Doxa. Cuadernos de Filosofía del Derecho, (2022), *45*
e-ISSN: 2386-4702 | ISSN: 0214-8676

¿Es posible una Inteligencia artificial respetuosa con la protección de datos?

Is Data Protection-friendly Artificial Intelligence Possible?

Celia Fernández-Aller y M.ª Mercedes Serrano Pérez

Autores:
Celia Fernández-Aller
Universidad Politécnica de Madrid, España
mariacelia.fernandez@upm.es
https://orcid.org/0000-0002-0642-2058

Mª Mercedes Serrano Pérez
Universidad de Castilla La Mancha, España
Mercedes.Serrano@uclm.es
https://orcid.org/0000-0001-6955-2063

Recibido: 8-6-2021
Aceptado: 24-9-2021

Citar como:
Fernández-Aller, Celia y Serrano Pérez, M.ª Mercedes (2022). ¿Es posible una Inteligencia artificial respetuosa con la protección de datos? Doxa. Cuadernos de Filosofía del Derecho, *45*, pp. 307-336. https://doi.org/10.14198/DOXA2022.45.11

Abstract

La Inteligencia artificial es un reto clave en la cuarta revolución industrial que vivimos. El artículo pretende analizar la compatibilidad de ésta con el derecho a la protección de datos. Para ello se analizan los principios que subyacen a la regulación y se estudia el encaje de las características de la tecnología en los mismos. Se destacan los principales derechos de los sujetos, así como las obligaciones de los responsables de los tratamientos.

Especial atención se presta a la discriminación algorítmica en el caso de los perfiles y el marketing político, que supone un reto a evitar por su vulneración del principio de igualdad y justicia en el tratamiento de la información personal. Resulta esencial salvaguardar el derecho a la explicabilidad algorítmica. Se concluye con un análisis de la necesaria regulación que está en fase de discusión en Europa.

Palabras clave: discriminación algorítmica; inteligencia artificial; privacidad

Abstract

Artificial Intelligence is a key challenge in the fourth industrial revolution we are living. The article aims to analyze its compatibility with the right to data protection. the principles underlying the regulation are analyzed and the fit of the characteristics of the technology with these principles is studied. The main rights of the subjects are highlighted, as well as the obligations of data controllers.

Special attention is paid to algorithmic discrimination in the context of profiles and political marketing, which is a challenge to be avoided due to its violation of the principle of equality and fairness in the processing of personal information. It is essential to safeguard the right to algorithmic explainability. It concludes with an analysis of the necessary regulation currently under discussion in Europe.

Keywords: algorithmic discrimination; artificial intelligence; privacy

1. INTRODUCCIÓN Y CONTEXTO: AMENAZAS DE LA IA PARA LA PROTECCIÓN DE DATOS

La Inteligencia Artificial (IA) implica el diseño y la implementación de sistemas que presentan capacidades de la mente humana como el razonamiento, el conocimiento, la percepción, la planificación, el aprendizaje y la comunicación. La IA engloba una serie de sub-disciplinas como aprendizaje automático, satisfacción de restricciones, búsqueda y los sistemas multi-agente, el razonamiento y la ingeniería y el procesamiento del lenguaje natural[1].

Las definiciones sobre inteligencia artificial son abundantes, aunque no armonizadas entre sí. Podemos utilizar aquella que define la IA como:

«un conjunto de tecnologías cuya finalidad es la de reproducir la inteligencia humana, es decir, el funcionamiento analítico de un cerebro. Para alcanzar este objetivo ha sido fundamental el desarrollo del machine learning o aprendizaje automático, donde por medio de algoritmos de aprendizaje y un gran volumen de datos, se pueden extraer patrones o insights y de esta forma resolver problemas y/o predecir comportamiento, entre otros» (Bonmatí Sánchez, J., y Gonzalo Doménech, J. J., 2020).

La definición contenida en la propuesta de Reglamento sobre inteligencia artificial define un sistema de inteligencia artificial como:

> «software que se desarrolla con una o más de las técnicas y enfoques enumerados en el Anexo I y pueden, para un conjunto dado de objetivos definidos por el ser humano, generar resultados como contenido, predicciones, recomendaciones o decisiones que influyen en los entornos con los que interactúan» (art. 3.1).

Los sistemas de Inteligencia Artificial (IA) reciben cantidades ingentes de datos, muchos de los cuales son de carácter personal. De hecho, algunos de los dispositivos basados en IA son recolectores masivos de datos personales que se infiltran de manera imperceptible nuestra vida cotidiana. Es decir, no todos los datos que se recogen cuentan con conocimiento y consentimiento de los y las interesadas. A esto hay que añadir el hecho de que los sistemas de IA procesan de manera cada vez más compleja toda esta información personal y consiguen resultados a menudo impensables a partir de unos datos desagregados y muchas veces anónimos. Por todo ello, la protección de datos es uno de los retos que suelen señalarse en el caso de los sistemas de IA.

Hay que tener en cuenta la capacidad que tienen los sistemas de IA para influir en las decisiones humanas a través del análisis de grandes cantidades de datos (a menudo personales) en muchos terrenos (desde decisiones comerciales a elecciones políticas). Por eso se propone un seguimiento de las diferentes tecnologías IA para evitar que la aplicación de la IA al tratamiento de los datos personales restrinja la libertad real o

1. Una buena introducción a la Inteligencia artificial puede encontrarse en Smith, B. C., (2019) *The promise of artificial intelligence. Reckoning and judgment*, The MIT Press, Cambridge (Massachusets). Además, es muy descriptivo el texto de Tegmark, M. (2017) *Life 3.0. Being human in the age of Artificial Intelligence*. Vintage.

percibida de las personas y se proteja de dichas influencias a colectivos especialmente vulnerables como es el caso de los niños. En general, la IA debiera guiarse por un enfoque basado en derechos humanos[2]. Esto implica poner a las personas más vulnerables en el centro de nuestras prioridades, utilizando los principios de los derechos humanos como guía de nuestras decisiones.

En todos estos procesos que implican el tratamiento de datos personales deben respetarse los requisitos fijados en la normativa de protección de los derechos de las personas en relación con el tratamiento de sus datos personales, en concreto en el Reglamento General de Protección de Datos –Reglamento (UE) 2016/679– (RGPD), así como la Ley de protección de datos en el caso español.

La privacidad supone un espacio de autonomía personal en la sociedad digital. Como cualquier derecho fundamental, la privacidad tiene límites[3], claro está, que son especialmente relevantes con ocasión de la IA.

De entre las posibles aplicaciones de la IA con impacto en la protección de datos (técnicas de reconocimiento fácil y utilización de datos biométricos, uso de sistemas de IA en los procesos de gestión de datos y de toma de decisiones, el uso de IA con funciones predictivas, la prevención de la discriminación algorítmica, la elaboración de perfiles), nos centraremos en la recopilación y tratamiento de datos por parte de las organizaciones políticas con fines de comunicación a través del uso de técnicas de big data e IA.

Este asunto fue especialmente relevante en el caso de *Cambridge Analytica*, en el que un test de personalidad llevado a cabo a usuarios de Facebook se usó para inferir perfiles psicológicos de cada uno de ellos. Así, la empresa logró saber cómo debía ser el contenido y el tono de un mensaje publicitario para cambiar la forma de pensar de los votantes de forma casi individualizada. El resultado es conocido.

Algo parecido sucedió con ocasión del Brexit en el Reino Unido, en donde la decisión final de los ciudadanos también se vio influida por técnicas de IA que elaboraron perfiles ideológicos específicos a los que enviaron publicidad segmentada. Este asunto ha puesto de manifiesto el potencial de la IA –mal utilizada– para remover los cimientos de la democracia[4], o amenazar la integridad de sus procesos de elección.

En España, Ley Orgánica 3/2018, de protección de datos y garantía de los derechos digitales, introdujo un nuevo artículo 58.bis en la Ley Orgánica 5/1985 del régimen electoral general que, en su apartado 1 establecía que:

---

2. Fernández-Aller, C. et al *(*2021*)* «An Inclusive and Sustainable Artificial Intelligence Strategy for Europe Based on Human Rights», in *IEEE Technology and Society Magazine*, vol. 40, no. 1, pp. 46-54, March 2021, doi: 10.1109/MTS.2021.3056283; Australian Human Rights Commission (2019) *Human Rights and Technology • Discussion Paper***.** ISBN 978-1-925917-15-4
3. Schermer, B (2011), «The limits of privacy in automated profiling and data mining» 27 *Computer Law & Security Review* 45. https://doi.org/10.1016/j.clsr.2010.11.009 (20-09-2021)
4. Christodoulou, Ee. y Lordanou, K. (2019) «Democracy Under Attack: Challenges of Addressing Ethical Issues of AI and Big Data for More Democratic Digital Media and Societies» *Frontiers in Political Science; https://www.frontiersin.org/articles/10.3389/fpos.2021.682945/full* (20-09-2021).

> «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».

De esta forma, se autorizaba a los partidos políticos la recogida de datos personales de cualquier fuente (internet, redes sociales, páginas web, tratamientos no automatizados…) con el fin de perfilar a las personas en función de sus opiniones políticas, y todo ello sin el consentimiento de los sujetos afectados. Al poco tiempo se declaró inconstitucional[5].

Fundamentalmente, el Tribunal Constitucional criticó el precepto impugnado por tres razones: i) no determinó la finalidad del tratamiento de las opiniones políticas. Únicamente mencionó el interés público de manera genérica, pero sin especificar ese interés que fundamenta la restricción del derecho fundamental y sin establecer claramente sus límites y su regulación. Por tanto, no puede valorarse la legitimidad de la finalidad ni la proporcionalidad de la medida; ii) No determinó reglas precisas sobre los supuestos y condiciones en los que puede restringirse el derecho fundamental a la protección de datos. La mención del «marco de sus actividades electorales» no es una condición suficientemente precisa para conocer la finalidad o el bien constitucional que justifica la limitación del derecho fundamental; iii) No incorporó garantías adecuadas ni remitió expresamente a fuentes externas con rango normativo adecuado.

