Cuadernos Europeos de Deusto

No. 68/2023

DOI: https://doi.org/10.18543/ced682023

ESTUDIOS

Inteligencia artificial y ética: hacia una aplicación de los principios éticos en el ámbito de la UE

Artificial intelligence and ethics: towards an application of ethical principles in the EU

Francisco Javier Martín Jiménez

doi: https://doi.org/10.18543/ced.2699

Recibido el 13 de septiembre de 2022 • Aceptado el 16 de enero de 2023 • Publicado en línea: abril de 2023

Inteligencia artificial y ética: hacia una aplicación de los principios éticos en el ámbito de la UE

Artificial intelligence and ethics: towards an application of ethical principles in the EU

Francisco Javier Martín Jiménez
Profesor Asociado
Universidad Pontificia de Salamanca
fjmartinji@upsa.es

doi: https://doi.org/10.18543/ced.2699

Recibido el 13 de septiembre de 2022
Aceptado el 16 de enero de 2023
Publicado en línea: abril de 2023

Resumen: La inteligencia artificial se muestra como la aplicación o conjunto de aplicaciones informáticas que pretenden emular las actividades cognoscitivas desarrolladas habitualmente por los seres humanos. Si bien esta realidad nos puede parecer extraordinariamente fructífera para el desarrollo de nuestra sociedad en cualquier ámbito, también entraña riesgos, entre otras cosas porque la inteligencia artificial puede llegar a tomar decisiones al margen, incluso, de las propias previsiones de su creador. Ante estos riesgos podemos aplicar principios éticos que puedan ser referentes para desarrolladores, distribuidores, aplicadores de estas tecnologías, estén o no incorporados a normas de obligado cumplimiento. Tanto los riesgos como las respuestas están siendo objeto de estudio, con aportación de propuestas, en el seno de la Unión Europea, asentándose, como punto de referencia, en los principios éticos. Acercarse al análisis de esta posición es el objetivo de este trabajo.

Palabras clave: Inteligencia Artificial, principios éticos, Unión Europea.

***Abstract*:** *Artificial intelligence is shown as the application or set of computer applications that aim to emulate the cognitive activities usually developed by human beings. While this reality may seem extraordinarily fruitful for the development of our society in any field, it also entails risks, among other things*

because artificial intelligence can make decisions outside even its creator's own foresight. In the face of these risks, we can apply ethical principles that can serve as benchmarks for developers, distributors and applicators of these technologies, whether or not they are incorporated into mandatory standards. Both the risks and the responses are being studied and proposals are being made within the European Union, based on ethical principles as a point of reference. The aim of this paper is to approach the analysis of this position.

Keywords: *Artificial Intelligence, ethical principles, European Union.*

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

I. **Introducción**

La acción del hombre, más aún el desarrollo tecnológico, tiene la capacidad de generar oportunidades y desarrollar efectos positivos y, también, negativos en la sociedad a la que se dirige. Pueden, por ello, verse afectados los principios y valores sociales y económicos imperantes en un contexto determinado.

Las nuevas tecnologías están nuevamente en disposición de mejorar la productividad, de avanzar en nuevos logros sociales y económicos, y, por ejemplo, avanzar en nuevas actividades laborales y profesionales, empleo eliminando tareas repetitivas o engorrosas, superar las condiciones medioambientales, evitando la congestión del tráfico y los contaminantes atmosféricos, mejorar el transporte reduciendo colas y optimizando rutas, aumentar la seguridad vial, reduciendo las posibilidades de error humano. Pero, a su vez, las nuevas tecnologías pueden conllevar riesgos para los derechos adquiridos. Será importante asegurar la falibilidad de los sistemas de inteligencia artificial (IA), que guarde la uniformidad de las distintas normativas de aplicación territorial y también el exceso de regulación, subsanando las lagunas jurídicas existentes. Los derechos fundamentales deben asegurarse, la libertad de expresión, del derecho a una información veraz, el pluralismo en libertad de los medios de opinión y comunicación, la seguridad jurídica, desde una adecuada aplicación de la norma, desde una interpretación normativa común en la Unión, también de los de las definiciones eminentemente tecnológicos, como algoritmos, programas informáticos, datos, etc.

La IA es el anticipo del desarrollo independiente de la inteligencia en las máquinas, como seres creados por el hombre. Pensar que aquellas puedan ser tanto o más inteligentes que los seres humanos más inteligentes nos puede producir preocupación pero también alivio, porque el arte de pensar no deja de ser un esfuerzo que puede ser comparable, salvando las distancias y sin tratar de herir sensibilidades, con el esfuerzo físico. Por ello, tan fácil puede comprenderse que una máquina haya podido acabar con actividades peligrosas o de gran deterioro físico, de igual forma puede hacerlo con tediosas y agotadoras actividades intelectivas, que podrían ser realizadas por las máquinas, sin mayor limitación que el coste energético. Así las cosas, BOSTROM[1], citando a NILSON[2] nos apunta que la IA podría llegar a nivel humano entre 2030 y 2100.

---

[1] Nick Bostrom. *Superinteligencia. Caminos, peligros, estrategias*. (Teell Editorial, S. L. 2016). 19.

[2] Nils Johan Nilsson. *The quest for artificial intelligence*. (Cambridge University Press, 2009).

Siguiendo a HAYES et al, 2021[3], la IA supone «una aproximación computacional a tareas asociadas a la inteligencia humana relacionadas con el aprendizaje continuo, resolución y explicación de problemas y creación de patrones».

Por su parte, la Comisión Europea[4] define a la IA como «un sistema basado en programas informáticos o incorporado en dispositivos físicos que manifiesta un comportamiento inteligente al ser capaz, entre otras cosas, de recopilar y tratar datos, analizar e interpretar su entorno y pasar a la acción, con cierto grado de autonomía, con el fin de alcanzar objetivos específicos».

La propuesta de Reglamento[5] es más precisa al definir el sistema de IA como «el software que se desarrolla empleando una o varias de las técnicas y estrategias y que puede, para un conjunto determinado de objetivos definidos por seres humanos, generar información de salida como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúa».

Analizar el entorno y tomar decisiones acertadas, entendiendo por estas las que producen un beneficio al individuo o a la colectividad, constituye, de forma simplista, la actividad inteligente. Es acertado relacionar el análisis y la decisión con el objetivo específico pues es el resultado lo que, a la postre, mostrará un mayor vigor inteligente, aunque los intentos también deban ser considerados.

La IA no es un producto que nace por sí solo sino que procede, como es evidente, de la labor del hombre. Para SUSSKIND, R. y SUSSKIND, D[6], nos encontramos ante «sistemas no pensantes de alto rendimiento». Estos podrán, en el futuro, elaborar soluciones a los problemas planteados con dosis de creatividad, incluso se podrían considerar ingeniosas o propias de

---

[3] Jameson L. Hayes,., *et al*. «Can social media listening platforms' artificial intelligence be trusted? Examining the accuracy of Crimson Hexagon's (now Brandwatch Consumer Research's) AI-Driven analyses». *Journal of Advertising*, (2021), vol. 50, n.º 1. 81-91. https://doi.org/g2vx

[4] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL)). Artículo 4, a). Adaptación de la definición que figura en la Comunicación de la Comisión Europea COM(2018)0237, de 25.4.2018, p. 1.Leído en https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_ES.html.

[5] Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadoras en materia de IA (Ley de IA) y se modifican determinados actos legislativos de la Unión, de 24 de abril de 2021. Leído en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=uriserv%3AOJ.C_.2017.288.01.0001.01.SPA&toc=OJ%3AC%3A2017%3A288%3ATOC

[6] Richard Susskind y Daniel Susskind. *El futuro de las profesiones. Cómo la tecnología transformará el trabajo de los expertos humanos*, trad. de J. C. Ruiz. (Teell Editorial, 2016). 272

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

capacidades de cognición y pensamiento. Se puede limitar, no sin reparos, a creer que la actividad pensante es exclusiva del ser humano, como característica intrínseca y necesaria, sin la cual este no sería tal. No seríamos humanos si no pensáramos, circunstancia que no va a ser nunca determinante de las máquinas.

La IA puede, como decíamos, abarcar un sinfín de posibilidades que beneficien a la sociedad, en campos tan dispares como los préstamos, la salud, la contratación, la información[7] y un largo, etc., y en un sentido lógico, en las mismas zonas de acción de la actividad humana pensante.

