UNIVERSITAT DE BARCELONA

LA PROTECCIÓN DE DATOS EN LOS TRATAMIENTOS QUE INCORPORAN INTELIGENCIA ARTIFICIAL

ANDREA SÁNCHEZ RODRÍGUEZ

NIUB:18011836

IRENE ARAGUÁS

TFG Dret-2n semestre
28/06/2020

SUMARIO

INTRODUCCIÓN

En la década de los 50, y después del éxito de los primeros ordenadores, algunos investigadores se dieron cuenta que aquellas nuevas máquinas eran capaces de hacer mucho más que procesar números, podían llevar a cabo procesos similares a los de ser humano. En consecuencia, el año 1956 se organizó una reunión en Dartmouth donde se debatieron estas inquietudes y dónde se usó por primera vez el termino *Inteligencia Artificial* (IA).

Pero no fue hasta principios del siglo XXI cuando la inteligencia artificial tomó fuerza en las economías de centro[1]. Actualmente, podemos afirmar que son muchos los procesos en los que se utiliza IA y que esta presente en nuestras vidas cuotidianas de forma casi imperceptible. Desde la utilización de un asistente virtual para organizar nuestro trabajo diario o nuestros emails, hasta nuestros teléfonos sugiriéndonos canciones, series, ropa o composiciones de nuestras fotos que nos pueden gustar. La IA es una realidad en las sociedades modernas. Pero las utilidades de la IA no se agotan en hacer nuestra vida más fácil, sino que nos ayudan a resolver algunos de los mayores desafíos con los que se enfrenta la humanidad en estos momentos. Así, por ejemplo, se ha utilizado para detectar y tratar enfermedades crónicas, acabar con trabajos repetitivos, eliminar riesgos laborales, reducir la tasa de mortalidad de accidentes de tráfico, predecir desastres naturales o hasta para anticipar amenazas cibernéticas.

Estos avances indudables que traen los sistemas de IA en algunos casos precisan datos de carácter personal que deberán tratarse de manera ética, responsable y transparente, de no ser así, podrían entrar en conflicto con los derechos fundamentales de los titulares de dichos datos. Es por este motivo que la Unión Europea, desde los años 90, se ha ocupado de ofrecer un marco normativo que proteja a los ciudadanos en materia de intimidad, y concretamente, en la parcela relacionada con la protección de sus datos (*habeas data*) promulgando normativa que ha culminado con la aprobación, en 2016, del vigente Reglamento General sobre Protección de Datos Personales (RGPD). Es este reglamento el que se aplica como norma general a todos aquellos sistemas de IA que incorporen datos personales. La casuística en este terreno y la complejidad que plantea ha llevado, sin embargo, a que esta normativa se haya desarrollado de manera sectorial para cubrir las necesidades específicas que plantean los distintos sectores donde se utiliza la IA. En realidad, es tal el impacto de la IA, que la Comisión Europea creó, en junio de 2018, el Grupo de Alto Nivel en Inteligencia Artificial para desarrollar una estrategia europea que promueva el desarrollo en *Inteligencia Artificial Fiable.*

Pese a los amplios beneficios que nos pueden aportar los sistemas de IA estos tienen también una parte oscura que no podemos obviar. Los riesgos que generen deberán ser reflexionados con responsabilidad y seriedad suficientes, tratado de encontrar el equilibrio entre el acelerado avance de las nuevas tecnologías, las nuevas necesidades sociales y científicas y el respeto de los derechos individuales y colectivos.

Este trabajo de investigación pretende llevar a cabo una revisión sistemática con el objeto analizar la legislación en materia de protección de datos que afecta a los sistemas de IA y estudiar los riesgos que estos tratamientos pueden generar para determinar cuál puede ser la gestión adecuada

[1] Según la teoría del sistema-mundo de Immanuel Wallerstein, este está dividido en grupos de países a los que denominó “centrales”, “semiperiféricos” y “periféricos”. Los países centrales son industrializados, desarrollados y ricos, con una posición dominante en el sistema-mundo moderno. Esta teoría ha sido reconfigurada en la actualidad con la incorporación de nuevas potencias al panorama internacional, siendo China el caso más evidente.

de los mismos. Creo necesario mencionar que, pese a que la cuestión está estrechamente relacionada con los derechos de los titulares de los datos, concedidos por los artículos 12 a 22 del Reglamento General de Protección de Datos 2016/679, en este trabajo no haré referencia a ellos.

1. ¿QUÉ ES UN SISTEMA DE INTELIGENCIA ARTIFICIAL?

La Comisión Europea emitió una Comunicación en abril de 2018 en el que definió la IA como aquellos *"sistemas que manifiestan un comportamiento inteligente, al ser capaces de analizar el entorno y realizar acciones, con cierto grado de autoridad, con el fin de alcanzar objetivos específicos*[2]*"*. Posteriormente aquel mismo año y debido al potencial cambio de paradigma en el que nos íbamos a ver inmersos a raíz de la IA, la Comisión Europea creo el Grupo de Alto Nivel en Inteligencia Artificial (IA-HLG) con el fin de asesorar y promover el desarrollo de sistemas de IA éticos. Este grupo reformuló la definición de IA aportada por la Comisión Europea y en uno de sus primeros pronunciamientos la definición que hoy en día rige a nivel europeo:

> "*Los sistemas de inteligencia artificial son sistemas de software (y en algunos casos también de hardware) diseñados por seres humanos* [3]*que, dado un objetivo complejo, actúan en la dimensión física o digital mediante la percepción de su entorno a través de la obtención de datos, la interpretación de los datos estructurados o no estructurados que recopilan, el razonamiento sobre el conocimiento o el procesamiento de la información derivados de esos datos, y decidiendo la acción o acciones óptimas que deben llevar a cabo para lograr el objetivo establecido. Los sistemas de IA pueden utilizar normas simbólicas o aprender un modelo numérico; también pueden adaptar su conducta mediante el análisis del modo en que el entorno se ve afectado por sus acciones anteriores."*[4]

El término de IA abarca un abanico muy extenso de posibles tratamientos, por lo que resulta necesario distinguir entre tres categorías en función del alcance y el ámbito de aplicación de ésta. La inteligencia artificial fuerte, general y débil.

La IA fuerte es aquella que puede resolver problemas que van más allá de las capacidades humanas. La IA general es la que puede resolver problemas que el ser humano también es capaz de solucionar. Por último, la IA débil es aquella que, en contraste con las otras dos, es capaz de resolver un problema concreto y acotado. Las aplicaciones de esta última son muy variadas, desde sistemas de defensa, pasando por el entorno biomédico, control industrial, robótica, buscadores de internet, asistentes personales, marketing, medio ambiente, etc.

El ámbito de la aplicación de la IA débil se extiende a todos los sectores y, por ende, es la que ha genera más problemas y mayor discusión. Tras el caso *Cambridge Analytica* el debate sobre los límites a los que debe someterse la IA, que hasta el momento únicamente había interesado a la comunidad científica, llegó a todos y cada uno los hogares de nuestra sociedad. En este conocido

[2]Comunicación (COM (2018) 237 final) de la Comisión, de 24 abril 2018, "*Artificial Intelligence for Europe*". Página 2. [En línea: https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe] [visitado el 28/06/2020]

[3] Los seres humanos diseñan sistemas de IA directamente, aunque también pueden emplear técnicas de IA para optimizar su diseño.

[4] High-Level Expert Group on Artificial Intelligence, 8 de abril de 2019, "*Definition developed for the purpose of the AI HLEG's deliverables".* Página 6. [En línea: https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top] [visitado el 28/06/2020]

caso[5] se utilizó un sistema dotado de IA combinado con la fuerza de las redes sociales y el tratamiento de *big data*, con el fin de manipular la información y así incidir en las preferencias políticas para beneficiar a Donald Trump en la campaña electoral a la presidencia de EE. UU. de 2016. También se empleó un artificio semejante para campaña *Leave EU* sobre la salida de Reino Unido de la Unión Europea (Brexit). Los periódicos estadounidenses[6] destaparon el escándalo *Cambridge Analytica* en marzo de 2018. Las noticias revelaron un traspaso indebido de datos de millones de usuarios de Facebook a dicha empresa, que utilizó para incidir sobre los electores mediante una estrategia de *microtargeting*[7] basada en técnicas militares de ataque psicológico. Posteriormente, se descubrió que dicha empresa también estuvo detrás de otros sucesos políticos como la victoria de Bolsonaro en las elecciones de Brasil y el mantenimiento en el poder de Narendra Modi en India.

Este caso puso en alerta a las democracias occidentales y, en España, permitió cuestionar la legalidad de algunas previsiones que se habían incorporado recientemente a la Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General (LOREG)[8], que ofrecían la posibilidad a los partidos políticos hacer acopio de datos personales de los ciudadanos amparados por el interés público.

1.1 . Tratamientos que incorporan sistemas de IA

El desarrollo de una solución IA se efectúa mediante distintos modelos basados en algoritmos matemáticos complejos. El campo científico de la IA, básicamente, podemos dividirla en dos subcampos o modelos, los sistemas de aprendizaje automático (*Machine Learning)*, y los sistemas o maquinas que razonan (*Machine Reasoning*). Los modelos *Machine Reasoning* son comúnmente utilizados para planificar, representar y razonar conocimiento introducido, optimizar procesos, búsquedas, etc. Estos utilizan algoritmos matemáticos como, por ejemplo, los sistemas basados en reglas que fundamentalmente resuelven los problemas que se le plantean a partir de conocimientos que previamente se han incorporado al sistema.

[5]KOZLOWSKA I., 30 de abril de 2018, *Facebook and Data Privacy in the Age of Cambridge Analytica.* The Henry Jackson School of International Studies. University of Washington [En línea: https://jsis.washington.edu/news/facebook-data-privacy-age-cambridge-analytica/] [visitado el 28/06/2020]

[6]CADWALLADR C. and GRAHAM-HARRISON E, "Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach: Whistleblower describes how firm linked to former Trump adviser Steve Bannon compiled user data to target American voters". *The Guardian*, 17 de marzo 2018. [En línea: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election] [visitado el 28/06/2020]

[7]CONFESSORE N, and HAKIM D." Data Firm Says, "Secret Sauce" Aided Trump, Many Scoff". *The New York Times.* 6 de marzo 2018. [En línea: https://www.nytimes.com/2017/03/06/us/politics/cambridge-analytica.html] [visitado el 28/06/2020]

[8]La disposición adicional 3ª de la Ley Orgánica 3/2018 de Protección de Datos Personales y garantía de los derechos digitales, modificaba la Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General añadiendo el artículo 58 bis, que en su apartado 1º estipulaba lo siguiente: "*1. La recopilación de datos personales relativos a las opiniones políticas de las personas que lleven a cabo los partidos políticos en el marco de sus actividades electorales se encontrará amparada en el interés público únicamente cuando se ofrezcan garantías adecuadas."* Este artículo fue objecto de un recurso de inconstitucionalidad el 5 de marzo de 2019 por parte del Defensor del Pueblo y, finalmente, mediante sentencia, se acabó declarando inconstitucional el apartado 1º del mismo. (STC 76/2019) [En línea: https://www.boe.es/buscar/doc.php?id=BOE-A-2019-9548] [visitado el 28/06/2020]

Frente a este sistema, los modelos de aprendizaje automático pueden usar algoritmos genéticos o algoritmos mediante reglas neuronales, deep learning, entre muchos otros. El modelo de IA más prometedor y que ha tenido más éxito en la praxis, es el sistema de Aprendizaje Automático, en concreto aquellos que utilizan algoritmos mediante redes neuronales. Estos sistemas intentan simular el modelo humano de aprendizaje, la manera en la que somos capaces de realizar determinadas tareas. *Los humanos responden y razonan mediante la asociación de ideas. Unas ideas conectan con otras y dan lugar a otras nuevas ideas y otras muchas son descartadas porque no interesan para esa situación a la que se busca respuesta. Este proceso es el que permite al humano desarrollar conocimiento, experiencias, creencias, etc.*[9]. En definitiva, se permite al sistema de IA desarrollar procesos cognitivos.

Para el desarrollo de este modelo IA son necesarios unos datos iniciales y un algoritmo capaz de crear asociaciones, apreciarlas o descartarlas. Eso le proporciona al sistema dotado de IA la capacidad de evolucionar y obtener soluciones por él mismo a los problemas que se le planten. Es más, la capacidad de sus algoritmos para tratar los enlaces entre asociaciones y la potencia de sus procesadores para hacer el análisis en un tiempo mínimo puede proporcionar soluciones que al cerebro humano se le podrían escapar, a primera vista. En este sentido podríamos decir que los sistemas de IA con Aprendizaje Automático son modelos predictivos que analizan conjuntamente las variables proporcionadas en un inicio, identifican patrones y establecen criterios de clasificación. Posteriormente, y una vez el modelo ha establecido los criterios, al proporcionar al sistema nuevos datos, éste es capaz de tomar decisiones. Esta técnica esta comúnmente relacionada con el *big data*[10], o el análisis masivo de datos. No obstante, no es necesariamente siempre así, ya que los sistemas predictivos también pueden prosperar mediante el análisis de "*small data*".

Como se ha mencionado, las técnicas de aprendizaje automático son muy utilizadas en la actualidad, con lo que es muy común que nos relacionemos en nuestro día a día con ellas. Éstas son utilizadas por grandes corporaciones como Google, en su algoritmo de búsqueda y en las recomendaciones personalizadas de los contenidos en línea, del mismo modo que Twitter, Amazon, Netflix, Facebook, etc. La revista electrónica Quaderns del CAC[11] estima que más del 10% de los contenidos de las redes sociales y un 62% de todo el tráfico en internet es generado por *bots*[12].

Los sistemas de IA que toman decisiones automatizadas pueden adoptar dos roles distintos: el de soporte y ayuda en la toma de decisiones, proporcionando información sobre la misma para que el ser humano ejecute la decisión final; o la toma y la ejecución de la decisión por el mismo

[9]PASCUAL ORTEGA M.ª T., *"*¿Inteligencia? Artificial", *Tribuna*, 2018, vol. 208, página 88. En línea: https://www.coit.es/archivo-bit/primavera-2018/inteligencia-artificial] [visitado el 28/06/2020]

[10] El término *big data* alude a las tecnologías, técnicas y metodologías relacionadas con el procesamiento de grandes y heterogéneos volúmenes de datos y que superan los límites y capacidades de herramientas de software habitualmente utilizadas para la gestión y procesamiento de estos datos.

[11] Quaderns de CAC. *Fake news, algoritmos y burbujas informativas* (44 Ed.), de julio de 2018, Vol. XXI. Barcelona: Consell d'Audiovisual de Catalunya. Página 6. [En línea: https://www.cac.cat/sites/default/files/2018-08/Q44_ES.pdf] [visitado el 28/06/2020]

[12]Los responsables del "*Project on Computational Propaganda*", Samuel Woolley y Philip Howard, investigadores de la interacción de los algoritmos, la automatización y la política de la Universidad de Oxford, definen los *bots* como "*identidades automatizadas que pueden llevar a cabo tareas rutinarias como recopilar información, pero también pueden comunicarse con personas y sistemas. Se pueden emplear para realizar trabajos legítimos como difundir noticias e información. También son usados para actividades maliciosas asociadas al spam y el acoso".* [En línea: https://comprop.oii.ox.ac.uk/] [visitado el 28/06/2020]

sistema de IA. Un modelo IA que use *Machine Learning* con tratamiento de datos personales, puede afectar a personas físicas en sus decisiones, y éstas pueden afectar a su entorno social, al acceso a contratos o servicios, la personalización de dichos servicios, o a una infinidad de cuestiones. Precisamente por ese motivo, los sistemas de IA que usan datos personales, como un modelo de perfilado de marketing o electoral, deben adecuarse a los dispuesto en el Reglamento (UE) 2016/679 y la normativa sectorial para asegurar la correcta protección de los derechos de los titulares de los datos. Al margen de dicha normativa quedarán, sin embargo, los sistemas de IA que no usan datos de carácter personal. Por ejemplo, un modelo industrial de control de calidad o los sistemas de toma de decisiones de compra y venta de productos financieros.

1.2. Ciclo de la vida de un sistema con IA

El ciclo de la vida de un sistema con IA, desde su desarrollo hasta su descarte es muy parecido a otros procesos tecnológicos y está compuesto generalmente por las siguientes fases. La primera es la *Concepción y el análisis*, donde se fijan los objetivos del sistema y el sector del Ordenamiento jurídico en el que se va a desenvolver. La segunda es el *Desarrollo*[13], esta etapa incluye la investigación, el prototipo, el diseño, las pruebas el entrenamiento y la validación del sistema. La tercera es la *Explotación*, en esta etapa se lleva a cabo la integración en un tratamiento, producción, despliegue, inferencia, decisión, evolución y mantenimiento. Por último, cuando este ciclo llega a su fin se procede a la *Retirada* del sistema de IA. Hay que tener presente que las etapas del ciclo de vida de un sistema IA no se desarrollan separada y ordenadamente, sino que en muchas ocasiones se solapan y se producen en paralelo.

En cuanto a los datos personales que puedan tratarse en la solución IA, estos pueden emplearse durante el transcurso de todo el ciclo de vida del sistema. Es por este motivo que es sumamente importante actuar en cada una de estas etapas con rigurosa cautela, asegurando el buen cumplimiento de la normativa y protegiendo adecuadamente los derechos de los titulares de esos datos, en todo momento.

Los sistemas de Aprendizaje Automático, por ejemplo, utilizan frecuentemente datos personales en la fase de entrenamiento. Posteriormente en la fase de validación pueden usarse también datos personales para comprobar la bondad del sistema y estos pueden ser distintos de los datos usados en la fase de entrenamiento.

Sin embargo, los modelos de IA no compuestos por algoritmos de Aprendizaje Automático, posiblemente no van a tratar datos en la frase de desarrollo porque en ésta tan sólo se integrará el conocimiento necesario para la resolución de problemas. Más adelante, en la fase de explotación, es posible que sí se traten datos personales al integrarse en un tratamiento o un proceso en el que la IA vaya a intervenir.

En fase de explotación de un sistema de IA, es donde probablemente los datos personales van a tener más presencia. En concreto, durante la evolución del sistema, la solución de IA podrá usar los datos de los interesados para, tomar o sugerir decisiones y a su vez mejorar el modelo. Por regla general, cuando esto suceda la actividad de la IA deberá desarrollarse bajo los mandatos del RGPD. No obstante, la normativa sobre protección de datos no debe cumplirse cuando el uso que

[13]Las posibles fases de esta etapa son muy variadas y dependiendo de la IA observamos unas u otras. Por ejemplo, en los sistemas de IA de Aprendizaje Automatizado encontraremos el entrenamiento en esta etapa, pero en los que no se desarrollan mediante este algoritmo, no contarán con esta fase.

se hace de los datos es doméstico (artículo 2.2.c). No es necesario que se someta a la normativa sobre protección de datos la recopilación de datos personales realizada por el propio usuario dentro en sus actividades personales y domésticas.

Pese a que los sistemas de IA estén tratando datos personales, entenderemos que la excepción doméstica se cumple cuando los datos que recopile la solución para cumplir su función y mejorar su sistema lo haga de forma aislada y autónoma. Es decir, cuando estos datos no salgan del mismo sistema y estén solo al abasto del propio usuario del producto de IA. Luego, si estos datos recopilados por la IA son enviados a terceros, es decir, encontramos comunicación y almacenamiento, por ejemplo, para posteriores tratamientos o comunicaciones de datos, los que ostentan estos datos deberán cumplir rigurosamente la normativa.

En último lugar, no debe olvidarse que la fase de retirada del sistema deberá efectuarse asegurando la correcta eliminación de los datos usados durante el ciclo de vida de la solución de IA.

Hablamos de *ciclo de vida* o "*ciclo de Deming*"[14] cuando nos referimos al desarrollo de soluciones IA. Una vez creado el modelo es necesario el mantenimiento y mejora de la calidad de éste y por ende a preservar, durante su desarollo, el estándar de cumplimiento normativo como compromiso de permanente actualización.

En suma, una solución de IA puede o no tratar datos personales, del mismo modo que si toman decisiones automatizadas, estas pueden o no afectar a personas físicas. De trabajar con datos personales y/o afectar a personas físicas deberá adecuarse a la normativa del sector del ordenamiento en el que incida y al RGPD. Ahora bien, cada etapa del ciclo de una solución de IA en la que se traten datos personales constituirá un tratamiento, tal como apunta la Agencia Española de Protección de Datos (AEPD):

> *"Durante el ciclo de vida de una solución IA pueden haberse usado datos personales de alguna forma, por ejemplo, en la etapa de desarrollo. En ese caso, dicha etapa constituye un tratamiento y está sujeta al cumplimiento del RGPD. En etapas posteriores del ciclo de vida de la solución IA, por ejemplo, cuando se integra en un tratamiento, hay que evaluar si se tratan datos personales para determinar si el tratamiento está sujeto al cumplimiento del RGPD, al menos con relación a la solución IA. Si se considera que no se tratan datos personales, por que estos se han eliminado o anonimizado, hay que demostrar que estos procesos han sido realmente efectivos y evaluar cuál es el riesgo de reidentificación"*[15].

Dicho esto, debe además tenerse en cuenta que generalmente una solución de IA se desarrolla para llevar a cabo alguna función dentro de un tratamiento o proceso, conformado por la solución de IA y otras muchas partes. En otras ocasiones los sistemas de IA son desarrollados separadamente del tratamiento o proceso, y posteriormente se integran a éste o incluso el sistema IA puede ser desarrollado y posteriormente integrado en distintos tratamientos o sistemas con diferentes responsables, fines y objetos. Por este motivo, aunque la solución IA no haya tratado datos durante su desarrollo, podrá hacerlo si el tratamiento o proceso al que se incorpora sí que trabaja con datos personales. Por todo ello, es muy importante que no solo los sistemas IA se

[14]Se trata de un sistema de control de calidad basado en el aprendizaje y la mejora del producto o proceso, es decir, la capacidad de aprender de lo observado. Por lo que se trata de un ciclo de mejoramiento continuo durante la vida del proceso o producto.

[15]Agencia Española de Protección de Datos, *"Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una Introducción"* febrero de 2020. Página 14 [En línea: https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia.pdf] [visitado el 28/06/2020]

ocupen de cumplir correctamente la normativa, sino que también deberán hacerlo los tratamientos que incorporan estos sistemas.