1.1. El derecho a la protección de datos

Los artículos 8 de la Carta de los Derechos Fundamentales de la Unión Europea (en adelante, CDFUE), 8 del Convenio Europeo de Derechos Humanos (en adelante, CEDH), 16 del Tratado de Funcionamiento de la UE (en adelante, TFUE) y 18.4 de la Constitución Española, (en adelante, CE) reconocen el derecho de la persona a decidir y controlar el uso de sus informaciones personales en manos de terceras personas, sean estas públicas o privadas, en la aceptación de que la intimidad, dignidad y personalidad se proyectan en forma de informaciones personales que revelan parte de nuestro modo de ser y de actuar. Dicho poder de decisión sobre los datos personales propios se concreta en un conjunto de facultades que permiten al sujeto ejercer dicho control para ser dueño de su propia libertad e identidad. El conjunto de facultades forma parte del contenido esencial del derecho fundamental a la protección de datos de carácter personal, de manera que la lesión en alguna de las facultades aludidas podría provocar una vulneración del derecho en cuestión. El contenido esencial es inalterable por el

5. Vid. Fernández-Aller, C. (2020),) «¿Vale todo para hacer propaganda electoral en internet? » *The Conversation* https://theconversation.com/profiles/celia-fernandez-aller-718032/articles; Vid. así mismo la Sentencia del Tribunal Constitucional 76/2019 del 22 de mayo de 2019 que declaró inconstitucional el apartado 1 del artículo 58 bis de la LOREG (Ley Orgánica del Régimen Electoral General).

legislador ordinario y permite identificar al derecho como tal, reconocerlo a través de la finalidad perseguida por dicha categoría jurídica: proteger al individuo a través de la protección de sus datos personales[6].

El derecho a la protección de datos constituye el principio de la relación entre el Derecho y la tecnología. De las primeras cuestiones que saltaron al nuevo escenario que ofrecía la tecnología fueron por un lado las ventajas que proporcionaba y por otro los peligros que suponía su uso para los derechos de la persona. La protección de datos surge así como el primer foco donde se advierte la necesaria intervención del Derecho, y no menor, pues pese a sus titubeantes inicios a la sombra de la intimidad se ha configurado como un auténtico derecho fundamental con sustantividad propia frente al derecho con el que conserva una relación sustancial, pero con el que sus diferencias justifican un contenido esencial propio y diferenciado de aquél[7]. La ciudadanía demanda espacios de protección individual propios en la sociedad digital que se han de añadir al contenido esencial de los derechos en su versión más clásica y que el Derecho no puede obviar. La modulación alcanza no solo a los derechos fundamentales, sino que la propia democracia y el mercado experimentan también la influencia de la tecnología, incluso la persona y sus características[8]. La vida individual pero también la vida social refleja y proyecta una profunda transformación de la mano de la tecnología.

La relación de la protección de datos con la IA es innegable. Por ello, la primera consecuencia que se deriva de esta conexión es que la IA que recurre a datos personales ha de someterse a las normas sobre protección de datos vigentes. La IA puede emplear la tecnología para extraer conclusiones a partir de datos por medio de algoritmos de una manera independiente de la voluntad humana y puede provocar por ello riesgos en los derechos de los ciudadanos que deben ser conjurados. Con las tecnologías que incorporan la IA, las máquinas pueden actuar y aprender por medio de una combinación de algoritmos cuya finalidad es asemejarse en todo lo posible a las capacidades del ser humano. Los algoritmos no sólo emplean datos personales de los sujetos, sino que a partir de estos datos se pueden elaborar más volúmenes de datos que aportan una información sobre el individuo totalmente desconocida para este último. Respecto de

6. Rebollo Delgado, L. y Serrano Pérez, M.ª M., *Manual de Protección de Datos*, Dykinson, S. L., Madrid 2019, pág. 137-139.
7. Así lo afirmó el TC en un análisis acertado del derecho, procurando una interpretación en consonancia con el reconocimiento del mismo en el contexto europeo e internacional y difícilmente reconocible en el entorno del art. 18.4 CE. El ámbito de la protección de datos es más amplio que el de la intimidad (limitado a las esferas más esenciales y cercanas a la personalidad del individuo).
8. De la Quadra-Salcedo Fernández del Castillo, T. (2019), «Derechos fundamentales, democracia y mercado en la edad digital», *Derecho Digital e Innovación*, núm. 1, enero-marzo. Este autor afirma que se protege el dato en su objetividad, aunque por su conexión con la persona humana y el libre desarrollo de la personalidad. Por ello la cláusula del pleno ejercicio de sus derechos cobra especial relevancia. *Ya no estamos solo en la protección del dato personal, sino que nos deslizamos hacia otros horizontes, donde la protección del menor, de la igualdad, de la libertad de expresión e información, del mercado y de la democracia misma comienzan a emerger.*

este tipo de datos habrá que determinar cómo se aplican el consentimiento, el derecho a la información, el principio de finalidad o el ejercicio de los derechos de la ciudadanía.

Habrá que preguntarse si la regulación originaria pensada para proteger al individuo y sus datos personales es suficiente para dispensar dicha protección en los sistemas de IA que manejen datos personales. La IA desde el enfoque de los derechos debe ser la premisa a tener en cuenta siempre. Esa orientación es la que sostiene la distinción contenida en la propuesta de Reglamento sobre inteligencia artificial que contempla los sistemas de IA que entrañan un riesgo inaceptable en función de la amenaza que suponen para los derechos de las personas y para los valores de la Unión Europea, así como un alto riesgo en el caso de que entrañen un peligro para los derechos fundamentales o para la seguridad.

El empeño de la Unión Europea a través de las iniciativas legislativas y de las propuestas sobre IA aspira a liderar este campo a nivel mundial bajo estándares comunes. No solo está en juego el valor económico del dato, y por tanto la economía europea, sino también un modo de vida que defiende el Estado social y democrático de Derecho que puede verse atacado por la IA y la generalización de la tecnología sin reglas y directrices concretas, en especial el disfrute de los derechos fundamentales por los ciudadanos. En este sentido, la propuesta de Reglamento por el que se establecen normas armonizadas sobre inteligencia artificial[9] prohíbe aquellos usos de la IA que directamente sean contarios a los valores de la Unión o a los derechos fundamentales; en la misma línea señala requisitos específicos para los sistemas de IA de alto riesgo y obligaciones para los operadores de dichos sistemas.

Para conjurar los riesgos derivados del uso de la IA es preciso acompañar las reglas jurídicas de pautas éticas centradas en convencer sobre la necesidad de respetar los derechos de las personas y defender un uso respetuoso de la tecnología. Proclamar y extender directrices éticas y crear y activar los instrumentos que efectivamente protejan la dignidad de la persona[10] son dos pilares esenciales en la consolidación adecuada de la IA. Junto a ello la elaboración de códigos de conducta (art. 40 RGPD) puede ser un mecanismo útil para intentar una adecuación de la normativa de protección de datos a los sistemas de IA.

El sometimiento de los sistemas de IA a la normativa de protección de datos[11] obliga al menos a preservar el contenido esencial que el derecho posee y que atribuye

9. Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, COM(2021) 206 final, disponible en https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence

10. De la Quadra-Salcedo Fernández del Castillo, T., «Las transformaciones del Derecho…», ob. cit., pág. 7.

11. Por ejemplo, un tratamiento puede incorporar procesos de IA, por ejemplo, que tomen decisiones sobre sujetos en base a la información aportada a través de los datos personales. Es el caso de una empresa que aplica la IA a datos de los interesados como clientes, empleados, etc., con el fin de diseñar políticas de marketing teniendo en cuenta las preferencias de sus clientes. Los datos personales de dichos

y garantiza al sujeto al control de sus datos personales y el uso que de los mismos se está haciendo. Es importante reconocer la necesidad de preservar un mínimo contenido esencial para proteger las facultades de control que incorpora el derecho. Así se desprende de la sentencia del Tribunal de Distrito de La Haya, de 5 de febrero de 2020, donde el tribunal reconoce que la elaboración de un informe de riesgo en las condiciones que recoge el proyecto SyRI (Sistema de Indicación de Riesgos) y la inclusión del informe en el registro es un factor significativo para el tribunal a la hora de evaluar que la normativa señalada se ajusta o no al art. 8, apartado 2, del CEDH y por tanto que realiza una interferencia en el respeto a la vida privada del sujeto. Para llevar a cabo dicho análisis, el tribunal sopesa que:

> «parte del derecho a la protección de datos personales es el derecho de toda persona a poder realizar un seguimiento razonable de sus datos personales y a ser informado sobre el procesamiento de sus datos. Aunque el inicio de un proyecto de Syri se publica en la Gaceta del gobierno, posteriormente se puede incluir un informe de riego en el registro por un período de dos años sin que el interesado lo sepa»[12].

Aceptando que se produce una injerencia en la vida privada de los ciudadanos y que dicha intromisión es necesaria en una sociedad democrática, es preciso que la legislación sea accesible y previsible, que exista un equilibrio entre el objetivo legítimo perseguido y la interferencia en la vida privada de los ciudadanos. La accesibilidad y previsibilidad de la ley que contemple un sistema de IA que emplea datos personales para cumplir sus objetivos debería precisar de forma exhaustiva las categorías de datos que pueden recogerse, los responsables, el periodo de mantenimiento de los datos, el principio de confidencialidad. Especialmente importantes son el principio de transparencia, de limitación de la finalidad, y el principio de minimización de datos.

1.2. La ética de la inteligencia artificial en la protección de datos

Que la IA pueda tomar decisiones de modo semejante a como lo podría hacer un ser humano, dependiendo de la combinación de algoritmos, no deja lugar a dudas para

clientes o empleados deberán gozar de protección. En estos casos el riesgo real de lesionar el derecho a la intimidad, a la protección de datos, a la igualdad… es permanente.

12. Sentencia de la Corte de Distrito de La Haya de 5 de febrero de 2020, (párrafo 6.60). La sentencia recupera la doctrina del TEDH sobre la calidad de la ley que aborde el tratamiento de datos a través de los sistemas de inteligencia artificial, y que debe proporcionar protección contra la arbitrariedad y contra la discrecionalidad de las autoridades públicas. Así resulta esencial «La existencia de reglas claras que especifiquen las medidas a adoptar, así como las garantías mínimas relativas al almacenamiento, duración, uso, acceso de terceros, procedimientos para preservar la integridad y confidencialidad de los datos y los procedimientos para su destrucción, aportando las garantías suficientes contra la amenaza de abuso y arbitrariedad», según se desprende del caso S. y Marper contra el Reino Unido. Un comentario a la sentencia en Lazcoz Moratinos, G., y Castillo Parrilla, J. A., «Valoración algorítmica ante los derechos humanos y el Reglamento General de Protección de Datos: el caso SyRI», *Revista Chilena de Derecho y Tecnología*, vol. 9, núm. 1 (2020), págs. 207-225. https://doi.org/10.5354/0719-2584.2020.56843

incorporar la ética a la misma. La ética como sostén de la forma de empleo de los sistemas de IA pretende reforzar desde otra perspectiva los principios jurídicos que presiden todo tratamiento de datos personales. La ética no contempla principios o comportamientos alejados del mundo jurídico. Al contrario, los postulados éticos aspiran a proteger la dignidad, la libertad, la igualdad. Para ello los sistemas de IA deben ser controlados desde el diseño y de forma continua. Además, es necesario ser «vigilantes tanto de la legitimidad ética de los tratamientos como de los efectos inesperados de estos»[13], así como los efectos colaterales de estos dependiendo del entorno social al que afectan[14].