Si hablamos, por ejemplo, de la salud, la IA puede proporcionarnos un mejor diagnóstico o propiciar una detección temprana de enfermedades que permita, de este modo, adoptar medidas preventivas. La IA puede utilizar la ingente información que genera la investigación biomédica, ordenarla, estructurarla o integrarla para obtener resultados que mejoren los resultados médicos, una mejor atención o una más eficaz asignación de los recursos[8].

Si hablamos de educación[9], la IA puede servir para hacer un seguimiento personalizado de cada alumno, en sí mismo considerado o en relación a otros alumnos, analizando el desempeño, las necesidades y adaptando, en cada caso, esa enseñanza a capacidades y habilidades de cada individuo. El uso, por ejemplo, de la realidad virtual o interactiva nos ofrece muchas posibilidades.

Si nos referimos al mundo empresarial, se pueden aportar soluciones que permitan una mayor eficiencia en las cadenas de producción o en el transporte de los productos elaborados a los clientes.

La defensa, la justicia, la administración tributaria, la seguridad ciudadana, la protección de menores[10], son campos en los que la IA puede desarrollarse con grandes posibilidades. HUESO COTINO[11] considera que tanto el big data como la IA pueden ser útiles y positivas en el ámbito de las

---

[7] Filippo A. Raso *et al*. «Artificial intelligence & human rights: Opportunities & risks». *Berkman Klein Center Research Publication*, n.º 2018(2018). 6.

[8] Lorenzo Cotino Hueso. «Inteligencia artificial, big data y aplicaciones contra la covid-19: Privacidad y protección de datos». *IDP: revista de Internet, derecho y política = revista d'Internet, dret i política*. n.º 31. (2020). 1.

[9] Celia Rangel. «Inteligencia Artificial como aliada en la supervisión de contenidos comerciales perjudiciales para menores en Internet». *Revista Mediterránea de Comunicación: Mediterranean Journal of Communication*, vol. 13, n.º 1 (2022). 26.

[10] Dana Lee Olstad y Joon Lee, «Leveraging artificial intelligence to monitor unhealthy food and brand marketing to children on digital media». *The Lancet Child & Adolescent Health*, vol. 4, n.º 6, (2020). 418-420. https://doi.org/gsd7. La IA es una herramienta válida para automatizar mecanismos que permitan detectar contenidos que pueda ser perjudicial a los menores.

[11] Lorenzo Cotino Hueso. «Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales». *Dilemata*, n.º 24 (2017) 143.

actuaciones policiales y judiciales siempre que pueda ser revisada o completada por el ser humano. También la IA permite extraer información relevante útil para identificar riesgos de corrupción en el ámbito del sector público[12].

Pero no todo son buenas noticias. Existen riesgos que provocan perjuicios. Estos se han puesto de manifiesto por el propio Consejo de Europa[13]: señala que están en juego un buen número de derechos y libertades, de expresión o reunión, de no discriminación por razón de sexo, raza u origen étnico, religión o credo, discapacidad, edad u orientación sexual, de protección de los datos personales y de intimidad, de tutela judicial efectiva.

TORRES JARRÍN[14] pone como ejemplo la utilización del reconocimiento facial o de voz, para generar discriminación hacia grupos sociales más vulnerables o para cercenar derechos en el ámbito de regímenes autoritarios. Ante lo anterior cabe un enfoque regulatorio duro (hardlaw) que precise la utilización de leyes o reglamentos capaces de determinar comportamientos permitidos desde la investigación, desarrollo o puesta en práctica de la IA. El autor cita la propuesta china IA, acostumbrada a dejar todo bien atado, en la denominada —«New Generation Artificial Intelligence Development Plan» (AIDP) del año 2017 y la «China New Generation Artificial Intelligence Development Report» del año 2019[15].

Frente a la primera tenemos la recomendación o soft law que utiliza las guías o propuestas que, sin la coactividad propia de la ley, quieren ser un mecanismo que conduzca por un lugar respetuoso con los derechos la investigación, desarrollo y uso de la IA. En este enfoque se encuentran los países de la UE (Directrices Éticas para un IA fiable) o EE. UU. (Preparation for the Future of Artificial Intelligence).

En nuestra opinión puede ser muy dañino para la investigación y su desarrollo crear normas limitativas puesto que la propia interpretación del jurista creará incertidumbre en el investigador. Es en el uso de la IA dónde se debe mover el Derecho.

Somos partidarios de la utilización de la recomendación o la autoregulación en la prospección y el descubrimiento de la IA. En cuanto al uso, cada derecho afectado será muy dueño de saberse defender con las armas

---

[12] Juli Ponce Solé. La prevención de riesgos de mala administración y corrupción, la inteligencia artificial y el derecho a una buena administración. *Revista internacional de Transparencia e Integridad*, n.º 6 (2018). 11.

[13] https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5

[14] Mario Torres Jarrín. «La UE & la gobernanza ética de la inteligencia artificial: Inteligencia artificial & diplomacia». *Cuadernos salmantinos de filosofía*, vol. 48 (2021) 219-220

[15] Jonathan Piedra Alegría. «Descolonizando la "Ética de la Inteligencia Artificial"». *Dilemata*, n.º 38 (2022) 247-248.

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

de qué disponga o que deban ser mejoradas. Por ello, la recomendación y la flexibilidad que conlleva, sirve mejor a la IA en las fases creativas que la rigidez de la Ley.

Desde la línea más suave, la segunda opción, como manifiesta ALEGRÍA[16], se percibe una gran proliferación de las denominadas «Ethical guidelines», que se suceden con aportaciones diversas, para tratar de resolver, dese un enfoque ético o filosófico la problemática que pudiera surgir de la implementación de la IA.

No participamos de la idea de que la existencia de estas guías parte de la iniciativa del sector privado con el objetivo de evitar la regulación[17], para distraer la atención con soluciones que no abordan verdaderamente el problema[18]. El sector privado prefiere siempre la iniciativa pero, no debemos entender que lo hace necesariamente de mala fe, sino para evitar las modificaciones intempestivas de las reglas del juego que desde el poder político puedan impulsarse, causando pérdidas en la inversión.

Tampoco se trata de eludir la responsabilidad que pueda devenir de la aplicación de la IA. Detrás de un efecto nocivo, debe haber una respuesta jurídica pues las aplicaciones defectuosas o perniciosas de la IA que puedan causar un mal ajeno pertenecen al inmovilizado, sea inmaterial o material, de las empresas, siendo estas últimas las responsables de sus efectos. El mecanismo jurídico de respuesta existe, aunque debe mejorarse.

II. **El Libro Blanco[19] sobre la inteligencia artificial: un enfoque europeo**

La IA muestra oportunidades y también riesgos, como indica la Comisión Europea en el citado documento. Y va a cambiar nuestras vidas. Tendrá ventajas en los ciudadanos, en una asistencia sanitaria más eficiente, unos servicios de transporte rápidos y seguros, una mayor utilidad y durabilidad de los productos, maquinas más eficientes, productos nuevos que satisfagan mejor las necesidades del consumidor, una agricultura y ganade-

---

[16] Piedra Alegría. «Descolonizando la "Ética de la Inteligencia Artificial"». 248

[17] Luciano Floridi. «Translating principles into practices of digital ethics: Five risks of being unethical». En *Ethics, Governance, and Policies in Artificial Intelligence*. Springer, Cham, (2021) 81-90., s13347-019-00354-x. https://doi.org/10.1007/s13347-019-00354-x

[18] Ben Wagner. Ethics as an escape from regulation. From «ethics-washing» to ethics-shopping?. 2018. En Lisa Janssens *et al*. *Being Profiled: Cogitas Ergo Sum*. (Amsterdam University Press, 2019) 84-89. https://doi.org/10.2307/j.ctvhrd092.18

[19] Comisión Europea (ed.). *Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza*. Oficina de Publicaciones de la Unión Europea, 2020. Leído en https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf

ría ecológica, amigable con el medio ambiente, agricultura, una reducción de los costes en bienes y servicios, mejor gestión energética y de residuos, sin que lo anterior suponga una disminución de los derechos y libertades de los ciudadanos, derechos sociales y políticos.

La IA ofrece grandes oportunidades pero también grandes retos en los que será preciso la actuación conjunta de los países de la UE, desde los valores europeos y el Estado de Derecho. Nos enfrentamos a nuevos riesgos[20], que afectan, esencialmente a los derechos fundamentales y a la forma en que los responsables de los daños en estos derechos se enfrentan a las opciones de resarcimiento.

