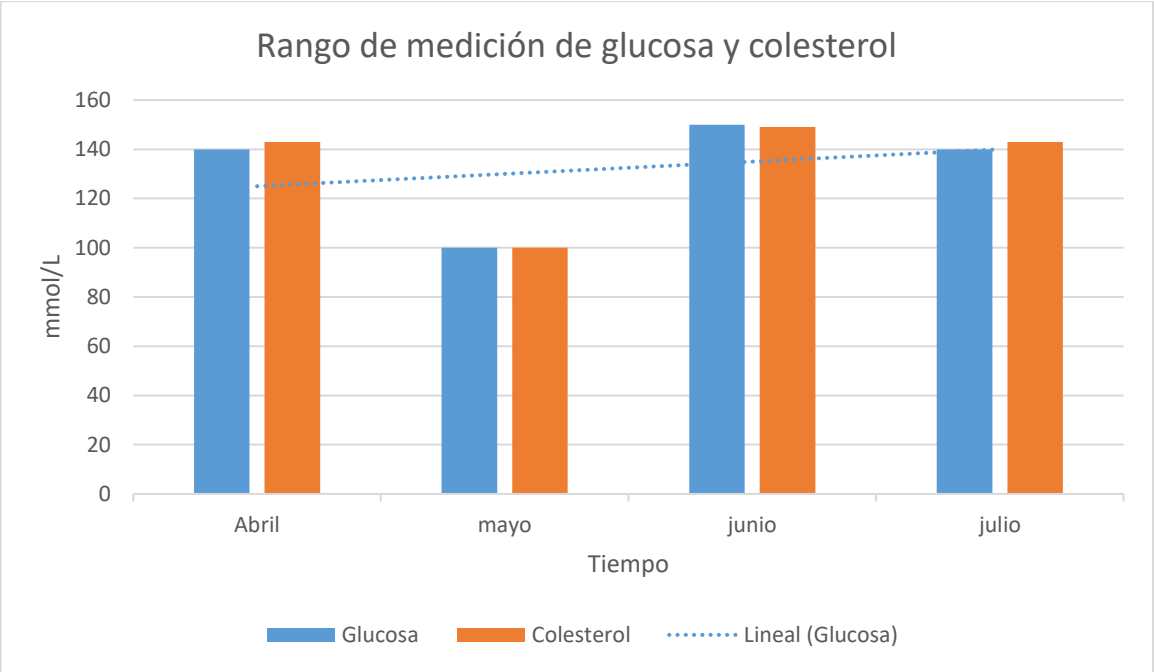

Gráfico 1. Datos de salud en tiempo real de los usuarios

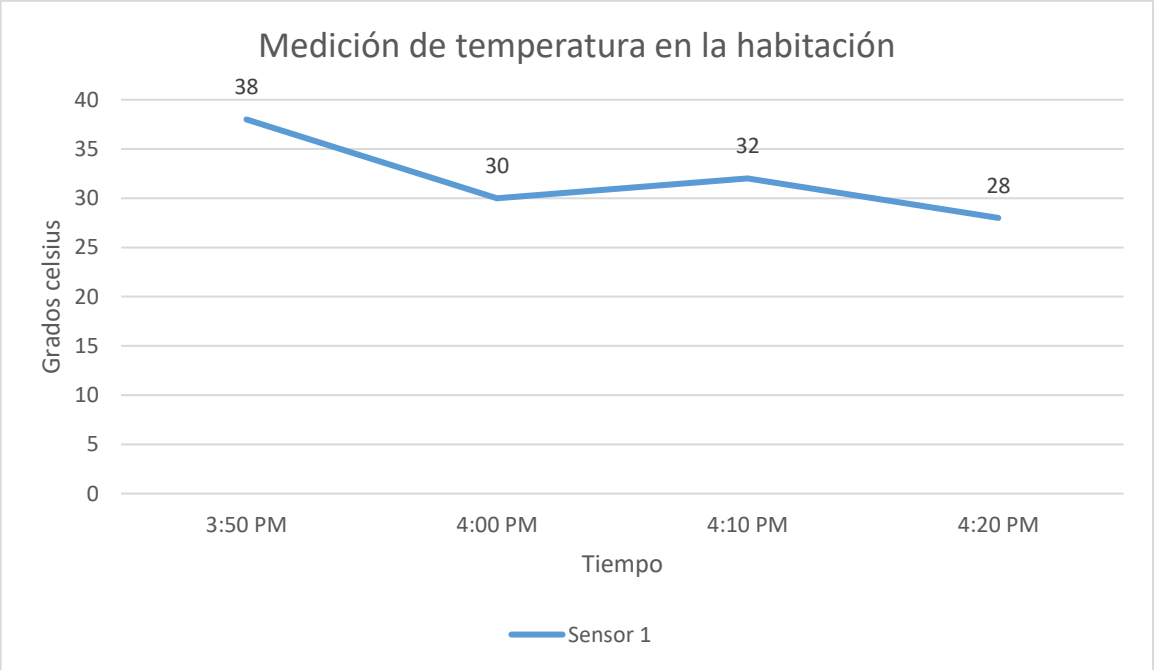

Gráfico 2. Datos de temperatura en las habitaciones

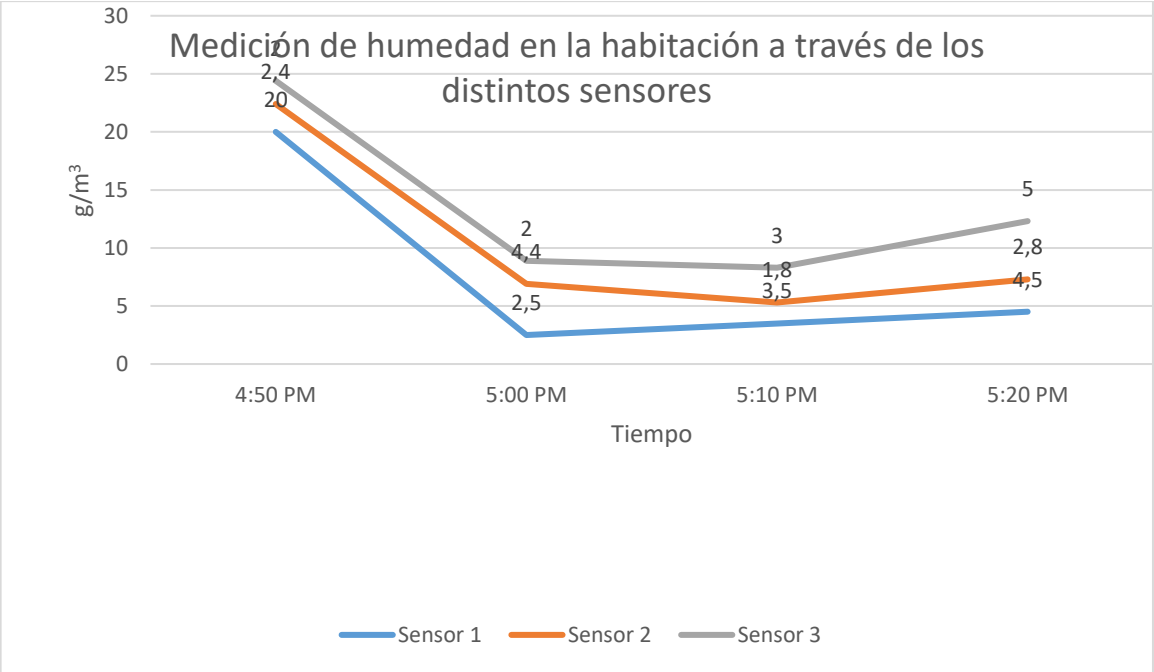

Gráfico 3. Datos de humedad en las habitaciones

Análisis en materia de privacidad y protección de datos:

La teleasistencia básica es regulada por la Ley 39 del 2006 de Promoción de la Autonomía Personal y Atención a las personas en situación de dependencia, como uno de los servicios sociales de promoción de la autonomía personal y de atención a las personas en situación de dependencia. A su vez, la Resolución 15 de enero de 2018, de la Secretaría de Estado de Servicios Sociales e Igualdad, por la que se publica el Acuerdo de Consejo Territorial de Servicios Sociales y del Sistema para la Autonomía y Atención a la dependencia, incorpora tanto la teleasistencia básica como avanzada, siendo ésta última definida como: "aquella que incluye, además, de los servicios de teleasistencia básica que la persona usuaria precise, apoyos tecnológicos complementarios dentro o fuera del domicilio, o en ambos casos, así como la interconexión con los servicios de información y profesionales de referencia en los sistemas sanitario y social, desarrollando procesos y protocolos de actuación en función de la situación de necesidad de atención detectada."