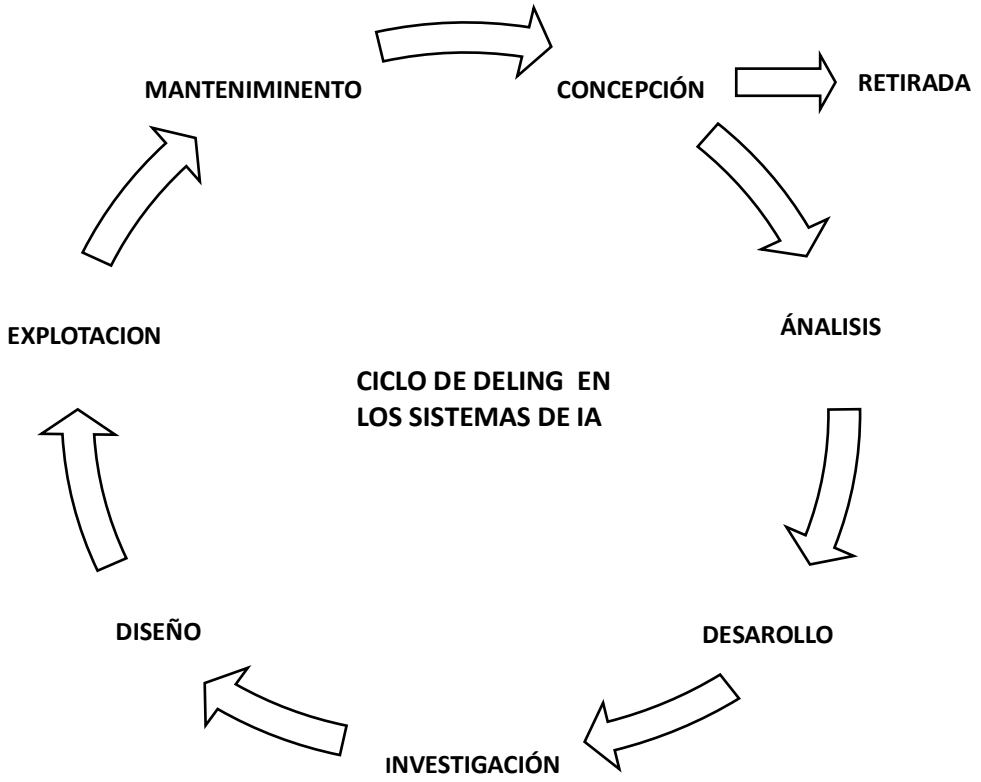

Figura 1º - Etapas básicas en el Ciclo de vida de un sistema de IA.
Fuente: *Elaboración propia*.

2. MARCO NORMATIVO.

El acelerado desarrollo de las nuevas tecnologías en los últimos diez años ha generado un gran volumen de datos, sin embargo, el 90% de ellos se han originado en los dos últimos años. Se estima que actualmente el volumen de datos está por encima del *Zettabyte*[16] y esta cifra crecerá exponencialmente año tras año. Este desmesurado volumen de datos ha provocado que los legisladores se vean en la necesidad de establecer medidas de amparo, protección y respeto a la privacidad de los individuos titulares de estos datos.

En la medida en que algunos Sistemas de IA tratan datos de carácter personal, analizándolos y hasta en algunas ocasiones tomando decisiones en base a estos, resulta necesario establecer cuáles son las condiciones generales que deben respetarse. Es decir, determinar cuáles son los principios y límites a los que debe sujetarse genéricamente el tratamiento de datos para después, en los apartados posteriores, concretar que especificidades presenta esta cuestión cuando se utilizan sistemas de IA como los descritos hasta el momento.

[16] Zettabyte (ZB) = 100000000000000000000 bytes = 10^{21} bytes.

Veremos entonces en este apartado cómo surge el derecho a la autodeterminación informática (*habeas data*) y cómo evoluciona a nivel internacional, europeo y nacional hasta llegar a ocupase de las cuestiones vinculadas con la IA.

2.1. La normativa sobre protección de datos a nivel europeo.

El derecho a la protección de los datos personales tiene sus orígenes en el derecho al respeto de la vida privada y la intimidad. Nació como una manifestación específica de éste. El derecho a la intimidad y al honor se reconocen por primera vez a nivel internacional en 1948 en el art.12 de la Declaración Universal de Derechos Humanos de las Naciones Unidas, donde se reconocen como derechos la intimidad y el honor. Posteriormente, en el art.8 del Convenio Europeo para la Protección de los Derechos Humanos y Libertades Fundamentales (CEDH) del año 1950, se estableció también el respeto por la vida privada de toda persona.

Sin embargo, esta correlación entre intimidad y derecho a la protección de datos personales no estuvo tan clara en un principio. Ciertamente ninguno de los instrumentos internacionales anteriormente citados hacía referencia explícita al derecho a la protección de datos como uno de los elementos imbricados en la intimidad. No fue hasta 1969, cuando empezaron a aparecer los primeros riesgos que entraña la informática y el tratamiento de datos personales, que la Asamblea Parlamentaria del Consejo de Europa planteó al Comité de Ministros la cuestión de sí el art.8 del CEDH resultaba suficientemente adecuado para proteger a los individuos del uso abusivo de la Informática. El órgano ministerial consideró que era insuficiente el precepto citado para tutelar adecuadamente el derecho a la protección de datos y sugirió que se implementara algún instrumento que amparase el derecho a la protección de los datos personales de forma específica.

Se aprueba así el primer instrumento internacional jurídicamente vinculante adoptado en el ámbito de la protección de datos, el Convenio n. º108 del Consejo de Europa, de 28 de enero de 1981, para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal. Ahora sí, se constituye éste como un derecho pariente del derecho a la vida privada, proclamando en su primer artículo, que el fin de la norma no es otro que *garantizar el respeto de los derechos y libertades fundamentales, concretamente su derecho a la vida privada, con respecto al tratamiento automatizado de los datos de carácter personal correspondientes a dicha persona.* Además de prohibir, si no se dan garantías jurídicas adecuadas, el tratamiento de los datos denominados "sensibles", tales como la raza, las opiniones políticas, la salud, la religión, la vida sexual o los antecedentes penales.

El año 1989, a raíz del *caso Gaskin*[17], el Tribunal Europeo de Derechos Humanos (TEDH) desautorizó al Consejo de Ministros al considerar que el derecho a la protección de los datos personales sí quedaba amparado por el artículo 8 de CEDH puesto que su contenido debía interpretarse en sentido amplio[18]. Según el TEDH esta interpretación se corresponde con el Convenio n. º108 del Consejo de Europa, de 28 de enero de 1981, para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, que viene a reforzar este derecho. A esta sentencia la siguieron muchas otras que terminaron por establecer definitivamente la relación entre ambos derechos.

[17]STEDH A 160 Gaskin v. Reino Unido, de 7 de julio de 1989 (1990) 12 E.H.R.R. 36 (n.° 37, p. 45)
[18]STEDH A 116 Leander v. Suecia, de 26 de marzo de 1987 (1987) 9 E.H.R.R. 433 (n.° 48, p. 450).

Simultáneamente, varios ordenamientos jurídicos europeos[19]y la legislación comunitaria hasta el momento así lo entendieron, y vincularon el derecho a la protección de los datos con el derecho a la vida privada y a la intimidad. La jurisprudencia comunitaria, sin embargo, eludió mencionar nunca esta conexión. Pese a los múltiplos pronunciamientos del TEDH, *el Tribunal de Justicia de la Unión Europea (TJUE) rehusó conectar expresamente el derecho a la protección de los datos personales con el que protege la vida privada[20] en ninguna de sus sentencias[21]*.

El derecho a la protección de los datos personales en la UE, tal y como lo conocemos actualmente, es un derecho independiente y autónomo que, a diferencia de otros, no fue introducido por la jurisprudencia, sino por la legislación. En concreto lo encontramos reflejado por primera vez en el artículo 16 del Tratado Fundacional de la Unión Europea (TFUE) y el posterior derecho derivado.

Con la firma del tratado de Maastricht y la creación de la Unión Europea (UE) el año 1993 empezó su trabajo de armonización legislativa. En materia de protección de datos, el año 1995 se aprobó la primera Directiva 95/46/CE del Parlamento europeo y el Consejo de 24 de octubre, sobre protección de datos y libre circulación de estos. Esta directiva vino a reproducir lo establecido en el artículo 16.1 TFUE, según el cual, toda persona tiene derecho a la protección de los datos de carácter personal que le conciernen. Además, y como veremos más adelante, esta Directiva fue transpuesta en el Ordenamiento Jurídico Español mediante la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal que entró en vigor en el año 2000.

Entrada la década de los dos mil, se producen tres sucesos que supusieron un cambio definitivo en la materia. En la Cumbre de Niza se aprueba la Carta de los Derechos Fundamentales de la Unión Europea en la que no solo se garantiza el respeto a la vida privada y familiar en el artículo 7, sino que, definitivamente, se reconoce el derecho a la protección de datos como derecho independiente en el artículo 8, digno de protección por sí mismo. En particular, éste se materializa en el Capítulo II, sobre las libertades, donde además de reconocerse este como derecho, se definen los principios sobre el tratamiento de lealtad, consentimiento, legitimidad y los derechos a acceso y rectificación de estos datos.

Simultáneamente, ese mismo año el TEDH en las Sentencias de los *casos Amann contra Suiza* y *Rotaru contra Rumanía[22]* reconoce que, pese a la estrecha relación con el derecho a la vida privada, el derecho a la protección de los datos de carácter personal, y, en definitiva, la autodeterminación informativa, debe tratarse como un derecho independiente.

El último suceso remarcable de ese año fue la aprobación del primer Reglamento de la UE en materia de protección de datos, dirigido las instituciones y organismos comunitarios, el Reglamento n.º 45/2001 de 18 de diciembre de 2000, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones y los organismos

[19]Como por ejemplo España, Finlandia y Holanda.

[20]RUIZ MIGUEL, C. *"*El derecho a la protección de los datos personales en la Carta de Derechos Fundamentales de la Unión Europea: análisis crítico*"*. *Revista de Derecho Comunitario Europeo*, 2003, n.º 7, n.º 14, páginas 13-18. [En línea: https://dialnet-unirioja-es.sire.ub.edu/servlet/articulo?codigo=635290] [visitado el 28/06/2020]

[21] Un ejemplo de esto es la Sentencia del Tribunal de Justicia, de 8 de abril de 1992 (C-62/90), Comisión v. Alemania [1992] Rec. 1-2575 (párrafo 23). En este importante caso el derecho a los datos personales no fue invocado directamente, pero era claramente subyacente y el tribunal eludió relacionarlo el derecho a la vida privada.

[22]SSTUDH. Amann c. Suiza, 16 febrero 2000, Rotaru c. Rumania, 5 mayo 2000. Ambos procesos dirigidos contra el almacenamiento de datos por las autoridades de esos países.

comunitarios y la libre circulación de estos datos. Tenía por finalidad garantizar en toda la UE una aplicación coherente y homogénea de las normas de protección de los derechos y las libertades fundamentales de las personas en lo que respecta al tratamiento de los datos personales a nivel público. El Reglamento n.º 45/2001 teniendo carácter prioritario sobre a la Directiva 95/46/CE, no la derogó por tener ámbito de aplicación distinto, constituyendo ambos, la normativa básica a seguir en materia de protección de datos.

En 2009[23] la UE empezó a cuestionarse la necesidad de establecer un marco jurídico en materia de protección de datos más sólido y coherente con el contexto tecnológico y globalizado, capaz de hacer frente a los nuevos riesgos surgidos en los últimos años. Una normativa que evitase la fragmentación entre ordenamientos jurídicos de los distintos estados miembros y que, de igual forma que la Directiva del 95/46/CE, consagrase *las ambiciones del proceso de integración europea: por una parte, la protección de los derechos y libertades fundamentales de las personas, en particular, del derecho fundamental a la protección de datos, y, por otra parte, la realización del mercado interior, es decir, en este caso, la libre circulación de datos personales*[24].

Se abrió un periodo de extenso debate y consulta que culminó en 2012, cuando la Comisión Europea emitió una comunicación donde se planteó, definitivamente, la necesidad de un cambio del *marco legislativo, que fuese sólido, coherente y que cubriese todas las políticas de la Unión, reforzara los derechos individuales, potenciara la dimensión de mercado único de la protección de datos y redujese los trámites burocráticos engorrosos para las empresas*[25]. La Directiva 95/46/CE era demasiado formalista y ya no podía dar respuesta a los importantes retos que estaban acompañando a la sociedad digital y el desarrollo tecnológico.

De esta manera se inició un proceso legislativo que finalizó con la publicación del Reglamento (UE) 2016/679 del Parlamento y Consejo Europeo, 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 en el Diario Oficial de la Unión Europea en fecha 4 de mayo de 2016.

Dos años después, en 2018, se aprobó el Reglamento (UE) 2018/1725 del Parlamento Europeo y del Consejo, de 23 de octubre de 2018, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones, órganos y organismos de la Unión, y a la libre circulación de esos datos, y por el que se derogan el Reglamento (CE) n.° 45/2001 y la Decisión n.° 1247/2002/CE. Este Reglamento se ocupa de la protección de los datos en el sector público, y acaba de completar, definitivamente, la normativa unificadora en materia de protección de datos que tanto anhelaba la UE.

[23]Se organizaron varias rondas de consultas públicas sobre la reforma de la protección de datos: la primera se desarrolló entre julio y diciembre de 2009 (http://ec.europa.eu/justice/news/consulting_public/news_consulting_0003_en.htm) y la segunda entre noviembre de 2010 y enero de 2011 (http://ec.europa.eu/justice/news/consulting_public/news_consulting_0006_en.htm). [visitado el 28/06/2020]

[24]Comunicación (COM (2010) 609 final) de la Comisión, de 4 noviembre 2010, "*Un enfoque global de la protección de los datos personales en la Unión Europea".*[En línea: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52010DC0609&from=es] [visitado el 28/06/2020]

[25]Comunicación (COM (2012) 9 final) de la Comisión, de 25 enero 2012, "*La protección de la privacidad en un mundo interconectado: Un marco europeo de protección de datos para el siglo XXI*". Página 4. [En línea: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52012DC0009&from=ES] [visitado el 28/06/2020]

El marco normativo que actualmente comprende la protección de los datos de carácter personal en relación con el estudio de este trabajo está formado por el Reglamento (UE) 2016/679 del Parlamento y Consejo Europeo, 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, en adelante RGPD, y se complementa con la Ley Orgánica 3/2018, de 5 de diciembre, de protección de datos personales y garantía de los derechos digitales, en adelante LOPDGDD. Juno a estos instrumentos generales también hay que tomar en consideración la normativa sectorial de aplicación vigente, publicada antes y después de la entrada en vigor del Reglamento.

2.2. El contenido del Reglamento (UE) 2016/679.

En este capítulo nos centraremos en hacer un breve análisis del derecho sobre protección de datos personales en áreas de afrontar adecuadamente el objeto de este trabajo, la gestión del riesgo de los sistemas de IA que incorporan datos personales. Desde 1986 España es un Estado Miembro de la UE y, en consecuencia, el Ordenamiento Jurídico de la UE es fuente de nuestro derecho. Por ello nos centraremos en el estudio del Reglamento (UE) 2016/679, puesto que, este es de aplicación directa, alcance general y es obligatorio en todos sus elementos para los estados miembros de la UE.

El **objeto** al que el legislador europeo aspira con el RGPD es el de establecer un marco normativo uniforme en todos los estados miembros, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y las normas relativas a la libre circulación de tales datos. Con el motivo de proporcionar el amparo suficiente a sus derechos y libertades fundamentales.

Este anhelo se materializa en el artículo 1 del RGPD en el que, simultáneamente, se proclama el primer principio de la norma por el cual "*la libre circulación de los datos personales en la Unión no podrá ser restringida ni prohibida por motivos relacionados con la protección de las personas físicas en lo que respecta al tratamiento de datos personales".* Este principio viene a declarar que la libre circulación de los datos no podrá ser limitada en ningún caso, ni por motivo de orden público. Por medio de este principio también se puede extraer la voluntad de la UE de proteger el mercado interior y, en el caso concreto, la libre circulación de los datos.

El principio de la libre circulación de los datos podría parecer a simple vista estar en contradicción con la primera parte del artículo 1, pero no es así. El derecho la protección de datos está tan relacionado con otros muchos derechos (libertad de información, libertad de expresión, libertad de empresa, protección de la propiedad…) que a menudo colisiona con ellos. Por este motivo, cunado esto suceda deberán ponderarse los derechos en conflicto y realizar un juicio de proporcionalidad. Ningún derecho es absoluto, por lo que, a priori, ninguno debe prevalecer sobre otro, y serán las circunstancias de cada caso las que determinarán cuál debemos proteger en cada momento.

El **ámbito de aplicación** se dispone en los artículos 2 y 3 del RGPD. Este se divide en dos campos, el ámbito de aplicación material y el territorial.

- **Ámbito de aplicación Material.**

Según el artículo 2 del RGPD el reglamento se aplicará a cualquier *tratamiento total o parcialmente automatizado de datos personales, así como al tratamiento no automatizado de datos personales contenidos o destinados a ser incluidos en un fichero.*

En la segunda parte del precepto se contemplan aquellos tratamientos de datos personales a los que la presente norma no les vinculará. Cabe destacar dos de estas excepciones. En primer lugar, como se ha mencionado brevemente en capítulos anteriores, el artículo 2.2.c) y su desarrollo en el artículo 18 RGPD declara que no se aplicará a aquellos tratamientos realizados por una persona física dentro de su esfera personal y, en general, fuera de su actividad profesional o comercial, como podría ser un sujeto que adquiere un dispositivo Google Home. Este asistente tecnológico es capaz de relacionarse con las personas para asistirla en una infinidad de actividades. Si el dispositivo recopila los datos para cumplir las expectativas del comprador y a su vez, mejorar sus funciones, sin además salir de la esfera personal o domestica del usuario, esta actividad no deberá atender al cumplimiento del Reglamento. En el caso que dispositivo Google Home recopile los datos para, además, enviarlos a la empresa Google, y esta los almacene, los incorpore en otros tratamientos, o hasta los comunique a un tercero, estaremos entonces, en el ámbito de aplicación del RGPD.

En segundo lugar, según el punto d) del mismo artículo, el reglamento tampoco se aplicará a *los tratamientos que lleven a cabo las autoridades competentes con fines de prevención investigación, detección o enjuiciamiento de infracciones penales, o de ejecución de sanciones penales, incluida la de protección frente a amenazas a la seguridad pública y su prevención.* El reglamento además recuerda que las instituciones, órganos y organismos, en lo que especta a los tratamientos que incorporan datos personales, están sujetas al Reglamento (CE) n.º 45/2001.

- **Ámbito de aplicación Territorial.**

En relación con las normativas anteriores, en este punto se produce un cambio sustancial. Así, el artículo 3 del RGPD se aplica al tratamiento de datos personales en el contexto de las actividades realizadas por un responsable[26] o encargado[27] con independencia de que el tratamiento se realice en la Unión. Además, el reglamento también se aplicará en aquellos tratamientos que, pese a que el responsable o encargado no resida en la Unión, los usuarios de esos datos sí que sean residentes de la Unión, cuando estos estén relacionados con: la oferta de bienes o servicios a dichos interesados en la Unión, independientemente de si a estos se les requiere su pago, o el control de su comportamiento, en la medida en que este tenga lugar en la Unión.

Llegados a este punto, creo necesario hacer un inciso y aclarar algunos conceptos a los que se alude en este reglamento y que son esenciales para abordar correctamente esta materia. El **dato personal** es definido por el artículo 4.1 del RGPD como:

> "*toda información sobre una persona física identificada o identificable («el interesado»); se considerará persona física identificable toda persona cuya identidad pueda*

[26] Según el artículo 4.7 del RGPD, *la persona física o jurídica, autoridad pública, servicio u otro organismo que, solo o junto con otros, determine los fines y medios del tratamiento; si el Derecho de la Unión o de los Estados miembros determina los fines y medios del tratamiento, el responsable del tratamiento o los criterios específicos para su nombramiento podrá establecerlos el Derecho de la Unión o de los Estados miembros.*

[27] Según el artículo 4.8 del RGPD, *la persona física o jurídica, autoridad pública, servicio u otro organismo que trate datos personales por cuenta del responsable del tratamiento*.

> *determinarse, directa o indirectamente, en particular mediante un identificador, como por ejemplo un nombre, un número de identificación, datos de localización, un identificador en línea o uno o varios elementos propios de la identidad física, fisiológica, genética, psíquica, económica, cultural o social de dicha persona"*

Esta definición es desarrollada con más profundidad por el Grupo de Trabajo del Artículo 29 (GT Art. 29)[28], en su Dictamen[29] 4/2007, sobre el concepto de datos personales. En términos generales el dictamen arguye que el legislador europeo se propuso adoptar un concepto amplio, aunque no debe interpretarse como ilimitado. Por este motivo el GT Art. 29 cree necesario analizar el significado dividiéndolo en cuatro conceptos fundamentales, y acotar el significado de cada uno. Estos son, "toda información", "sobre", "identificada o identificable" y "persona física".

Para que estos datos puedan ser tratados, según el RGPD, será estrictamente necesario que se cumpla las condiciones de licitud del artículo 6, y en concreto, es importante mencionar el consentimiento del artículo 7. La incorporación de este concepto ha supuesto uno de los cambios más significativos del Reglamento respecto a la Directiva 95/46/CE. Actualmente el consentimiento[30] debe proporcionarse informada e inequívocamente. El interesado debe aceptar mediante una declaración o una clara acción afirmativa, ya no es suficiente deducir éste del silencio o de la inacción.

