

LA NECESARIA PROTECCIÓN DE LOS DERECHOS FUNDAMENTALES COMO PUNTO DE PARTIDA EN LAS PROPUESTAS DE REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL

THE NECESSARY PROTECTION OF FUNDAMENTAL RIGHTS AS A STARTING POINT IN PROPOSALS FOR THE REGULATION OF ARTIFICIAL INTELLIGENCE

Susana Álvarez González¹

¹Universidade de Vigo, Vigo, Espanha.
E-mail: sualvago@uvigo.gal

Resumen: La Inteligencia Artificial (IA) es un concepto complejo que abarca diferentes tipos de soluciones, siendo el aprendizaje automático uno de los más destacados. Aunque la definición de IA como la creación de máquinas inteligentes está arraigada, actualmente se debate la dificultad de determinar qué constituye un comportamiento inteligente. La Comisión Europea define la IA como la combinación de tecnologías que utiliza datos, algoritmos y capacidad informática para manifestar comportamiento inteligente. Existen diferentes categorías de IA, como la fuerte, general y débil, con aplicaciones diversas. El desarrollo de la IA está transformando la investigación tecnológica y plantea interrogantes sobre los derechos individuales. Aunque estos avances técnicos ofrecen beneficios, también pueden representar amenazas para las personas y sus derechos, lo que podría requerir nuevas protecciones y derechos. Existe una correlación entre los avances técnicos y los cambios jurídicos cuando tienen un amplio impacto social.

Palabras clave: Inteligencia Artificial, definición, controversia, derechos individuales, avances técnicos, protección.

Abstract: Artificial Intelligence (AI) is a complex concept that encompasses different types of solutions, machine learning being one of the most prominent. Although the definition of AI as the creation of intelligent machines is well established, the difficulty of determining what constitutes intelligent behavior is currently debated. The European Commission defines AI as the combination of technologies that uses data, algorithms and computing power to manifest intelligent behaviour. There are different categories of AI, such as strong, general, and weak, with diverse applications. The development of AI is transforming technological research

DOI: <http://dx.doi.org/10.20912/rdc.v18i44.1357>

Autora convidada



Esta obra está licenciada com uma Licença Creative Commons
Atribuição-NãoComercial-SemDerivações 4.0 Internacional.

and raising questions about individual rights. While these technical advances offer benefits, they can also pose threats to individuals and their rights, which may require new rights and protections. There is a correlation between technical advances and legal changes when they have a broad social impact.

Keywords: Artificial Intelligence, definition, controversy, individual rights, technical advances, protection.

1 Nuevos retos, ¿nuevos derechos?

La aproximación al concepto de inteligencia artificial (IA) resulta compleja pues abarca muchos tipos de soluciones que emplean distintas técnicas, destacando entre ellas las basadas en el aprendizaje automático. Conocida es la precursora definición de McCarthy que hace referencia a la ciencia e ingenio de hacer máquinas inteligentes (2007). En estos términos, la Real Academia de la Lengua Española define la inteligencia artificial como la “disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico”¹. Sin embargo, en la actualidad, dicha definición no está exenta de polémica, especialmente por la dificultad de determinar lo que se considera conducta inteligente. Estas definiciones, según Pinto Fontanillo (2020, p. 17,) presentan el problema de asimilar la IA al “modo de operar de la mente humana”, lo que puede dar lugar a determinadas teorías conspiratorias. Así, autores con Kaplan, señalan que “hay pocas razones para creer que la inteligencia maquina tenga demasiada relación con la inteligencia humana” (2017, p. 1).

La Comisión Europea la define en sus textos como la combinación de tecnologías que agrupa datos, algoritmos y capacidad informática. En su Comunicación de 25 de abril de 2018, sobre *Inteligencia Artificial para Europa* señala que el término se aplica a los sistemas que manifiestan un comportamiento inteligente, pues son “capaces de analizar su entorno y pasar a la acción –con cierto grado de autonomía– con el fin de alcanzar objetivos específicos”, destacando que los sistemas basados en la IA pueden consistir simplemente en un programa informático, pero también ésta puede estar incorporada en dispositivos de hardware. El alcance y el ámbito de aplicación de la IA puede ser muy diverso por lo que se suelen diferenciar tres categorías: las inteligencias artificiales fuertes, generales y débiles. La Agencia Española de Protección de Datos, en su documento sobre *Adecuación al Reglamento General de Protección de Datos de tratamientos de Inteligencia Artificial* de febrero de 2020, señala: “la IA general podría resolver cualquier tarea intelectual resoluble por un ser humano; la IA fuerte o superinteligencia iría más allá de las capacidades humanas. Pero el tipo de IA que ha disparado la aplicación práctica de esta disciplina es la que se conoce como IA-débil (AIweak) y que, en contraste con la IA fuerte y general, se caracteriza por desarrollar soluciones capaces de resolver un problema concreto y acotado”. Sobre el desarrollo de la IA general y fuerte existen grandes expectativas que resurgen tras ciertos periodos de hibernación de la disciplina por falta de avances y que dan lugar a un cierto debate social, político y jurídico sobre su regulación. Así, dicha clasificación no está exenta