Sin embargo, no pueden soslayarse las dificultades para reconciliar perspectivas muy heterogéneas a la hora de abordar los dilemas éticos de la IA. Asistimos a una proliferación de códigos éticos, elaborados por organizaciones dispares (empresas, asociaciones profesionales, instituciones públicas, etc), que poco están contribuyendo a incluir las inquietudes éticas desde el diseño de la tecnología[15].

La característica esencial de la IA ha de ser la fiabilidad. La IA fiable[16] es el requisito previo de su extensión y adopción por parte de las sociedades y de las personas. Solo con la generalización de la IA (fiable) se podrán conseguir todos los beneficios posibles. Pero la fiabilidad no se consigue bajo cualquier condición sino con la base sólida de la ética. La ética es uno de los componentes de una IA fiable, junto con la licitud y la robustez tanto técnica como social, según señala el documento Directrices éticas para una Inteligencia Artificial fiable[17], elementos que han de estar presentes en el diseño, desarrollo y utilización de los sistemas de IA. La finalidad del citado documento es convertir la ética en un pilar esencial que permita desarrollar la IA desde un enfoque único. En lo que respecta a la IA fiable desde la perspectiva ética hay que tener en cuenta cuatro principios básicos: el respeto a la autonomía humana, prevención del daño, la equidad y la explicabilidad[18].

La inteligencia artificial explicable es un conjunto de procesos y métodos que permite a los usuarios humanos comprender y confiar en los resultados generados por

13. AGDP (2020*) Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*, pág. 7.
14. *Adecuación al RGPD de…*, ob. cit., La AEPD señala las relaciones del tratamiento con el entorno desde la perspectiva cultural, teniendo en cuenta el contexto en el que se despliega dicho servicio y en lo que tiene que ver con la interconexión masiva de componentes en la sociedad de la información, pág. 7.
15. El proyecto europeo SIENNA está promoviendo muy activamente este concepto del *Ethics by design,* https://sienna-project.eu/public-consultation/ai/ethics-by-design/
16. Los requisitos de la IA fiable son: acción y supervisión humanas, solidez técnica y seguridad, gestión de la privacidad y de los datos, transparencia, diversidad, no discriminación y equidad, bienestar social y ambiental, rendición de cuentas., Directrices éticas para una Inteligencia Artificial fiable, Grupo independientes de expertos de alto nivel sobre inteligencia artificial, creado por la Comisión Europea en junio de 2018, publicado en abril de 2019, disponible en https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1, pág. 18.
17. Directrices éticas para…, ob. cit., pág. 6.
18. Directrices éticas para…, ob. cit., pág. 10.

algoritmos de machine learning. La IA explicable ayuda a caracterizar la precisión, la imparcialidad, la transparencia y los resultados de los modelos en la toma de decisiones basada en IA. La explicabilidad de la IA[19] apoya la adopción de enfoques responsables de esta tecnología.

El campo de la protección de datos está fuertemente regulado. Podríamos decir que la normativa de la protección de datos está consolidada. Sin embargo, las novedades de la IA han provocado dificultades en la aplicación de determinados principios propios de la protección de datos, que han de estar presentes en los sistemas de IA que emplean datos personales, como por ejemplo la transparencia, el derecho de acceso, el derecho de información. La ética, bajo la fórmula de la prevención del daño y del respeto a los derechos fundamentales debe presidir estas dificultades de aplicación. La falta de la misma erosionaría la fiabilidad de la IA y por tanto provocaría una regresión en su implantación y aplicación, bajo el objetivo de mejorar la vida colectiva y social[20].

El empleo de cualquier técnica que provoque falta de equidad o no evite un posible daño debe ser rechazado por falta de ética, además de poder ser contrario a las leyes.

Dicho esto, no hay que desatender los peligros que suponen las formulaciones éticas de las empresas, que suelen ser muy amplias. Existe una sobreabundancia de códigos éticos, que hacen difícil el consenso en los valores fundamentales. En otras palabras, habrían de evitarse los peligros de los que alerta Floridi[21]: a) *Ethics shopping*, que supone que una organización elija, entre las muchas iniciativas que hay de códigos éticos muy dispersos, el que mejor se adapte a su forma de hacer, justificando así sus intenciones, poco coherentes con la ética en ocasiones. b) *Ethics dumping*, que consiste en la conducta de exportar prácticas no éticas a países donde hay más laxitud o diferencia de criterios. c) *Ethics lobbying*, o la práctica de algunos actores privados de usar autorregulación en temas como la ética de la Inteligencia Artificial para hacer lobby en contra de la introducción de normas con fuerza jurídica, sometidas estas últimas a mecanismos más exigentes en caso de incumplimiento. d) *Bluewashing*[22], que es la mala práctica de una organización pública o privada que busca aparecer socialmente

19. Arya, V., Bellamy, R. Chen, P., Dhurandh,A., et. Al (2019) «One explanation does not fit all: A toolkit and taxonomy of ai explainability techniques» arXiv preprint. En este texto se explica el kit de explicabilidad de IBM, *AI Explainability 360.*
20. Las Directrices éticas para…, ob. cit., pág. 13, señalan que *la reflexión ética sobre la tecnología de la IA puede servir para proteger a las personas y a los grupos en el nivel más básico…estimular nuevos tipos de innovaciones que busquen fomentar valores éticos, como las que contribuyen a lograr los Objetivos de Desarrollo Sostenible de las Naciones Unidas… para mejorar el bienestar individual y el bienestar colectivo mediante la generación de prosperidad, la creación de valor y la maximización de la riqueza…contribuir a construir una sociedad justa ayudando a mejorar la salud y el bienestar de los ciudadanos de manera que promuevan la igualdad en la distribución de las oportunidades económicas, sociales y políticas.*
21. FLORIDI, L. (2019), «Translating principles into Practices of Digital Ethics: Five Risks of Being Unethical». *Philosophy and Technology,* 32, 185-193. https://doi.org/10.1007/s13347-019-00354-x
22. Vid. Yeung,K., Howes, Pogrebna (2020), «AI Governance by Human Rights-Centered Design, Deliberation and Oversight: and End to Ethics Washing», in *The Oxford Handbook of Ethics of AI*, Ed. Oxford University Press.

como más verde, sostenible y comprometida de lo que en realidad es. Pero en realidad, el compromiso social es mucho menor de lo que aparenta ser.

Otros temas que suscitan preocupación desde el punto de vista ético son los siguientes, según un análisis *focus group*[23] en el que participaron 63 personas:

- Reconciliar las diferentes y heterogéneas perspectivas éticas.
- El auge de la polarización y el populismo, muy agudizadas por el uso de IA, que erosionan la democracia.
- La falta de sensibilidad ética de los ingenieros que desarrollan la IA.
- Falta de presión sobre los productores de IA para incorporar las cuestiones éticas (escasas consecuencias en caso de incumplimientos éticos)
- Los usuarios eligen a veces su conveniencia por encima de la privacidad u otros valores éticos.
- Retos en relación a las legislaciones insuficientes

1.3. Retos que presenta el uso generalizado de algoritmos: el caso de la publicidad política personalizada

Existe un debate prolongado en relación con el tratamiento automatizado de datos y los algoritmos y su impacto en el derecho a la intimidad y la protección de datos. Los algoritmos facilitan la recopilación de grandes cantidades de datos e imágenes, lo que puede tener graves consecuencias para el disfrute del derecho a la vida privada y familiar, incluido el derecho a la protección de datos.

Los algoritmos se utilizan en el seguimiento y la elaboración de perfiles en línea de las personas[24] cuyos patrones de navegación se registran mediante *cookies* y tecnologías similares, como la huella digital, agregada a las consultas de búsqueda (motores de búsqueda/asistentes virtuales). Por otro lado, los datos de comportamiento se procesan a partir de dispositivos inteligentes, como los datos de localización y otros datos de los sensores a través de aplicaciones en los dispositivos móviles, lo que trae consigo desafíos para la privacidad.

Las aplicaciones de seguimiento y elaboración de perfiles en línea se han utilizado en la publicidad dirigida, por parte de los partidos políticos, basada en el perfil de los presuntos intereses de una persona. En este caso, el consentimiento del usuario es un requisito imprescindible[25]. En los casos comentados anteriormente, el *micro targeting* se

23. C, E. y Lordanou, K., (2019), «Democracy Under Attack…», ob. cit., https://www.frontiersin.org/articles/10.3389/fpos.2021.682945/full (20-09-2021).
24. Vid. Ebert, N. (2020) *Algorithms and Law*. Ed. Cambridge University Press, pag. 70; Woodrow, B. (2020), *The Cambridge Handbook of the Law of Algorithms.* University of Washington, Ed. Cambridge University Press, p. 632.
25. No lo sería en los casos en que el ciudadano haya hecho públicos esos datos con carácter previo. Sin embargo, esta excepción ha de interpretarse de forma muy restrictiva, tal y como ha puesto de

ha llevado a cabo sin la preceptiva solicitud de consentimiento. Ninguna de las causas que eximen de este requisito justificaban la utilización de inteligencia artificial para perfilar a los ciudadanos y dirigirles campañas de publicidad política personalizadas: ni interés legítimo[26], ni interés vital del interesado, ni contrato existente justificaron dichas actuaciones.

Se está intentando actualizar el Convenio del Consejo de Europa para la Protección de las Personas con respecto al Tratamiento Automatizado de Datos de Carácter Personal de 1981 (Convenio 108) en consonancia con la evolución tecnológica, y para definir con mayor precisión los derechos del interesado con respecto a las implicaciones para la privacidad de las herramientas contemporáneas de recogida, tratamiento, reutilización y elaboración de perfiles de datos. El artículo 8 del proyecto de Convenio modernizado establece el derecho explícito de toda persona a no ser sometida a una decisión que le afecte significativamente, basada únicamente en un tratamiento automatizado de datos, sin que se tome en consideración su opinión; el derecho a conocer los motivos del tratamiento de datos cuando se le apliquen los resultados de dicho tratamiento; y a oponerse en cualquier momento, por motivos relacionados con su situación, al tratamiento de los datos personales que le conciernen, a menos que el responsable del tratamiento demuestre que existen motivos legítimos para el tratamiento que prevalecen sobre sus intereses o derechos y libertades fundamentales. Las propuestas de modernización pretenden, además, ofrecer garantías complementarias en materia de transparencia (artículo 7bis) y la necesidad de examinar el impacto probable del tratamiento de datos sobre los derechos y libertades fundamentales de la persona antes de iniciar dicho tratamiento (artículo 8bis). Es decir, el derecho a la explicabilidad y estudios de impacto en la privacidad.