1. *Riesgos para los derechos fundamentales*[21]

El uso de la IA puede conculcar derechos fundamentales[22], valores sobre los que se fundamenta la UE, como la dignidad personal, la no discriminación por razón de sexo[23], raza u origen étnico[24], religión o credo, discapacidad, edad u orientación sexual la libertad de expresión, la libertad de reunión, el derecho a la intimidad y la protección de los datos personales, el derecho a una tutela judicial efectiva. Un buen ejemplo de lo anterior se puede producir con actuaciones de vigilancia masiva, sobre los ciudadanos en general, empleados, consumidores, atentando al derecho a la protección de datos personales. O por ejemplo, si, utilizando técnicas de IA se desanonimizan datos personales previamente anonimizados para conculcar derechos.

Las técnicas de IA tienen un riesgo adicional, que los sistemas «aprendan» con el funcionamiento. Esta circunstancia haría infructuosa cualquier control en la fase de diseño. Si a esta complejidad que se produce por la impredecibilidad del comportamiento de los sistemas de IA le sumamos la opacidad de los algoritmos utilizados, los riesgos y su corrección se tornan más difíciles.

---

[20] El RGPD y la Directiva sobre la privacidad y las comunicaciones electrónicas (nuevo Reglamento sobre la privacidad y las comunicaciones electrónicas en fase de negociación) aborda estos riesgos, aunque los sistemas de IA plantean riesgos adicionales.

[21] Comisión Europea (ed.). *Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza*. *Ob cit*,

[22] Según el trabajo de investigación del Consejo de Europa, un gran número de derechos fundamentales podría verse afectado por el uso de la IA (https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5).

[23] Songül Tolan, *et al*. «Why machine learning may lead to unfairness: Evidence from risk assessment for juvenile justice in catalonia». En *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*. (2019) 83-92.

[24] Joy Buolamwini y Timmit Gebru. «Gender shades: Intersectional accuracy disparities in commercial gender classification». En *Conference on fairness, accountability and transparency*. PMLR (2018) 77-91.

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

2. *Riesgos para la seguridad y el funcionamiento del régimen de responsabilidad civil*

Aparecen nuevos riesgos de seguridad para los usuarios de productos y servicios. Por ejemplo, si un vehículo autónomo detecta por error un objeto en medio de la carretera, habrá que determinar quién es el responsable. Determinar la causa no será tarea fácil, puesto que puede tener su origen en el diseño del programa, pero también de la calidad de los datos recibidos que han hecho reaccionar la conducción autónoma sin motivo. Para resolver este asunto deben existir normas claras aplicables que permitan a las personas damnificadas obtener de forma ágil la correspondiente compensación. También proporcionan seguridad a las empresas productoras y comercializadoras de bienes en las que se utilice IA.

En el marco de la Directiva sobre responsabilidad por los daños causados por productos defectuosos[25], es el fabricante el responsable si el producto tiene defectos. Pero en el caso de un sistema que utilice IA, volviendo al ejemplo del vehículo autónomo, no es tan fácil demostrar el nexo causal entre el este y el daño. Tampoco resulta claro si sería aplicable la mencionada Directiva, por ejemplo, si se ha producido un fallo de ciberseguridad en el uso del producto.

Por si fuera poco, el perjudicado no tendrá fácil obtener las pruebas necesarias para demostrar en los tribunales quien es el causante o los causantes y en qué medida deben uno y otros reparar los daños.

El Libro Blanco[26], analiza repercusiones en materia de responsabilidad civil de la IA en el marco jurídico existente.

De acuerdo con él, «la legislación vigente sobre seguridad de los productos ya recoge un concepto amplio de protección de la seguridad». Sin embargo, se propone, asegurar la supervisión humana a lo largo de todo el ciclo de vida de los productos y sistemas de IA, no solo en el diseño, y requerir recurrentes evaluaciones de riesgo, medidas preventivas ante la posible utilización de datos incorrectos en las fases de diseño y usos de los productos que utilicen IA, eliminando la opacidad que pueden ocultar los algoritmos, afinar estas medidas para los programas comercializados separadamente de un producto, cuando afecten a la seguridad, clarifi-

---

[25] Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos. Leído en https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:31985L0374&from=ES

[26] Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica contenido en el LIBRO BLANCO sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza. 20-21.

car para determinar la trazabilidad de que nos lleve al daño producido en el damnificado, esencialmente cuando la presumible responsabilidad civil no provenga de los productores. Incluso se apuesta por incluir normas que alteren la carga de la prueba, en beneficio de los usuarios de sistemas de IA, exigida por las normas de derecho interno en relación a la responsabilidad civil.

No obstante, conviene afirmar, con RAMÓN FERNANDEZ[27] que, de ninguna forma, conviene dotar a los robots de unos cuasi derechos que derivarían en pantallas ficticias utilizadas para limitar esa responsabilidad.

III. **Desde la Ética o desde el Derecho**

1. *Desde la Ética*

La ética puede ser un elemento plausible en su estudio y eficaz en su plasmación para evitar la excesiva proliferación de normas que, en sí mismas, son más rígidas por su mayor precisión y concreción. La ética también puede ser útil mientras se acometen las necesarias reformas normativas de adaptación a las nuevas aportaciones tecnológicas[28].

Y la UE puede liderar propuestas éticas que sean asumidas voluntariamente por las partes implicadas, aunque la tendencia apunta a la incorporación de nuevas normas de obligado cumplimiento[29]. En este sentido, señala COTINO HUESO[30], que «es inteligente la estrategia de la UE de situarse a la vanguardia del mundo aplicando una IA ética y confiable» y que pueda servir como producto exportable. Estos principios éticos, pueden servir de guía de las conductas y procesos implicados. No obstante, tendremos siempre la dificultad que surge de enfrentar una tecnología en constante evolución con una lista de principios éticos seguramente estáticos y ajenos a la novedad de los nuevos contextos[31].

---

[27] Francisca Ramón Fernández. «Robótica, inteligencia artificial y seguridad: ¿Cómo encajar la responsabilidad civil?» *La Ley (Online)*, n.º 9365 (2019) 8.

[28] Jorge Castellanos Claramunt. «La gestión de la información en el paradigma algorítmico: inteligencia artificial y protección de datos». *MÉI: Métodos de Información*, vol. 11, n.º 21 (2020). 76.

[29] En este sentido, las propuestas de Reglamento van en esa línea.

[30] Lorenzo Cotino Hueso. «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*, , vol. 58 (2019) 35.

[31] Piedra Alegría. «Descolonizando la "Ética de la Inteligencia Artificial"». 251.

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

COTINO HUESO[32], desgrana algunos principios éticos que deberían ser los pilares de una IA confiable avalados por Europa. El hecho de que estos principios éticos puedan ser el fruto de una cultura y moral europeas y que pueda ser más o menos ajenos a las culturas de otros órdenes geográficos, no debe preocuparnos en tanto nuestro objetivo es garantizar una utilización eficiente de acuerdo con los criterios de justicia occidentales. Tratar de buscar principios éticos universales no está entre los objetivos del trabajo. No obstante, más adelante tendremos ocasión de desgranar los principios que, imbricados en nuestro entorno jurídico y social, pueden ser considerados.

2. *Desde el Derecho*

Entendemos que es en el Derecho donde se pueden arbitrar los mecanismos que efectivamente se preserven los derechos, desde la seguridad jurídica que proporciona su protección delegada en el poder coercitivo del Estado y amparada en la fuerza del poder judicial. De esta forma también el Parlamento Europeo hace alusión a este asunto al considerar que «los principios éticos solo son eficaces cuando están también asentados en Derecho»[33] y están perfectamente determinados quiénes deben valorar, controlar y garantizar el ajuste a la norma.

Pero, además, hay que advertir que la los conceptos técnicos y los conceptos jurídicos tienen que estar interrelacionados en la norma[34]. El aprendizaje automático en la IA puede generar discriminaciones por asociación y por su capacidad de inferir correlaciones entre datos[35], máxime si se considera la opacidad de sus algoritmos ante los mecanismos jurídicos existentes de control antidiscriminatorias. Y el Derecho debe establecer límites que traten de preservar los derechos que pudieran verse implicados. Su establecimiento supone fijar reglas de conducta estáticas frente a la movilidad intrínseca de las nuevas tecnologías como la IA.