Además, el reglamento distingue entre dos categorías distintas de datos, los generales y los especiales. El tratamiento de los **datos de carácter especial** necesitará además observar la licitud del artículo 6, cumplir lo establecido en el artículo 9 del RGPD. El legislador europeo encontró pertinente proporcionar una especial protección a estos últimos por considerarlos especialmente sensibles. Estos datos son aquellos que *revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos y datos biométricos[31] dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida o las orientaciones sexuales de una persona física.* El RGPD establece la prohibición genérica de su tratamiento salvo cuando se den las circunstancias descritas en los apartados 2, 3 y 4 del mismo artículo. En concreto, el consentimiento de la cesión de estos datos es más estricto, añadiendo que este deberá ser "explícito", mediante algún tipo de acción positiva. Sin embargo, hay supuestos en los que el tratamiento de estos datos especiales no requerirá del consentimiento de los titulares. Por ejemplo, los apartados g) y i) del artículo 9.2 del RGPD permiten la trata de esos datos sensibles cunado el tratamiento es necesario para garantizar el interés público esencial o cuando este sea necesario

[28] El Grupo de Trabajo del artículo 29 fue un ente consultivo independiente introducido por la Directiva 95/46/CE integrado por las Autoridades de Protección de Datos de todos los Estados miembros, el Supervisor Europeo de Protección de Datos y la Comisión Europea. Este se ocupaba de cuestiones relacionadas con la protección de la privacidad y los datos personales hasta el 25 de mayo de 2018 que entró en vigor el RGPD y lo sustituyo por el *European Data Protection Board (EDPB*). [En línea: https://edpb.europa.eu/about-edpb/about-edpb_en] [visitado el 28/06/2020]

[29] Grupo de Trabajo del artículo 29, *Dictamen 4/2007 sobre el concepto de datos personales WP 136*, de 20 de junio de 2007. [En línea: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_es.pdf][visitado el 28/06/2020]

[30]Grupo de Trabajo del artículo 29, *Directrices sobre el consentimiento en el sentido del Reglamento (UE) 2016/679 WP259*, 28 de noviembre de 2017. [En línea: https://apdcat.gencat.cat/web/.content/03-documentacio/Reglament_general_de_proteccio_de_dades/documents/wp260rev01_es-transparencia.pdf] [visitado el 28/06/2020]

[31] El reconocimiento de estos datos como sensibles constaba en nuestra constitución y en la Directiva 95/46/CE, pero con la entrada en vigor del reglamento se añadieron dos tipos más, los datos genéticos i biomédicos.

para evitar peligros graves en el ámbito de la salud pública. En ambos casos, el tratamiento de esos datos sin el consentimiento de los titulares deberá realizarse una vez ponderada su necesidad (juicio de proporcionalidad).

A menudo estos datos surgirán derivados o inferidos de otros datos que no son catalogados como especiales. Cuando esto suceda, también tendrá que respetarse lo dispuesto en el artículo 9. Por ejemplo, podrían extraerse datos derivados de la salud de un grupo de personas mediante el estudio de su compra en el supermercado comparados con datos sobre la calidad y el contenido energético de los alimentos.

Por último, resulta conveniente examinar otro concepto estrechamente relacionado con el objeto de la norma y esencial para su adecuado análisis. El **tratamiento de los datos personales y sus distintas formas**. Un tratamiento, según el artículo 4.2 del RGPD, es *cualquier operación o conjunto de operaciones realizadas sobre datos personales o conjuntos de datos personales, ya sea por procedimientos automatizados o no, como la recogida, registro, organización, estructuración, conservación, adaptación o modificación, extracción, consulta, utilización, comunicación por transmisión, difusión o cualquier otra forma de habilitación de acceso, cotejo o interconexión, limitación, supresión o destrucción.*[32]

Este concepto no debe confundirse con el de **perfilado o la elaboración de perfiles**, puesto que, pese a ser una forma de tratamiento, es preciso darle una atención especial. En general, el perfilado implica la recogida de información sobre una persona o colectivo de personas para la posterior evaluación de sus características o determinación de sus patrones de comportamiento con el fin de catalogarla. Así, el artículo 4.4 del RGPD define el perfilado como *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*.

Luego, ¿qué es **decisión automatizada**? Este concepto no se encuentra definido como tal en el reglamento, por este motivo, el GT Art. 29 elaboró un Dictamen[33] sobre las decisiones individuales automatizadas y la elaboración de perfiles a los efectos del Reglamento 2016/679. En este trabajo es definida como la capacidad de tomar decisiones por medios tecnológicos sin la participación del humano. Las decisiones automatizadas pueden basarse en cualquier tipo de datos, ya sean datos proporcionados por las personas, datos observados directamente[34] o datos derivado o inferidos.

Ha de tenerse en cuenta que las decisiones automatizadas pueden llevarse a cabo con o sin elaboración de perfiles y a su vez la elaboración de perfiles puede llevarse a cabo mediante decisiones automatizadas o no. Las decisiones automatizadas tienen un ámbito de aplicación distinto y pueden solaparse parcialmente con la elaboración de perfiles o derivarse de esta.

[32]Este artículo constituye un listado *numerus apertus*, por lo que podemos encontrarnos tratamientos que no se citen en él y sí que deban regirse por la normativa de Protección de datos comunitaria.
[33]Grupo de Trabajo del artículo 29, *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679 (WP251)*, de 06 de febrero de 2018. [En línea: https://www.aepd.es/sites/default/files/2019-12/wp251rev01-es.pdf] [visitado el 28/06/2020]
[34] Como, por ejemplo, la geolocalización mediante las aplicaciones de los móviles

El artículo 22 del RGPD se dirige únicamente a los tratamientos automatizados estipulando que todo *interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar*. Pese a estar redactado como un derecho, al ponerlo en relación con el n.º 71 de la Exposición de Motivos nos damos cuenta de que en realidad debe ser tratado como una prohibición general. Un imperativo legal por el cual quedan prohibidas las decisiones basadas únicamente en tratamiento automatizado que producen efectos jurídicos y que afectan significativamente a los propietarios de esos datos.

Pese a lo anterior, esta exclusión se romperá si se cumple alguna de las excepciones del apartado 2 del mismo artículo: Cuando se celebre un contrato, cuando esté autorizada por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento y que establezca asimismo medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, o se base en el consentimiento explícito del interesado. Respecto a este último, el responsable deberá acreditar que los interesados entienden exactamente para qué están consintiendo.

Cuando la decisión incida en las categorías de datos del artículo 9 del RGPD, el responsable además deberá garantizar que se hayan tomado medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado.

Por último, es necesario referirnos a **los principios** proclamados en el artículo 5 del RGPD que vinculan a todos los tratamientos de datos personales, y, por lo tanto, inspiran todo el Reglamento. En esta materia se han introducido dos novedades respecto a la regulación contenida en la Directiva 95/46/CE. Se incluyen dos principios más, el principio de transparencia del tratamiento, requisito fundamental del RGPD, y el principio de responsabilidad.

- **Los principios de Licitud, lealtad y transparencia.**

El apartado 1º del artículo 5 del RGPD establece que los datos personales serán tratados de manera lícita, leal y transparente en relación con el interesado. El primer paso, entonces, para poder iniciar un tratamiento es instituir una base jurídica legitimadora. Ésta debe determinarse en la fase de concepción del tratamiento y en el caso de los sistemas de IA, ésta podrá ir mutado durante el transcurso del ciclo de la vida, pudiendo ser distinto en cada etapa.

Para que un tratamiento sea lícito deberá cumplir al menos alguna de las condiciones del artículo 6.1 del RGPD:

- El tratamiento es necesario para la ejecución de un contrato en el que el interesado es parte, o para la aplicación de medidas precontractuales a petición de éste.
- Cumplir una obligación legal aplicable al responsable del tratamiento[35]. Por ejemplo, la que puede efectuar un empleador para obtener cobertura de la seguridad social de sus trabajadores,

[35]Otro ejemplo, menos común, pero con importante interés respecto a la situación que España está viviendo por el Covid-19, sería la trata y comunicación de datos que debe hacer el empleador en la prevención de riesgos laborales a fin de evitar contagios. No obstante, no todo está permitido, con lo que no todos los tratamientos podrán justificarse en este precepto, y en particular, aquellos que traten datos sensibles. Será esencial que se presten las garantías necesarias de seguridad, información, limitación de finalidad. La AEPD expresado especial preocupación por la toma de temperatura masiva que están efectuando las empresas tanto a sus trabajadores como clientes, al ser estos datos, además, catalogados según el art.9 del

puesto que la legislación le obliga a proporcionar datos personales de estos a la autoridad pertinente.

- Proteger intereses vitales del interesado o de otra persona física. Según el n.º46 de la Exposición de Motivos del RGPD, los datos personales únicamente deben tratarse sobre la base del interés vital de otra persona física cuando el tratamiento no pueda basarse manifiestamente en una base jurídica diferente.

- El tratamiento es necesario para el cumplimiento de una misión realizada en interés público[36] o en el ejercicio de poderes públicos conferidos al responsable del tratamiento. Para acogerse a esta base jurídica será necesario efectuar *ex ante* un juicio de ponderación ente los derechos que se ven implicados antes de determinar la limitación de alguno de ellos.

- La satisfacción del interés legítimo[37], siempre que sobre dichos intereses no prevalezcan los intereses o los derechos y libertades fundamentales. El interés legítimo es una alternativa de legitimación del tratamiento que en todo caso exigirá al responsable un mayor grado de compromiso, formalidad y competencia. Si el tratamiento se basa en un interés legítimo, no será necesario recabar el consentimiento de los titulares de los datos, pero si deberá cumplir los deberes de información[38], comunes para todos los tratamientos.

- O, como se ha mencionado antes, el consentimiento de los interesados.

El principio de lealtad supone la persistencia durante todo el tratamiento o proceso de un compromiso con los derechos y libertades de las personas implicadas y el beneficio que éste les pueda a portar a ellas directamente o al conjunto de la sociedad. Encarna la honestidad del tratamiento. Luego, si en algún momento de éste dejase de ser beneficioso para para los titulares de los datos, y, por lo tanto, hubiese un conflicto con sus derechos fundamentales, por ejemplo, el tratamiento fuese discriminatorio, sería desleal.

Por otra parte, el principio de transparencia, desarrollado en la Exposición de Motivos 39, 58 y 78, supone el derecho que tiene el titular de los datos a supervisar el tratamiento al que están sometidos. En particular, la transparencia está ligada con la información de los artículos 13 y 14[39],

RGPD. [En línea: https://www.aepd.es/es/prensa-y-comunicacion/notas-de-prensa/comunicado-aepd-temperatura-establecimientos] [visitado el 28/06/2020]

[36] La lucha contra de la pandemia del Covid-19 ha generado un fuerte control sobre los ciudadanos mediante la trata de gran cantidad de datos por sistemas dotados de IA en diversos países alrededor del mundo como son China o Corea del Sur. El importante éxito que han tenido estos sistemas ha conducido al Gobierno Español a aprobar la Orden SND/297/2020 por la que se pretende geolocalizar a la población utilizando IA dotada de Machine Learning, así como también programas de detección automática de síntomas mediante datos denominados sensibles, legitimados por el Interés público de los artículos 6.1 e) y el artículo 9.2 g), h), i), y j) del RGPD. Esta cuestión ha suscitado importantes dudas en algunos sectores de la sociedad. En este sentido es interesante la *Nota en relación a los tratamientos de datos personales relacionados con las medidas para hacer frente al COVID-19* que publicó la ACPD el 15 de marzo de 2020. [En línea: https://apdcat.gencat.cat/es/actualitat/noticies/noticia/Nota-en-relacio-amb-els-tractaments-de-dades-personals-relacionats-amb-les-mesures-per-fer-front-al-COVID-19] [visitado el 28/06/2020]

[37] El concepto de Interés legítimo ha sido ampliamente trabajado por el GT del Art. 29 en el *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*, de 9 de abril de 2014. [En línea: https://www.aepd.es/sites/default/files/2019-12/wp217_es_interes_legitimo.pdf] [visitado el 28/06/2020]

[38] Según el artículo 12.1 del RGPD, el responsable del tratamiento debe facilitar a los interesados información concisa, transparente, inteligible y de fácil acceso sobre el tratamiento de sus datos personales.

[39] El deber de información de los datos que no se hayan obtenido directamente del interesado no será necesario según el artículo 14.5 del RGPD cuando: el interesado ya posea la información, cuando el registro o la comunicación de los datos personales estén expresamente establecidos por ley, o cuando facilitar la

la eficacia y las capacidades y limitaciones del tratamiento de estos datos o, en este caso, del sistema de IA. Constituye el deber de proporcionar información sobre los datos que están siendo tratados. El mandato vinculará tanto a los responsables que estén tratando datos recabados directamente del titular como si los datos se obtienen de otra fuente. En los tratamientos que incorporan IA, el cumplimiento del deber de transparencia es cíclico cuando este llega a la fase de evolución.

La incorporación de este último principio supone reforzar este deber de información además de augmentar la cantidad de ésta que el responsable tiene obligación de facilitar.

- **El principio de Limitación de la Finalidad.**

Una base jurídica lícita, leal y transparente no habilita para el uso de datos para cualquier propósito. Es más, es imprescindible definir un propósito, no siendo posible tratar datos para propósitos inciertos. Una vez acotados los fines, estos deberán conservarse durante el transcurso del tratamiento, o en el caso de sistema de IA, durante todo el ciclo de vida de este.

- **El principio de Minimización de los datos.**

Este es el proceso orientado a garantizar que los datos personales son adecuados, pertinentes y limitados a lo necesario en relación con el fin o fines del tratamiento. Además, según el texto Preliminar n.º 39 los datos personales sólo deben tratarse si la finalidad del tratamiento no puede lograrse razonablemente por otros medios.

La minimización persigue optimizar el tratamiento desde el punto de vista de la protección de los datos, analizado las necesidades de proceso de distintos datos en las distintas fases[40] En definitiva, la minimización consiste en: limitar la extensión de las categorías de datos[41] que se utilizan en cada fase del tratamiento a aquellas que son estrictamente necesarias y relevantes, limitar el grado de detalle o precisión de la información, limitar la extensión en el número de interesados de los que se tratan los datos, y por último, limitar la accesibilidad de las distintas categorías de datos al personal del tratamiento.

En concreto, para los tratamientos de IA existen distintas técnicas de minimización ampliamente desarrolladas por la AGEP en Guía de Privacidad desde el diseño, en particular algunos ejemplos comúnmente usados son:

- La identificación y supresión, durante el entrenamiento, de aquellos datos que no son útiles en el aprendizaje del sistema de IA.
- Supresión de aquella información recogida que no es necesaria.
- La aplicación de estratégicas de privacidad diferencial[42].

información al interesado resulte imposible o exija un esfuerzo desproporcionado. Luego, nos podríamos encontrar con una situación así si el tratamiento se realizara con fines de archivo en interés público, fines de investigación científica o histórica o fines estadísticos.

[40] Agencia Española de Protección de Datos, "*Adecuación al RGPD…"*, ob-cit. página 38 a 39.

[41] Extensión de la categoría de datos se refiere a el número de campos de datos asociados a una persona física: nombre, direcciones físicas y lógicas, campos sobre su salud, situación laboral, social, relaciones, gustos, creencias, ideología, …

[42] Según la AEPD las estratégicas de privacidad diferencial son patrones que modifican el resultado de las consultas añadiéndoles nuevos datos (ruido) extraídos aleatoriamente de una distribución generada a partir de los datos originales de modo que estadísticamente dicha modificación tiene un efecto insignificante en los resultados del algoritmo que analiza los datos y sin embargo permite preservar la privacidad de los

- El entrenamiento con datos cifrados utilizando técnicas Homomórficas[43].
- Una de las más usadas en la praxis es la Anonimización y Seudonimización[44], no sólo en la comunicación de datos, sino también en los datos usados durante el entrenamiento.

- **El principio de Exactitud**.

Los responsables deberán tener presente el principio de exactitud durante todas las fases del tratamiento o del ciclo de la vida del sistema de IA. Esto quiere decir que deberán revisar, y responsabilizarse de que los datos que intervienen en cada etapa del ciclo no son inexactos. Las inexactitudes pueden provocar la toma de soluciones o decisiones incorrectas, que, si afectan a personas físicas, pueden inferir en sus derechos y libertades.

Los responsables deberán introducir medidas para garantizar de modo continuo que los datos tratados son precisos y aquellos datos que se hayan obtenido de forma derivada o inferida sean veraces y estén actualizados.

- **El principio de Limitación del plazo de conservación**

Aunque preservar datos puede tener muchas ventajas, los tratamientos que incorporan datos personales no podrán disponer de ellos infinitamente. Así, el RGPD articula el principio de limitación de conservación de los datos personales por el cual no se podrá disponer de los datos durante un tiempo superior al necesario para alcanzar los fines definidos originalmente. Del mismo modo el Preámbulo nº39 del RGPD contempla este principio y aclara que el responsable del tratamiento establecerá plazos para su supresión o revisión periódica. No obstante, Los artículos 5.1 b) y 89.1 del RGPD permiten conservar estos datos más allá de lo necesario *siempre que se traten exclusivamente con fines de archivo en interés público, fines de investigación científica o histórica o fines estadísticos.*

- **El principio de Integridad y confidencialidad**

Los datos deberán ser tratados de tal manera que se garantice una seguridad adecuada, incluida la protección contra el tratamiento no autorizado o ilícito y contra su pérdida, destrucción o daño accidental, mediante la aplicación de medidas técnicas u organizativas apropiadas.

En este sentido el nº49 del Preámbulo refuerza este principio al señalar que el responsable deberá velar por que el tratamiento de datos personales se lleve a cabo en una medida estrictamente necesaria y proporcionada.

- **El principio de la Responsabilidad proactiva.**

El principio de responsabilidad proactiva es un ámbito esencial del RGPD, una de las mayores novedades respecto a la Directiva 95/46/CE. Este resultará un concepto central en capítulos ulteriores, y fundamental para abordar el análisis de la gestión del riesgo de los sistemas de IA.

---

individuos. "*Guía de Privacidad desde el diseño"* de octubre de 2019. página 24 [En línea: https://www.aepd.es/sites/default/files/2019-11/guia-privacidad-desde-diseno.pdf][visitado el 28/06/2020]

[43] Según la AEPD las técnicas homomórficas son operaciones realizadas sobre la información cifrada donde es posible realizar tratamientos trabajando con datos agregados sin tener acceso a la información individual. *Ídem.*

[44] El artículo 4.5 define la seudonimización como *el tratamiento de datos personales de manera tal que ya no puedan atribuirse a un interesado sin utilizar información adicional, siempre que dicha información adicional figure por separado y esté sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos personales no se atribuyan a una persona física identificada o identificable.*

El artículo 5 del RGPD sitúa la responsabilidad proactiva como principio que rige el reglamento y es posteriormente desarrollada por el artículo 24 y 25. Así, el Reglamento describe este principio en el n.º 74 de la Exposición de Motivos, como la necesidad que tiene el responsable del tratamiento de buscar las estrategias que le permitan el cumplimiento del reglamento mediante el diseño de medidas técnicas y organizativas apropiadas a fin de garantizar y poder demostrar que el tratamiento es conforme al RGPD. De tal manera que cuando hablamos de responsabilidad proactiva nos referimos a dos conductas indispensables: la necesidad de que el responsable del tratamiento adopte medidas eficaces para aplicar los principios del RGPG y proteger los derechos y libertades de los titulares de los datos y la necesidad de demostrar que se han adoptado estas medidas.

Una política de responsabilidad proactiva supone la adopción de medias que se anticipen a las amenazas, identificando las debilidades del sistema, con el objetivo de neutralizar o minimizar los riesgos en lugar de centrarse en el diseño de medidas correctivas. Esta reconfiguración de la responsabilidad en los tratamientos de datos huye de la política de subsanación y pretende adelantase a ésta para que simplemente, no suceda.

Además, el artículo 35 del RGPD refuerza esta responsabilidad cuando el tratamiento utilice nuevas tecnologías que entrañen un alto riesgo para los derechos y libertades de las personas físicas mediante la obligación de realizar una evaluación ex ante sobre el impacto que las operaciones pueden tener. Cuando estos tratamientos sean automatizados de toma de decisiones y se utilizan datos personales de personas físicas que puedan afectarles significativamente deberá realizarse una evaluación ex ante sistemática y exhaustiva de los riesgos que entrañan estas decisiones.

En el Capítulo IV del Reglamento se desarrolla el modelo de responsabilidad y cumplimiento que debe llevarse a cabo en un tratamiento cuyos elementos rectores según la AEPD[45] son:

a) La identificación de una responsabilidad en el tratamiento

b) El análisis del riesgo de los derechos y libertades

c) El estudio de la necesidad y la proporcionalidad de las operaciones del tratamiento con respeto a la finalidad inicialmente reconocida

d) El despliegue de medidas de gestión del riesgo, medidas de privacidad por defecto des del diseño, medidas de seguridad, de gestión de incidentes, etc.

Una importante novedad en esta área, y que está estrechamente relacionada con la implementación de la responsabilidad proactiva en el Reglamento, es que se elimina la obligación de comunicar el tratamiento de datos personales a las autoridades de control. Ahora será el responsable quien tendrá la obligación de llevar un registro de actividades que deberá contener una información de mínimos. El responsable, para demostrar el cumplimiento de todas las obligaciones que establece el Reglamento deberán hacer uso de códigos de conducta y mecanismos de certificación según disponen los artículos 40 a 43 del RGPD.

[45]Agencia Española de Protección de Datos, "*Adecuación al RGPD…"*, ob-cit. página 38-40

La Figura nº.2 articula brevemente cómo funcionan algunos de estos principios aplicados a un tratamiento que incorpora datos personales:

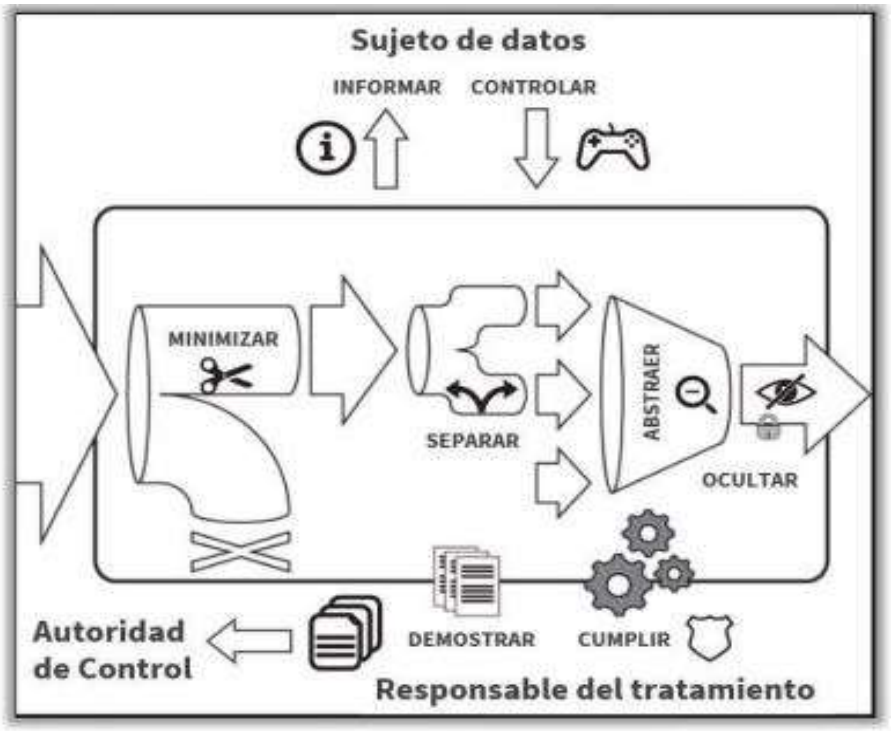

Figura 2º - Cumplimiento desde el diseño.
Fuente: *Gaia de Privacidad des del diseño. (AEPD). Octubre del 2019*

El RGPD establece un nuevo marco europeo en materia de protección de datos basado en unos principios tradicionales, pero cambia radicalmente la forma de afrontar el cumplimento de estas bases. Un giro fundamental sobre la base de la responsabilidad proactiva y la gestión del riesgo, en que cada entidad tendrá que valorar como afecta ese tratamiento a los titulares de esos datos, obligándolas a plantearse cuál es la manera más adecuada para reducir ese impacto, asegurado el buen cumplimiento de la normativa y de sus principios.