Las *Directrices sobre la protección de las personas físicas en lo que respecta al tratamiento de datos personales en un mundo de Big data* adoptadas recientemente por el Comité del Convenio para la Protección de las Personas con respecto al Tratamiento Automatizado de Datos Personales proporcionan un marco general para aplicar políticas y medidas adecuadas para seguir haciendo efectivos los principios de protección de datos en el contexto de Big Data.

Los marcos normativos de protección de datos a nivel de la UE, como el RGPD ya mencionado, también establecen normas para el uso de algoritmos en la recopilación de datos, incluyendo posiblemente un derecho limitado a la información.

---

manifiesto el Comité Europeo de Protección de Datos en su Resolución 2/2019 de 13 de marzo. Esta excepción, defiende, no puede legitimar el procesamiento de este tipo de datos derivados de los primeros.

26. Algunos autores defienden que para legitimar los tratamientos de big data hay que acudir al interés legítimo del responsable. Moerel, E.M.L. «Big Data Protection. How to Make the Draft EU Regulation on Data Protection Future Proof» (February 14, 14). Available at SSRN: https://ssrn.com/abstract=3126164 or http://dx.doi.org/10.2139/ssrn.3126164

La elaboración de perfiles, en sí misma, significa la extrapolación de los datos disponibles en Internet mediante procesos de recopilación automatizada de información y la posterior construcción y aplicación de perfiles. Las técnicas de elaboración de perfiles pueden beneficiar a los individuos y a la sociedad, por ejemplo, al permitir una mejor segmentación del mercado o un análisis de los riesgos y el fraude. Sin embargo, el uso de esta técnica en el ámbito de la propaganda electoral también suscita importantes preocupaciones. La Recomendación del Consejo de Europa sobre la elaboración de perfiles[27] aborda el riesgo de que los perfiles basados en datos atribuidos a un sujeto permitan generar nuevas informaciones, incluso mediante la agregación de datos. Esta información puede entonces ser explotada mediante el uso de algoritmos, lo que crea un riesgo de vigilancia a gran escala por parte de entidades privadas y gobiernos por igual. El Consejo de Derechos Humanos de las Naciones Unidas se hace eco de esta opinión y, el 22 de marzo de 2017, señaló con preocupación:

> «que el tratamiento automático de datos personales para la elaboración de perfiles individuales puede dar lugar a discriminaciones o a decisiones que podrían afectar al disfrute de los derechos humanos, incluidos los derechos económicos, sociales y culturales».

La principal preocupación de utilizar los datos de los perfiles para diferentes fines a través de algoritmos es que los datos pierden su contexto original. La reutilización de los datos puede afectar a la autodeterminación informativa de una persona. Generalmente, los/las interesados/as no serán conscientes de la posterior reutilización de los datos más allá de su contexto original. Además, los resultados obtenidos a través de los algoritmos de búsqueda pueden ser incompletos, inexactos o desfasados, extrayendo conclusiones que pueden ser perjudiciales.

En este sentido, resulta clave el cumplimiento del deber de información al interesado: si los ciudadanos conociesen que sus datos iban a tratarse con IA para elaborar campañas de publicidad política personalizadas, ¿darían su consentimiento? Parece claro que no, al menos una gran mayoría.

A pesar de la importancia del consentimiento y la información en el caso de la utilización de IA para la elaboración de perfiles que permitan publicidad electoral personalizada, la realidad impide que estos requisitos puedan llevarse a cabo. La naturaleza de las tecnologías trae consigo que los resultados no puedan predecirse, por lo que difícilmente puede informarse con carácter previo.

27. Consejo de Europa. *Protección de las personas en lo que respecta al tratamiento automático de datos personales en el contexto de la elaboración de perfiles*. Recomendación CM/Rec (2010)13 y exposición de motivos. Consejo de Europa, 23 de noviembre de 2010. https://www.coe.int/t/dghl/standardsetting/cdcj/CDCJ%20Recommendations/CMRec(2010)13E_Profiling.pdf

Con respecto al consentimiento, ¿cómo podría garantizarse la revocabilidad del mismo? Esta solicitud limitaría el número de los datos utilizados, y pondría en serias dificultades a las empresas de IA de tamaño reducido[28].

2. EL PRINCIPIO DE IGUALDAD, LA PROTECCIÓN DE DATOS Y LA INTELIGENCIA ARTIFICIAL

Tal y como se ha explicado, el tratamiento de datos personales a través de la IA puede originar la elaboración de un perfil[29] de los individuos que puede convertirse en una amenaza contra el derecho fundamental a la igualdad[30]. En la medida en que diseñemos un proceso de selección que valore ciertos datos y ciertas informaciones por encima del resto podemos estar realizando una discriminación mecánica que puede tener su influencia en los ciudadanos/as, independientemente de los errores de sesgo que acompañan a la IA. Incluso sin información sesgada, el efecto de la aplicación de la IA puede provocar una discriminación real que sufrirían los sujetos sobre los que se ha elaborado un determinado perfil en base a la IA; es obvio, sin embargo, que también se pueden obtener ventajas, objetivo final de la tecnología.

La discriminación puede venir de cualquiera de las circunstancias específicas contempladas en el art. 14 CE o de la genérica. Quizá a estas alturas del siglo XXI merecería introducir esta causa de discriminación de forma particular en el art. 14, tal y como señala Quadra-Salcedo[31], en una futurible reforma constitucional que adaptara el precepto a las nuevas circunstancias de discriminación. La elaboración de perfiles puede por tanto agravar los estereotipos fomentando las diferencias sociales, provocar la denegación de determinados productos o incluso bienes o servicios. Es más, todo puede estar basado en una predicción inexacta o errónea, en la que si no está contemplada su corrección a través de la intervención humana puede perpetuarse indefinidamente.

28. Mitrou, L., «Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof?» (December 31, 2018). Available at SSRN: https://ssrn.com/abstract=3386914 or http://dx.doi.org/10.2139/ssrn.3386914, pág. 40.
29. El art. 4 apartado 4 RGPD define la elaboración de perfiles 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».
30. De La Quadra Salcedo Fernández del Castillo, F., «Retos, riesgo y oportunidades de la sociedad digital», *Sociedad digital y Derecho,* De la Quadra-Salcedo, T. y Piñar Mañas, J. L., (directores), Ministerio de Industria, Comercio y Turismo, Madrid 2018, pág. 57 y ss.
31. De La Quadra Salcedo Fernández del Castillo, F., «Las transformaciones del Derecho…», ob. cit., afirma que, aunque se considerarse por algunos que estos <<prejuicios>> con soporte estadístico no encajan tampoco en la razón genérica, residual, innominada y final del artículo 14 CE, la cuestión es si merecería a la altura del tiempo en que estamos introducirla expresamente como una causa de discriminación que no pudo ser considerada en 1978, pues no existía la tecnología ni la capacidad para suponer la real existencia de los perfiles.

La aplicación del RGPD a estas situaciones se regula en el art. 22, donde se recoge el derecho del individuo a no ser objeto de una decisión basada en un tratamiento automatizado; además, el RGPD se aplica a la recogida de los datos para la creación de los perfiles y a la aplicación de dichos perfiles a los individuos[32]. Este tipo de casos pueden darse si se trata de un algoritmo que toma una decisión automatizada, en base a un perfil o no, pero sin intervención previa y determinante de un ser humano. Es verdad que este apartado se excepciona en las circunstancias del apartado segundo, donde destaca el consentimiento del sujeto, que la decisión sea necesaria en el contexto de la celebración o ejecución de un contrato entre el sujeto y el responsable del tratamiento o bien si dicha decisión está contemplada en el Derecho de la Unión o de los Estados miembros al que se somete el responsable del tratamiento, y siempre que se adopten las medidas adecuadas para la protección de los derechos y libertades del interesado. Esto en todos los casos admitidos por la norma, señalando expresamente el derecho a obtener la intervención humana en el proceso de decisión a través de la IA, a poder expresar su punto de vista y a impugnar la decisión. De la elaboración de perfiles a través del manejo de datos se excluyen las categorías especiales de datos del art. 9.2 RGPD (entre los que figuran los datos relativos a opinión política, salvo si se ha dado consentimiento o por motivos de interés público). Además de las garantías específicas que recoge el art. 22 RGPD, los arts. 13 y 14 relativos al derecho a la información recogen la obligación de comunicar al interesado/a, para asegurar un tratamiento de datos transparente y leal, «la existencia de decisiones automatizadas incluida la elaboración de perfiles a que se refiere el artículo 22, apartado 1 y 4 y, al menos en tales casos, información significativa sobre la lógica aplicada, así como las consecuencias previstas de dicho tratamiento para el sujeto». Esta información ha de proporcionarse tanto si los datos se recaban del interesado como si no se recaban directamente de él. Por otro lado, la comunicación que se suministre al sujeto en relación con el art. 22 RGPD, ha de ser, según el art. 12.1 RGPD, concisa, transparente, inteligible y de fácil acceso, con un lenguaje claro y sencillo, aspectos que ha destacado la sentencia de la Corte de Distrito de La Haya en relación a la norma. Por ello es necesario respetar los principios del art. 5 RGPD, la licitud, la lealtad y la transparencia (art. 5.1 a) RGPD). Igualmente, el principio de finalidad para evitar que los datos recogidos para un fin sean empleados para obtener un perfil del individuo que nos aporte información sobre, por ejemplo, su estilo de vida, finalidad no prevista en la recogida inicial (para lo que será necesario recabar un nuevo consentimiento o bien informar en la recogida primera de los datos de esta posibilidad).

32. El RGPD distingue entre perfiles (tratamiento automatizado de datos personales que arrojan una evaluación de la persona, nos ofrecen una determinada conclusión sobre ella) y decisiones automatizadas son decisiones que se toman, puede ser en base a un perfil ya existente o en base a datos de las personas, fundamentadas solamente en un tratamiento automatizado sin la intervención del ser humano. Las decisiones automatizadas pueden efectuarse sin la elaboración de perfiles, pero puede ocurrir que la toma de decisiones automatizadas se transforme en un proceso de elaboración de perfiles, dependiendo del uso de los datos.