Entre la tecnología que permite la Ley para el servicio de teleasistencia avanzada dispone lo siguiente:

- Detectores de caídas, escapes de gas, de agua, de fuego, convulsiones, enuresis y otras).

- Pulsera que contemple como funciones principales la localización y alertas de zona.

- Dispositivo móvil específico que integra las funcionalidades de la pulsera y añade la posibilidad de emisión y recepción de comunicaciones. Entre las posibles modalidades para la emisión y recepción de mensajes se encuentra la videocomunicación.

- *Smartphone/tablet* adaptados para la recepción de mensajes y eventos, agenda personal o compartida con la persona cuidadora, comunicación de incidencias, o envío de pictogramas y mensajes preestablecidos.

- Aplicación móvil de teleasistencia (app) para gestionar y/o solicitar los servicios contemplados.

En efecto, esta tecnología se encuentra aceptada para prestar servicios sociales de asistencia al colectivo vulnerable de personas dependientes, dado al crecimiento de la población demográfica envejecida en España.

En consecuencia, si se decide utilizar dicha tecnología, es importante minimizar el impacto en los derechos y libertades de los interesados mediante las siguientes recomendaciones en relación con la seguridad y privacidad de los datos:

Con el fin de garantizar la seguridad de la información obtenida y procesada se debe, entre otras medidas:

- Aplicar medidas de anonimización cuando los datos se almacenen en la plataforma de *big data*.
- Aplicar medidas de seudonimización siempre que sea posible, minimizando así el riesgo de brechas de seguridad.
- Garantizar la seguridad y cifrar los datos cuando estos se encuentren en reposo, en tránsito o en comunicaciones punto a punto. Tanto la información obtenida de sensores como las conversaciones de videollamadas.
- La destrucción segura y garantizada o anonimización de la información al final de su ciclo de vida.

- El análisis de los datos personales presentado en el "Dashboard" sólo podrá ser accedido por los terceros estipulados previamente en el contrato. En particular, solo deberían poder acceder el responsable de teleasistencia a quien se le haya adjudicado el caso, los familiares o cuidadores escogidos (solo los necesarios) y el personal de la empresa que precise acceder a la información para mantener el servicio de manera adecuada. Por lo tanto, el acceso a los datos, independientemente del canal utilizado debe contar con un control de accesos que permita discriminar el nivel de acceso y hacer un seguimiento de las acciones realizadas por el usuario.

Con el fin de garantizar unos altos niveles de privacidad:

- Debería incorporarse una funcionalidad con la capacidad de desconectar el sistema de monitorización en cualquier momento por el usuario, con el fin de proteger su derecho de intimidad y el libre desarrollo de su personalidad.
- En base al principio de limitación del tratamiento, los datos personales recogidos no podrán ser utilizados para finalidades diferentes a las que fueron recabadas. Por ejemplo, utilizar los datos para enviar comunicaciones comerciales o elaborar perfiles comerciales. Para ello se deberá informar apropiadamente los usuarios de la información.
- Se recomienda que, si en la vivienda habitan varias personas, por el principio de minimización de datos, sólo debería localizarse los sensores en la habitación de la persona mayor y no ubicarlos en las habitaciones de las otras personas que habitan con él/ella.
- Los dispositivos usados deben garantizar la privacidad y confidencialidad de las comunicaciones. Para ello, debemos asegurarnos de que únicamente los datos recogidos son transmitidos a la empresa, la aseguradora y/o el cuidador. No a terceros como el proveedor de servicios o el del sistema operativo.
- Hay que asegurar que cualquier proveedor con acceso a la información obtenida ha proporcionado garantías suficientes de cumplimiento.

- Se deberá garantizar que los datos se conservan únicamente mientras sean necesarios para el servicio y una vez ha finalizado deberán eliminarse o se anonimizados. Sin perjuicios de los plazos legales de conservación.
- Con el fin de garantizar el ejercicio de derechos, la información debe ser clasificada adecuadamente y permitir su borrado, bloqueo (sin borrado) y portabilidad en los plazos legales establecidos.
- Con el fin de asegurar que se mitiga cualquier posible amenaza para la privacidad de las personas mayores, se debe realizar una Evaluación de Impacto sobre la Protección de Datos de manera que se mitiguen los riesgos de privacidad y de seguridad a los que estén expuestos. Durante el proceso piloto se podría realizar una encuesta a los usuarios del proyecto con el fin de evaluar el impacto que tiene este sobre su privacidad y qué medidas considerarían suficientes para reducirlo.