2.3. Evolución normativa española en materia de datos personales.

El Constituyente en 1978 se anticipó a los peligros y abusos que podían surgir de las nuevas tecnologías y, del mismo modo que lo hará posteriormente el TEDH en el *caso Gaskin*, interpretó el derecho a la protección de los datos personales como una consecuencia o derivación del derecho a la intimidad[46]. Así este se introdujo en el artículo 18.4 de la CE estableciendo que "*La Ley limitará el uso de la informática para garantizar el honor y la intimidad personal y familiar de los ciudadanos y el pleno ejercicio de sus derechos".* Este precepto fue redactado precisamente,

[46]La Exposición de motivos de la LO 5/1992 define de forma muy acertada el derecho a la intimidad como aquel que "*protege la esfera en que se desarrollan las facetas más singularmente reservadas de la vida de la persona -el domicilio donde realiza su vida cotidiana, las comunicaciones en las que expresa sus sentimientos, por ejemplo-, la privacidad constituye un conjunto, más amplio, más global, de facetas de su personalidad que, aisladamente consideradas, pueden carecer de significación intrínseca pero que, coherentemente enlazadas entre sí, arrojan como precipitado un retrato de la personalidad del individuo que éste tiene derecho a mantener reservado".*

cuando comenzaban a surgir las primeras alarmas sobre los riesgos que podía entrañar el archivo y uso ilimitado de datos, siendo este artículo la primera medida en materia de protección de datos en nuestro derecho interno.

Es preciso que la interpretación del contenido del artículo 18.4 de la CE comprende también la obligación del Estado de apoyar y ofrecer medios a los operadores jurídicos para impulsar este sector de la forma más segura posible, tal como apunta Martínez[47]:

> *"Hoy el precepto debe interpretarse en el sentido de atribuir al legislador la tarea de ordenar el adecuado uso de la informática. Y ordenar no solo implica limitar, sino que muchas ocasiones podría llegar a significar una labor de impulso".*

Pese a lo pionero de la previsión constitucional, el art. 18.4 CE no fue bien acogido por la doctrina ni posteriormente por la jurisprudencia.[48] Al tiempo también fue criticado por su ambigüedad y por vincularlo al derecho al honor, a la intimidad personal y familiar de los ciudadanos. No se le otorgó un estatus independiente, aunque también se encontraba emparentado con el derecho a la dignidad humana del artículo 10.1 de la CE, sin embargo, se prefirió incluirlo en el ámbito de la intimidad, como una parte del derecho a *la privacy* (*habeas data*).

Es por este motivo, que la doctrina española, inspirada en la jurisprudencia del Tribunal Constitucional Alemán, empieza a configurar lo que hoy en día es conocido como la autodeterminación informativa (*habeas data)*. Un derecho fundamental que pese a estar íntimamente ligado con el derecho del artículo 18 de la CE, debe ser tratado de forma autónoma. Este concepto pretende *satisfacer la necesidad, sentida por las personas en las condiciones actuales de la vida social, de preservar su identidad controlando la revelación y el uso de los datos que les conciernen y protegiéndose frente a la ilimitada capacidad de archivarlos, relacionarlos y transmitirlos propia de la informática, y de los peligros que esto supone*.[49] Lucas Murillo ha definido la autodeterminación informativa como:

> *"el control que a cada uno de nosotros nos corresponde sobre la información que nos concierne personalmente, sea íntima o no, para preservar de este modo y en último extremo la propia identidad, nuestra dignidad y libertad. En su formulación como derecho, implica necesariamente poderes que permitan a su titular definir los aspectos de su vida que no sean públicos, que desea que no se conozcan, así como facultades que le aseguren que los datos que de su persona manejan terceros informáticamente son exactos, completos y actuales, y que se han obtenido de modo leal y lícito"*[50]

[47]MARTÍNEZ, R. Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo. *Revista Catalana de Dret Públic*. 2019. n. º58. pág. 64-81. [En línea: https://doi.org/10.2436/rcdp.i58.2019.3317] [visitado el 28/06/2020]

[48]Una interpretación estricta puede llevar a considerar este derecho como una especificación del derecho a la intimidad, no obstante, el Tribunal Constitucional ha interpretado reiteradamente que se trata de un derecho autónomo, pese a estar estrechamente relacionado con el derecho a la intimidad (*SSTC 254/1993, de 20 de julio, 202/1999 de 8 de noviembre, 290/2000, de 30 de noviembre y 292/2000 de 30 de noviembre*).

[49] LUCAS MURILLO DE LA CUEVA, P. "*El derecho a la autodeterminación informativa*" Madrid: Tecnos. 1991. Págs. 173-174. ISBN**:** 84-309-1937-6

[50] LUCAS MURILLO DE LA CUEVA, P. "*Informática y protección de datos personales (estudios sobre la Ley Orgánica 5/1992 de Regulación del Tratamiento Automatizado de los Datos de Carácter Persona*l)". Madrid: Centro de Estudios Políticos y Constitucionales. 1993. Págs. 32 y 51. ISBN: 84-259-0940-6

Habrá que esperar al año 1992 para que se apruebe la Ley Orgánica 5/1992, de 29 de octubre, de Regulación del Tratamiento Automatizado[51]de los Datos de Carácter Personal (LORTAD). Esta nueva previsión legal vino impuesta, no tanto por las amenazas existentes del rápido desarrollo de las nuevas tecnologías ni tampoco por los reiterados pronunciamientos de la doctrina española sobre la autodeterminación informativa, sino, más bien, por la ratificación el año 1990 del Convenio Schengen. Precisamente, para que España pudiese entrar en ese grupo de países se imponía a los que querían incorporarse, el control de ciertas bases de datos. *De no haber sido por esta circunstancia, posiblemente habíamos tenido que esperar mucho más tiempo, y se habría perpetuado una situación en la que la única previsión normativa sobre la materia era la que ofrecía la disposición transitoria primera de la Ley Orgánica1/1982, de 5 de mayo, de protección civil del derecho al honor, a la intimidad personal y familiar y a la propia imagen*.[52]

Lo importante es que la Ley Orgánica 5/1992, reguló por primera vez en España la protección de los datos de carácter personal, proclamando en la Exposición de Motivos, la existencia de un nuevo derecho fundamental y la autodeterminación y consentimiento como como principios sobre los que descansaba esa regulación. Además, se creó la Agencia Española de Protección de Datos (AEPD)[53] como garantía de la protección de estos derechos y el cumplimento de la normativa en materia de protección de datos, que comenzó a funcionar en 1994.

Unos meses después de la entrada en vigor de la LORTAD y antes del pronunciamiento del TEDH en los *casos Amann contra Suiza* y *Rotaru contra Rumanía*, el Tribunal Constitucional, que hasta entonces no se había pronunciado al respecto, acogió el planteamiento doctrinal español en su Sentencia 254/1993 de 20 de julio. En ésta, el Tribunal se refirió a una nueva garantía constitucional a la que denominó "libertad informática" que poco a poco, fue perfilando en sucesivas sentencias. Así se abrió una línea jurisprudencial que culminó con la STC 292/2000 de 30 de noviembre, que instauró explícitamente el derecho fundamental a la protección de datos de carácter personal como categoría autónoma y distinta del derecho a la intimidad[54] y, además, recalcó la estrecha relación que existe entre el Convenio n. º108 y el artículo 18.4 de la CE llenando este último de contenido.

El año 1995 la UE aprueba la primera Directiva en materia de protección de datos, la Directiva 95/46/CE. El legislador español, al ser España un estado miembro estaba obligado a la transposición de ésta, surgiendo así la Ley Orgánica 15/1999, del 13 de diciembre, de Protección de Datos de Carácter Personal (LPD). En un inicio se optó por una mera reforma de la LORTAD, pero finalmente se decantaron por la derogación de la Ley Orgánica 5/1992, y la aprobación de una ley que fuera más allá que la directiva, dotando de un nivel superior de protección a los sujetos.

Finalmente, y como se ha visto en el capítulo anterior, en 2016, la UE aprobó el RGPD que entró en vigor el 25 de mayo de 2018. Este reglamento ordena la normativa sobre protección de datos personales y, además, pretende unificar las legislaciones de los estados miembros en esta materia. Las características de este instrumento legislativo lo dotan de alcance general y eficacia directa,

[51]Hay que tener presente que las decisiones automatizadas no tienen porque siempre desarrollarse mediante soluciones IA, y que no toda la IA está compuesta de un sistema de toma de decisiones automatizadas.

[52]LUCAS MURILLO DE LA CUEVA, P. "La Constitución y el derecho a la autodeterminación informativa". *Cuadernos de derecho público.* 2003. n. º19-20. Páginas 27-44. ISSN 1138-2848

[53] En 2010 y mediante mandato del artículo 65 del Estatuto de Autonomía de Catalunya, se aprobó la Ley 32/2010, de 1 de octubre, de la Autoridad Catalana de Protección de Datos. Así surgió la autoridad de control autonómica, Autoridad Catalana de Protección de datos y, que, a día de hoy, sigue vigente.

[54]STC 292/2000 de 30 de noviembre. FJ n.º 5. pág. 15. [En línea: http://hj.tribunalconstitucional.es/es-ES/Resolucion/Show/4276] [visitado el 28/06/2020]

por lo que no requerirá de su transposición para generar derechos y obligaciones. No obstante, y si éste lo permite, los Estados podrán transponerlo a su ordenamiento jurídico en la forma que crean conveniente, ahora bien, sin modificar su contenido. El RGPD unas veces impone y otras habilita a los Estados miembros a concretar determinadas materias. En este sentido, la nota de aplicabilidad directa que rige los reglamentos no impide que, en aras de depurar el ordenamiento interno y de complementar el RGPD, los Estados miembros aprueben normas sobre la materia, como así han hecho la mayoría de los estados miembros.

De esta manera, en 2018 se aproó la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD). Parece que el legislador español optó por una práctica legislativa ciertamente cuestionable, al incluir en el mismo instrumento jurídico de protección de datos personales, la protección a los derechos digitales. *Sin perjuicio de la relación que hay entre ambos, dada su importancia, deberían haberse contemplado en una norma independiente*[55]. En definitiva, esta nueva normativa dejó sin efecto la LPD y el Reglamento que la asistía.

Las principales novedades incorporadas en la LOPDGDD respecto al RGPD y, por supuesto, a las que se ha habilitado al legislador son:

1º- Se excluye en el artículo 2.2 b) del ámbito de aplicación de la LOPDGDD a las personas fallecidas, pero se permite que las personas vinculadas al fallecido, por razón familiar o de hecho o sus herederos puedan solicitar el acceso a esos datos para su rectificación o supresión, tal como estipula extendidamente el artículo 3 de la ley.

2º- Se disminuye a 14 años la edad de los menores para el consentimiento lícito para el tratamiento de sus datos de carácter personal, en el artículo 7 de la LOPDGDD.

3º- Se introducen nuevas especificaciones en el derecho de acceso del artículo 13.2 de la LOPDGDD. Se reconoce la posibilidad de que se pueda crear un módulo el que el titular de los datos pueda acceder a su información de forma remota, directa, simple y segura.

4º- Se contempla en el artículo 24 de la LOPDGDD un sistema de información de denuncias internas que incluso pueden ser anónimas.

5º- Mientras que el RGPD no concreta las situaciones por las que debe efectuarse ex ante una evaluación de impacto, la LOPDGDD realiza una enumeración *numerus apertus* de estos en el artículo 28. Así, detalla una serie de supuestos como discriminación, daño para la reputación, pérdida de confidencialidad de datos sujetos al secreto profesional, reversión no autorizada de la Seudonimización, perjuicio moral o social significativo para los afectados, cuando conlleve la recogida masiva de datos personales, disminución de los derechos y libertades de los afectados, grupos de afectados en situación de especial vulnerabilidad y, en particular, de menores de edad y personas con discapacidad… De esta manera aclara las situaciones a las que hace referencia el artículo 35 del RGPD a modo de ejemplos más usuales, sin llegar a limitar o cerrar la interpretación de este.

6º- Hay un importante cambio de régimen jurídico de la AEPD en el artículo 44 de la LOPDGDD. Se configura ahora como una autoridad administrativa independiente de ámbito estatal, de las

[55]MAYOR GÓMEZ, R. "Principales novedades de la Ley Orgánica 3/2018, de 5 de diciembre, de protección de datos personales y garantía de los derechos digitales". *Gabilex. revista del gabinete jurídico de Castilla-la Mancha.* 2018. nº16. páginas 201-246 [En línea: https://dialnet-unirioja-es.sire.ub.edu/servlet/articulo?codigo=6901855] [visitado el 28/06/2020]

previstas en la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, con personalidad jurídica y plena capacidad pública y privada, que actúa con plena independencia de los poderes públicos en el ejercicio de sus funciones.

7º- En la Disposición adicional decimoséptima se contemplan tratamientos específicos de datos en el ámbito de la salud, proporcionado a las autoridades sanitarias e instituciones públicas con competencias en vigilancia de la salud pública podrán llevar a cabo estudios científicos sin el consentimiento de los afectados en situaciones de especial relevancia y gravedad para la salud pública.[56]

Ahora bien, debido al objeto de ese trabajo, no se dará más importancia a las novedades que incorpora la LOPDGDD, por cuanto no introduce modificaciones respecto a la gestión del riesgo en los tratamientos de datos personales y simplemente vine a reproducir las normas aprobadas en el RGPD.

En síntesis, podría decirse que la normativa sobre protección de datos persigue la adecuación, exactitud y la veracidad de las bases de datos, así como la cancelación de estos cuando dejen de ser necesarios. Proporcionando, del mismo modo que lo hace el Convenio nº.108 del Consejo de Europa, una especial protección a los datos denominados sensibles, reconocidos en el artículo 16 de la Constitución Española. Este grupo de datos está compuesto por aquellos que afectan a la ideología, religión o creencias y los relativos a la salud. A su vez también pretende el conocimiento y la posibilidad de acceso por parte de los titulares de estos datos pudiendo decidir cuáles van a proporcionar a terceros. *El contenido de este derecho consiste en disponer y controlar de los datos personales, tanto frente al Estado como frente a un particular*[57].

[56]Recientemente las administraciones públicas y, sobre todo, las sanitarias, han estado tratando de forma masiva los datos personales a raíz de la pandemia del Covid-19 que ha llevado a la declaración del estado de alarma en España y el confinamiento de todos los ciudadanos. Es por esta razón que se ha solicitado a la AEPD que resuelva sobre la legalidad de esta situación. Ciertamente el n.º 46 de la Exposición de Motivos del RGPD reconoce que hay situaciones excepcionales *en las que la trata de datos puede justificarse tanto sobre el importantes motivos de interés público como el interés vital, como por ejemplo cuando el tratamiento es necesario para fines humanitarios, incluido el control de epidemias y su propagación, o en situaciones de emergencia humanitaria, sobre todo en caso de catástrofes naturales o de origen humano*, sin prejuicio de que pueden existir otras bases. Respeto a la situación de pandemia por el Covid-19, las administraciones podrán tratar datos sin consentimiento respaldándose en los artículos 6.1, e) y d), 9.2 g), h), i), y j) del RGPD, así como en la Disposición adicional decimoséptima de la LOPDGDD y numerosas leyes sectoriales en materia de salud Pública. Respecto al deber de información sobre el tratamiento de dichos datos, podrán, en determinados casos prescindir de él en base al artículo 14.5 del RGPD. No obstante, y tal como señala la AEPD, el estado de alarma no anula el derecho a la protección de datos y, en general, los derechos fundamentales, por lo que las autoridades sanitarias y el Gobierno deberán efectuar un correcto ejercicio de ponderación entre los distintos derechos afectados con el objetivo de tomar las medidas menos limitativas posibles y, en todo caso, deberá cumplirse la LOPDGDD, y en particular, los principios proclamados en ella. En mi opinión, es precisamente esto último lo que se está poniendo en tela de juicio estos días, es decir, no la legitimidad para tomar ciertas medidas sino cómo estas medidas se adoptando. [En línea el *Informe 0017/2020 del Gabinete Jurídico de la AEPD*: https://www.aepd.es/es/documento/2020-0017.pdf] [visitado el 28/06/2020]

[57]POUS DE LA FLOR, Mª P. La protección de datos personales. In Mª D Díaz-Ambrona (Eds.), *Derecho Civil de la Unión Europea.* Madrid: Tirant lo Blanch, 2017. páginas 125 -168. [En línea: https://biblioteca-nubedelectura-com.sire.ub.edu/cloudLibrary/ebook/show/9788491431978] [visitado el 28/06/2020]

3. DIMENSIÓN ÉTICA.

La Inteligencia Artificial ciertamente cambiará nuestras vidas y la forma en la que entendemos las relaciones interpersonales. Conquistará todos los sectores de la sociedad desde el transporte, hasta la educación, la sanidad, marketing, finanzas, sector de seguros, las ciencias, el mundo del entretenimiento, la agricultura, la industria, etc. Si bien es cierto que la IA trae consigo innumerables beneficios, pudiendo llegar a solucionar inconvenientes con los que llevamos años combatiendo, también acarrea nuevos retos y peligros. Algunos de los cuales son bien conocidos, pero se manifiestan de una forma distinta y, hasta en ocasiones, magnificados. Entre estos intereses en peligro se encuentra la privacidad o la propia igualdad. Pero también ciertas situaciones que hasta ahora eran desconocidas como pueden ser el sesgo, la conexión entre datos o los problemas de atribución de responsabilidad de aquellos tratamientos o procesos que son capaces de tomar decisiones automatizadas. Peores o mejores, estos riesgos son desconocidos, por lo que es necesario hacerlos frente adecuada y proporcionalmente e intentar anticiparse a ellos, ya que ahora tenemos la oportunidad de minimizarlos o evitarlos.

En un principio los retos éticos que plantea la IA únicamente interesaban en el plano teóricos, técnicos y académicos, sin embargo, hoy han pasado a ocupar el centro del debate. En los últimos cuatro años tanto los gobiernos como las empresas privadas de las principales potencias geopolíticas, entre ellas los Estados Unidos, China, la Unión Europea y la India, están promoviendo declaraciones y políticas relativas a la IA y la ética.

El Gobierno de Obama en 2016 ya publicó un informe sobre el futuro de la IA[58] y le siguió, al año siguiente China y otros importantes actores mundiales. Todos estos informes coincidieron en la importancia de abordar una estrategia de IA que incluye la ética. Por ejemplo, China tiene un plan de desarrollo en el que se recomendaba reducir al mínimo los riesgos[59]. A nivel europeo son muchos los países que han publicado informes y estrategias acerca de este tema, proponiendo una IA fiable y transparente, lo que se ha reflejado en el trabajo supranacional sobre la política de la IA y la formación del HLEG AI[60].

Pese a que no existir un principio ético explícitamente respaldado por todas las directrices existentes, la mayoría de las propuestas relativas a la ética en los sistemas de IA parecen coincidir en la necesidad de actuar conforme a criterios de transparencia, la justicia y la equidad. Todo apunta a una *prioridad moral emergente que exige procesos transparentes en todo el proceso de la IA (desde la transparencia en el desarrollo y el diseño de algoritmos hasta las prácticas transparentes para el uso de la IA), y advierte a la comunidad mundial del riesgo de que la IA*

[58]National Science and Technology Council Committee on Technology, '*Preparing for the Future of Artificial Intelligence'* (Executive Office of the President, Office of Science and Technology Policy (OSTP) 2016). [En línea: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf] [visitado el 28/06/2020]

[59] State Council of China 2017 'New Generation Artificial Intelligence Development Plan. (新一代人工智能发展规划) [En línea: https://Flia.Org/Notice-State-Council-Issuing-New-Generation-Artificial-Intelligence-Development-Plan/] [visitado el 28/06/2020]

[60]El Grupo Europeo sobre la Ética de la Ciencia y las Nuevas Tecnologías (GEE) en el 'Statement on Artificial Intelligence, Robotics and "Autonomous" Systems' 2018. Esta declaración sobre avance de las nuevas tecnologías también se propone una serie de principios que posteriormente inspiraron los principios propuestos por el HLEG AI. [En línea: http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf] [visitado el 28/06/2020]

pueda aumentar la desigualdad si no se abordan adecuadamente las consideraciones de justicia y equidad.[61]

Una cuestión clave, entonces, es definir cuáles son las normas éticas que deben atenderse en el desarrollo de sistemas de IA. Dada la divergencia significativa en las distintas soluciones propuestas para hacer frente a los desafíos éticos de la IA, surge otro problema. El carácter transnacional de las tecnologías digitalizadas, el papel fundamental de las empresas privadas en el desarrollo y la aplicación de la inteligencia artificial y la economía globalizada plantean *interrogantes sobre qué jurisdicciones o actores decidirán las normas jurídicas y éticas a las que puede adherirse la inteligencia artificial*[62].

En definitiva, la ética de la IA debe plantearse como una ética de la dignidad humana centrada en la garantía de los derechos fundamentales. Es decir, el desarrollo de la IA debe orientarse a las personas y fundamentarse en el compromiso de utilizar dichos sistemas al servicio de la humanidad y del bien común, con el objetivo de mejorar el bienestar y la libertad de los seres humanos. *La IA no es un fin en sí mismo, sino un medio prometedor para favorecer la prosperidad humana y, de ese modo, mejorar el bienestar individual y social y el bien común, además de traer consigo progreso e innovación*[63]. Este integra, por lo que podríamos definir así, una primera barrera al desarrollo de esta tecnología.