Junto a ello la minimización de los datos y la limitación de su conservación deben estar garantizados por el responsable del tratamiento que originará la elaboración de perfiles. Muy importante es la exactitud del dato pues, si los datos empleados en un proceso de decisión automatizada o de elaboración de perfiles son inexactos en origen, la decisión final o el perfil serán erróneos.

Además, el responsable del tratamiento ha de facilitar el ejercicio del derecho contemplado en el art. 22 RGPD, esto es, impugnar la decisión. Para la impugnación de la decisión es fundamental haber cumplimentado el derecho de información. La falta de observancia del derecho a la información podría provocar el desconocimiento de la existencia de decisiones y de perfiles sobre la base de datos personales, lo que impediría llevar a cabo al menos un seguimiento razonable de los datos, tal y como ya hemos señalado, y podría provocar una lesión en el derecho a la protección de datos personales. El responsable del tratamiento ante el que se solicite la impugnación de la decisión adoptada o el perfil, en virtud del art. 22, y según el art. 12.3 RGPD, deberá atender dicha solicitud en el plazo de un mes desde la recepción, aportando la información necesaria. El plazo podrá prorrogarse hasta dos meses en atención a la complejidad del asunto o del número de solicitudes.

Puede concluirse que, ninguna de estas previsiones del legislador, están siendo aplicables en el caso de los perfiles y el marketing político. Ningún ciudadano conoce que está siendo perfilado, ni ejercita los derechos previamente mencionados.

La situación es grave, puesto que se puede lesionar el derecho a la igualdad en función del perfil que tengamos del sujeto tras el tratamiento de los datos recogidos. Los datos, que por otra parte será un volumen elevado, aportan la información, pero el diseño del perfil puede excluir a sujetos con base en una solución estadística. El riesgo está en la elaboración del perfil y en el uso que del mismo se haga.

Algunos retos que están pendientes en torno a la discriminación algorítmica son los siguientes:

1) el factor humano y el desafío de los estereotipos y los sesgos cognitivos explican cómo los sesgos implícitos, los estereotipos nocivos y los prejuicios discriminatorios que tienen los humanos corren el riesgo de infectar los algoritmos que crean los humanos y cómo los sesgos de automatización y anclaje refuerzan estos riesgos;
2) el reto de los datos, puesto que los datos encarnan las pautas de discriminación históricamente consolidadas que estructuran la sociedad y el entrenamiento de los algoritmos con esos datos puede ser un factor de riesgo.
3) el reto de la correlación: el reto de la correlación explica cómo los algoritmos pueden dar lugar a correlaciones discriminatorias (por ejemplo, el género puede correlacionarse negativamente con el rendimiento laboral, no por una relación causal, sino porque las mujeres han sido históricamente evaluadas más negativamente que los hombres por el mismo rendimiento laboral) al tratarlas como

causalidades y utilizándolas como base para posteriores decisiones, recomendaciones o predicciones;

4) el reto de la transparencia y la explicabilidad: se refiere a las dificultades para supervisar y probar ciertos tipos de algoritmos (incluso para los informáticos) y a la falta de información sobre las características protegidas;
5) el reto de la escala y la velocidad describe cómo la discriminación algorítmica puede extenderse a mayor escala y a un ritmo mucho más rápido que la discriminación humana, ya que los algoritmos aceleran y amplían la toma de decisiones; y
6) el reto de la responsabilidad y la rendición de cuentas: se refiere a la dificultad de identificar a quién hay que considerar responsable de un resultado discriminatorio en el contexto de las complejas relaciones entre el hombre y la máquina, dado que, en el diseño, la comercialización y el uso de los algoritmos intervienen muchas partes diferentes.

Se necesita mayor profundización en las relaciones entre privacidad y no discriminación algorítmica[33]. Por otro lado, es llamativa la ausencia del género entre los datos sensibles que prevé el RGPD en su artículo 10.

Como conclusión, debemos decir que existe una necesidad de lo que se denomina la transparencia algorítmica[34], que según la definición de la Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL)) consiste en que «siempre ha de ser posible justificar cualquier decisión que se haya adoptado con ayuda de la inteligencia artificial y que pueda tener un impacto significativo sobre la vida de una o varias personas; siempre debe ser posible traducir los cálculos del sistema de inteligencia artificial a una forma comprensible para los humanos».

La Recomendación del Comité de Ministros del Consejo de Europa sobre *Los impactos en los derechos humanos de los sistemas algorítmicos*[35] establece algunas recomendaciones aplicables al caso que comentamos sobre publicidad política personalizada. Así, por ejemplo, hace énfasis en el conjunto de datos que se utilizan en el diseño, el desarrollo, el despliegue continuo y la adquisición de sistemas algorítmicos. Los Estados deben evaluar cuidadosamente qué normas de derechos humanos y de no discriminación pueden verse afectadas como resultado de la calidad de los datos que se introducen y extraen de un sistema algorítmico, ya que estos a menudo contienen

33. Le Clainche, J. y Le Métayer, D. (2012), «Données personnelles, vie privée et non-discrimination: des protections complémentaires, une convergence nécessaire» (Personal data, privacy and non-discrimination: complementary protections, a necessary convergence), *Revue Lamy Droit immatériel,* 90
34. Vid., además The Alan Turing Institute /ICO (2020) *Explaining decisions made with AI.* https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/explaining-decisions-made-with-ai/
35. Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the *human rights impacts of algorithmic systems (Adopted by the Committee of Ministers on 8 April 2020 at the 1373rd meeting of the Ministers' Deputies).*

sesgos[36] y pueden sustituir a clasificadores como el género, la raza, la religión, la opinión política o el origen social.

3. DERECHOS DE LOS INTERESADOS Y DEBERES DEL RESPONSABLE

3.1. Los derechos de los ciudadanos que reciben publicidad personalizada posterior a un tratamiento de IA

Los/las responsables de los sistemas de IA deben facilitar el ejercicio de los derechos de los sujetos en cada momento.

Resulta esencial, puesto que el sistema de IA puede incorporar a varios responsables durante toda su vida, que deben quedar perfectamente identificados para que el sujeto pueda dirigirse a ellos en caso de ejercitar los derechos correspondientes. El derecho de información al interesado y la transparencia del tratamiento vuelven a ser fundamentales para la protección de los derechos de los ciudadanos.

El RGPD recoge determinados aspectos referidos a la información sobre el tratamiento de datos personales. Entre ellos se encuentran:

a) el derecho a ser informado sobre determinadas cuestiones cuando se recojan datos personales, como el periodo en que se almacenarán los datos personales
b) el derecho a ser informado cuando se traten datos personales que no hayan sido obtenidos directamente del interesado
c) el derecho de acceso, que incluye el derecho a preguntar al responsable del tratamiento si sus datos personales están siendo tratados o no, y sobre determinados aspectos de ese tratamiento, como su finalidad.

Sin embargo, tal y como ya se comentó más arriba, ni el principio de consentimiento ni el de información son de fácil cumplimiento en el caso de los perfiles de ciudadanos a los que se envía publicidad política.

Como hemos mencionado, el artículo 22 del RGPD establece derechos específicos para una persona sometida a una toma de decisión individual automatizada, incluida la elaboración de perfiles. Estos derechos incluyen: el derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos o que le afecten significativamente de forma similar.

36. Los sesgos pueden estar relacionados con los métodos (por ejemplo, el sesgo de medición, el sesgo que afecta a las metodologías de las encuestas); con el objeto de la investigación (por ejemplo, el sesgo social debido al sesgo histórico o la infrarrepresentación de algunas categorías); con sus fuentes de datos (por ejemplo, el sesgo de selección) o con la persona responsable del análisis (por ejemplo, el sesgo de confirmación).

Cuando una persona haya dado su consentimiento explícito para el uso de sus datos personales en un tratamiento automatizado, la persona conserva al menos el derecho a obtener una intervención humana por parte del responsable del tratamiento, a expresar su punto de vista y a impugnar la decisión. Este derecho ha sido ya implementado en algunas leyes nacionales de la UE.

Algunos expertos sostienen que el RGPD no recoge un derecho explícito a la explicabilidad. Más bien el derecho de acceso proporciona un derecho a la explicación de la funcionalidad del sistema o un derecho a ser informado, que está restringido por los intereses del responsable del tratamiento. Otros sostienen que el RGPD apoya un derecho de acceso a la explicabilidad de una decisión que afecta a los datos personales de un individuo. Andrew Selbst y Julia Powles[37], por ejemplo, sostienen que leídos conjuntamente, los artículos 13-15 del RGPD establecen el derecho a una información significativa sobre la lógica implicada en una decisión automatizada. Aunque los autores no concluyen con precisión cómo debe ser esa explicación, afirman: *Creemos que el derecho a la explicabilidad debe interpretarse de forma funcional y flexible, y debe como mínimo, permitir que el interesado ejerza sus derechos en virtud del RGPD y de la legislación sobre derechos humanos.*

Otros expertos se han centrado en un concepto más amplio de transparencia, más que de explicabilidad. Margot Kaminski, por ejemplo, concluye que el RGPD regula el derecho a la transparencia al dar una importante capacidad de acceder a la información sobre el algoritmo; exigiendo a las empresas el establecimiento de regímenes internos de responsabilidad interna y de divulgación, incluida la realización de evaluaciones de impacto de la protección de datos; y recomendando el uso de auditores externos, con acceso a toda la información necesaria sobre el funcionamiento interno del sistema de aprendizaje automático sistema o algoritmo de aprendizaje automático.

3.2. Deberes del responsable que lleva a cabo perfilado de ciudadanos con IA para realizar campañas de marketing político

Un cumplimiento más riguroso de los deberes del responsable evitaría las desastrosas consecuencias que el perfilado con fines de marketing político está trayendo consigo.

3.2.1. Seguridad

El enfoque de responsabilidad proactiva[38] que caracteriza el RGPD (véase en especial su artículo 24), según el cual al responsable del tratamiento se le exige una valoración

37. Selbest, A. D. and Powles, J. (2017), «Meaningful information and the right to explanation» *International Data Privacy Law 233*, 7(4), 20 https://doi.org/10.1093/idpl/ipx022
38. Hernández Corchete, J.A, «Expectativas de privacidad, tutela de la intimidad y protección de datos», *Sociedad Digital y Derecho*, ob. cit., págs. 279-300.

general del riesgo que el mismo conlleva, pone el énfasis más en el derecho fundamental del ciudadano que en el mero cumplimiento de una reglamentación técnica por el responsable del tratamiento. Este nuevo enfoque requiere para su correcto funcionamiento de algunas opciones regulatorias complementarias. Destaca entre ellas que el incumplimiento de determinadas previsiones concretas de la normativa de protección de datos no necesariamente conlleve la imposición de sanciones, debiendo valorarse y justificarse cuándo la adopción de otras medidas correctoras no se considera como medida adecuada y suficiente.