1 <https://dle.rae.es/inteligencia>.

de polémica generando literatura científica dividida entre partidarios y detractores de la misma (Churnin, 2012).

Al margen de discusiones doctrinales en torno a su definición, no cabe duda de que el desarrollo de la IA está marcando en los últimos tiempos la investigación tecnológica, de forma muy similar a lo que sucedió años antes con el Proyecto Genoma Humano, si bien con parámetros y consecuencias distintas. Al igual que sucedía entonces, este avance técnico produce una paradoja en relación con los derechos, pues se trata de proyectos fruto de la libertad, investigación y producción científica que abren el camino a múltiples beneficios pero que pueden constituir una amenaza para la persona y sus derechos. Cabe esperar, por tanto, que genere, como mínimo, “la necesidad de nuevas protecciones y de nuevos derechos” (Peces-Barba Martínez, 1994, p.202), pues suele existir correlación entre “avance técnico-cambio jurídico” cuando éste tiene una proyección social amplia.

La consecuencia jurídica del avance técnico puede consistir, como señala Díez-Picazo, en la aparición de una nueva legislación o en la reacción del Ordenamiento jurídico en sus vías de aplicación e interpretación, siendo frecuente que la primera reacción sea la de *“tratar de ajustar o englobar el conjunto de los nuevos hechos que el nivel técnico determina en los moldes antiguos, preconizando de este modo el mantenimiento de la misma solución anteriormente establecida, por medio de lo que puede ser una forzada interpretación extensiva de los textos”* (1987, p.90). De esta forma, como el Derecho surge de la realidad es necesario que intervenga con reglas de juego que se impongan en la sociedad en transformación.

Conscientes las distintas autoridades políticas de la enorme repercusión de los avances en IA y de las consecuencias de los mismos, la Comisión Europea publicó en febrero de 2020 el *Libro Blanco sobre IA - un enfoque europeo orientado a la excelencia y la confianza* (Libro blanco sobre IA). Su punto de partida ha sido el análisis de la dicotomía que presenta la IA, en el marco del liderazgo tecnológico, señalando los posibles beneficios, pero destacando también los riesgos potenciales, entre los que señala, de forma específica, la opacidad en la toma de decisiones, su uso con fines discriminatorios, la incidencia en la vida privada o su utilización con fines delictivos.

Teniendo en cuenta lo señalado, parece que la tendencia, por lo menos de forma más inmediata, es intentar ajustar los problemas que puedan surgir a las categorías jurídicas existentes en el momento actual. Partiendo de esa base, y al margen de los posibles ilícitos penales, los derechos fundamentales con más riesgo de afección por el uso de esta tecnología son el derecho a la igualdad y no discriminación y el derecho a la privacidad. Al respecto, cabe recordar que este último, también denominado derecho a la autodeterminación informativa o protección de datos, constituye una nueva modalidad de libertad personal, tendente a proteger jurídicamente la “identidad personal” (Pérez Luño, 1992 p. 107). Es la respuesta del Estado de Derecho frente a las nuevas necesidades derivadas de la revolución tecnológica y que garantiza a la persona un poder de control y disposición sobre sus propios datos personales. Este actúa al mismo tiempo como instituto o garantía del pleno disfrute de los restantes derechos fundamentales, como pueden ser la igualdad y la no discriminación².

Se reitera también como riesgo irremediamente unido a la IA la opacidad en la toma de las decisiones, que debe ser abordada en la legislación de forma específica y cuyo tratamiento

2 Fundamento jurídico séptimo, sentencia del Tribunal Constitucional Español 290/2000, de 30 de noviembre.

deberá ir vinculado a la protección de los derechos mencionados; cuestión estrechamente relacionada con la “elaboración de perfiles” y con las decisiones individuales automatizadas – considerando 71 Reglamento UE 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (RGPD)-, resulta evidente, en este contexto, que los riesgos de discriminación en la toma de decisiones pueden incrementarse a través de la IA y sus efectos pueden ser mucho más amplios.