3.1. Principales debates.

Los sistemas dotados de IA y, en particular, los modelos de aprendizaje automático, como se ha visto, implican un proceso de recopilación, trata el intercambio de datos. Al ser estos datos de carácter personal y/o al poder afectar las decisiones que se tomen a partir de su tratamiento a nuestra esfera personal, la probabilidad que esta tecnología colisione con nuestros derechos resulta verdaderamente alta. El escenario al que se enfrenta el derecho con la irrupción de la IA se encuentra cargado de incertidumbres. Para enfrentarnos a esta problemática deberemos recurrir a la ética. En el RGPD no se contienen las soluciones a todas las incógnitas, como veremos a continuación.

La ética no es más que un contrato social y, por lo tanto, es necesario que se produzca un debate respecto a las implicaciones que tiene la IA. Pero para que suceda esto, primero, habrá que dotar a la sociedad de conocimientos. Actualmente estamos acostumbrados a las mejoras constantes que aportan las nuevas tecnologías sin cuestionarnos que comporta aquello que se está usando, cómo funciona y qué comporta más allá de lo que ve.

Los riesgos directamente relacionados con el uso de IA y el tratamiento de datos personales son en realidad, muy frecuentes. Por ejemplo, un incidente de seguridad que puede poner en peligro la confidencialidad de los datos, el uso inadecuado de la información con fines delictivos, etc. Por

[61]JOBIN A. IENCA M. VAYENA E. "Artificial Intelligence: the global landscape of ethics guidelines". *Health Ethics & Policy Lab, Zurich, Switzerland.* 2019. Página 14. [En línea: https://arxiv.org/ftp/arxiv/papers/1906/1906.11668.pdf] [visitado el 28/06/2020]

[62]The Chinese University Of Hong Kong Faculty Of Law (Research Paper nº 2019-15) "*Artificial Intelligence Governance and Ethics: Global Perspectives"* [En línea: https://ssrn.com/abstract=3414805]

[63]High-Level Expert Group on Artificial Intelligence, abril de 2019, "*The Ethics Guidelines for Trustworthy Artificial Intelligence*". Página 5-9. [En línea: https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines] [visitado el 28/06/2020]

otra parte, surgen también otro tipo de riesgos, quizá no tan obvios, como son los que provienen de la aplicación concreta que se hace de los datos. Un buen ejemplo de esta otra clase de riesgos son las noticias personalizadas. Podríamos decir que, en esos casos, cuando buscamos una información y se nos muestra sólo una parte de la misma, se está restringiendo, limitando nuestro campo de visión de la realidad. *Es el fenómeno de filtrar la bombilla. Además, hay que tener presente que quien controla lo que vemos, de ese modo pueden por terminar controlando lo que pensamos, sentimos o hacemos*[64].

La agencia de contratación de trabajo finlandesa *DigitalMinds* tiene una veintena de grandes corporaciones como clientes. Como recibe millones de currículos cada mes, han recurrido a que sea un sistema de IA de decisiones automatizadas el que haga la selección. Este sistema, además de hacer una selección de los mejores currículos, ha sido diseñado para acceder a los perfiles de los candidatos en las redes sociales como Facebook y Twitter y a email como el de Gmail o el de Microsoft Office. Analiza la actividad de cada uno de los aspirantes y como éstas interactúan y se relacionan entre ellas. Con estos resultados, la empresa es capa de saber si una persona es introvertida o atenta, así como otros múltiples aspectos de su personalidad. La pregunta es ¿Cuántos de nosotros estaríais dispuestos a dar el consentimiento? ¿Realmente los candidatos tienen otra opción?

Este ejemplo es real, elaborado en el informe de *Algorithm Watch Automating Society*[65]. En este capítulo pretendo abordar brevemente algunas de las cuestiones éticas que están siendo debatidas respecto a los sistemas de IA, que difícilmente pueden resolverse recurriendo únicamente a la ley.

Una primera cuestión, compartida con todas las tecnologías desde hace ya años[66], es la cuestión del **respeto a la privacidad de las personas** o incluso el conocimiento de éstas sobre si sus datos están siendo usados. A menudo los usuarios no saben que la IA está detrás de una aplicación o que sus datos están siendo tratados, ni cómo se está usando. Es una práctica muy frecuente, además, pedir el consentimiento a los usuarios para la tarta de sus datos mediante largos textos llamados "términos y condiciones" establecidos como *condición sine qua non* para el uso de esos aplicativos. Es habitual, también, que estos datos cedidos en un dominio sean utilizados posteriormente en otro contexto y en otro dominio sin que los usuarios tengan conocimiento real de esta situación.

Los gigantes que hoy dominan el mundo, Facebook, Apple, Twitter y Google, facturan miles de millones de dólares cada año y responden con páginas y páginas de farragosas explicaciones en letra pequeña escritas en lenguaje de leguleyo. Insisten en afirmar que sus compañías no venden a terceras partes la información personal del usuario, pero eso no es exactamente así. Disponen de esa información porque se la hemos dado gustosamente. Y a ciegas.[67] Los datos recabados, en ocasiones, son añadidos en sistemas de IA programados para analizar el comportamiento de las

[64]Autoridad Catalana de Protección de Datos. *"Inteligencia Artificial. Decisiones Automatizadas en Catalunya"* enero 2020. Página 11 y 12. [En línea: https://apdcat.gencat.cat/es/documentacio/jornades_i_congressos/apdcat/2020/] [visitado el 28/06/2020]

[65]Algorithm Watch, Automating society. "*Taking stock of Automated Decision-Making in the EU*" 2019. Páginas 62 y 63. [En línea: https://algorithmwatch.org/en/automating-society/] [visitado el 28/06/2020]

[66]Las nuevas tecnologías han suscitado desde siempre inseguridades en materia de privacidad. Las primeras preocupaciones en esta materia datan de 1980. Warren y Brandeis en la revista "The Right to Privacy" Harvard Law Review, expresaron su preocupación por la intromisión en la privacidad de las de grabadoras de voz y las fotografías instantáneas.

[67]MIGUEL ARIZA L. "Lo saben todo de usted". *El País* 9 de diciembre 2015. [En línea: https://elpais.com/elpais/2015/12/04/eps/1449252033_849371.html] [visitado el 28/06/2020]

personas, capaces de identificar no sólo las preferencias, la edad y el sexo de los individuos, sino también su orientación sexual y sus opiniones religiosas o políticas.

Estos datos, no obstante, no sólo son recabados por las grandes corporaciones, es más, nos sorprenderíamos al ver la cantidad de sistemas de IA que a diario tratan nuestros datos sin que nos demos cuenta. Supermercados, bancos, compañías de seguros, tiendas de ropa, administraciones públicas, la lista es infinita. Un ejemplo que personalmente me llamó la atención es un caso que ocurrió en Estados Unidos en 2010. Un padre entró en un gran supermercado quejándose que el supermercado estaba enviando, a su hija menor de edad, anuncios sobre productos premamá, muebles para habitaciones de bebés, etc. Resulta que el supermercado había adquirido un sistema de IA que analizaba las compras de los clientes y podía establecer predicciones sobre su comportamiento de compra. El sistema era capaz de saber con bastante precisión si una mujer estaba embarazada mediante el análisis de 25 clases de productos en un periodo de tiempo corto. Al final, la chica resultó estar embarazada, pero el padre no lo sabía.

Es necesario marcar una línea entre lo que es de interés comercial o científico y la ética. ¿Es ético que una compañía obtenga nuestros movimientos o nuestras conversaciones mediante los dispositivos móviles para, posteriormente, bombardearnos con publicidad? ¿Es ético que los bancos y compañías de seguros sepan nuestros hábitos, consumo, gastos y preferencias de ocio para posteriormente aprovecharse de esta información para generar más beneficios? ¿Es ético la tarta masiva de datos sanitarios por las administraciones públicas durante la pandemia del covid-19, hasta de aquellas personas que no tienen el virus?

Ciertamente es difícil encontrar el equilibrio ente el avance de las nuevas tecnologías de inteligencia artificial y la privacidad. Creo que el primer paso hacia una solución es la puesta a disposición de información más clara y veraz sobre el funcionamiento, alcance y fines de estos sistemas. No tienen sentido los derechos proclamados por el RGPD como, el derecho de rectificación, supresión, capacidad de restringir el procesamiento de los datos etc. si no se tiene conocimiento real de cómo estos datos están siendo tratados en primer lugar. Así la *40ª Conferencia Internacional de Comisionados de Protección de Datos y Privacidad* proclamo la necesidad de:

> "*toda creación, desarrollo y utilización de sistemas de inteligencia artificial deberá respetar plenamente los derechos humanos, en particular los derechos a la protección de los datos personales y a la privacidad, así como la dignidad humana, la no discriminación y los valores fundamentales, y deberá ofrecer soluciones que permitan a las personas mantener el control y la comprensión de los sistemas de inteligencia artificial*[68]*"*

La segunda cuestión integra de uno de los conflictos más discutidos, **la parcialidad o el sesgo de los sistemas de IA**. Se trata de una desviación inadecuada en el proceso de inferencia. Significa un error en la programación del algoritmo[69], generando que algunos individuos o grupos queden en situación de desventaja por el resultado del sistema. Por este motivo será altamente necesario

[68] 40th International Conference of Data Protection and Privacy Commissioners. "*Declaration On Ethics And Data Protection In Artifical Intelligence".* October 2018. Brussels. Página 3. [En línea: https://edps.europa.eu/sites/edp/files/publication/icdppc-40th_ai-declaration_adopted_en_0.pdf] [visitado el 28/06/2020]

[69] Una mala elección del modelo de IA, es decir del algoritmo adecuado en relación con el objetivo también puede generar sesgos en el resultado.

asegurar la calidad de los datos, es decir, la pulcritud técnica de las muestras de datos y la representatividad estadística del conjunto de datos (el parámetro).

El sesgo suele producirse involuntariamente y, puede darse en diversas etapas del proceso de aprendizaje de los sistemas de IA. Puede surgir en la captación de los datos, selección del conjunto de estos, en los algoritmos utilizados, etc. Los sesgos son especialmente graves cunado producen discriminación ente grupos sociales. Un ejemplo para entenderlo es el caso que vivió Amazon en 2014 al estrenar un algoritmo de IA para recaudar nuevos trabajadores en sus almacenes. El sistema puntuaba del uno al cinco a los posibles candidatos ahorrando muchas horas al departamento de recursos humanos. Un año más tarde, la multinacional se percató que no se había contratado a ninguna mujer en los puestos técnicos. No podía ser que no hubiese ninguna candidata apta. Resulta que los datos inferidos en el sistema de IA durante la fase de aprendizaje fueron los currículums recibidos hacía unos años y, en esa época, la mayoría del personal cualificado eran hombres. De esta manera, cuando el sistema detectaba la palabra mujer directamente penalizada el currículum poniendo menos puntuación.

El sesgo de los sistemas de IA puede influir en sectores del más variado tipo. La aplicación de la IA en el sistema judicial es un tema muy delicado. En Estados Unidos se han puestos en práctica sistemas[70] de IA, usualmente diseñados por empresas privadas, que son capaces de calcular la reincidencia criminal. Mediante estos sistemas se toman decisiones tan importantes como decretar la prisión preventiva, o facilitar permisos de libertad condicional. Uno de ellos demostró tener sesgos que marcaban una probabilidad más alta a las personas negras, atribuía un mayor riesgo de comisión de delitos a los negros frente a los blancos.

Están apareciendo, además, en el mercado programas que pueden afectar principios esenciales del sistema jurisdiccional, como el de tutela judicial efectiva. Ofrecen la posibilidad de valorar el porcentaje de éxito de una demanda concreta según el tribunal ante la que se presente. *Desde el punto de vista de la garantía del juez ordinario predeterminado por la ley, y de la independencia de criterio del juzgador en el ejercicio de la función jurisdiccional, resulta preocupante pensar que existan máquinas capaces de predecir un determinado resultado*[71].

No obstante, es discutible si un determinado sesgo o discriminación es injusto o improcedente puesto que ciertamente hay situaciones en las que determinados grupos de personas están más cualificados que otras por el mero hecho de pertenecer a ese grupo. Por ejemplo, podría considerarse que las mujeres tienen más cualidades empáticas para desarrollar trabajos como el de enfermeras o, que los hombres, están más cualificados para trabajar en el sector de la construcción, por sus cualidades físicas. *La respuesta a esa pregunta no es una cuestión meramente técnica, sino ética y política, depende de nuestra visión de la justicia y del tipo de sociedad que queremos*[72].

[70]El programa estadounidense que suscito problemas fue el llamado "*Correctional Offender Management Profiling for Alternative Sanctions*" (COMPASS). Por muy marciano que esto suene, en Catalunya, existe un programa similar llamado *RisCanvi* que, de hecho, actualmente se encuentra en funcionamiento. Este programa se puso en marcha en el 2009 y ya va por la versión 3.0. Su objetivo es estimar que probabilidad tiene una persona de volver a delinquir una vez sale de la prisión. Hasta el momento, según varios expertos, parece que funciona bastante bien.

[71]MARTÍNEZ, R. ob-cit. página 64-81.

[72]COECKELBERGH M. "Ethics of artificial intelligence: Some ethical issues and regulatory challenges". *Technology and Regulation*. 2019. Página 31-34. [En línea: https://doi.org/10.26116/techreg.2019.003] [visitado el 28/06/2020]

Por otra parte, el sesgo es algo con lo que ya estamos familiarizados. La parcialidad y discriminación han estado presente en los sistemas sociales y culturas, posiblemente, desde la edad de piedra y contra lo que llevamos años luchando. Esencialmente lo que preocupa es que los sistemas de IA puedan perpetuarlos aumentado su impacto. Lo cierto es que existe otro segundo sesgo que puede ser aún más peligroso que el sesgo que se genera con la interpretación de los datos: este es el sesgo humano. Éste consistente en aceptar, sin espíritu crítico, los resultados que proporcionan los sistemas de IA derivados de las expectativas creadas por dichos sistemas. Este fenómeno es denominado comúnmente por la psicología como *biax de confirmación*[73]. Es precisamente este último el que inquieta, pudiendo llegar a perpetuar y agravar las discriminaciones generadas por el sesgo en los sistemas de IA.

Mas allá del conocimiento de las personas sobre el tratamiento de sus datos, existe un tercer dilema, y es el de **la transparencia de los sistemas**, especialmente en los sistemas de decisiones automatizadas. Estos algoritmos tienen la habilidad de aprender durante todo el ciclo de su vida mientras tratan datos con los que toman decisiones que en ocasiones afectarán a personas físicas.

Hemos subrayado anteriormente la importancia de informar a los usuarios del funcionamiento y alcance de los sistemas de IA que tratan sus datos. Pero el problema es que, incluso los expertos no siempre lo saben todo. Y esto nos lleva la cuestión de los sistemas de "*Caja Negra*"[74]. No siempre está claro lo que sucede en algunos de los sistemas de aprendizaje automático que utiliza redes neuronales como las aplicaciones dedicadas al procesamiento del lenguaje natural, la diagnosis médica, la bioinformática, o la detección de fraude financiero. No es posible rastrear la decisión ni la cadena de decisiones que ha efectuado el sistema, a diferencia de los modelos de árbol de decisiones. Por lo que estos sistemas son muy opacos[75]. Se trata de uno de los mayores dilemas que existe en la actualidad con los sistemas de IA.

En nuestro sistema democrático, las personas tienen derecho a saber la razón de la toma de decisiones que les incumben. Si no se puede justificar una decisión, ¿Cómo se pueden defender los afectados por ella? ¿Y si está decisión deriva de un sesgo en el sistema? La explicabilidad de los sistemas es, entonces, un requisito moral y necesario.

Por último, el cuarto problema al que quero hacer referencia, y que adquiere especial relevancia en el caso de los sistemas de IA, es la **atribución de responsabilidad**. Dado que los robots, es decir, los sistemas integrados con IA, no pueden ser agentes morales y por lo tanto tampoco responsables, la única manera de garantizar la responsabilidad de sus acciones es hacer responsables a los seres humanos. Concretar la responsabilidad pueda que pueda derivarse de un error es una cuestión compleja para cuya resolución se han formulado infinidad de teorías. De entre ellas la que ha alcanzado mayor predicamento ha sido la de considerar al diseñador como responsable o la empresa que provee el servicio. Para así, atribuirles la responsabilidad de los daños causados.

[73]Este término también es comúnmente utilizado por la criminología para denominar el sesgo que se puede producir en el transcurso de las investigaciones policiales por el que estas solo se dirigen únicamente hacia un sujeto o un grupo de sujetos sin tener en cuenta otras variables igualmente posibles.
[74]Autoridad Catalana de Protección de Datos. ob-cit. página 11 y 12.
[75]En ocasiones la opacidad de los sistemas no viene dada por el desconocimiento del funcionamiento del sistema, sino que, más bien, se produce por cuestiones de propiedad intelectual. La mayor parte de los sistemas, hasta los utilizados por las Administraciones públicas, son fabricados por empresas privadas. Estas no diseñan los sistemas de IA con la posibilidad reconstruir la decisión para que esta sea accesible al público, no porque no sepan cómo hacerlo, sino por temas de competencia y propiedad intelectual.

Los robots, al igual que los animales, son sujetos *sui generis*, no obstante, cabría plantearse la necesidad de atribuir una responsabilidad distinta cuando las decisiones que toma el sistema provienen del análisis de su entorno. Es decir, la existencia de sistemas de IA que interactúen con personas puede generar problemas jurídicos más complejos. Una máquina de IA que interacciona con terceros para prestar servicios podría devenir xenófoba si aprende en un ambiente racista o el ambiente que la rodeara privilegiase a los nacionales. En esos casos, la atribución de responsabilidad a un sujeto no es tan sencilla de determinar.

La dificultad de atribuir la responsabilidad en muchas ocasiones surge del llamado problema de *"muchas manos"*[76]. Este implica que muchas personas están involucradas en la acción de la que derivan los daños. Además, la IA suele formar parte de un sistema tecnológico muy fragmentado, con gran entrada y salida de datos, cuestión que dificulta averiguar si ha sido la propia IA la que causó el problema.

Otra cuestión, es la posibilidad, en ocasiones, de atribuir la responsabilidad a los usuarios de los sistemas de IA. Este tipo de atribución de responsabilidad genera problemas cuando las personas que hacen uso de los sistemas de IA ignoran el funcionamiento de la mayor parte del sistema y su historia, como, por ejemplo, cómo se han generado y combinado los datos. *Se supone que las personas que utilizan los sistemas deben asumir la responsabilidad, pero esto se hace difícil si no saben lo que están haciendo*.[77]

3.2. La Inteligencia Artificial Fiable. Ética de la Unión Europea.

En este último título se analizarán los esfuerzos del HLEG AI, compuesto por 52 expertos de distintas áreas, en definir que es un sistema de IA ético, y como deben llevarse a cabo los tratamientos que incorporan IA para llegar a este fin. Como se ha mencionado en diversas ocasiones, se ha denominado a este fenómeno *IA Fiable*.

La Unión Europea aboga por una IA que gravita en torno a tres componentes. El primero es la *Licitud*, aquella que cumple la normativa de aplicación, el segundo es la *Ética*, aquella que garantice el respeto a los principios y los derechos de las personas, por último, la *Robustez*, es aquella IA diseñada mediante procesos de prevención de accidentes.

Estos tres conceptos deben interpretare conjuntamente puesto que actúan en armonía, y se complementan recíprocamente, debiéndose entender como un todo y no separadamente. Sin embargo, en aras de la metodología seguida en este documento, sólo se analizarán la ética y la robustez de los sistemas puesto que ya se ha hecho referencia en anteriores capítulos a la licitud.

3.2.1. La ética de la Inteligencia Artificial.

La ética de los sistemas de IA es un subcampo de la ética que analiza los problemas que plantean el desarrollo despliegue y utilización de la IA. Como punto de partida el HLEG AI remite a una serie de derechos fundamentales y principios éticos, algunos de ellos bien conocidos por la bioética. Estos derechos y principios deben entenderse como deberes morales y legales de los

[76]COECKELBERGH M. ob-cit. página 31-34.
[77]*Ídem.*

tratamientos que incorporan IA. Estos, constituyen no sólo una obligación legal sobre, que podemos hacer, sino también una orientación para identificar aquello que debemos hacer.

El grupo de **derechos fundamentales** que deben inspirar todo el ciclo de vida de un sistema de IA queda recogido en el documento *The Ethics Guidelines for Trustworthy Artificial Intelligence*[78] y pueden resumirse como:

- **Derecho a la dignidad Humana**

En el contexto de la IA este principio implica que todas las personas deben ser tratadas con el respeto que merecen como sujetos morales con valor intrínseco que jamás podrá ser dañado. De ello se infiere que nunca las personas podrán ser tratadas como meros objetos, ni ser filtrados, ordenados, dirigidos, puntuados, condicionados o manipulados. Es, por tanto, necesario que los tratamientos que incorporen IA se desarrollen en este sentido, sin olvidar que estos sistemas están al servicio *de la integridad física y mental de los seres humanos, el sentimiento de identidad personal y cultural y la satisfacción de sus necesidades esenciales.*

- **Derecho a la Libertad individual**

Las personas tienen derecho a tomar decisiones vitales por sí mismos, sin que terceros puedan inferir en éstas. De hecho, *la libertad individual entraña un compromiso de permitir que los individuos ejerzan un control aún mayor sobre su vida, incluidos (entre otros derechos) la protección de la libertad de empresa, la libertad de las artes y de las ciencias, la libertad de expresión, el derecho a la privacidad y la vida privada y la libertad de reunión y asociación*.

En este sentido, los estados deberán garantizar la libertad de los individuos y, en especial, que las personas en riesgo de exclusión tengan la oportunidad de disfrutar en igualdad de condiciones de las oportunidades y beneficios que ofrece la IA.

- **Derecho al Respeto a la Democracia, la justicia y el Estado de Derecho**

Los tratamientos de IA deberán mantener e impulsar los procesos democráticos, garantizando que todo poder gubernamental está autorizado legalmente y limitado por el Ordenamiento Jurídico. También deberán garantizar el respeto al pluralismo social. Asimismo, *los sistemas de IA deben incluir un compromiso de garantizar que su funcionamiento no menoscabe los compromisos esenciales en los que se fundamenta el estado de Derecho y de asegurar el respeto de las garantías procesales y la igualdad ante la ley*.