En general, la mejora de la seguridad de los sistemas frente a ciberataques y otras amenazas es un reto clave de la IA[39]. El coste de las medidas de seguridad preventivas es menor que la gestión de los incidentes producidos. Hay que tener en cuenta que la IA ha creado nuevos riesgos y amenazas que habrá que considerar[40].

Acerca de la seguridad, comenta Vida Fernández[41], que la seguridad interna es un elemento común a todos los sistemas que deben funcionar de forma correcta para lo que han de estar adecuadamente diseñados. Asimismo, ese funcionamiento debe ser estable y resistente frente a alteraciones imprevistas en las condiciones de desarrollo, ya sean fortuitas o provocadas, por ejemplo, a través de un ciberataque. En el caso de la IA estas exigencias son más difíciles de cumplir en la medida que se trata de sistemas especialmente complejos que dependen de muchas variables para su correcto funcionamiento, por lo que se mantienen en un equilibrio muy precario que puede verse fácilmente alterado y dar lugar, no ya a una interrupción de su funcionamiento sino a desviaciones que pueden derivar en daños incluso mayores.

En cuanto a la seguridad externa, los sistemas de IA plantean unos problemas específicos ya que pueden plantearse dudas en cuanto a las consecuencias de su actividad. Los algoritmos que generan la capacidad de autoaprendizaje deben responder no solo a las condiciones de funcionamiento normal sino también a circunstancias extraordinarias para evitar que puedan verse superados ante su incapacidad para improvisar. Asimismo, debe garantizarse que la totalidad de los posibles resultados de su funcionamiento sean siempre seguros para la sociedad, lo cual introduce el problema del carácter autónomo que se analizará en el siguiente apartado.

39. González Espejo, M. J. y Pavón, J. (editores); Barrio Andrés, M. [et al.] (2020), *An Introductory Guide to Artificial Intelligence for Legal Professionals*. Ed. Kluwer Law International
40. Damiani, E. Artificial Intelligence Cybersecurity Challenges, ENISA, Diciembre 2020, https://www.enisa.europa.eu/publications/artificial-intelligence-cybersecurity-challenges.
41. Vida Fernández, J. «Los retos de la regulación de la inteligencia artificial: algunas aportaciones desde la perspectiva europea», en *Sociedad digital y Derecho,* ob. cit., pág. 202.

3.2.2. El papel del Consejo Europeo de la IA

Hay que destacar el papel del Consejo Europeo de la Inteligencia Artificial, propuesto en la normativa europea sobre IA que se encuentra en fase de proyecto, para asegurar que esta regulación se implemente y se cumpla.

La nueva regulación establece la creación de un Consejo Europeo de Inteligencia Artificial (el *Consejo*). El Consejo estará compuesto por un representante de cada autoridad nacional de control y un representante de la Comisión Europea. Cada autoridad nacional de control designará en el Consejo a un representante competente para desempeñar diversas funciones. El Consejo adoptará su propio reglamento interno por mayoría simple de sus miembros.

Las competencias de este Consejo Europeo serán:

a) supervisión de la aplicación coherente del Reglamento por parte de los Estados miembros, incluso mediante la emisión de dictámenes o documentos de orientación interpretativos;
b) recopilación y divulgación de las mejores prácticas entre los Estados miembros;
c) participación en el desarrollo de normas armonizadas o especificaciones comunes relacionadas con la inteligencia artificial;
d) asesoramiento y conocimientos técnicos a la Comisión y a otras instituciones, agencias y organismos de la Unión sobre cuestiones específicas relacionadas con la inteligencia artificial;
e) seguimiento continuo de la evolución técnica y del mercado en relación con la inteligencia artificial, incluidas sus repercusiones en la salud y la seguridad y en los derechos y libertades fundamentales de las personas
f) garantía de la coherencia y la coordinación en el funcionamiento de los compartimentos estancos reguladores de la inteligencia artificial.

La nueva regulación europea establece, además, el mantenimiento de una base de datos de aplicaciones de IA de alto riesgo, que facilitará un control sobre el alcance de estas aplicaciones y su impacto en los derechos.

3.2.3. Evaluaciones de impacto en la protección de datos

La evaluación de impacto en la protección de datos (EIPD) es un elemento clave de la ley de protección de datos que se centra en la responsabilidad y la protección de datos desde el diseño.

No se deben entender las EIPD como un simple ejercicio de cumplimiento de requisitos. Pueden servir de hoja de ruta para identificar y controlar los riesgos para los derechos y las libertades que puede plantear el uso de la IA. También son una oportunidad ideal para considerar y demostrar la responsabilidad por las decisiones que los responsables toman en el diseño o la adquisición de sistemas de IA.

Las EIPD son necesarias, en la gran mayoría de los casos, porque el uso de la IA implicará un tipo de tratamiento que puede suponer un alto riesgo para los derechos y las libertades de las personas y, por lo tanto, dará lugar a la obligación legal de realizar una EIPD. Se deberá realizar esta evaluación caso por caso. En los supuestos en los que considere que un uso concreto de la IA no implica un tratamiento de alto riesgo, deberá documentarse cómo se ha realizado esta evaluación.

Si el resultado de la evaluación indica que existe un riesgo residual elevado para las personas que no puede reducirse suficientemente, deberá consultarse a la autoridad de protección de datos antes de iniciarse el tratamiento.

Además de llevar a cabo una EIPD, también se puede recomendar a los responsables de los tratamientos que realicen otros tipos de evaluaciones de impacto o que los haga voluntariamente. Por ejemplo, evaluaciones de impacto de algoritmos.

A este tipo de evaluación se refiere el art. 35 RGPD y ahora también el art. 28 de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

La EIPD será obligatorio si el uso de la IA implica: a) una evaluación sistemática y amplia de aspectos personales basada en el tratamiento automatizado, incluida la elaboración de perfiles, sobre la que se toman decisiones que producen efectos legales o de importancia similar; b) un tratamiento a gran escala de categorías especiales de datos personales; c) o una vigilancia sistemática de zonas de acceso público a gran escala.

La EIPD debe describir la naturaleza, el alcance, el contexto y los fines del tratamiento de los datos personales. Tiene que dejar claro cómo y por qué va a utilizar la IA para procesar los datos y debe detallar:

a) cómo va a recoger, almacenar y utilizar los datos;
b) el volumen, la variedad y la sensibilidad de los datos
c) la naturaleza de su relación con los individuos;
d) y los resultados previstos para los individuos o la sociedad en general.

El hecho de que un sistema que utilice IA sea, en general, más o menos arriesgado que un sistema que no la utilice depende de las circunstancias específicas. Por lo tanto, debe evaluarse en función de su propio contexto. La EIPD debe mostrar pruebas acerca de la consideración de alternativas menos arriesgadas, si las hay, que logren el mismo objetivo del tratamiento, y argumentar por qué no se han elegido.

La agencia noruega de protección de datos ha analizado algunas herramientas técnicas que pueden facilitar la inclusión de la protección de datos en la IA[42].

42. Norwegian Data Protection Authority (2018) *Artificial Intelligence and privacy,* pág. 26. https://www.datatilsynet.no/en/regulations-and-tools/reports-on-specific-subjects/ai-and-privacy/

3.2.4. Auditorías

Muchos expertos defienden la auditabilidad de los algoritmos[43] como un sistema que permitirá evitar la puesta en el mercado de sistemas de inteligencia artificial que conlleven riesgos inasumibles.

Los algoritmos deben desarrollarse teniendo en cuenta no sólo su funcionalidad sino también la licitud de sus premisas, la regularidad de su funcionamiento y la posibilidad de que un tercero pueda auditar todo ello.

El software debe estar diseñado para satisfacer estas exigencias de transparencia (limitada) y auditabilidad (plena). Esta auditabilidad respetará la confidencialidad del *know how,* secreto integrado en el sistema y la privacidad de los datos personales de los interesados.

El sistema tiene también debilidades: no sirve para auditar algoritmos *no deterministas* (los que trabajan con un componente esencial de aleatoriedad) ni para algunos sistemas de *machine learning*. Para estos hay otras soluciones.

Cotino (2019) recuerda además que «para determinar la intención, la responsabilidad por los actos de los instrumentos inteligentes, se necesitan *AI guardians*, esto es, programas de IA para examinar los propios programas de IA». La idea es que una IA que pueda vigilar los algoritmos.

En España, la recién aprobada Carta de Derechos Digitales recoge esta cuestión en la mención que hace de los derechos ante la Inteligencia Artificial: «se establecerán condiciones de transparencia, auditabilidad, explicabilidad, trazabilidad, supervisión humana y gobernanza».

3.2.5. Evaluación de conformidad para aplicaciones de alto riesgo

Antes de comercializar un sistema de IA de alto riesgo en el mercado de la Unión o de ponerlo en servicio de otra forma, los proveedores deberán someterlo a una evaluación de la conformidad. Esto les permitirá demostrar que su sistema cumple los requisitos obligatorios de una IA digna de confianza (por ejemplo, calidad de los datos, documentación y trazabilidad, transparencia, supervisión humana, exactitud y solidez). En caso de que el sistema en sí o su finalidad se modifiquen sustancialmente, deberá repetirse la evaluación. Para determinados sistemas de IA, también deberá participar en este proceso un organismo notificado independiente. Los sistemas de IA que sean componentes de seguridad de productos contemplados en la legislación sectorial de la Unión siempre se considerarán de alto riesgo cuando sean objeto de una evaluación de la conformidad por terceros con arreglo a dicha legislación sectorial. También en el

43. Kazim, E., Mendes, D., Denny, D. M. T. Y Koshiyama, A. (2021), «AI auditing and impact assessment: according to the UK information commissioner's office», *AI and Ethics.* https://doi.org/10.1007/s43681-021-00039-2

caso de los sistemas de identificación biométrica, siempre hará falta una evaluación de la conformidad por terceros.