---

[32] Cotino Hueso. «Ética en el diseño para el desarrollo de una inteligencia artificial„…» 36 a 40. Aquellos principios señalados por COTINO son los siguientes: «beneficencia» (hacer el bien), «no maleficencia» (no hacer daño), autonomía o acción humana (respeto por la autodeterminación) y justicia (trato justo y equitativo para todos). A estos podemos añadir el principio de «explicabilidad o transparencia»

[33] Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas(2020/2012(INL)). Introducción. Leído en https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_ES.html

[34] Pilar Rivas Vallejo. «Sesgos de género en el uso de inteligencia artificial para la gestión de las relaciones laborales: análisis desde el derecho antidiscriminatorio». *e-Revista Internacional de la Protección Social*, , vol. 7, n.º 1 (2022) 9-10.

[35] Sandra Wachter.: «Affinity profiling and discrimination by association in online behavioural advertising», Berkeley Technology Law Journal, n.º 35 (2020) 371.

La aportación más importante viene de la mano de la Constitución y desde la base sólida que de la interpretación de la misma conforma la Jurisprudencia del TC. Cualquier programa informático que desarrolle IA tendrá que tener en cuenta la doctrina arraigada del Alto Tribunal, que deberá ser aplicada por los órganos jurisdiccionales ordinarios, en relación, por ejemplo, al artículo 14 de la Constitución Española (CE) (no discriminación) o al artículo 18 CE (intimidad[36] y protección de datos personales[37]).

Respecto a estos últimos, podemos apuntar que, con PIÑAR MAÑAS[38], la mayor parte de las innovaciones tecnológicas que surgen actualmente tienen relación con el tratamiento de datos de carácter personal y en particular, la IA.

También, fuera de la norma constitucional, existen normas que protegen los derechos afectados como extensión o desarrollo de la norma fundamental. El RGPD, norma básica en el tratamiento de datos, contiene mecanismos de gran potencia para controlar el cumplimiento de la norma. Y los datos son la gasolina que mueve buena parte de las aplicaciones de IA. A estos hay que añadir los algoritmos, cuya opacidad y complejidad lo ponen más difícil.

En este sentido, los arts. 15 al 23 del RGPD presentan los derechos reconocidos por el uso de sus datos personales que, en el ámbito de la IA se mueve en actividades como la elaboración de perfiles o la toma de decisiones automatizadas, donde las obligaciones y responsabilidades se muestran más evidentes y también más difíciles de determinar[39].

El derecho a la transparencia y la obligación de lealtad en el tratamiento de datos personales que se espera del responsable a lo largo del tiempo[40], configuran una limitación extraordinaria frente a la IA.

---

[36] SSTC 57/1994, de 28 de febrero y 143/1994, de 9 de mayo. El derecho a la intimidad, por tanto, en la medida que no es un derecho absoluto, puede «ceder ante intereses constitucionalmente relevantes, siempre que el recorte que aquel tenga que experimentar y se revele como necesario para lograr el fin legítimo previsto, proporcionado para alcanzarlo y, en todo caso, sea respetuoso con el contenido esencial del derecho»

[37] STC 11/1981, de 8 de abril (FJ 8). , el contenido constitucional del derecho estará conformado por las facultades atribuidas a los interesados que permiten ejercer un control sobre sus datos personales, sin las cuales se estaría conculcando ese derecho fundamental

[38] José Luis Piñar Mañas *et al*. *Derecho e innovación tecnológica: retos de presente y futuro*. (Madrid: CEU Ediciones, 2018) 14 y 15.

[39] María Alvarez Caro *El derecho de rectificación, cancelación, limitación del tratamiento, oposición y decisiones individuales automatizadas*. Editorial Reus. 2016. En José Luis Piñar Mañas. *Reglamento General de Protección de Datos. Hacia un modelo europeo de privacidad*. (Editorial Reus. 2016). En Artemi Rallo Lombarte. *Tratado de protección de datos. Actualizado con la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos personales y Garantía de los Derechos Digitales*. (Valencia, Tirant lo Blanch. 2019).

[40] Jose Luis Dominguez Álvarez. «Inteligencia Artificial, derecho administrativo y protección de datos personales. Entre la dignidad de la persona y la eficacia administrativa». *Ius et Scientia*. Vol. 7. n.º 1. (2021) 319 https://editorial.us.es/es/revistas/ius-et-scientia https://dx.doi.org/10.12795/IETSCIENTIA.

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

Los responsables que utilicen IA deben minimizar el riesgo para los derechos y libertades fundamentales de los sujetos objeto de tratamiento[41]. Para ello será preciso diseñar políticas de protección de datos, que deben incluir mecanismos de garantía como, en ocasiones, evaluaciones de impacto en protección de datos[42].

El panorama, en la práctica, no es sencillo y, como señala DOMINGUEZ ALVAREZ[43], podemos perder el equilibrio deseado entre el necesario desarrollo tecnológico y el aseguramiento de la protección los derechos fundamentales de la persona.

El Consejo de Europa, por su parte, también ha aprobado declaraciones en esta línea: las Directrices sobre Inteligencia Artificial y Protección de Datos de enero de 2019[44], la Declaración del Comité de Ministros sobre las capacidades manipuladoras de los procesos algorítmicos de febrero de 2019[45] y Unboxing Artificial Intelligence: 10 steps to protect Human Rights de mayo de 2019[46], son una muestra de lo dicho.

En la primera se relaciona la IA con el denominado «Convenio 108+[47] del Consejo de Europa». Este protocolo que entrará en vigor cuando lo ratifiquen todos los miembros firmantes del Convenio[48], pretende reforzar su aplicación en aras a un mayor respeto de los derechos humanos, en especial, al derecho a la protección de datos personales, lo que da a entender la presencia de un mayor riesgo con el uso de la IA que exige una mayor protección de los derechos humanos[49].

---

[41] Miguel Recio Gayo. *Protección de datos e innovación: ¿(in) compatibles?* (Editorial Reus, 2018).

[42] Carlos Alberto Sáiz Peña. Seguridad de los datos, evaluación de impacto, códigos de conducta y certificación. En *Tratado de protección de datos: actualizado con la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales*. Tirant lo Blanch, (2019) 387-430.

[43] Domínguez Álvarez. «Inteligencia Artificial, derecho administrativo y protección de datos personales…». 322.

[44] Disponibles en https://rm.coe.int/guidelines-on-artificial-intelligence-and-dataprotection/168091f9d8

[45] Disponible en https://search.coe.int/cm/pages/result_details.aspx?objectid=090000168092dd4b

[46] Disponible en https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rightsreco/1680946e64

[47] Propuesta de Decisión del Consejo por la que se autoriza a los Estados miembros a ratificar, en interés de la Unión Europea, el Protocolo que modifica el Convenio del Consejo de Europa para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal (STCE n.º 108)

[48] Convenio para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, hecho en Estrasburgo el 28 de enero de 1981.

[49] The protection of human dignity and safeguarding of human rights and fundamental freedoms, in particular the right to the protection of personal data, are essential when

En la segunda, se advierte que las nuevas utilidades tecnológicas permiten predecir las decisiones que tomarán las personas, influir en las emociones y los pensamientos y alterar la acción, a veces de manera subliminal, por lo que serán preciso tener marcos de protección adicionales a los ya existentes.

En la tercera se indican 10 pasos para proteger los derechos fundamentales frente a la IA. En relación al derecho a la protección de datos se apunta garantías suplementarias en relación con datos sensibles como los datos genéticos, datos sobre infracciones administrativas o penales, sus sanciones y condenas, datos biométricos, datos relacionados con el origen racial o étnico, la afinidad política, las creencias religiosas, la orientación sexual, la afiliación sindical.

La Comisión Europea, en su Libro Blanco, muestra las directrices que pueden ser asimiladas a principios y que ya fueron recogidas en la Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones de 08.04.2019[50] como conclusión a las aportaciones del grupo de expertos. En estas directrices podemos incorporar los principios éticos ya referidos. Estas 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 medioambiental;
— rendición de cuentas»

La directriz *«acción y supervisión humanas»* puede asumir el principio ético de autonomía o acción humana. Esta acción y supervisión se traduce en los enfoques de la participación humana (human-in-the-loop), la supervisión humana (human-on-theloop), o el control humano (human-in-command). La citada Comunicación explica esos enfoques como sigue: «Human-in-the-loop (HITL) implica la participación humana en el momento de la decisión del sistema; Human-on-the-loop (HOTL) se relaciona con la participación en el diseño y en su funcionamiento. Human-in-command (HIC) se centra en la supervisión de la actividad global del sistema pudiendo tomar decisiones en las que la intervención humana sus-

developing and adopting AI applications that may have consequences on individuals and society.