- **Derecho a la Igualdad, no discriminación y solidaridad**

Para los tratamientos que integran IA, la igualdad implica que el funcionamiento del sistema no debe generar resultados injustos o sesgados. Esto significa que deberá hacerse un estudio de los datos con el fin de comprobar que estos son inclusivos, suficientemente variados, actualizados, etc., proporcionando una adecuada protección a aquellos datos considerados sensibles y respetando a aquellas personas o grupos de personas potencialmente vulnerables.

[78]High-Level Expert Group on Artificial Intelligence, abril de 2019. ob-cit. página 13 - 16

- **Derechos de los ciudadanos**

Los sistemas de IA tienen un elevado potencial para mejorar el alcance y eficacia de los derechos otorgados de los ciudadanos, pero al mismo tiempo algunas aplicaciones también pueden llegar a incidir en ellos negativamente.

Finalmente, HLEG AI efectúa una lista de 4 **principios básicos vinculados a los derechos fundamentales** que también deberían inspirar todo el ciclo de vida de un sistema de IA. Estos principios se inspiraron en los principios propuestos por el Grupo europeo de Ética de la Ciencia y las Nuevas Tecnologías en 2018. El HLEG AI califica estos principios como imperativos éticos que los desarrolladores de IA deben esforzarse en todo momento por observar.

Como veremos, varios de estos principios se encuentran recogidos en el RGPD, por lo que están establecidos como requisitos legales de obligado cumplimiento, si bien el HLEG AI insiste en que la observancia de estos últimos trasciende del mero cumplimento legal.

- **Principio de autonomía humana**

Se trata de un principio derivado de un derecho fundamental en el que se basa el consentimiento informado[79] como valor. Los sistemas de IA deberán diseñarse con el objetivo de ayudar a las personas, de forma que aumenten, complementen y potencien las aptitudes cognitivas, sociales y culturales de las personas.

- **Principio de prevención del daño o no maleficencia**

Los sistemas de IA no deben causar daños o agravar los ya existentes. No deben perjudicar a los humanos de ningún modo. Por ende, los sistemas de IA deberán garantizar la integridad física y moral de las personas asegurándose que operan en entornos seguros.

Deberá garantizar que los sistemas de IA han sido dotados de un sistema técnico robusto, no pudiendo ser destinados a usos malintencionados. Se deberá prestar especial atención a situaciones en las que los sistemas de IA *puedan provocar efectos adversos debido a asimetrías de poder o de información, por ejemplo, entre empresarios y trabajadores, entre empresas y consumidores o entre gobiernos y ciudadanos. La prevención del daño implica asimismo tener en cuenta el entorno natural y a todos los seres vivos.*

- **Principio de Equidad**

Este principio está compuesto por dos dimensiones, la sustantiva y la procedimental. La dimensión sustantiva implica el compromiso de garantizar una distribución justa de los beneficios y costes, asegurado que las personas implicadas no sufren sesgos injustos, discriminación o estigmatización. Se deberá fomentar la igualdad de oportunidades en términos de acceso a la educación, sanidad y la tecnología.

El uso de tratamientos con IA no deberá conducir a engaño a los usuarios y a limitar su libertad de elección. proporcionándoles, en todo caso, la oportunidad de oponerse a las decisiones tomadas por los sistemas y por las personas que los manejan. Por este motivo será imprescindible poder explicar los procesos de toma de decisiones en el tratamiento. *Asimismo, la equidad implica que*

[79]Es el conocimiento y comprensión de todas las consecuencias del procesamiento de datos.

los profesionales de la IA deberían respetar el principio de proporcionalidad entre medios y fines, y estudiar cuidadosamente cómo alcanzar un equilibrio entre los diferentes intereses y objetivos contrapuestos.

- **Principio de Explicabilidad**

Se trata de un principio esencial para garantizar la confianza de los usuarios en los sistemas de IA. Para este fin, los procesos deberán ser transparentes, siendo necesario comunicar las capacidades y los fines de los sistemas de IA ya que, como se ha visto, las decisiones deben poderse explicar. El HLEG AI constata que no siempre resulta posible explicar por qué un modelo ha generado un resultado o una decisión. En tales circunstancias puede ser necesario adoptar otras medidas de seguimiento del sistema con el fin de augmentar su transparencia. Por ejemplo, podrían tomarse medidas de la trazabilidad, auditabilidad y comunicación transparente sobre las prestaciones del sistema. El grado de necesidad de explicabilidad dependerá en gran medida del contexto y la gravedad de las posibles consecuencias que podrían derivar de un resultado erróneo o inadecuado.

En definitiva, el HLEG AI concluye que *para garantizar la fiabilidad de la IA es necesario, más allá de desarrollar un conjunto de normas, crear y mantener una cultura y una mentalidad éticas a través del debate público, la educación y el aprendizaje práctico.*

3.2.2. La robustez de la Inteligencia Artificial.

Para que un tratamiento que incorpore IA sea considerado robusto, será necesario que se dote de una gestión del resigo adecuada con el fin de evitar escenarios imprevistos y perjudiciales para los principios, derechos y valores fundamentales anteriormente expuestos. El HLEG AI ofrece una lista de siete requisitos obligatorios en los tratamientos de IA con el fin de lograr una IA fiable. A continuación, se detallarán los requisitos que enumera el HLEG AI aplicables a las diferentes partes interesadas que participan en algún momento del ciclo de la vida del sistema IA:

1. Acción y supervisión humanas. Los sistemas de IA deberán respaldar la autonomía y la toma de decisiones de las personas, garantizando los derechos fundamentales, la acción y la supervisión humana. La autonomía del usuario debe ocupar un papel central en el sistema, dotándole de suficientes conocimientos y herramientas para comprender los sistemas de IA. Además, la supervisión humana es un punto clave para garantizar que un sistema de IA no reduzca o imposibilite la autonomía humana o provoque otros efectos adversos.

2. Solidez técnica y seguridad. Estrechamente relacionado con la prevención del daño, la solidez requiere que los sistemas de IA se desarrollen mediante un enfoque de prevención de riesgos, de modo que se comporten siempre según lo esperado, minimizando los daños imprevistos. Esto incluye, en primer lugar, la capacidad de resistencia a los ataques y la seguridad. Los sistemas de IA deberán protegerse de los ataques exteriores, como, por ejemplo, de los piratas informáticos. Para que los sistemas se consideren seguros, es necesario tener en cuenta las posibles aplicaciones imprevistas o maliciosas de la IA. También se deberán incorporar medidas para prevenir y mitigar los riesgos. En segundo lugar, debe confeccionarse un plan de repliegue y seguridad general en caso de que surjan problemas. El nivel de las medidas de seguridad requeridas depende de la magnitud del riesgo que plantee un sistema de IA. En tercer lugar, será necesario indicar la probabilidad de error que existe en el sistema IA, asegurar que hay un alto nivel de precisión. Por

último, los sistemas deberán ser fiables y reproducibles. La reproducibilidad se valora a partir de la comprobación de que un sistema con IA muestra el mismo comportamiento cuando se repite varias veces en las mismas condiciones. Esto permite describir con exactitud lo que hacen los sistemas de IA.

3. Gestión de la privacidad y de los datos. La privacidad, como se ha ido viendo, es un derecho que se ve especialmente afectados en los tratamientos de IA que utilizan datos de carácter personal. Es crucial una adecuada gestión de estos datos, que asegure la calidad y la integridad de estos. La recopilación de datos que contengan sesgos o la manipulación malintencionada de estos datos puede alterar su comportamiento, sobre todo de aquellos sistemas de aprendizaje automático. Los procesos y conjuntos de datos utilizados deben ponerse a prueba y documentarse en cada paso, hasta en aquellos sistemas de IA que no hayan sido desarrollados internamente, sino adquiridos externamente. Es necesario, además, garantizar la protección de la intimidad de los datos al largo de todo el ciclo de la vida del sistema IA.

Los protocolos que limiten el acceso pueden ser una herramienta muy útil para evitar manipulaciones ilícitas, debería describirse quién puede acceder a los datos y en qué circunstancias.

4. Transparencia. Este requisito guarda estrecha relación con el requisito de explicabilidad y está compuesto por tres elementos: la trazabilidad, la explicabilidad y la comunicación.

Es necesaria la documentación del proceso del sistema de IA a fin de posibilitar la trazabilidad y aumentar la transparencia. Esto permitirá identificar los motivos de una decisión errónea por parte del sistema y permitirá prevenir futuros errores. La trazabilidad, por tanto, facilita la auditabilidad y la explicabilidad. A su vez la explicabilidad concierne a la capacidad de explicar tanto los procesos técnicos de un sistema de IA como las decisiones humanas asociadas. Aún más cuando la decisión del sistema de IA tenga un impacto significativo en la vida de las personas. Por último, la comunicación hace referencia al derecho de las personas a saber si están interactuando con un sistema de IA o una persona.

5. Diversidad, no discriminación y equidad. Hay que tener en cuenta a todos los afectados por el sistema de IA y garantizar su participación en todo el proceso. También es necesario garantizar la igualdad de acceso mediante procesos de diseño inclusivos, sin olvidar la igualdad de trato. Este requisito está estrechamente relacionado con el principio de equidad. La contratación de personas procedentes de diversos contextos, culturas y disciplinas puede garantizar la diversidad de opiniones y debería fomentarse.

6. Bienestar social y ambiental. De acuerdo con los principios de prevención del daño y equidad se deberá tener en cuenta la sociedad en conjunto, el respeto a los animales y al medio ambiente como partes interesadas al largo del ciclo de vida de la IA. Se deberá apostar por el desarrollo de sistemas de IA con Objetivos de Desarrollo Sostenible.

8. Rendición de cuentas. Todos los requisitos anteriores se supervisarán mediante la rendición de cuentas, estrechamente relacionada con el principio de equidad. Esta rendición deberá llevarse a cabo garantizando la auditabilidad, la minimización de efectos negativos y la notificación de estos, la búsqueda de equilibrios y las compensaciones.

En primer lugar, la auditabilidad es la capacidad para evaluar los algoritmos, los datos y los procesos de diseño. La evaluación por parte de auditores internos o externos y la disponibilidad de los correspondientes informes de evaluación pueden contribuir a esta fiabilidad. Cuando los

sistemas de IA afecten a los derechos fundamentales deberían poderse someter a auditorias independientes.

Es igualmente importante informar sobre las acciones y decisiones del sistema de IA así como responder a las consecuencias de dicho resultado. Es, por tanto, necesaria la identificación, evaluación y minimización de los efectos negativos que puedan tener los sistemas de IA. Pueden resultar útiles las evaluaciones de impacto antes y después del desarrollo del sistema de IA. Los tratamientos que incorporen IA deberán identificar los intereses y valores subyacentes al sistema. Cuando surjan conflictos con derechos fundamentales deberán explicitar cómo se ha intentado buscar el equilibrio entre ellos y evaluar dicho equilibrio en términos del riesgo, de respeto de los principios éticos y derechos fundamentales.

Por último, deben preverse mecanismos accesibles que aseguren la compensación adecuada a aquellos que hayan sufrido daños por efectos adversos e injustos del sistema de IA. Se debería prestar atención a las personas o grupos vulnerables.

3.3. Retos éticos para la Inteligencia Artificial del futuro.

Pese a los esfuerzos de los diferentes actores, tanto públicos como privados, en definir o trazar unos valores en el desarrollo de sistemas de IA éticos, no está claro si estas expresiones de preocupación por la ética de la IA conducirán a acciones concretas por parte de los desarrolladores. Existe el riesgo de que la ética proclamada en estos informes se utilice como una forma de justificar cualquier desarrollo de la IA ayudando a garantizar la aceptabilidad de la tecnología y el beneficio económico.

Así, el trabajo del HLEG AI, pese a sus buenas intenciones, describe un reducido e insuficiente grupo de derechos y una serie de principios éticos genéricos, y por lo general, ya muy asentados. Preocupa que las directrices en este documento sean demasiado vagas y abstractas, abiertamente interpretables por los diseñadores de los sistemas IA y difíciles de poner en práctica efectivamente.

En primer lugar, sorprende que no se hayan incluido en la lista de derechos tales como el del derecho a la vida privada y familiar, del art.7, o el derecho a la protección de datos del art. 8 de la CDFUE. Como se ha visto, estos dos derechos han tenido una relación crucial con las tecnologías desde su inicio y teniendo en cuenta que los sistemas de IA se nutren de datos, en ocasiones de carácter personal para su desarrollo y funcionamiento, resulta fundamental que se incluyan en este listado.

En segundo lugar, si bien el HLEG AI advierte de la posibilidad de que estos derechos, en ocasiones, entren en conflicto y la necesidad de que se faciliten métodos que posibiliten un debate responsable sobre dichas tensiones, no se facilitan las herramientas necesarias para dar solución a estas situaciones. Así, principios como el de prevención del daño y autonomía humana podrían entrar en conflicto, tal como lo explica el ente europeo. Por ejemplo, en tratamientos con IA para la actuación policial preventiva, cuando estos sistemas pretenden ayudar a reducir la delincuencia pueden, al mismo tiempo, vulneran la libertad y la privacidad de toda la ciudadanía con sus técnicas de vigilancia masiva. Pero ¿qué parámetros hay que tener en cuenta para llevar acabo el balance entre los derechos en conflicto? Es necesario en este caso, llevar a cabo *un juicio de ponderación* similar al que se realiza en los supuestos de conflicto entre derechos fundamentales

y, así, valorar si los beneficios que se vayan a obtener con un sistema de IA son superiores y justifican los riesgos o daños que puedan producirse a otros derechos. Sin embargo, el documento no recoge la jerarquía o la relación que existe entre los derechos fundamentales y los principios, y se limita a mencionar que existen algunos derechos que son de carácter absoluto, como por ejemplo el de la dignidad humana, que no pueden ser balanceados ni sacrificados.

Según el HLEG AI, los profesionales deberán afrontar los dilemas éticos y *analizar las ventajas e inconvenientes a través de un proceso de reflexión razonada con base empírica, en lugar de guiarse por la intuición o por criterios aleatorios*[80]. Estos conflictos, necesariamente, deberán ser confrontados por los desarrolladores *ad hoc* y el documento no parece proporcionar ninguna solución. ¿Deben observarse parámetros objetivos o subjetivos? Mas bien, el documento parce que se limita a ofrecer un catálogo de derechos y principios a escoger, trasladando a los desarrolladores la obligación de interpretarlos correcta y fundamentadamente. Sería fundamental hacer referencia a conceptos, incluso a principios, que ya han sido precisamente identificados por los propios Tribunales nacionales y de la UE, como criterios a seguir también en esta labor de búsqueda de equilibrio entre derechos y principios que pueden entrar en conflicto.

En definitiva, cabe concluir que el proyecto del HLEG AI establece un respecto a los derechos fundamentales y cumplimiento de la normativa aplicable como requisito previo para alcanzar un tratamiento con IA adecuado. Sin embargo, no se relaciona en ningún momento el proyecto con las disposiciones de la normativa aplicable. Esto genera la sensación de que el desarrollo, aplicación y uso de la IA apenas se encuentra regulado, una sensación de vacío jurídico, cuando realmente no es así. Muchas de las directrices y requisitos que se establecen en el proyecto son requisitos legales, reales y aplicables por la legislación europea y nacional. Así, el centro de investigación *CISPA Helmholtz center for information security*[81] sugiere al Grupo de expertos europeo que *consulte con expertos en el campo de la protección de datos y la regulación de la tecnología de la UE para garantizar que los requisitos legales aplicables (como el artículo 22 y el párrafo 1 del artículo 25 de la Ley de Protección de Datos) se reflejen en las directrices.*

Pese a todo, el documento del grupo de expertos europeo es el que más ha avanzado hacia la puesta en práctica de una IA ética y, todavía queda mucho trabajo por hacer en este sentido. En la actualidad, parece que sigue siendo un enorme desafío tender un puente entre los principios abstractos y las prácticas concretas.

Existe, además, una brecha entre la comprensión y cohesión de los trabajos que provienen de las ciencias sociales y los que tienen una formación más técnica. Es necesario que se unan fuerzas y se apueste por la interdisciplinariedad para la formación de políticas enfocadas a la ética en el desarrollo de las nuevas tecnologías en general. La reglamentación que surja no tiene por qué limitarse únicamente a la prohibición de conductas, sino que también será necesario una ética positiva y una reflexión sobre donde queremos que nos lleve la IA.

Sería positivo apostar por un enfoque proactivo de la ética en los sistemas de IA, tal como está configurado en los artículos 5.2 y el artículo 24 del RGPD respecto a la responsabilidad proactiva referente al cumplimiento de la legislación durante todo el ciclo de la vida del sistema de IA. Las nuevas tecnologías requieren que la ética no sólo venga después, sino que se cree un marco

[80] High-Level Expert Group on Artificial Intelligence. ob-cit. página 13 – 16.

[81] CISPA Helmholtz center for information security. "*Comments on the "Draft Ethics Guidelines for Trustworthy AI" by the High-Level Expert Group on Artificial Intelligence*." 2019. Página 3 [En línea: https://publications.cispa.saarland/2798/] [visitado el 28/06/2020]

regulador para estimular y asegurar que la ética se tenga en cuenta durante todas las fases de desarrollo del sistema de IA desde su inicio.

4. GESTIÓN DEL RIESGO EN LOS SISTEMAS DE IA.

La IA permite a los humanos desarrollar tareas más complejas o repetitivas y fijar procesos con elaborados diagramas que de otra manera llevarían mucho tiempo a los humanos. *Las maquinas nunca convertirán a una persona, y la persona tampoco se convertirá en una máquina. Ambas son complementarias y la adhesión es favorable al desarrollo de nuestra sociedad*[82]. Sin embargo, las herramientas de IA pueden generar riesgos que es importante conocer, medir y controlar, ello, además, resulta necesario para poder alcanzar una *IA robusta*.

En este contexto el RGPD exige, como punto clave de la responsabilidad proactiva, que se implementen medidas de análisis y gestión de los riesgos con el fin de establecer las medidas de seguridad y control para garantizar los derechos y libertades de las personas, tal como se establece en el artículo 24.1 y 25.1 y 2 del RGPD.

La gran variedad de posibles tratamientos de datos imposibilita ofrecer unos parámetros exactos y obligatorios, comunes para todos los tratamientos. Como ya se ha visto, la IA es una tecnología que entraña multitud de retos y riesgos que además variaran en gran medida en función del campo en el que los vayas a implementar. Así, un sistema de IA en el ámbito sanitario se llevará a cabo de forma muy distinta a uno que pueda usarse en el mundo financiero, generando un abanico de riesgos que necesariamente deberán tratarse *ad hoc.* Por este motivo, el legislador europeo optó por proporcionar un amplio grado de flexibilidad, sin establecer de forma específica los elementos que hay que implementar para llevar acabo la gestión del riesgo o la forma en la que hay que hacerlo. Sin embargo, sí se establece que dichas medidas deben tratarse efectuando un análisis basado en el riesgo[83]. En particular, el riesgo para los derechos y libertades de las personas y datos personales implicados en el tratamiento de IA con el fin de *aplicar medidas técnicas y organizativas apropiadas al fin de garantizar y poder demostrar que el tratamiento es conforme al reglamento*[84].

Podemos diferenciar tres fases comunes en todos los sistemas de IA en cuanto al análisis del riesgo. Una primera fase consiste en la identificación de las amenazas y riesgos que entraña el tratamiento con IA y la evaluación los mismos. La segunda fase está dirigida a la gestión del riesgo mediante la implementación de medidas técnicas y organizativas, adecuadas y proporcionales, para eliminarlo o mitigarlo. Por último, la tercera fase consiste en evaluar el riesgo residual[85] e implementar medidas para tenerlo controlado durante todo el ciclo de vida de la IA.

[82]BOGROFF A., GUÉGAN D. "*Artificial Intelligence, Data, Ethics An Holistic Approach for Risks and Regulation".* Department of Economics. Ca' Foscari University of Venice. 2019. No. 19. Página 6. ISSN 1827-3580.

[83] Risk based Thinking (ISO 31000:2018). Se trata de un grupo de normas relativas a la gestión de riesgos realizado por la Organización Internacional de Normalización. El propósito de este código es el de proporcionar los principios y directrices genéricas sobre la gestión de riesgos. [En línea: https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-1:v1:en] [visitado el 28/06/2020]

[84]Artículo 24.1 del RGPD, última parte.

[85]Se define riesgo residual como aquel riesgo que subsiste, después de haber implementado los controles o medidas preventivas, directas o correctivas destinadas a minimizar el riesgo.

Tanto si es para determinar el resigo como para implementar las medidas adecuadas para gestionarlo, es importante que se analice fase por fase el ciclo de vida del sistema de IA, asegurado que en cada una de ellas de toman las medidas pertinentes y adecuadas para tratar los peligros que entrañan cada una de ellas.

4.1. Evaluación del nivel de riesgo.

La identificación del riesgo[86] de un tratamiento resulta esencial para poder posteriormente gestionarlo con eficacia e, implementar medidas para minimizarlo o hasta suprimirlo. Una vez identificados estos riesgos, hay que proceder a su evaluación, es decir, determinar el nivel de impacto que pueden llegar a generar. Evaluar un riesgo, según la AEPD[87] *consiste en valorar el impacto de la exposición a la amenaza, junto a la probabilidad de que esta se materialice*. Es decir, si esta amenaza, por ejemplo, puede llegar a ocasionar daños a los derechos y libertades de los interesados o, simplemente, afectará a los diseñadores del mismo tratamiento. *Según la probabilidad y el impacto, asociados a las amenazas, es posible determinar el nivel de riesgo inherente*[88]. Se trata entonces de establecer cuáles pueden ser las vulnerabilidades intrínsecas del sistema de IA y las amenazas a las que se enfrenta.