Los proveedores de sistemas de IA de alto riesgo también tendrán que aplicar sistemas de calidad y gestión de riesgos para garantizar su conformidad con los nuevos requisitos y minimizar los riesgos para los usuarios y las personas afectadas, incluso después de que un producto se haya comercializado. Las autoridades de vigilancia del mercado apoyarán el seguimiento posterior a la comercialización mediante auditorías y ofreciendo a los proveedores la posibilidad de informar sobre incidentes graves o violaciones de los derechos fundamentales de que hayan tenido conocimiento.

Los Estados miembros desempeñan un papel clave en la aplicación y el cumplimiento del Reglamento de Inteligencia artificial (en proyecto). A este respecto, cada Estado miembro deberá designar una o varias autoridades nacionales competentes para supervisar la aplicación y ejecución, así como para realizar actividades de vigilancia del mercado. A fin de aumentar la eficiencia y establecer un punto de contacto oficial con la población y otros homólogos, cada Estado miembro deberá designar una autoridad nacional de supervisión, que también representará al país en el Comité Europeo de Inteligencia Artificial.

3.2.6. Sanciones por incumplimiento

Cuando se comercialicen o utilicen sistemas de IA que no respeten los requisitos del Reglamento, los Estados miembros deberán establecer sanciones efectivas, proporcionadas y disuasorias, incluidas multas administrativas, por las infracciones y notificarlas a la Comisión.

El Reglamento fija umbrales que deberán tenerse en cuenta:

- hasta treinta millones de euros o el 6 % del volumen de negocios anual total a escala mundial del ejercicio financiero anterior, optándose por la de mayor cuantía, por las infracciones por incumplimiento o prácticas prohibidas en relación con los requisitos sobre los datos;
- hasta veinte millones de euros o el 4 % del volumen de negocios anual total a escala mundial del ejercicio financiero anterior por el incumplimiento de cualquier otro requisito u obligación del Reglamento;
- hasta diez millones de euros o el 2 % del volumen de negocios total anual a escala mundial del ejercicio anterior por el suministro de información incorrecta, incompleta o engañosa a los organismos notificados y a las autoridades nacionales competentes en respuesta a una solicitud.

Con el fin de armonizar las normas y prácticas nacionales en materia de establecimiento de multas administrativas, la Comisión formulará directrices con el asesoramiento del Comité.

Como las instituciones, agencias y organismos de la UE deben dar ejemplo, también estarán sujetas a las normas y a las posibles sanciones; el Supervisor Europeo de Protección de Datos estará facultado para imponerles multas.

4. REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL

4.1. Consideraciones generales

Los esfuerzos reguladores de la UE en otros ámbitos digitales como la economía de las plataformas sugieren que se utilicen alternativas a la legislación como la autorregulación o los métodos de co-regulación[44].

La autorregulación puede ser una parte importante de un proceso de regulación junto con los esfuerzos legislativos convencionales, pero este nuevo enfoque de gobernanza debe garantizar que todos los actores relevantes sean invitados a participar, no sólo los proveedores de la industria.

Según el enfoque de los derechos humanos, los organismos y foros de autorregulación no deberían consistir en reuniones en las que los actores de la industria discuten cuestiones de confianza, seguridad y protección de los ciudadanos sin contar con estos últimos. La gobernanza y la regulación de los algoritmos son cuestiones importantes que necesitan un amplio consenso en el que un debate democrático debe tener en cuenta la opinión de la comunidad en general.

La ética debe introducirse en la estructura de gobernanza, y no debe excluirse la regulación de ningún mecanismo de control. Actividades como la elaboración de perfiles y el *microtargeting*, que suponen graves amenazas para el pluralismo político y cultural, probablemente no se abordarían en un organismo de autorregulación compuesto únicamente por miembros corporativos.

Las regulaciones de la privacidad de los datos, que ya están en vigor en varios países de todo el mundo, ya afectan a la actividad algorítmica, pero no de forma completa. Aunque el RGPD establece las normas relativas a los algoritmos que utilizan datos personales, hay otros algoritmos que utilizan otros tipos de datos, que mal manejados pueden afectar a los derechos de las personas. El RGPD no es suficiente para un sistema de algoritmos basado en los derechos humanos.

La autorregulación es una de las alternativas, y aquí es donde actualmente nos encontramos. Deberíamos hacer una reflexión amplia para conocer si está funcionando. Aparentemente no: las grandes empresas tecnológicas americanas están publicando sus principios genéricos de IA sin especificar claramente un conjunto de comprobaciones reales para certificar el cumplimiento de estos principios.

44. Clarke, R. (2019) «Australia Regulatory alternatives for AI», *Computer Law & Security review*, num. 35, págs. 398-409 https://doi.org/10.1016/j.clsr.2019.04.008

En algunos casos, se han encontrado ejemplos notables que contravienen su propio conjunto de principios. Estas incoherencias denotan la desconexión entre los comités de ética de las empresas y los equipos de desarrollo y puesta en marcha de productos. Es esencial un enfoque multidisciplinar que incluya la ética y que esté en el centro del sistema de decisión de las empresas.

La integración de los conocimientos de expertos y expertas, en particular de las ciencias sociales e incluyendo métodos cualitativos, puede ayudar a superar algunas limitaciones. Estos conocimientos pueden utilizarse, por ejemplo, para informar al modelador de variables que pueden ser influyentes, pero no observadas (dos conjuntos independientes, el racismo estructural y el racismo del jurado, pueden conducir injustamente a un veredicto de culpabilidad).

Algunas voces muy autorizadas (Vinuesa, Azizpour, Leite, Balaam, Dignum, Domish, Fellander, Langhans, Tegmark, Nerini, 2020[45]) han afirmado que:

> «creemos que es imperativo desarrollar una legislación relativa a la transparencia y la responsabilidad de la IA, así como decidir el impulso que están dando iniciativas como la del IEEE (Instituto de Ingenieros Eléctricos y Electrónicos) y las nuevas directrices éticas de la UE para una IA fiable, Ethical Guidelines for trustworthy AI».

Sin embargo, el margen de maniobra del RGPD en el ámbito de los sistemas de decisión algorítmica (ADM) es bastante restringido. En cuanto a los objetivos sociales y de grupo, como la no discriminación y la participación, el RGPD tiene poco que ofrecer. Por este motivo, es necesario debatir sobre herramientas normativas complementarias más allá del RGPD.

La aplicación de la prohibición de los sistemas ADM del RGPD está, por diversas razones, muy restringida. El RGPD solo prohíbe la toma de decisiones totalmente automatizada.

Los sistemas que sólo preparan la base de las decisiones humanas y dan recomendaciones pueden seguir utilizándose. Para que la prohibición entre en vigor, los sistemas ADM deben tomar decisiones totalmente automatizadas sobre la base de datos personales. Además, las decisiones deben tener consecuencias legales o afectar de forma significativa al interesado. Si falta uno de estos tres criterios los sistemas deben ser lo más robustos posible para garantizar que –aunque se respete el propósito ético– la IA no cause daños involuntarios.

Hasta el momento, solo pueden mencionarse algunos ejemplos de regulación de justicia algorítmica. Uno es la Ley Básica sobre Robots (Corea del Sur) de 2019, y la propuesta de Algorithmic Accountability Act de Estados Unidos de 2019, que busca regular el sesgo en los sistemas automáticos de toma de decisiones. No especifica una definición de imparcialidad, permitiendo flexibilidad. Se exige a las empresas que

45. Tegmark et al. (2020) «The role of artificial intelligence in achieving the Sustainable Development Goals», *Nature Communications.*

auditen su aprendizaje automático para detectar sesgos y discriminación en una evaluación de impacto.

Es interesante conocer la regulación del uso de la IA en el ámbito público, por ejemplo, en el caso de uso de datos de salud. Contamos con algunas regulaciones comparadas, como el caso francés y el australiano[46]. En Australia, el uso de la inteligencia artificial para la toma de decisiones administrativas se está extendiendo, con habilitaciones normativas mucho más simples que las francesas.

4.2. Regulación europea de la IA en fase de discusión

Las exigencias legales que existen en este momento para la IA son comunes a cualquier tecnología que trate datos personales, no existiendo por el momento normativa específica aprobada. Como se ha mencionado, existe una propuesta a nivel europeo[47] que merece especial detenimiento por su pertinencia. Los sistemas de IA plantean retos adicionales gracias a su capacidad de procesamiento de los datos, por lo que pueden dar lugar a formas de tratamientos desconocidos hasta ahora, cuyos resultados pueden desbordar las medidas de protección dispuestas en la normativa común. Por este motivo, una norma que se adapte a los retos que presenta la IA es una necesidad clara de nuestras sociedades.

La regulación europea es el primer marco legal sobre esta tecnología, que además llega acompañada de otra normativa sobre maquinaria y robots. Una nueva normativa sobre IA que quiere garantizar la seguridad y fortalecer la inversión en IA en la Unión Europea, creando varios niveles de riesgo y prohibiendo el reconocimiento facial en determinadas situaciones**.**

La propuesta de la Comisión Europa para regular la inteligencia artificial informa de que las normas deberán ser implementadas por todos los Estados miembros por igual, quedando excluidos de la normativa los usos de la IA a nivel militar.

Los riesgos se clasifican en cuatro niveles. El mayor es el riesgo inaceptable, el que constituye una amenaza para la seguridad, los medios de vida y los derechos de las personas**.** Estos sistemas de IA estarán prohibidos, como el caso de la IA diseñada para manipular comportamientos y los sistemas de puntuación social, que dan una valoración social en función del comportamiento digital de los ciudadanos.

En un segundo punto de alto riesgo se incluyen usos de la IA en infraestructuras críticas que puedan afectar a la salud de los ciudadanos, usos de IA aplicada en la

46. Ponce Solé, J. *Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico*. http://laadministracionaldia.inap.es/usuarios/noticia.asp?id=1509505#nota75

47. Propuesta de regulación de la Inteligencia artificial de abril de 2021: https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682

educación, componentes en cirugía, sistemas de reclutamiento de personal, servicios públicos, legislación, inmigración o IA para la administración o justicia.

En todos estos ámbitos, la IA deberá estar sujeta a obligaciones estrictas, entre las que se incluye un análisis de riesgos, trazabilidad de resultados, documentación detallada, supervisión humana[48] y un alto nivel de robustez.

En un nivel más bajo, de riesgo limitado se incluyen los sistemas como *chatbots*, que deberán tener un mínimo nivel de transparencia y donde los usuarios deberán ser advertidos que están hablando con una máquina**.**

En el riesgo mínimo se engloba el resto de usos como videojuegos, aplicaciones de imagen u otros sistemas de IA que no impliquen riesgos. En estos casos, la nueva normativa no especifica ninguna medida.