[50] Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Bruselas, 8.4.2019. Leído en file: ///C:/Users/Usuario/Downloads/COM(2019)168_0.pdf

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

tituye a la IE. Se pone de manifiesto que la supervisión humana es necesaria en la utilización de sistemas de IA y su objetivo último debe ser mejorar nuestra sociedad sin que eso suponga un perjuicio al ejercicio de los fundamentales[51].

La directriz «*solidez técnica y seguridad*» absorbe el principio de no maleficencia, en el sentido antedicho, de prevenir, aplicando medidas técnicas de seguridad, ante cualquier afectación nociva. En este caso, esa seguridad permitirá rechazar cualquier tentativa de ataque no solo sobre los datos sino, también, sobre los algoritmos y tener habilitados planes para el caso de que se produzca. La robustez y seguridad precisas deben asegurar la utilización de algoritmos fiables y sólidos que se mantengan en el tiempo, a lo largo de la vida de las aplicaciones.

La directriz «*gestión de la privacidad y de los datos*» puede encajar en el principio ético de justicia. Esto se traduce en el pleno ejercicio de los derechos amparados en la normativa sobre protección de datos y que puedan ser utilizados para limitarlos o dañarlos.

La directriz de «trasparencia» es una réplica del principio ético de explicabilidad y transparencia. Con ella, debe garantizarse la trazabilidad de las aplicaciones de IA y el acceso a los algoritmos con suficiente claridad e inteligibilidad.

La directriz «*diversidad, no discriminación y equidad*» se sitúa dentro del principio ético de justicia: los sistemas de IA deben garantizar la no discriminación de las personas y el trato equitativo de las mismas.

La directriz «*bienestar social y medioambiental*, también puede ir unida al principio ético de justicia, al promover la vinculación a los sistemas de IA a la sostenibilidad y la responsabilidad ecológicas.

Por último la «*rendición de cuentas*», que puede relacionarse con el principio ético de no maledicencia y de explicabilidad y transparencia. Así las cosas deben implantarse mecanismos que garanticen la exigencia de responsabilidad civil o penal que vean la luz como consecuencia de esa rendición de cuentas.

IV. **El riesgo como elemento director**

El elemento vehicular de la regulación propuesta por la UE[52] es el análisis del riesgo en el presente y orientado al futuro. La utilización del riesgo

---

[51] Esteban, Luis Miguel Pedrero Esteban y Ana Pérez Escoda. «Democracia y digitalización: implicaciones éticas de la IA en la personalización de contenidos a través de interfaces de voz». *Recerca. Revista de pensament i anàlisi*, , vol. 26, n.º 2 (2021). https://doi.org/gsd8

[52] Resolución del Parlamento Europeo de 20 de octubre de 2020.

como punto de partida ya presente en otras regulaciones europeas[53]; entre otros, en el Reglamento General de Protección de Datos (RGPD[54]). En el artículo 35 de esta norma se establece una evaluación de impacto de protección de datos en función del riesgo del tratamiento[55] —o en el Reglamento de Productos Sanitarios[56] (artículo 51 y anexos de dicha norma).

Desde esta perspectiva, se parte de una lista bastante completa de sectores reconocidos como «de alto riesgo» que debe reevaluarse periódicamente. Este riesgo se perfila en la medida que pueda causar lesiones en los derechos fundamentales de las personas o en las normas referidas a cuestiones de prevención y seguridad establecidas en el Derecho europeo.

El riesgo debe evaluarse ex ante, de forma imparcial, determinada por criterios previamente definidos en su desarrollo, despliegue y uso. Debe calibrarse según sea el sector en el que las tecnologías se implementan, según los fines perseguidos y la gravedad del ataque al derecho o interés protegido que pueda producirse. Este último se determinará teniendo en cuenta «la magnitud de la lesión o daño potencial, el número de personas afectadas, el valor total del perjuicio ocasionado y el daño a la sociedad en su conjunto» [57]. Así se crea una lista de sectores de alto riesgo y de usos o fines de alto riesgo como recoge el anexo[58]. Para ello la Resolución, en la lí-

---

[53] Guillermo Lazcoz Moratinos. «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*, vol. 6, n.º 2 (2020). 29.

[54] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas

[55] Katerina. Demetzou. «Data Protection Impact Assessment: A tool for accountability and the unclarified concept of 'high risk'in the General Data Protection Regulation». *Computer Law & Security Review*, 2019, vol. 35, n.º 6, p. 105342. https://doi.org/10.1016/j.clsr.2019.105342

[56] Reglamento (UE) 2017/745 del Parlamento Europeo y del Consejo, de 5 de abril de 2017, sobre los productos sanitarios, por el que se modifican la Directiva 2001/83/CE, el Reglamento (CE) n.° 178/2002 y el Reglamento (CE) n.° 1223/2009 y por el que se derogan las Directivas 90/385/CEE y 93/42/CEE del Consejo

[57] Resolución del Parlamento Europeo de 20 de octubre de 2020. Considerando 11. *Ob. cit.* Leído en https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020IP0275&from=EL

[58] La referida Resolución de 20 de octubre de 2020, en su anexo, bajo el título de «Lista exhaustiva y acumulativa de sectores de alto riesgo y de usos o fines de alto riesgo que conllevan un riesgo de violación de los derechos fundamentales y las normas de seguridad» determina que son sectores de alto riesgo: el empleo, la educación, la asistencia sanitaria, el transporte, la energía, el sector público (asilo, migración, controles fronterizos, sistema judicial y servicios de seguridad social), la seguridad y defensa y las finanzas, bancos, seguros y como usos o fines de alto riesgo, la contratación, la clasificación y evaluación de estudiantes, la asignación de fondos públicos, la concesión de préstamos, el comercio, corretaje, fiscalidad, los tratamientos y procedimientos médicos, los procesos electorales y campañas políticas, las decisiones del sector público que tienen un impacto significativo y directo en los derechos y las obligaciones de las personas físicas o jurídicas, la conducción automatizada, la

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

nea indicada, recomienda la utilización de los instrumentos de evaluación ya existentes como la evaluación de impacto recogida en el Reglamento (UE) 2016/679.

Estos aspectos, como indica MORATINOS LAZCOZ[59] deben estar nítidamente perfilados en el futuro Reglamento y en las normas internas de desarrollo a nivel de Ley, para garantizar la seguridad jurídica necesaria y suficiente.

Por su parte, la propuesta de Reglamento de IA[60] perfila aún más los sectores de alto riesgo. En su artículo 6 diferencia los sistemas de IA de alto riesgo:

— Si reúnen las dos condiciones siguientes[61]:

- que sirva «como componente de seguridad de uno de los productos contemplados en la legislación de armonización» de la Unión o es en sí mismo uno de dichos productos;
- que, de acuerdo con las normas armonizadoras incorporadas en el anexo II del texto, «el producto del que el sistema de IA es componente de seguridad[62], o el propio sistema de IA como producto, debe someterse a una evaluación de la conformidad» determinada.

— Si figuran en el anexo III de la propuesta.

Estos sectores de alto riesgo deberán cumplir una serie de obligaciones incluidas en el Capítulo 2 de la propuesta, entre ellas:

— Establecerán, implantarán, documentarán y mantendrán un sistema de gestión de riesgos.
— Velarán para que los datos utilizados cumplan los criterios de calidad requeridos

---

gestión del tráfico, los sistemas militares autónomos, la producción y distribución de energía, la gestión de residuos y el control de emisiones. Resulta anecdótica la incorporación en la redacción de una clausula abierta al indicar en Usos o fines de alto riesgo al apartado a el «Comercio, corretaje, fiscalidad, etc». Esta indicación «etc» además de ser difícilmente descifrable, no debe incluirse en una lista que debe ser cerrada en aras de la seguridad jurídica que debe presidir este tipo de normas restrictivas.

[59] Para Moratinos Lazcoz, en 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. *ob cit*., p. 40, resultaría conveniente que el tercer criterio («la gravedad de los posibles daños o lesiones causados») establecido por el artículo 14, tuviera la concreción mostrada en el Considerando 11.