Datos de usuarios:

Mediante el uso de esta aplicación se tratarán datos identificativos personales, de geolocalización y datos de salud, así como de comportamiento.

En el Reglamento General de Protección de datos (en adelante RGPD), se advierte que los datos de salud no son sólo aquellos que hace alusión de forma directa con la salud física y mental, sino que incorpora de igual modo, información relativa al estado de salud.

Por lo tanto, toda aquella información captada o procesada a través de los determinados dispositivos (como puede ser sensores que determinan el número de pasos dados, o las calorías ingeridas) pueden terminar revelando un determinado estado o un concreto peligro.

En general los datos de salud serán aquellos que engloben determinados aspectos como son los que a continuación se enumeran:

- Datos que posean de forma inherente o claramente manifiesta un carácter médico
- Datos personales en bruto procedentes de sensores, que pueden ser utilizados individualizadamente o en combinación con otros para emitir conclusiones relativas al estado de salud o riesgo para la salud de un determinado interesado

- Determinadas conclusiones extraídas en relación con el estado de riesgo para la salud del interesado, más allá de que las mismas sean o no ciertas, procedentes o legítimas

En efecto, mediante el uso de esta tecnología, se recogerá en todo momento datos de estado de salud procedente de sensores, por lo que se debe determinar su legitimación para el tratamiento mediante la aplicación de una de las excepciones del artículo 9 del RGPD. En este caso, al considerar que estamos ante un servicio de carácter social como se ha explicado anteriormente, se podrá aplicar la base legitimadora del artículo 6. 1 literal b (el tratamiento es necesario para la ejecución de un contrato en el que el interesado es parte), en consonancia con la excepción contenida en el literal h del artículo 9.2 del RGPD:

> *El tratamiento es necesario para fines de medicina preventiva o laboral, evaluación de la capacidad laboral del trabajador, diagnóstico médico, prestación de asistencia o tratamiento de tipo sanitario o social, o gestión de los sistemas y servicios de asistencia sanitaria y social, sobre la base del Derecho de la Unión o de los Estados miembros o en virtud de un contrato con un profesional sanitario y sin perjuicio de las condiciones y garantías contempladas en el apartado 3.*

Cualquier tratamiento adicional que no esté incluido en los servicios del contrato requerirá el consentimiento de los interesados o de otra base legitimadora adecuada al tratamiento. En el caso del consentimiento, será preciso que el interesado pueda retirarlo de la misma manera que se otorgó o similar.

Adicionalmente se recomienda que, al dirigirse a un colectivo vulnerable como pueden ser los usuarios del servicio, la información debe proporcionarse en un lenguaje claro y sencillo con el fin de cumplir el principio de transparencia. Por lo tanto, las cláusulas informativas deben ser redactadas evitando hacer referencias legales e informar de manera íntegra de los tratamientos de datos personales de manera concisa, inteligible y de fácil acceso para los mayores y los familiares. Ahora bien, en este caso, dicho requisito se deberá aplicar con un nivel de exigencia importante, dado que, que el interesado forma parte del colectivo de personas dependientes, y en virtud del principio de responsabilidad proactiva, se considera que sería la manera más adecuada de cumplir con las disposiciones del RGPD.

En efecto, esta tecnología de telemedicina permitirá la elaboración de perfiles y análisis predictivos, por lo tanto, es necesario considerar la aplicación del artículo 22 del RGPD, en

este caso, dicha decisión automatizada es necesaria para la celebración de un contrato entre el interesado y el responsable, lo cual permitiría su aplicación de acuerdo con el numeral 2 de dicho artículo. No obstante, como en este caso se tratarán datos de categorías especiales se deberá tener en cuenta especial consideración salvaguardando los derechos y libertades mediante la implementación de medidas de seguridad adecuadas, como la privacidad por diseño, la intervención del personal sanitario que es notificado cuando ocurre alguna incidencia en la salud del paciente, y la aplicación del literal 9 letra g). del RGPD como se mencionó anteriormente, atendiendo al mandato del art. 22 literal 4 del RGPD.