Si bien los derechos señalados son los que pueden verse directamente *más* afectados por la creación y/o usos de la IA según el estado actual de la técnica, ésta presenta también un riesgo directo o indirecto para otros derechos que pueden verse comprometidos en la deriva de un uso torticero de la tecnología. Así, si los riesgos inmediatos deben abordarse con normas para proteger la privacidad, la igualdad y la no discriminación, resulta también esencial la valoración de los posibles daños para otras libertades. Entre los derechos a los que debe prestarse una especial atención ante un mal uso o creación de IA, cabe señalar la libertad de expresión, la libertad de reunión, la libertad de pensamiento, el derecho a la tutela judicial efectiva y a un juicio justo, el derecho a la educación e incluso el derecho a la vida y a la salud. Muchos de estos derechos ya están siendo afectados por el uso de ciertas aplicaciones con fines diversos entre los que destacan los vinculados al control masivo de poblaciones.

El Consejo de Europa aprobó el 8 de abril de 2020 una serie de *Recomendaciones sobre el impacto de los algoritmos en los derechos humanos*. Se insta a los Estados miembros a garantizar el derecho a la autodeterminación informativa en el diseño, desarrollo y puesta en marcha de los sistemas algorítmicos, evaluando con detalle los derechos humanos que pueden verse afectados y, en particular, el derecho a la no discriminación. Se insta, de forma específica, al sector privado a garantizar los derechos humanos tomando conciencia de la responsabilidad hacia los valores de una sociedad democrática. En este sentido, cabe recordar, con el profesor Peces-Barba, que, si bien los derechos fundamentales se presentan en su fórmula clásica como mecanismos para preservar la libertad del individuo frente al poder de Estado, para proteger a los hombres “de los maleficios del poder”, la quiebra de la convicción igualitaria ha promovido la extensión de los derechos fundamentales al ámbito privado, de tal forma que ha permitido preguntarse si los derechos no deben ser también preferentes o, en todo caso, protegibles en las mismas condiciones en los dos ámbitos (1995, p.72).

Aunque es imposible predecir cómo va a evolucionar el estado, los conocimientos y aplicaciones vinculadas a la IA en los próximos años, parece indudable que el Derecho deberá enfrentarse a nuevos retos derivados del análisis de grandes cantidades de datos de carácter personal y su conexión, incluso con relación a conjuntos de datos que, en sí mismos, no contienen datos personales. Previsiblemente, si se logran grandes avances en la inteligencia artificial general y fuerte, el encaje de los riesgos en las categorías actuales de derechos resulte extremadamente complejo, siendo preciso quizás una extensión de su contenido que dé cobertura a nuevas situaciones o tal vez la creación de nuevos derechos³. La utilización de los adverbios de duda que deberían limitarse en los escritos jurídicos constituyen, sin embargo, en esta temática una obligación en un intento de poner de manifiesto un hecho futurible que, de producirse, tendrá

3 Sobre los nuevos derechos y tecnologías, Vid. DE ASÍS ROIG, R.; *Derechos y tecnologías*, Dykinson, Madrid, 2022, pp. 43-51.

grandes repercusiones para la humanidad y ante el cual el legislador y los poderes públicos deben estar preparados para abordar con “el nivel superior de maestría necesario” la competencia tecnológica (Bostrom, 2016, p. 11). En este sentido, como señala Llano Alonso, “debemos y podemos salvar nuestra circunstancia con las tecnologías, pero a la vez debemos librarnos de su fuerza de atracción centrípeta para evitar ser absorbidos por el agujero negro del posthumanismo” (Llano Alonso, 2018, p.192).

2 La confianza individual y social: el principio de transparencia como requisito esencial en la regulación de la IA

Teniendo en cuenta los problemas que se han señalado en el epígrafe anterior, se presenta como necesaria la implantación de un requisito cuya relevancia pudiera parecer menor pero que resulta esencial en la puesta en marcha de los dispositivos de IA: la necesidad de generar tanto confianza individual como social.