[60] Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadoras en materia de IA (Ley de IA) y se modifican determinados actos legislativos de la Unión, de 24 de abril de 2021. Leído en file:///C:/Users/Usuario/Documents/Fran/articulos%20investigaci%C3%B3n/verano%202022/ia%20bibliograf%C3%ADa/textos%20ue/propuesta%20ceur%20alto%20riesgo.pdf

[61] que indica en el anexo II de la Propuesta

[62] Conforme a la legislación de armonización de la Unión que se indica en el anexo II

— Se preparará la documentación técnica antes de que llegue al mercado o sea utilizada y tendrá que ser actualizada en todo momento.
— Precisarán archivos de registro mientras están en funcionamiento que permitan reconocer la trazabilidad del funcionamiento de la aplicación de IA durante todo el periodo de aplicación.
— Garantizarán un nivel de transparencia que permita que los usuarios puedan conocer qué realizan realmente los sistemas de IA.
— Podrán ser vigilados por personas y no por máquinas mientras sean utilizados.
— Alcanzarán un nivel suficiente de seguridad durante todo su ciclo de vida.

V. **Sobre los principios éticos en relación a la inteligencia artificial**

Los principios éticos comunes deben estar asentados en Derecho, no basta un mero pronunciamiento doctrinal. Esta es una cuestión fundamental que presupone un acercamiento del concepto de principio ético al de principio general del Derecho.

La norma, para que tengan virtualidad y efectividad, debe establecer quiénes son los las personas que deben evaluar, inspeccionar o garantizar la conformidad con esos principios éticos durante toda la vida útil de las aplicaciones, incluidos algoritmos y datos utilizados o producidos y el aseguramiento de la responsabilidad si se producen perjuicios. En todo caso, debe reconocerse que estos principios ya están imbricados, con enfoques diferentes, en los textos constitucionales y otros de gran fuerza normativa en infinidad de normas de Derecho europeo y nacional, por lo que tendríamos oportunidad de hacerlos valer, en muchas ocasiones, si fuera preciso.

Estos principios, de acuerdo con lo ya descrito en los siguientes:

— *El principio de respeto de la dignidad humana, la autonomía y la autodeterminación de la persona* que evite que su uso conlleve un sometimiento directo o indirecto, que conculque la autonomía psicológica desde la manipulación, el engaño o la vigilancia injustificada. Tiene que ver con saber en todo momento cuándo se interactúa con una máquina y cuándo con un ser humano. También sobre la necesidad de reservar tareas exclusivamente para humanos[63].

[63] Grupo Europeo sobre Ética de la Ciencia y las Nuevas Tecnologías. Declaración sobre Inteligencia artificial, robótica y sistemas «autónomos». Marzo (2018). 15. http://www.bioeticayderecho.ub.edu/archivos/pdf/EGE_inteligencia-artificial.pdf

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

— *El principio de respeto a la diversidad, a la no discriminación a la equidad y el fortalecimiento del bienestar social y medioambiental***.** Con la dignidad humana viajan un conjunto de derechos adquiridos desde los fundamentales dentro de una sociedad de bienestar. Deben garantizarse la no discriminación[64] y la igualdad de género, la igualdad de oportunidades, las mejoras de la salud y la educación, la protección de la infancia, los derechos de los trabajadores, la alfabetización digital, la protección de los marginados o en situación de vulnerabilidad[65], como las personas con discapacidad, sean tenidos en cuenta y estén representados debidamente. Incluye eliminar sesgos, estigmatización y cualquier aspecto que mantenga discriminaciones o cree otras nuevas[66]. Se pretende alcanzar la igualdad de acceso a la tecnología y que esta suponga una efectiva distribución equitativa de beneficios y oportunidades.

La defensa del medio ambiente nos lleva al impulso de la sostenibilidad. El desarrollo, el despliegue y el uso de estas tecnologías deben contribuir a la «transición verde», que intente minimizar los daños que puedan sufrir los ecosistemas, por ejemplo, como consecuencia de la extracción de materiales o las emisiones de gases nocivos para el medio ambiente en la generación y consumo de energía.

— *El principio de intervención humana, del control democrático y de la recuperación del control humano*. Ante los riesgos posibles del uso de la IA, Se precisa salvaguardar los pilares de nuestra sociedad democrática basada en el Estado de Derecho, la independencia de los medios de comunicación y la garantía de acceso libre a una información objetiva y que preserve una mayor cohesión social. Supondrá la eliminación de los sesgos[67] que, utilizados de forma automatizada o no, creen formas de discriminación, en particular cuando nos referimos

---

[64] *Montreal Declaration for a Responsible Development of Artificial Intelligence*. 2017. [Pronunciada en la clausura del Forum on the Socially Responsible Development of AI]. https://www.montrealdeclaration-responsibleai.com/_files/ugd/ebc3a3_506ea08298cd4f8196635545a16b071d.pdf

[65] Celia Rangel. «Inteligencia Artificial como aliada en la supervisión de contenidos comerciales perjudiciales para menores en Internet». *Revista Mediterránea de Comunicación: Mediterranean Journal of Communication*, vol. 13, n.º 1 (2022). 26.

[66] Asilomar *AI Principles*. Future of Life Institute, 2017. Leído en https://futureoflife.org/2017/08/11/ai-principles/

[67] Moratinos, Guillermo Lazcoz. 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. Ob cit, propone la siguiente definición de sesgo: en Mireille Hildebrandt. The issue of bias: the framing powers of ML. *Machine Learning and Society: Impact, Trust, Transparency, MIT Press forthcoming*, (2020) propone la siguiente definición: l) «sesgo», toda

a grupos de personas homogéneos o con características similares. No olvidemos que una IA desarrollada puede llegar a crear y reforzar los sesgos en la línea equivocada. El control democrático se establece con la posibilidad de legislar, estableciendo normas que encaucen el uso y desarrollo de estas tecnologías. También deben ser el fruto de la participación de la Sociedad también a través de los cauces de participación democrática La recuperación del control humano tiene verdadera notoriedad cuando nos referimos a la IA. No basta el control a priori, sino que es preciso un control humano permanente en la medida en que la IA es capaz de «aprender» con el uso, cambiando durante el mismo sus propios algoritmos. Debe respetar la capacidad del ser humano de decidir cuándo y cómo tomar decisiones de control y para ello es fundamental que los sistemas de IA sean previsibles y transparentes para intervenirlos cuando así se estime.

— *El principio de fiabilidad, solidez y precisión*. Solidez para evitar las vulneraciones de la seguridad, los ciberataques y los usos indebidos de los datos personales. Precisión que socave las posibles fugas de datos, las noticias falsas y la desinformación. Ambas respetuosas con la privacidad de los datos, dentro del acerbo jurídico de la Unión en esta materia. Fiabilidad, basada en la calidad de los datos empleados, sin sesgos que dañen los datos acumulados y con unos algoritmos limpios, sin estrategias ocultas.

— *El principio de transparencia y explicabilidad de las tecnologías*. La IA se sirve de los algoritmos que son definidos como secuencia(s) finita(s) de operaciones e instrucciones lógicas que hacen posible obtener un resultado a partir de la entrada de información. Necesitamos, pues, un código fuente o secuencia de instrucciones y un conjunto de datos o «librerías», que puedan ser explicables[68], que muestren la motivación de la estrategia que conduce a los resultados. Debe garantizarse el derecho a la información a los interesados, suministrada de manera comprensible, accesible cuando sea solicitada, normalizada, completa. Debe contener el razonamiento utilizado, los resultados o repercusiones posibles. Debe mostrar las formas de verificación, impugnación y, en su caso, corrección y revisión del sistema. Para RIVAS VALLEJO[69], la positivización

desviación o representación estadística que reconfigura la distribución de los bienes, servicios, riesgos y oportunidades de una o varias personas físicas o jurídicas.

[68] Rivas Vallejo. «Sesgos de género en el uso de inteligencia artificial para la gestión de las relaciones laborales…». 83.

[69] Rivas Vallejo. «Sesgos de género en el uso de inteligencia artificial para la gestión de las relaciones laborales…». 30.

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

explícita de los distintos derechos que puedan verse heridos por la IA deben ser objeto de tutela con la explicabilidad y el acceso al algoritmo oculto. Para ello será preciso mitigar la fuerza de la propiedad intelectual para evitar que, amparándose en ella, pueda encubrirse decisiones opacas malintencionadas.

Se podría hablar de la transparencia algorítmica[70] y su falta resultará ser el enemigo a batir si queremos asegurar que los usos del big data y la IA se quedan del lado del Derecho.