El RGPD establece una **protección desde el diseño por defecto[89]**. Es decir, según el artículo 25, los riesgos y amenazas de los tratamientos de IA que incorporan datos personales se encuentran desde su inicio o puesta en marcha, evolucionando en función de las variaciones del contexto en el que se implemente y de los factores y elementos que intervengan en él durante todo su ciclo de vida. Por este motivo, se establece una obligación de evaluación y gestión del riesgo desde el inicio del sistema IA y posteriormente durante todo su ciclo de vida. *El responsable del tratamiento aplicará […] medidas técnicas y organizativas apropiadas, como la seudonimización, concebidas para aplicar de forma efectiva los principios de protección de datos, como la minimización de datos, e integrar las garantías necesarias en el tratamiento, a fin de cumplir los requisitos del presente Reglamento y proteger los derechos de los interesados.*[90]

El concepto "por defecto", del artículo 25 hace referencia a la obligación del responsable de garantizar que los datos personales que se están tratando son, por defecto, que sean estrictamente necesarios para cada uno de los fines específicos del tratamiento. *Esta obligación se aplicará a la cantidad de datos personales recogidos, a la extensión de su tratamiento, a su plazo de conservación y a su accesibilidad. Tales medidas garantizarán en particular que, por defecto, los*

[86] El GT29 (WP 248) define el riesgo en sus "*Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679"* de 4 de abril de 2017, como un escenario que describe un evento y sus consecuencias, estimado en términos de impacto y probabilidad. Por tanto, la gestión de riesgos es el conjunto de aquellas actividades y tareas realizadas en una organización para monitorizar y controlar su exposición ante los riesgos. [En línea: https://www.aepd.es/sites/default/files/2019-09/wp248rev01-es.pdf] [visitado el 28/06/2020]

[87] Agencia Española de Protección de Datos. "*Guía Práctica De Análisis De Riesgos En Los Tratamientos De Datos Personales Sujetos Al RGPD"* 2019. Página 3 a 4.

[88] *Ídem.*

[89] Cuando los datos personales del tratamiento estén correctamente anonimizados, y pueda demostrarse, el tratamiento ya no tendrá que sujetarse a los establecido en el RGPD.

[90] Artículo 25.1 del RGPD, última parte.

datos personales no sean accesibles, sin la intervención de la persona, a un número indeterminado de personas físicas[91].

El RGPD establece la obligación de fijar procedimientos de control que garanticen cumplir con los principios de protección desde el diseño y por defecto. También deberán suprimirse o mitigarse estos riesgos mediante medidas de control, gestión y seguridad de las actividades del tratamiento.

Según el artículo 32 del RGPD, el responsable, teniendo en cuenta la evaluación de los riesgos y amenazas del tratamiento concreto, deberá aplicar las medidas técnicas y organizativas para adoptar un nivel de seguridad adecuado. Seguidamente, proporciona un listado no exhaustivo de técnicas de seguridad que deberán implementarse tales como: *la seudonimización y el cifrado de datos personales; la capacidad de garantizar la confidencialidad, integridad, disponibilidad y resiliencia permanentes de los sistemas y servicios de tratamiento; la capacidad de restaurar la disponibilidad y el acceso a los datos personales de forma rápida en caso de incidente físico o técnico, y por último, un proceso de verificación, evaluación y valoración regulares de la eficacia de las medidas técnicas y organizativas para garantizar la seguridad del tratamiento*[92]. Para establecer este nivel de seguridad deberán tenerse particularmente en cuneta los riesgos como el de destrucción, perdida o alteración accidental o ilícita de los datos del tratamiento con sistema de IA. Además, el mismo artículo, en el punto 4, establece la obligación al responsable de controlar la base de datos del tratamiento mediante la limitación de esta todos aquellos que puedan tener acceso por estar bajo su autoridad.

Toda actividad en un tratamiento que incorpora datos personales deberá registrarse desde el inicio y contener toda la información indicada en el artículo 30.1 del RGPD. Este registro de actividades será el que permitirá a la Agencia Española de Protección de Datos o al órgano correlativo a nivel autonómico revisar que el tratamiento se está efectuando conforme a la normativa y que, sobre todo, el responsable está cumpliendo con el principio de responsabilidad proactiva.

Ahora bien, definir la actividad del tratamiento resulta un paso fundamental que requiere acotar los fines del tratamiento de datos personales y, en este caso, los fines del tratamiento que incorpora IA. *Corresponde a cada organización, de acuerdo al principio de responsabilidad proactiva, decidir el nivel de agregación o segregación para elaborar el registro de actividades de tratamiento y deberá valorar hasta qué punto esa agregación o segregación corresponde con finalidades, bases jurídicas y grupos de individuos distintos*[93]. Esto permitirá cumplir con el requisito de transparencia, dejando trazabilidad de todas las actividades y decisiones tomadas en el tratamiento. Definir la actividad o actividades del tratamiento nos ayudará, también, a determinar el tipo de riesgo asociado y seguir la hoja de ruta más adecuada para establecer las medidas de control considerando los riesgos a los que están expuestas las actividades de tratamiento.

Los riesgos asociados a los tratamientos que incorporan datos personales pueden ser divididos en: escasos riesgos, riesgo medio o alto riego. Los tratamientos que sean considerados de alto riesgo para los derechos y libertades de los interesados cuyos datos tratan, según el artículo 35 del RGPD, requerirán una **evaluación de impacto relativa a la protección de los datos** (EIPD). Por

[91] Artículo 25.2 del RGPD, última parte.
[92] Artículo 32.1 RGPD.
[93] Agencia Española de Protección de Datos. "*Guía Práctica De…"* ob-cit. página 19.

este motivo es imprescindible realizar un análisis previo para determinar de forma preliminar el nivel de riesgo al que puede estar expuesto un tratamiento y tomar la decisión adecuada en base a ello.

Si se entiende que las actividades del tratamiento no entrañan riesgos relevantes que motiven la necesidad de realizar una EIPD será necesario documentarlo adecuadamente con los motivos por los cuales se ha llegado a esa conclusión. En cualquier caso, se debe mantener evidencia de que se ha llevado a cabo este análisis.

En definitiva, el responsable del desarrollo, mantenimiento y/o distribución de un componente IA, así como el responsable de un tratamiento que incluya componentes IA, ha de tomar, en cada una de las respectivas etapas y responsabilidades, las medidas oportunas para minimizar o eliminar los factores de riesgo.

4.2. Evaluación del Impacto en la Privacidad.

Puesto que los sistemas de IA que incorporan datos personales, sobre todo aquellos dotados de aprendizaje automático, muy probablemente van a requerir una Evaluación del impacto de la privacidad (EIPD). La EIPD es una herramienta de carácter preventivo que se define como aquel *proceso concebido para describir el tratamiento, evaluar su necesidad y proporcionalidad y ayudar a gestionar los riesgos para los derechos y libertades de las personas físicas derivados del tratamiento de datos personales evaluándolos y determinando las medidas para abordarlos*[94]. La EIPD, además, debe entenderse como un proceso de mejora continua, de forma que siempre que se modifique el sistema o se introduzcan novedades de revisarse, generando un nuevo informe y plan de acción.

El artículo 35.3,4 y 5 RGPD nos ofrecen un listado de supuestos *numerus apertus* en los cuales es obligatorio realizar la EIPD sin necesidad de realizar análisis de riesgos, por lo tanto, la comprobación de las listas y supuestos contenidos en la regulación, debe ser el primer paso para determinar la necesidad de llevar a cabo una EIPD.

El artículo 35.3, tal y como se ha indicado en capítulos anteriores, recoge tres casos en los que es obligatorio hacer la EIPD y el artículo 28.2 LOPDGDD ofrece un listado mucho más extenso de posibles situaciones. Según el RGPD será obligatorio cuando se produzca:

- Una evaluación sistemática y exhaustiva de aspectos personales de personas físicas que se base en un tratamiento automatizado, como la elaboración de perfiles, y sobre cuya base se tomen decisiones que produzcan efectos jurídicos para las personas físicas o que les afecten significativamente de modo similar[95];

[94] El GT29 (WP 248) ob-cit. página 36

[95] Segú el GT29 en el documento (WP 248) ob-cit. página 24, podría llegar a considerase que afectan significativamente: *el seguimiento de las personas en diferentes sitios web, dispositivos y servicios, alterar las expectativas y deseos de las personas afectadas; alterar la forma en que se presenta un anuncio; o el uso de conocimientos sobre las vulnerabilidades de los interesados*.

- Un tratamiento a gran escala de las categorías especiales de datos a que se refiere el artículo 9, apartado 1, o de los datos personales relativos a condenas e infracciones penales a que se refiere el artículo 10, o

- Una observación sistemática a gran escala de una zona de acceso público. El RGPD no define que es un tratamiento de datos a gran escala, no obstante, podemos encontrar alguna orientación en el nº91 de la Exposición de Motivos del RGPD. Para determinar si nos encontramos ante esta situación, el GT29[96] nos recomienda tener en cuenta los cuatro factores siguientes: *el número de interesados afectados, bien como cifra concreta o como proporción de la población correspondiente; el volumen de datos o la variedad de elementos de datos distintos que se procesan; la duración, o permanencia, de la actividad de tratamiento de datos; el alcance geográfico de la actividad de tratamiento*.

Adicionalmente, en los artículos 35.4 y 35.5 del RGPD se prevé la existencia de listas[97] elaboradas por las autoridades de control que determinen en que casuísticas es obligatoria la realización de una EIPD, así como, también, en que situaciones los responsables quedan excluidos[98] de esta obligación.

Si se considera que el tratamiento queda excluido por alguno de los listados realizados por la Agencia Española de Protección de Datos, es necesario garantizar que efectivamente el tratamiento encaja en esa casuística sin lugar a duda documentándose adecuadamente para poder concluir que no es necesario llevar a cabo una EIPD.

Es importante destacar que los listados de la normativa, como se ha dicho, no son exhaustivos, por lo que, si el tratamiento no está incluido en los supuestos y en ninguna de las listas, no implica que no sea necesario llevar a cabo la EIPD. En todo caso, será necesario llevar a cabo una evaluación de riesgos para determinar si según el artículo 35.1 del RGPD y el nº76 de la Exposición de Motivos, no entraña *un alto riesgo para los derechos y libertades de las personas físicas. Si así fuese, el responsable del tratamiento realizará, antes del tratamiento, una evaluación del impacto de las operaciones de tratamiento en la protección de datos personales.* Para evaluar las características y riesgos de las actividades del tratamiento deberá analizarse la naturaleza, el alcance y fines del tratamiento. Podemos hacernos preguntas tales como ¿Se tratan categorías especiales de datos? ¿Se combinan diferentes conjuntos de datos? ¿Se realiza un proceso de toma de decisiones con efectos jurídicos? ¿Se valora la exclusión de beneficios sociales o fiscales?

El GT29 introduce criterios que pueden ayudar a evidenciar el elevado riesgo descrito en el artículo 35.1 del RGPD y que pueden necesitar de la elaboración de un EIPD en el documento WP248 "*Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679"*.

[96] El GT29 (WP 248) ob-cit. página 11

[97] La Agencia Española de Protección de Datos ha publicado una lista de tratamientos que precisan de una EIPD en https://www.aepd.es/sites/default/files/2019-09/listas-dpia-es-35-4.pdf [visitado el 28/06/2020]

[98] La Agencia Española de Protección de Datos ha publicado una lista de tratamientos que quedan excluidos de la realización de una EIPD, no obstante, será necesario que se justifiquen que el tratamiento debe ser catalogado como algo de la lista. En https://www.aepd.es/sites/default/files/2019-09/ListasDPIA-35.5l_0.pdf [visitado el 28/06/2020]

En cualquier caso, cuando estemos ante tratamientos que utilizan algoritmos de Machine Learning que, o bien elabore perfiles o tome decisiones automatizadas, los responsables deberán identificar todas las decisiones en cada una de las etapas del ciclo de vida del sistema, *detallarlas, y analizar los parámetros de funcionamiento* incluyendo *los márgenes de error del sistema y evaluar cuidadosamente que efectos tiene sobre los interesados*[99].

El RGPD establece las características mínimas de una EIPD en los artículos 35.7 y n.º 89 y 90 de la Exposición de Motivos:

- Una **descripción de las operaciones del tratamiento** previstas y los fines de este
- Una **evaluación de necesidad** y **proporcionalidad del tratamiento**. Deberá valorarse si el objeto del tratamiento no puede ser conseguido utilizando otro tipo de solución que alcance la misma finalidad, con un margen de rendimiento aceptable y un nivel de riesgo menor. Es decir, la disponibilidad o novedad de la nueva tecnología no justifica su utilización, sino que debe ser objeto de ponderación.
- Una **evaluación de riesgos para los derechos y libertades** de los propietarios de los datos y las personas que interactúan con el sistema
- Las **medidas preventivas** para afrontar los riesgos y demostrar la conformidad con el Reglamento.

No obstante, el RGPD ofrece flexibilidad a los responsables del tratamiento para determinar la estructura y forma precisas de la EIPD con el fin de permitir que esta se ajuste a las prácticas de trabajo ya existentes[100].

El artículo 35.2 establece la obligación de recabar el asesoramiento del delegado de protección de datos, si ha sido nombrado, al realizar la evaluación de impacto relativa a la protección de datos. El delegado de protección de datos supone una herramienta ofrecidas a garantizar el principio de transparencia y se trata de una de las mejores medidas a las que puede optar el responsable para orientar las políticas de explicabilidad del sistema IA. El artículo 37.1 del RGPD y artículo 34 de la LOPDGDD establecen una serie se situaciones en las que será obligatorio y no potestativo contar con un Delegado de Protección de Datos (DPD).

La utilización se un sistema de IA no implica siempre la obligación de un delegado de protección de datos. Sin embargo, es común que en solución de IA las operaciones que en ella de realicen precisen, por su naturaleza, alcance o fines, *una observación habitual y sistemática de interesados a gran escala* o que las actividades principales del sistema de IA consistan en el *tratamiento a gran escala de categorías especiales de datos personales con del al artículo 9 del RGPD o de datos relativos a condenas e infracciones penales a que se refiere el artículo 10 del RGPD*. En estos casos, y en virtud del artículo 37.1 b) y c), será obligatorio disponer de un DPD.

También es importante tener en cuenta el artículo 35.9 del RGPD que establece que, cuando proceda, *el responsable recabará la opinión de los interesados*[101] *o de sus representantes en*

[99] Agencia Española de Protección de Datos, "*Adecuación al RGPD…"*, ob-cit. página 30.
[100] El GT29 (WP 248) ob-cit. página 19
[101] El concepto de interesado ha de entenderse extendido tanto a los operadores humanos que interpretan o supervisan los resultados de la IA como a los sujetos sometidos a su tratamiento.

relación con el tratamiento previsto, sin perjuicio de la protección de intereses públicos o comerciales o de la seguridad de las operaciones de tratamiento.

El artículo 35.8 del RGPD establece que el cumplimiento de los códigos de conducta de los artículos 40 y 41 se tendrán en cuenta a la hora de evaluar el impacto de la operación de tratamiento de datos. Los códigos de conducta resultan útiles para demostrar que se ha elegido o aplicado un buen paquete de medidas, siempre que el código de conducta sea adecuado a los fines del tratamiento.

La EIPD debe estar documentada y, según el artículo 36 del RGPD, cuando de ella de deduzca un alto riesgo residual, se exigirá al responsable del tratamiento que eleve la consulta a la Autoridad de Protección de Datos antes de proceder con el tratamiento.

4.3. Aspectos determinantes en la gestión del riesgo de los sistemas de Inteligencia Artificial.

Hasta ahora hemos estado viendo cómo debe realizarse el análisis del riesgo en cualquier tratamiento que incorpore datos personales según el RGPD. Si bien es cierto que los tratamientos que incorporan IA comparten numerosos riesgos y amenazas con otras muchas nuevas tecnologías, existen algunos aspectos que sólo inciden cuando se realizan tratamientos con IA y a los que deberemos estar especialmente atentos. En esta última sección pretendo mencionarlos brevemente, ya que no dejan de ser las mismas preocupaciones a las que se hicieron referencia en el capítulo de ética de los sistemas de IA.

Antes de proseguir, creo interesante mencionar que el HLEG IA han creado una lista no exhaustiva (*check list)*[102] de 147 preguntas para la evaluación de la fiabilidad de un sistema IA, en particular para aquellos que interactúan con individuos, dirigida principalmente a los desarrolladores y responsables de los sistemas. La utilización de esta lista sirve como apoyo y guía en el desarrollo de tratamientos que incorporan IA, pero no constituye una prueba del cumplimiento legal ni sirve como guía para el cumplimento de la legislación vigente.

1º Especial vigilancia de la exactitud.

Como ya hemos vistos, los modelos de IA pueden tener sesgos, que pueden afectar a la exactitud del sistema y sus decisiones. En particular, el nº 71 de la Exposición de Motivos contiene una obligación dirigida a los responsables de los tratamientos de utilizar *procedimientos matemáticos o estadísticos adecuados para […] 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*.

Los factores que pueden influir en la exactitud del sistema o de los datos y, que finalmente deriven en sesgos son tres, en tres diferentes etapas de la concepción del sistema. En primer lugar, la elección del algoritmo o el modelo de IA adecuado para los fines o el tratamiento en el que queremos introducir el sistema. Un algoritmo o modelo inadecuado puede generar sesgos en sus decisiones o resultados al no tratar los datos de la forma más adecuada. Incluso podría suceder

[102] High-Level Expert Group on Artificial Intelligence. ob-cit. página 30 a 41.

que el sesgo proviniera de un error en la programación del sistema. En segundo lugar, el sesgo puede producirse porque el conjunto de datos que han sido escogidos para el entrenamiento o la validación del sistema son erróneos o no son representativos. Por último, la evolución sesgada también puede ser un detonante de la inexactitud del sistema. Hay que cuidar el ambiente en el que implementamos un sistema que incorpora aprendizaje automático o técnicas adaptativas, ya que, como hemos visto en ejemplos anteriormente dados, esto puede afectar a sus decisiones o resultados.

Actualmente, se están desarrollado los llamados "*IA guardianes*"[103]. Estas son técnicas orientadas a examinar y determinar de forma independiente la existencia de sesgos en los algoritmos de los sistemas de IA como, por ejemplo, *Algorithmic Impact Assessment* (AIA). Estos sistemas además deberán estar diseñados de tal manera que también puedan ser supervisados y monitorizados por humanos. No obstante, esta funcionalidad no evita completamente los sesgos.

2º La importancia de la explicabilidad del sistema

La explicabilidad es un factor fundamental para los sistemas de IA, sobre todo aquellos modelos de aprendizaje automático. Hoy en día sigue siendo uno de los mayores desafíos de los sistemas de IA que utilizan algoritmos de redes neuronales. Como se ha visto en capítulos anteriores, que no se pueda desglosar y explicar las decisiones que un sistema de IA puede afectar al derecho de defensa de los individuos incidiendo en gran medida sobre sus derechos y libertades. Es por este motivo sumamente importante desarrollar un sistema trazable y mantener un registro de las actividades que en él se desarrollen.

Se ha visto que se debe procurar informar a los afectados por el sistema de IA sobre su funcionamiento, alcance, fines, riesgos, etc. de modo que les permita entender el comportamiento del sistema. Esto generará una mayor transparencia del sistema y de los derechos y libertades de las personas. Sin embargo, a menudo, nos olvidamos de que es igualmente importante mantener al personal que trabaja con el sistema informado sobre las limitaciones de este. Aún más cuando el modelo toma decisiones automatizadas. *Será necesario en estos casos tomar medidas para gestionar el riesgo de que el elemento humano se comporte como una mera correa de transmisión de las inferencias realizadas por la solución IA*.[104] Estas medidas pueden estar compuestas por la información y formación adecuada a todos aquellos que emplean el sistema y posteriores auditorias de su comportamiento.

En el caso que, la solución sea transmitida a un tercero, el responsable ha de proporcionar la información suficiente para que pueda gestionar los riesgos en el nuevo ambiente en el que vaya a ser puesta en marcha.

3º Indicadores de calidad del sistema y certificación

Este tercer factor se encuentra muy estrechamente relacionado con la explicabilidad, pero debido a su importancia, creo necesario situarlo como punto separado. El artículo 42 del RGPD brinda

[103] COTINO HUESO L. "Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho". *Revista Catalana de Dret Públic*. (58). 2019. Página 43 a 44. [En línea: http://revistes.eapc.gencat.cat/index.php/rcdp/article/view/10.2436-rcdp.i58.2019.3303] [visitado el 28/06/2020]

[104] Agencia Española de Protección de Datos, "*Adecuación al RGPD…"*, ob-cit. página 35 a 38.

la posibilidad de desarrollar mecanismos de certificación o indicadores de calidad específicos con el fin de garantizar se han ensayado, validado y desarrollado de forma segura y acorde con la normativa aplicable en cada sector.

Teniendo en cuenta que no todas las personas implicadas en un tratamiento que incorpora IA van a entender plenamente su funcionamiento y efectos que este pueda generar, puede ser positio crear un sistema de certificación que pueda acreditar ante el público que el sistema de IA es trasparente, responsable y equitativo[105]. Además, puede ser útil y necesario para proteger los derechos conferidos por la propiedad industrial de los sistemas de IA. Ciertamente, así lo estipula el nº100 de la Exposición de Motivos del RGPD:

> *A fin de aumentar la transparencia y el cumplimiento del presente Reglamento, debe fomentarse el establecimiento de mecanismos de certificación y sellos y marcas de protección de datos, que permitan a los interesados evaluar con mayor rapidez el nivel de protección de datos de los productos y servicios correspondientes.*

No obstante, hay que tener presente que certificación tanto del sistema de IA como del tratamiento en el que se incorpora nunca puede substituir la responsabilidad.

4º Incorporación de medidas de privacidad de los datos desde el diseño.

La privacidad de los datos es una preocupación compartida con muchas tecnologías, pero por su importancia, he creído conveniente hacer mención en este apartado. La incorporación de medidas de privacidad debe seguir los siguientes principios básicos según la AEPD[106]:

- **Minimizar** los datos, tanto en volumen como en el tamaño de la población de estudio, así como el tiempo en el que estos datos se utilizan. Es decir, partir de una selección o muestra de individuos relevante en relación con el objeto del estudio, estableciendo un criterio de selección conservador y tratar solo aquellos datos que cumplan con el criterio. Una vez los datos ya no sean útiles, o relevantes con el objeto del sistema de IA.

- **Ocultar** los datos personales y sus interrelaciones para limitar su exposición y mantener su confidencialidad y desvinculación. De las distintas técnicas que se ofrecen para llevar a cabo esta función, en los sistemas de IA parece resultar útil la técnica de **Agregación**. Esta consiste en agrupar los datos personales relativos a varios sujetos utilizando técnicas de generalización y supresión para evitar el nivel de detalle y posibles correlaciones.