Europa quiere impulsar el desarrollo de estándares para la IA y propone a las distintas organizaciones nacionales que supervisen esta normativa. Adicionalmente, desde la Comisión invitan a la creación de códigos voluntarios de conducta para los sistemas de IA sin riesgos.

Todos los sistemas de identificación biométrica remota serán considerados de alto riesgo. La Comisión Europea ha decidido no prohibir directamente los sistemas de reconocimiento facial**,** aunque sí aplicará requisitos estrictos. Se prohibirá su uso en zonas públicas, pero se contemplan algunas excepciones.

Es decir, las autoridades no podrán utilizar el reconocimiento facial para prevenir posibles delitos, aunque hay excepciones: «cuando sea estrictamente necesario para buscar un niño desaparecido, para prevenir una amenaza terrorista específica e inminente o para detectar, localizar, identificar o enjuiciar a un perpetrador o sospechoso de un delito grave». Estos usos concretos estarán sujetos a la autorización de un órgano judicial u otro organismo independiente y a los límites adecuados en el tiempo, el alcance geográfico y las bases de datos buscadas.

Respecto al reconocimiento facial, algunas organizaciones como *European Digital Rights* (EDRi)[49] opinan que cada vez era más difícil de justificar el no prohibir la vigilancia masiva biométrica. Aplauden la prohibición, a pesar de que no va suficientemente lejos. Entre las dudas que surgen está la definición de *remoto* o *en tiempo real***,** que podría abrir la puerta a determinados sistemas de reconocimiento como el uso de drones cerca de los manifestantes, cámaras en los policías o sistemas que no sean en directo.

48. LAZCOZ MORATINOS, Guillermo (2021) «Análisis de la propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas». *IUS ET SCIENTIA* N.º 2 • ISSN 2444-8478, p. 40: En el apartado primero del artículo 7 que define la clase de intervención humana requerida para una IA de alto riesgo, sería oportuno que se añadiese a la garantía de una *supervisión humana integra*l el criterio cualitativo *y significativa*.
49. Muchos de los usos más dañinos no están prohibidos, como la vigilancia policial predictiva, los usos de la IA para el control de la migración, la categorización biométrica de raza, género, sexualidad y también la vigilancia de los trabajadores, que siguen siendo de alto riesgo.

CONCLUSIONES

El análisis realizado en este artículo ha pretendido establecer las relaciones y posibles colisiones entre el derecho a la protección de datos y la elaboración de perfiles de ciudadanos a través de la IA, con la finalidad de enviar publicidad política. Se están produciendo muchas colisiones con los derechos y los desarrollos futuros de esta tecnología plantearán nuevos retos.

Por el momento, la recomendación clave radica en reforzar los deberes del responsable de los tratamientos de datos personales a través de IA, que incluyen la realización de auditorías específicas y un exhaustivo deber de información que debe llegar a abarcar el principio de explicabilidad, superando las limitaciones impuestas por las *black-boxes.*

Valorando la necesidad de acudir a principios éticos consensuados, se aplaude muy especialmente la propuesta de regulación que la Comisión europea ha hecho pública recientemente. Dicha futura regulación debe incluir el enfoque de derechos humanos, asegurando un marco de protección reforzado para las personas, especialmente las que se encuentran en una situación de mayor vulnerabilidad. Todos y cada uno de los principios de derechos humanos deben incluirse en la legislación, con especial énfasis en el principio de transparencia y rendición de cuentas.

BIBLIOGRAFÍA

AGDP (2020*) Adecuación al RGPD de tratamientos que incorporan Inteligencia Artificial. Una introducción*. https://www.aepd.es/es/media/guias/adecuacion-rgpd-ia.pdf.

Australian Human Rights Commission (2019) *Human Rights and Technology Discussion Paper***.** ISBN 978-1-925917-15-4.

Barfield, W. (2020), *The Cambridge Handbook of the Law of Algorithms.* University of Washington, Ed. Cambridge University Press.

Bonmatí Sánchez, J., y Gonzalo Doménech, J. J., (2020) «La gestión de riesgos y su encaje legal en la regulación de la inteligencia artificial», *Era digital, Sociedad y Derecho*, Tirant lo Blanch, Valencia.

Christodoulou, E., Kalypso I. (2019), «Democracy Under Attack: Challenges of Addressing Ethical Issues of AI and Big Data for More Democratic Digital Media and Societies» *Frontiers in Political Science*.

Clarke, R. (2019) «Australia Regulatory alternatives for AI» *Computer Law & Security review*, num. 35, pp. 398-409.

Comisión Europea. Propuesta de regulación de la Inteligencia artificial de abril de 2021: https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682

Consejo de Europa (2020) *Protección de las personas en lo que respecta al tratamiento automático de datos personales en el contexto de la elaboración de perfiles*. Recomendación CM/Rec (2010)13 y exposición de motivos. Consejo de Europa, 23 de noviembre de 2010. https://www.coe.int/t/dghl/standardsetting/cdcj/CDCJ%20Recommendations/CMRec(2010)13E_Profiling.pdf

Cotino Hueso, L. (2019). «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), pp. 29-48.

Damiani, E. *Artificial Intelligence Cybersecurity Challenges*, ENISA, Diciembre 2020, https://www.enisa.europa.eu/publications/artificial-intelligence-cybersecurity-challenges

De la Quadra Salcedo Fernández del Castillo, T., (2018) «Retos, riesgo y oportunidades de la sociedad digital», *Sociedad Digital y Derecho,* de la Quadra-Salcedo, T., Piñar Mañas, J. L., (directores), Ministerio de Industria, Comercio y Turismo, Madrid

Dubber, P. (2020), *The Cambridge Handbook of the Law of Algorithms.* University of Washington, Ed. Cambridge University Press.

Ebert, N. (2020), *Algorithms and Law*. Ed. Cambridge University Press.

Fernández-Aller, C. (2020), «¿Vale todo para hacer propaganda electoral en internet?» *The Conversation* https://theconversation.com/profiles/celia-fernandez-aller-718032/articles

Fernández Aller, C. *et. al* (2021), «An Inclusive and Sustainable Artificial Intelligence Strategy for Europe Based on Human Rights», in *IEEE Technology and Society Magazine*, vol. 40, no. 1, pp. 46-54, March 2021, DOI: 10.1109/MTS.2021.3056283.

Floridi, L (2019), «Translating principles into Practices of Digital Ethics: Five Risks of Being Unethical». *Philosophy and Technology,* 32.

González-Espejo, M.J y Pavón, J. (editores); Moisés Barrio Andrés…[et al.] (2020) *An Introductory Guide to Artificial Intelligence for Legal Professionals*. Ed. Kluwer Law International.

Grupo de expertos de alto nivel sobre inteligencia artificial, creado por la Comisión Europea (2019) *Directrices éticas para una Inteligencia Artificial fiable*, https://op.europa.eu/es/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1

Hernández Corchete, J.A. (2018). «Expectativas de privacidad, tutela de la intimidad y protección de datos», *Sociedad Digital y Derecho,* de la Quadra-Salcedo, T., Piñar Mañas, J. L., (directores), Ministerio de Industria, Comercio y Turismo, Madrid (2018).

Kazim, E., Danielle M., Thame D., Adriano K. (2021), «AI auditing and impact assessment: according to the UK information commissioner's office», *AI and Ethics.* https://doi.org/10.1007/s43681-021-00039-2

Lazcoz Moratinos, G., y Castillo Parrilla, J. A., (2020) «Valoración algorítmica ante los derechos humanos y el Reglamento General de Protección de Datos: el caso SyRI», *Revista Chilena de Derecho y Tecnología*, vol. 9, núm. 1.

Lazcoz Moratinos, G. (2021) «Análisis de la propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas» *IUS ET SCIENTIA* N.º 2 • ISSN 2444-8478

Le Clainche, J. y Le Métayer, D. (2012), «Données personnelles, vie privée et non-discrimination: des protections complémentaires, une convergence nécessaire (Personal data, privacy and non-discrimination: complementary protections, a necessary convergence)», *Revue Lamy Droit immatériel,* 90.

Mitrou, L. (2018), «Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'? » Available at SSRN: https://ssrn.com/abstract=3386914 or http://dx.doi.org/10.2139/ssrn.3386914

Moerel, E.M.L. (2014), «Big Data Protection. How to Make the Draft EU Regulation on Data Protection Future Proof» (February 14, 14). Available at SSRN: https://ssrn.com/abstract=3126164 or http://dx.doi.org/10.2139/ssrn.3126164

Norwegian Data Protection Authority (2018), *Artificial Intelligence and privacy*

Ponce Solé, J. (2019), *Inteligencia artificial, Derecho administrativo y reserva de humanidad: algoritmos y procedimiento administrativo debido tecnológico*. http://laadministracionaldia.inap.es/usuarios/noticia.asp?id=1509505#nota75

Quadra-Salcedo Fernández del Castillo, T. (2019), «Derechos fundamentales, democracia y mercado en la edad digital», *Derecho Digital e Innovación*, núm. 1, enero-marzo.

Rebollo Delgado, L., Serrano Pérez, M.ª M., *Manual de Protección de Datos*, 3.ª ed., Dykinson, S. L., Madrid 2019.

Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, COM(2021) 206 final, disponible en https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence

Selbst, A. and Powles, J. (2017), «Meaningful information and the right to explanation» *International Data Privacy Law 233*, 7(4), 20.

Schermer, B (2011), «The limits of privacy in automated profiling and data mining» 27 *Computer Law & Security Review* 45.

Smith, B. C., (2019), *The promise of artificial intelligence*. *Reckoning and judgment*, The MIT Press, Cambridge (Massachusets).

Tegmark *et al*. (2020), «The role of artificial intelligence in achieving the Sustainable Development Goals» *Nature Communications.*

Tegmark, Max (2017), *Life 3.0. Being human in the age of Artificial Intelligence*. Vintage.

The Alan Turing Institute /ICO (2020) *Explaining decisions made with AI.* https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/explaining-decisions-made-with-ai/

Vida Fernández, J. (2018), «Los retos de la regulación de la inteligencia artificial: algunas aportaciones desde la perspectiva europea» en *Sociedad Digital y Derecho,* de la Quadra-Salcedo, T., Piñar Mañas, J. L., (directores), Ministerio de Industria, Comercio y Turismo, Madrid (2018).

Yeung, Howes, Pogrebna (2020), «AI Governance by Human Rights-Centered Design, Deliberation and Oversight: and End to Ethics Washing», in *The Oxford Handbook of Ethics of AI,* Ed. Oxford University Press.