Para COLCELLI y BURZAGLI, un sistema de IA será comprensible, predecible y controlable ex post, si también es capaz de garantizar el principio de transparencia y el Estado de derecho ex ante, es decir, en el momento del diseño algorítmico[71]. Deben conocerse, también, las actividades de mantenimiento de los sistemas, cómo se reparan o mejoran, por ejemplo, en particular, mediante las actualizaciones para corregir sus vulnerabilidades.

— *El principio de privacidad y de protección de datos*. En la medida que la IA se nutre de datos, muchos de ellos personales, son aplicables todos los principios ya establecidos para el tratamiento de los datos personales recogidos fundamentalmente en el artículo 5 del RGPD. Estos son los principios de licitud, lealtad y transparencia, de limitación de la finalidad, de minimización de datos, de exactitud, de limitación del plazo de conservación, de integridad y de confidencialidad. Ni los robots físicos, ni aquellos que operan a través las redes deben recopilar, tratar o transferir datos sin consentimiento. Más aún, no deben influir el desarrollo personal o en las relaciones interpersonales, libres de vigilancia. Aparecen dos nuevos derechos[72]: el derecho al contacto humano significativo y el derecho a no ser perfilado, medido, analizado, aconsejado o provocado.

— *El principio de rendición de cuentas*. Se debe garantizar a las personas, como un derecho, las vías de recurso o revisión imparciales, accesibles, asequibles y efectivas, efectuada por personas físicas, tanto en el sector público como en el privado. Deben desarrollarse soluciones sólidas que

[70] Resolución del Parlamento Europeo de 14 de marzo de 2017, sobre las implicaciones de los macrodatos en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley (2016/2225(INI)).

[71] Valentina Colcelli y Laura Burzagli. «Elementos para una cultura europea de desarrollo de herramientas de inteligencia artificial: el libro blanco sobre la inteligencia artificial y las directrices éticas para una IA fiable». Revista Justicia & Derecho, vol. 4, n.º 2 (2021) 4.

[72] Grupo Europeo sobre Ética de la Ciencia y las Nuevas Tecnologías. Declaración sobre Inteligencia artificial, robótica y sistemas «autónomos». Marzo 2018. Página 17. http://www.bioeticayderecho.ub.edu/archivos/pdf/EGE_inteligencia-artificial.pdf

asignen responsabilidades de forma clara y justa desde una legislación vinculante eficiente. Esta redención de cuentas no debe ser solo formal, sino material, que culmine, en su caso, con la reparación del daño. Tiene que ver con la necesidad de entender cómo funciona la IA y determinar quién es el responsable de ese funcionamiento»[73]. Podríamos hablar de sistemas inteligentes supervisados por humanos o por otros sistemas de IA que permitan controlar a la propia IA.

— *El principio de prevención de daños y de responsabilidad***.** La necesidad de prevenir daños debe impulsar y desarrollar procesos de verificación, validación y control de los sistemas de IA que atribuyan certificados o etiquetados de buenas prácticas[74]. Igualmente la elaboración de códigos deontológico o de buena conducta a los que se deban o puedan adherir las partes implicadas y que pongan en valor la protección de los principios de dignidad humana, integridad, libertad, privacidad, no discriminación y, en general, los derechos humanos fundamentales. Pero si la prevención no ha producido sus efectos, se derivará la responsabilidad a los causantes que deberán, en su caso, asumir las consecuencias resarcitorias e indemnizatorias. Esa responsabilidad también incorpora el deber ser diseñados de tal forma que su diseño sea respetuoso con los valores y los derechos fundamentales Participo con RAMÓN FERNANDEZ[75] que no es acertado atribuir la responsabilidad civil a los robots, ni dotarlos de semiderechos que servirían para crear pantallas cuyo objetivo evidente sería sortear la responsabilidad civil. Detrás de un robot, un programa informático, un producto, está la mente creadora del ser humano.

— *El principio de proporcionalidad y justificación de la necesidad de limitación de derechos*. En la utilización de medidas que sean restrictivas de derechos por causa debidamente justificada en aras del interés público, deberán mitigarse, por ejemplo, los efectos adversos sobre el derecho a la intimidad, a la protección de datos y a la no discriminación, como la utilización de tecnologías de reconocimiento facial que debe ser restringido en el tiempo, sujetos a control judicial que eviten abusos como la vigilancia masiva en actuaciones

---

[73] Andrés Boix Palop y Lorenzo Contino Hueso (coords.). «Monográfico Derecho Público, derechos y transparencia ante el uso de algoritmos, inteligencia artificial y big data». *Revista General de Derecho Administrativo*». (2019).

[74] Dictamen del Comité Económico y Social Europeo (CESE) sobre la «Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad» (Dictamen de iniciativa) (2017/C 288/01). Leído en https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=uriserv%3AOJ.C_.2017.288.01.0001.01.SPA&toc=OJ%3AC%3A2017%3A288%3ATOC

[75] Ramón Fernández. «Robótica, inteligencia artificial y seguridad…». 8.

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

policiales predictivas. La IA puede necesitarse, por razones de interés público, para limitar algunos derechos, por ejemplo, en el curso de investigaciones que tengan su causa en la comisión de delitos.

Todos estos principios éticos deben ser normativizados para una mejor aplicación que prevea sanciones para el caso de su incumplimiento. En este sentido, MORATINOS LAZCOZ[76], expresa que es necesario un régimen sancionador si queremos un efectivo cumplimiento de los principios éticos recogidos por la propuesta de Reglamento, referida en el trabajo, y que resulta imprescindible su existencia en sectores de alto riesgo.

VI. **Sobre la necesidad de un nuevo Reglamento**

La posibilidad de que algunos Estados miembros estén trabajando sobre iniciativas legislativas nacionales para resolver los problemas que conlleva la IA, nos lleva directamente a la posibilidad de que se produzca una fragmentación normativa. Esta fragmentación, ineludiblemente, conlleva problemas a las empresas que utilicen aplicaciones de IA utilizables en los distintos países de la UE. Para garantizar la competividad de las mismas será preciso un marco normativo común. Por ello, la Comisión pretende actualizar la legislación vigente y abordar un nuevo marco jurídico de la UE adaptado a la evolución tecnológica presente y futura.

La Comisión, en su Libro Blanco[77], detalla aspectos que hay que mejorar a nivel normativo, destacamos dos:

— el problema de la opacidad o falta de trasparencia de la IA que dificulta detectar y demostrar los incumplimientos normativos y que conllevan lesiones en los derechos fundamentales. Es preciso determinar quiénes son los verdaderos responsables de esas acciones, en la producción, en el suministro o en la aplicación efectiva.
— la necesidad de ampliar la exigencia de seguridad de los productos, programas o servicios que no tengan normas específicas[78]. Es aún más necesario en el caso de las aplicaciones que requieren actualizaciones informáticas continuadas o que sean capaces de un aprendizaje autónomo.

---

[76] Lazcoz Moratinos. «Análisis de la propuesta de reglamento sobre los principios éticos…» 40.

[77] Europea, Comisión (ed.). *Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza*. Oficina de Publicaciones de la Unión Europea, 2020.

[78] Por ejemplo, de acuerdo con el Reglamento sobre los productos sanitarios [Reglamento (UE) 2017/745], los programas informáticos destinados a fines médicos por el fabricante se consideran productos sanitarios

Además, existe ya un conjunto de normas de la legislación en vigor en la UE en materia de seguridad de los productos y responsabilidad civil[79], especialmente en las normas sectoriales que, completadas también por la normativa de cada país, es posible aplicarla a los sistemas de IA.

En lo que se refiere a la protección de los derechos fundamentales la UE contiene legislación relacionada y aplicable: la Directiva sobre igualdad racial[80], la Directiva sobre igualdad de trato en el empleo y la ocupación Directiva 2000/78/CE[81], las Directivas relativas a la igualdad de trato entre mujeres y hombres con relación al empleo y el acceso a los bienes y servicios[82], normas de protección de los consumidores[83] y normas sobre la protección de los datos personales y la privacidad, especialmente el Reglamento General de Protección de Datos y otras en este campo como la Directiva sobre protección de datos en el ámbito penal[84]. Además, a partir de 2025, también las normas sobre los requisitos de accesibilidad de bienes y servicios establecidas en el Acta Europea de Accesibilidad[85].