- **Separar los contextos del tratamiento** para dificultar la correlación, así como la posibilidad de inferir en los datos. Es decir, preocupar el almacenamiento separado de los datos implicados en cada contexto del tratamiento en vez de llevar a cabo un almacenamiento único de los mismos.

[105] Así lo defiende el IEEE Advancing Technology for Humanity en "*Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*" (EADe) Página 16 y 27. Estas son un grupo de normas de *soft law* que, al igual que las ISO, ayudan a los desarrolladores a garantizar que los tratamientos tecnológicos respeten desde el diseño los derechos fundamentales mediante códigos éticos y profesionales o de acreditación. [En línea: https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf] [visitado el 28/06/2020]

[106] *Ídem.* página 31 a 33.

- Otra técnica utilizada en áreas de preservar la privacidad de los individuos y fomentar la Transparencia y Explicabilidad del sistema es **permitir a los interesados tomar decisiones informadas** sobre el tratamiento de sus datos.

- Proporcionar medios a los propietarios de los datos para que puedan controlarlos mediante la implementación de mecanismos que les permitan efectuar el ejercicio de los derechos que le son conferidos por el RGPD.

A demás, el HLEG IA[107] recomienda la inclusión en los sistemas de IA de un mecanismo de apagado a prueba de fallos, que además posibilite la reanudación del funcionamiento del sistema tras el apagado forzoso.

5º Verificación y Validación

La realización de pruebas antes de la puesta en marcha del sistema y la posterior validación de este es una parte capital en su desarrollo. El responsable ha de tener presente que la inclusión de un sistema de IA verificado en un tratamiento específico no comporta que éste sea válido, ni mucho menos que aquel componente de IA sea adecuado en el ambiente de ese tratamiento. Por lo que la verificación del sistema de IA se extrae de la realización de las pruebas pertinentes para demostrar el funcionamiento seguro y robusto del sistema.

Por otro lado, la validación es el resultado positivo del análisis del sistema de IA introducido en el tratamiento específico, que garantiza que dicha combinación cumple con los resultados planificados para el producto o servicio concreto. Pretende garantizar que el sistema se comporta según lo previsto a lo largo de todo su ciclo de vida, y en particular, durante el despliegue. *Se asegura de que los productos y servicios resultantes satisfagan los requisitos relativos a una aplicación específica o uso previsto*[108]. En definitiva, es esencial que la validación de un tratamiento que incluya IA se realice en las condiciones del entorno real en el que se pretende desplegar. Como todo el análisis y gestión del riesgo, la validación de los sistemas de un tratamiento deberá validarse periódicamente, en la medida que evolucionen dichas condiciones o el tratamiento en sí.

5. CONCLUSIONES.

1) La IA es un sistema tecnológico de comportamiento inteligente diseñado por seres humanos con el fin de alcanzar unos fines específicos y que actúan en la dimensión física o digital mediante la percepción de su entorno a través de la obtención de datos y su interpretación.

2) La IA está suponiendo un cambio de paradigma y plantea un gran reto para el derecho y los juristas. Si bien es cierto que la IA trae consigo innumerables beneficios, pudiendo llegar a solucionar problemas con los que llevamos años combatiendo, también acarrea nuevos retos y peligros que necesariamente deberán ser abordados cuanto antes.

3) En particular preocupan aquellos sistemas de IA que utilizan o tratan datos de carácter personal y sobre todo aquellos que toman decisiones que afectan a las personas físicas tanto

[107] High-Level Expert Group on Artificial Intelligence, abril de 2019. ob-cit. página 26 a 27.
[108] ISO 9001:2015 Sistemas de gestión de la calidad. Requisitos.

en su entorno social como laboral, al acceso a contratos o servicios, la personalización de dichos servicios, etc. Precisamente por ese motivo, los sistemas de IA que usan datos personales, como un modelo de perfilado de marketing o electoral, deben adecuarse a los dispuesto en el Reglamento (UE) 2016/679 y la normativa sectorial para asegurar la correcta protección de los derechos de los titulares de los datos. Al margen de dicha normativa quedarán, sin embargo, los sistemas de IA que no usan datos de carácter personal. Por ejemplo, un modelo industrial de control de calidad o los sistemas de toma de decisiones de compra y venta de productos financieros.

4) El derecho fundamental a la autodeterminación informativa es fruto de una evolución normativa europea que acabara influyendo en el Ordenamiento Jurídico Español. Éste se inicia como un derecho derivado del derecho a la vida privada y la intimidad y finaliza con su proclamación en el artículo 16.1 TFUE como un derecho fundamental independiente y autónomo. Así, el derecho a la protección de los datos personales se ha cristalizado en una normativa que finalmente ha alcanzado un estatus normativo completo con el RGPD de 2016 fue transpuesto en España en la LOPDGDD de 2018.

5) La disponibilidad de una tecnología o su novedad no es razón suficiente para su explotación y comercialización indiscriminada. Es necesario un debate acerca de los riesgos que acarrea esta nueva tecnología con el fin de estar preparados adecuadamente para enfrentarnos a ellos. En este trabajo se han identificado cuatro de los mayores dilemas y preocupaciones que acompañan a todos y cada uno los sistemas de IA, se integren en el ambiente que se integren. Estos han sido, en primer lugar, la privacidad de las personas y de sus datos. En segundo lugar, el sesgo de los sistemas de IA y, en particular, el sesgo humano en relación con decisiones de los sistemas IA (*biax de confirmación*). En tercer lugar, otra de las grandes preocupaciones es la transparencia de los sistemas, los problemas de la caja negra. En cuarto y último lugar, la atribución de responsabilidad de los sistemas en caso de daños y, en particular, de aquellos sistemas de aprendizaje automático o Machine Learning.

6) El HLEG AI elabora una serie de documentos entorno al desarrollo de sistemas de IA más éticos, proporcionando mecanismos a los desarrolladores para alcanzar dicho objetivo. El grupo de 52 expertos parte del concepto de *IA fiable*, compuesto por tres componentes (licitud, ética y robustez) que, a la vez, giran entorno de los derechos fundamentales (respeto a la dignidad humana, a libertad del individuo, respeto de la democracia, la justicia y el estado de derecho, y los derechos de los ciudadanos) y de una serie de principios éticos, algunos de ellos ya bien conocidos por la bioética (el principio respeto a la autonomía humana, prevención del daño, equidad y transparencia). No obstante, quedan abiertas algunas cuestiones que no quedan suficientemente resueltas por el HLEG AI. El documento no proporciona las herramientas y mecanismos a los desarrolladores de los proyectos de IA para efectuar un correcto juico de proporcionalidad o ponderación entre los derechos fundamentales afectados. Mas bien, el documento parce que se limita a ofrecer un catálogo de derechos y principios a escoger, trasladando a los desarrolladores la obligación de interpretarlos correctamente.

7) El objetivo de este trabajo ha sido ofrecer una base para poder abordar el desarrollo de un sistema de IA, pero no cubre todos los posibles peligros que se pueden derivarse de un sistema concreto. Para poder identificar los riesgos de un sistema IA en particular, tendremos que analizar el entorno determinado en el que va a desarrollarse y analizar el cumplimiento de la normativa sectorial aplicable. En este contexto, el RGPD exige como punto clave de la

responsabilidad proactiva, que se implementen medidas de análisis y gestión de los riesgos con el fin de establecer las medidas de seguridad y control para garantizar los derechos y libertades de las personas, tal como se establece en el artículo 24.1 y 25.1 y 2 del RGPD. La gran variedad de posibles tratamientos de datos que pueden llevarse a cabo en un sistema de IA imposibilita ofrecer unos parámetros exactos y obligatorios, comunes para todos los tratamientos. Como ya se ha visto, la IA es una tecnología que entraña multitud de retos y riesgos que además variaran en gran medida en función del campo en el que los vayas a implementar. Así, un sistema de IA en el ámbito sanitario se llevará a cabo de forma muy distinta a uno que pueda usarse en el mundo financiero, generando un abanico de riesgos que necesariamente, deberán tratarse *ad hoc*.

8) Sin embargo, he podido identificar cinco aspectos determinantes que deberían tenerse en cuenta en la elaboración de la gestión del riesgo de un sistema de IA. Se ha hecho referencia a la importancia de la exactitud de los datos que se introducen en el sistema de IA con el fin de evitar sesgos. Hemos visto que los sesgos pueden surgir de una mala elección en el algoritmo del sistema en relación con los fines a alcanzar, de una mala programación del sistema, puede surgir cuando la muestra de datos escogida no es suficientemente representativa o errónea y, por último, puede suceder que el sistema evolucione sesgadamente y esto derive en una inexactitud de los datos ex post. Con el fin de combatir la inexactitud de los datos se han creado los llamados "IA guardianes". Estas son sistemas de IA orientados a examinar y determinar de forma independiente la existencia de sesgos en los algoritmos de los sistemas de IA. Pero como bien señala la AEPD, el principal problema del sistema de IA no son los mismos sino como los usan y los conciben las personas. *En particular, es necesario prestar especial atención a atribuir responsabilidades a componentes IA sin supervisión y sin adoptar una posición crítica. La delegación de la toma de decisiones a máquinas no es nueva; ya ocurre con los algoritmos deterministas, pero el sesgo de atribuir una autoridad o peso superior a un resultado inferido por una solución de IA puede hacer incrementar los riesgos derivados de esta delegación de responsabilidad.*[109]

9) También se menciona la importancia de la explicabilidad de los sistemas y de su trazabilidad. El problema de las Cajas Negras sigue siendo un de los mayores desafíos de los sistemas de IA. Que no se pueda desglosar y explicar las decisiones que un sistema de IA puede afectar al derecho de defensa de los individuos incidiendo en gran medida sobre sus derechos y libertades. Por este motivo se debe procurar informar a los afectados por el sistema de IA sobre su funcionamiento, alcance, fines y riesgos, de modo que les permita entender el comportamiento del sistema, así como al personal que trabaja con el mismo. Otra forma de mejorar la explicabilidad y la transparencia de los sistemas de IA es mediante los sistemas de certificación o los indicadores de calidad que a su vez ayudaran a combatir los problemas que puede derivarse de la propiedad industrial.

10) Existe, además, una brecha entre la comprensión y cohesión de los trabajos que provienen de las ciencias sociales y los que tienen una formación más técnica. Es necesario que se unan fuerzas y se apueste por la interdisciplinariedad para la formación de políticas enfocadas a la ética en el desarrollo de las nuevas tecnologías en general. La reglamentación que surja no tiene por qué limitarse únicamente a la prohibición de conductas, sino que también será necesario una ética positiva y una reflexión sobre donde queremos que nos lleve la IA. No hay que olvidar que pese a los innumerables riesgos que suponen los distintos usos y

[109] Agencia Española de Protección de Datos, "*Adecuación al RGPD…"*, ob-cit. página 49.

aplicaciones de la IA en nuestras vidas, un buen uso de las mismas puede suponer un gran avance para la humanidad y hasta un apoyo para proteger la privacidad e implementar nuevos mecanismos que aseguran la protección de los datos.

6. BIBLIOGRAFÍA.

40TH INTERNATIONAL CONFERENCE OF DATA PROTECTION AND PRIVACY COMMISSIONERS. "*Declaration On Ethics And Data Protection In Artifical Intelligence".* (2018) [En línea: https://edps.europa.eu/sites/edp/files/publication/icdppc-40th_ai-declaration_adopted_en_0.pdf] [visitado el 28/06/2020]

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS "*Guía de Privacidad desde el Diseño*" (2019) [En línea: https://www.aepd.es/sites/default/files/2019-11/guia-privacidad-desde-diseno.pdf] [visitado el 28/06/2020]

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. "Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción" (2020) [En línea: https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia.pdf] [visitado el 28/06/2020]

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. "*Guía Práctica De Análisis De Riesgos En Los Tratamientos De Datos Personales Sujetos Al RGPD*" (2019) [En línea: https://www.aepd.es/sites/default/files/2019-09/guia-evaluaciones-de-impacto-rgpd.pdf] [visitado el 28/06/2020]

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. "*Guía práctica para LAS Evaluaciones de Impacto en la Protección de LOS datos sujetas al RGPD*" (2019) [En línea: https://www.aepd.es/sites/default/files/2019-09/guia-evaluaciones-de-impacto-rgpd.pdf] [visitado el 28/06/2020]

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. "*Listas De Tipos De Tratamientos De Datos Que Requieren Evaluación De Impacto Relativa A Protección De Datos (Art 35.4)"* (2018) [En línea: https://www.aepd.es/sites/default/files/2019-09/listas-dpia-es-35-4.pdf] [visitado el 28/06/2020]

AGENCIA ESPAÑOLA DE PROTECCIÓN DE DATOS. "*Orientaciones y garantías sobre los procedimientos de anonimización de datos personales*" (2016) [En línea: https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf] [visitado el 28/06/2020]

ALGORITHM WATCH, AUTOMATING SOCIETY. "*Taking stock of Automated Decision-Making in the EU*" (2019) [En línea: https://algorithmwatch.org/en/automating-society/] [visitado el 28/06/2020]

AUTORIDAD CATALANA DE PROTECCIÓN DE DATOS. *"Inteligencia Artificial. Decisiones Automatizadas en Catalunya"* (2020) [En línea: https://apdcat.gencat.cat/es/documentacio/jornades_i_congressos/apdcat/2020/] [visitado el 28/06/2020]

BOGROFF A., GUÉGAN D. "*Artificial Intelligence, Data, Ethics An Holistic Approach for Risks and Regulation".* Department of Economics. Ca' Foscari University of Venice. 2019. No. 19. ISSN 1827-3580

CADWALLADR C. and GRAHAM-HARRISON E., "Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach: Whistleblower describes how firm linked to former Trump adviser Steve Bannon compiled user data to target American voters". *The Guardian*, 17 de marzo 2018. [En línea: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election] [visitado el 28/06/2020]

CISPA HELMHOLTZ CENTER FOR INFORMATION SECURITY. "*Comments on the "Draft Ethics Guidelines for Trustworthy AI" by the High-Level Expert Group on Artificial Intelligence*." (2019) [En línea: https://publications.cispa.saarland/2798/] [visitado el 28/06/2020]

COECKELBERGH M. "Ethics of artificial intelligence: Some ethical issues and regulatory challenges". *Technology and Regulation.* (2019) [En línea: https://doi.org/10.26116/techreg.2019.003] [visitado el 28/06/2020]

COMUNICACIÓN (COM (2010) 609 final) de la Comisión, de 4 noviembre 2010, "*Un enfoque global de la protección de los datos personales en la Unión Europea".*[En línea: https://eur-lex.europa.eu/legal content/ES/TXT/PDF/?uri=CELEX:52010DC0609&from=es] [visitado el 28/06/2020]

COMUNICACIÓN (COM (2012) 9 final) de la Comisión, de 25 enero 2012, "*La protección de la privacidad en un mundo interconectado: Un marco europeo de protección de datos para el siglo XXI*". [En línea: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52012DC0009&from=ES] [visitado el 28/06/2020]

COMUNICACIÓN (COM (2018) 237 final) de la Comisión, de 24 abril 2018, "*Artificial Intelligence for Europe*". [En línea: https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe] [visitado el 28/06/2020]

COMUNICACIÓN (COM (2020) 65 final) de la Comisión, de 19 de febrero 2020, "*WHITE PAPER. On Artificial Intelligence - A European approach to excellence and trust*". [En línea: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf] [visitado el 28/06/2020]

CONFESSORE N., and HAKIM D." Data Firm Says, "Secret Sauce" Aided Trump, Many Scoff". *The New York Times.* 6 de marzo 2018. [En línea: https://www.nytimes.com/2017/03/06/us/politics/cambridge-analytica.html] [visitado el 28/06/2020]

COTINO HUESO L. "Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho". *Revista Catalana de Dret Públic.* (58), 2019. [En línea: http://revistes.eapc.gencat.cat/index.php/rcdp/article/view/10.2436-rcdp.i58.2019.3303] [visitado el 28/06/2020]

GRUPO DE TRABAJO DEL ARTÍCULO 29 (*WP 136)*. "*Dictamen 05/2014 sobre técnicas de anonimización"*. 10 de abril de 2014 [En línea: https://www.aepd.es/sites/default/files/2019-12/wp216-es.pdf] [visitado el 28/06/2020]

GRUPO DE TRABAJO DEL ARTÍCULO 29 (*WP 136)*. "*Dictamen 4/2007 sobre el concepto de datos personales*, de 20 de junio de 2007. [En línea: https://ec.europa.eu/justice/article-29/documentation/opinion recommendation/files/2007/wp136_es.pdf] [visitado el 28/06/2020]

GRUPO DE TRABAJO DEL ARTÍCULO 29 (WP 217) "*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".* 9 de abril de 2014. [En línea: https://www.aepd.es/sites/default/files/2019-12/wp217_es_interes_legitimo.pdf] [visitado el 28/06/2020]

GRUPO DE TRABAJO DEL ARTÍCULO 29 (WP 248) "*Directrices sobre la evaluación de impacto relativa a la protección de datos (EIPD) y para determinar si el tratamiento «entraña probablemente un alto riesgo» a efectos del Reglamento (UE) 2016/679"* de 4 de abril de 2017. [En línea: https://www.aepd.es/sites/default/files/2019-09/wp248rev01-es.pdf] [visitado el 28/06/2020]

GRUPO DE TRABAJO DEL ARTÍCULO 29 *(WP251)*. "*Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679".* 6 de febrero de 2018. [En línea: https://www.aepd.es/sites/default/files/2019-12/wp251rev01-es.pdf] [visitado el 28/06/2020]

GRUPO DE TRABAJO DEL ARTÍCULO 29 (WP259). "*Directrices sobre el consentimiento en el sentido del Reglamento (UE) 2016/679".* 28 de noviembre de 2017. [En línea: https://www.avpd.euskadi.eus/contenidos/informacion/20161118/es_def/adjuntos/wp259rev01_es20180709.pdf] [visitado el 28/06/2020]

GRUPO EUROPEO SOBRE LA ÉTICA DE LA CIENCIA Y LAS NUEVAS TECNOLOGÍAS. "Statement on Artificial Intelligence, Robotics and "Autonomous" Systems" (2018) [En línea: http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf] [visitado el 28/06/2020]

HAGENDORFF T. "The Ethics of AI Ethics: An Evaluation of Guidelines" *Minds and Machines*. 1 de octubre de 2019. [En línea: https://doi.org/10.1007/s11023-020-09517-8] [visitado el 28/06/2020]

HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE "*The Ethics Guidelines for Trustworthy Artificial Intelligence*". (2019) [En línea: https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines] [visitado el 28/06/2020]

HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE. (2019). "*Definition developed for the purpose of the AI HLEG's deliverables".* [En línea: https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top] [visitado el 28/06/2020]

JOBIN A. IENCA M. VAYENA E. "Artificial Intelligence: the global landscape of ethics guidelines". *Health Ethics & Policy Lab, Zurich, Switzerland.* 2019. [En línea: https://arxiv.org/ftp/arxiv/papers/1906/1906.11668.pdf] [visitado el 28/06/2020]

KOZLOWSKA I. "Facebook and Data Privacy in the Age of Cambridge Analytica*".* 30 de abril de 2018. *The Henry Jackson School of International Studies. University of Washington* [En línea: https://jsis.washington.edu/news/facebook-data-privacy-age-cambridge-analytica/] [visitado el 28/06/2020]

LUCAS MURILLO DE LA CUEVA, P. "*El derecho a la autodeterminación informativa*" Madrid: Tecnos. 1991. ISBN**:** 84-309-1937-6.

LUCAS MURILLO DE LA CUEVA, P. "*Informática y protección de datos personales (estudios sobre la Ley Orgánica 5/1992 de Regulación del Tratamiento Automatizado de los Datos de Carácter Personal*)". Madrid: Centro de Estudios Políticos y Constitucionales. 1993. ISBN: 84-259-0940-6.

LUCAS MURILLO DE LA CUEVA, P. "La Constitución y el derecho a la autodeterminación informativa". *Cuadernos de derecho público.* 2003. n. º19-20. ISSN 1138-2848.

MARTÍNEZ, R. "Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo" *Revista Catalana de Dret Públic*. n. º58. 2019 [En línea: https://doi.org/10.2436/rcdp.i58.2019.3317] [visitado el 28/06/2020]

MAYOR GÓMEZ, R. "Principales novedades de la Ley Orgánica 3/2018, de 5 de diciembre, de protección de datos personales y garantía de los derechos digitales". *Gabilex. revista del gabinete jurídico de Castilla-la Mancha.* nº16. 2018. [En línea: https://dialnet-unirioja-es.sire.ub.edu/servlet/articulo?codigo=6901855] [visitado el 28/06/2020]

MIGUEL ARIZA L. "Lo saben todo de usted". *El País* 9 de diciembre 2015. [En línea: https://elpais.com/elpais/2015/12/04/eps/1449252033_849371.html] [visitado el 28/06/2020]

PASCUAL ORTEGA M.ª T., *"*¿Inteligencia? Artificial", *Tribuna.* vol. 208. 2018. [En línea: https://www.coit.es/archivo-bit/primavera-2018/inteligencia-artificial] [visitado el 28/06/2020]

POUS DE LA FLOR, Mª P. La protección de datos personales. In Mª D Díaz-Ambrona (Eds.), *Derecho Civil de la Unión Europea.* Madrid: Tirant lo Blanch, 2017. [En línea: https://biblioteca-nubedelectura-com.sire.ub.edu/cloudLibrary/ebook/show/9788491431978] [visitado el 28/06/2020]

QUADERNS DE CAC. *Fake news, algoritmos y burbujas informativas* (44 Ed.), Vol. XXI. (2018). Barcelona: Consell d'Audiovisual de Catalunya. [En línea: https://www.cac.cat/sites/default/files/2018-08/Q44_ES.pdf] [visitado el 28/06/2020]

RUIZ MIGUEL, C. *"*El derecho a la protección de los datos personales en la Carta de Derechos Fundamentales de la Unión Europea: análisis crítico*". Revista de Derecho Comunitario Europeo*, n.º 7, n.º 14. (2003) [En línea: https://dialnet-unirioja-es.sire.ub.edu/servlet/articulo?codigo=635290] [visitado el 28/06/2020]

THE CHINESE UNIVERSITY OF HONG KONG FACULTY OF LAW "*Artificial Intelligence Governance and Ethics: Global Perspectives"*. nº15 (2019) [En línea: https://ssrn.com/abstract=3414805] [visitado el 28/06/2020]