En este sentido RIVAS VALLEJO[86], afirma que el derecho de la UE (y, de la misma forma el español), «no están preparados para dar respuesta a la discriminación algorítmica», en la medida en la que las decisiones automatizadas escapan del marco legal, tanto en el caso de discriminación múltiple como en el de discriminación interseccional[87].

---

[79] El marco jurídico de la UE sobre la seguridad de los productos lo componen la Directiva sobre seguridad general de los productos (Directiva 2001/95/CE), y varias normas sectoriales en razón de las categorías de productos, como las máquinas, los aviones, los vehículos o los juguetes y los productos sanitarios.

[80] Directiva 2000/43/CE del Consejo, de 29 de junio de 2000, relativa a la aplicación del principio de igualdad de trato de las personas independientemente de su origen racial o étnico.

[81] Directiva 2000/78/CE del Consejo, de 27 de noviembre de 2000, relativa al establecimiento de un marco general para la igualdad de trato en el empleo y la ocupación.

[82] Directiva del Consejo 2004/113/CE, de 13 de diciembre de 2004, por la que se aplica el principio de igualdad de trato entre hombres y mujeres al acceso a bienes y servicios y su suministro; Directiva 2006/54/CE del Parlamento Europeo y del Consejo, de 5 de julio de 2006, relativa a la aplicación del principio de igualdad de oportunidades e igualdad de trato entre hombres y mujeres en asuntos de empleo y ocupación (refundición).

[83] Directiva sobre las prácticas comerciales desleales (Directiva 2005/29/CE) y la Directiva sobre los derechos de los consumidores (Directiva 2011/83/CE).

[84] Directiva (UE) 2016/680 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos

[85] Directiva (UE) 2019/882, sobre los requisitos de accesibilidad de los productos y servicios.

[86] Rivas Vallejo *et al*. Sesgos de género en el uso de inteligencia artificial para la gestión de las relaciones laborales…» 11

[87] La definición legal se encuentra en la Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación. artículo 3. Discriminación múltiple e interseccional.

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

En línea con lo anterior, encaja la ya referida Propuesta de Reglamento del Parlamento Europeo y del Consejo, o «Ley de inteligencia artificial» de 21 de abril de 2021, cuyos objetivos[88] quedan reflejados en la exposición de motivos y que muestran el camino a seguir en el ámbito normativo de la UE en relación a la IA y que serían, en resumen los siguientes:

— garantizar que los sistemas de IA sean seguros y respeten la normativa europea en materia de derechos fundamentales.
— garantizar la seguridad jurídica que amplifique la inversión e innovación en IA;
— evitar la fragmentación del mercado que permitan un uso seguro y fiable de las aplicaciones de IA.

En definitiva, muestran el objetivo último de conseguir, en el ámbito de la UE, una IA segura, confiable y ética[89].

VII. **Conclusiones**

Los sistemas de IA constituyen un conjunto de herramientas tecnológicas de futuro, que afectan y van a afectar a la mayor parte de los sectores económicos y sociales (si no a todos) de nuestro entorno. Esta circunstancia obliga, en el ámbito de la UE, a regular, con el instrumento de mayor peso normativo a nivel legislativo (aparte los propios Tratados) que es el Reglamento, los requisitos mínimos de uso que deben exigirse a los productos y aplicaciones desarrolladas con IA. La fuerza expansiva, las oportunidades y los riesgos de la IA son de tal magnitud que dejarlo al arbitrio de las legislaciones nacionales supondría una ineludible fragmentación, de mayores consecuencias que propició la aprobación del RGPD.

Y, además de lo dicho, hay que constatar que el uso indebido de la IA puede afectar a una buena muestra de derechos fundamentales, que sin ánimo de repetir lo dicho, pasan, entre otros, por el derecho a la dignidad personal, de no discriminación, a la intimidad, de protección de datos, de información veraz o de tutela judicial efectiva.

---

a) Se produce discriminación múltiple cuando una persona es discriminada de manera simultánea o consecutiva por dos o más causas de las previstas en esta ley. b) Se produce discriminación interseccional cuando concurren o interactúan diversas causas de las previstas en esta ley, generando una forma específica de discriminación.

[88] Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadoras en materia de IA (Ley de IA) y se modifican determinados actos legislativos de la Unión, de 24 de abril de 2021. *Ob. cit*. p. 2

[89] Torres Jarrín. «La UE & la gobernanza ética de la inteligencia artificial…» 231

Esta necesidad de una norma fuerte exige también la necesaria incorporación de principios éticos que, una vez insertados en la norma de obligado cumplimiento pasarían a ser verdaderos principios generales del Derecho. Estos principios han sido expuestos de una forma sucinta en este trabajo que se concretan en el principio de respeto de la dignidad humana, la autonomía y la autodeterminación de la persona, el principio de respeto a la diversidad, a la no discriminación a la equidad y el fortalecimiento del bienestar social y medioambiental, el principio de intervención humana, del control democrático y de la recuperación del control humano, el principio de fiabilidad, solidez y precisión, el principio de transparencia y explicabilidad de las tecnologías, el principio de privacidad y protección de datos, el principio de rendición de cuentas, el principio de el principio de prevención de daños y de responsabilidad y el principio de proporcionalidad y justificación de la necesidad de limitación de derechos.

La tecnología que da vida a la IA tiene dos características que la hacen especialmente compleja a los ojos del Derecho. Primero, que sus algoritmos, esto es, el conjunto de pautas, operaciones e instrucciones lógicas, que les permite desarrollarse pueden permanecer ocultos y para descubrir su entramado es precisa la colaboración de expertos y sortear los derechos de propiedad intelectual que quedarían al descubierto. Estas circunstancias deben sopesarse para dar una respuesta jurídica adecuada. El práctico del Derecho, privado o de servicio público, precisará, por tanto, la ayuda de verdaderos expertos en esta tecnología.

En segundo lugar, algo significativamente novedoso, y es que la propia IA puede, por sí misma, modificar sus algoritmos previos por otros después del «aprendizaje» que puede producirse con su uso. Esta «autonomía» de la IA, impide validar las aplicaciones solamente a priori y, por ello, exigen un control permanente de su desarrollo. No bastaría una certificación inicial, sino que esta siempre sería provisional.

Por último, debemos resaltar que los principios éticos indicados que ya están contenidos en los documentos referenciados de las instituciones de la UE, y que tienen como objetivo trasladarse a una norma de obligado cumplimiento, gozan de virtualidad y valor suficiente para conformar códigos de conducta asumidos por los desarrolladores de IA.

También es cierto que estos principios ya están imbricados en los textos constitucionales y otros de gran fuerza normativa en infinidad de normas de Derecho europeo y nacional, por lo que tendríamos oportunidad de hacerlos valer, en muchas ocasiones.

Cuadernos Europeos de Deusto
ISSN: 1130-8354 • ISSN-e: 2445-3587, No. 68/2023, Bilbao, págs. 89-115
doi: https://doi.org/10.18543/ced.2699 • http://ced.revistas.deusto.es

Sobre el autor

Francisco Javier Martín Jiménez es Licenciado en Derecho por la USAL. Doctor en Derecho desde el año 1999. Licenciado en Ciencias Empresariales por la UNED. Graduado en Relaciones Laborales y Recursos Humanos por la Universidad de Murcia. Técnico de Hacienda en la AEAT. Profesor Asociado de Derecho Tributario en USAL desde el año 1995 hasta el 2016 y Profesor Asociado de Derecho de Empresa y Sistema Fiscal en la Facultad de Informática de la Universidad Pontificia de Salamanca desde 2016 hasta la actualidad. Especialista Universitario en Auditoría Financiera y Máster en Acceso a la Abogacía. Director del Curso de formación AEPD-DPD. Autor de diversas publicaciones (monografía y artículos en revistas especializadas) esencialmente en el ámbito del Derecho Tributario.

About the author

Francisco Javier Martín Jiménez holds a degree in Law from the USAL. Doctor in Law since 1999. He holds a degree in Business Sciences from UNED and a degree in Labor Relations from the University of Murcia. Finance Technician at the AEAT. Associate Professor of Tax Law at the USAL from 1995 to 2016 and Associate Professor of Business Law and Fiscal System at the Faculty of Informatics of the Pontifical University of Salamanca from 2016 to the present. University Specialist in Financial Audit and Master's Degree in Access to the Legal Profession. Director of the course for certification in the AEPD-DPD. Author of various publications (monograph and articles in specialized magazines) essentially in the field of Tax Law.