Tras ciertos escándalos por todos conocidos derivados de un uso masivo ilícito de información de carácter personal con aficción directa a los derechos fundamentales, parece que, en gran medida, la confianza en sistemas vinculados al uso de la IA, dentro del marco del Estado de Derecho, vendrá dada y quedará condicionada por la forma en la que se realice su regulación y por las garantías que se implementen para la protección de estos derechos. La Comisión europea ha señalado como uno de los objetivos centrales en el abordaje de la materia la necesidad de generar un “ecosistema de confianza” con un enfoque antropocéntrico centrado en el ser humano. Dicho reto pasa por la necesidad previa de una base mínima de alfabetización digital de la ciudadanía, pues malamente podrá valorar un individuo las bondades o maldades de la implantación tecnológica o de la toma de decisiones basada en la misma si no le resulta posible su comprensión. Desde algunos sectores se ha apelado a una cierta “responsabilidad de la ciudadanía” que no se estaría ejerciendo en relación con los avances tecnológicos. No obstante, la misma no parece exigible en un contexto actual en el que la educación y la falta de comprensión de esta materia no permiten este control o responsabilidad en su utilización. Al respecto, la Unión Europea, dentro del marco de su *Estrategia sobre IA*, aprobada en abril de 2018, ya destacó la necesidad de concienciación en torno a la misma en todos los niveles de la educación con la finalidad de crear habilidades sobre la toma de decisiones. Evidentemente, dicha tarea resultará de enorme complejidad y de difícil extensión a ciertos sectores concretos de la población.

Los recelos sobre el uso tecnológico mal intencionado o sobre los posibles efectos perversos parecen forzar un debate técnico o pretécnico. En este contexto, la Comisión Europea, de forma previa a entrar en la regulación, ha señalado, siete requisitos a tener en cuenta en relación con la IA: acción y supervisión humanas; solidez técnica y seguridad; gestión de la privacidad y protección de datos; transparencia; diversidad, no discriminación y equidad; bienestar social y medioambiental y rendición de cuentas. Asume las *Directrices éticas para una IA fiable* elaboradas por el grupo independiente de expertos de alto nivel sobre inteligencia artificial, creado en junio de 2018.

Entre ellos, la transparencia se presenta como un requisito esencial no solo respecto a los datos de entrenamiento o a la conservación de los datos sino también para la creación de esa confianza social e individual y, en su caso, para reparar el daño causado. Según el *Libro Blanco*

de IA, este principio de transparencia conlleva necesariamente: la información respecto a las capacidades y limitaciones de la IA e información sobre la si el afectado se está relacionado con un ser humano o con un sistema de IA.

Este principio de transparencia está ya presente en la normativa de protección de datos en el ámbito de la Unión Europea, considerándolo como una medida de privacidad por defecto para que los ciudadanos puedan supervisar el tratamiento de su información de carácter personal y señalándolo como medida imprescindible para garantizar un tratamiento leal de los datos personales y del derecho a recibir una información completa, clara y sencilla relativa a todos los aspectos relevantes de un tratamiento de datos personales y a sus posibles consecuencias (Garriga Domínguez, 2018, p. 134).

No obstante, dicho principio ha de complementarse, en el marco de la IA y en aras a generar la confianza social en torno a la misma, con su extensión en el diseño y en la puesta en marcha de los dispositivos que usen este tipo de tecnología. Así, “este derecho cobra una nueva y mayor relevancia en el mundo de los datos masivos, de la computación ubicua y del Internet de las cosas” (Garriga Domínguez, 2018, p. 136). Esta exigencia implicará que la aplicación de dicho principio no puede reducirse a un instante puntual, sino que afectará a todos y cada uno de los momentos de creación y puesta en marcha de los dispositivos, así como a todos los elementos y participantes que intervienen en la solución tecnológica. Por los posibles riesgos para la ciudadanía y sus derechos, estos requisitos deben ser evaluados a lo largo de todo el ciclo de vida de un sistema de IA, tal y como ha señalado la Agencia Española de Protección de Datos en su documento ya citado sobre adecuación al RGPD de tratamientos que incorporen inteligencia artificial. En este sentido, como señala Tomás Mallén, la exigencia de transparencia debe ser considerada como “un derecho prestacional que requiere una actuación positiva por parte de las autoridades públicas” (2015, p.832).

Este principio debe complementarse necesariamente con la supervisión humana con el objetivo de no socavar la autonomía humana y protegerla de los efectos adversos. Esta supervisión planteará especial problema en aquellos sistemas de IA que evolucionen y aprendan de la experiencia en los que necesariamente se tendrán que adoptar mecanismos para evitar la pérdida de control de la IA por el ser humano. En el marco de las *Directrices Éticas para una IA fiable* se apuesta por una IA centrada en las personas, en el servicio a la humanidad y en el bien común y apoyada en tres componentes esenciales: lícita, que respete el marco normativo existente, prestando especial atención a los textos de derechos humanos, ética y robusta desde el punto de vista técnico y social para que los sistemas funcionen de manera segura y fiable y determinación, en su caso, los posibles daños involuntarios. Dentro de los principios éticos de la IA, basados en los derechos humanos, el citado grupo de expertos destaca los siguientes: respeto a la autonomía, prevención del daño, equidad y explicabilidad.

El Parlamento Europeo, el Consejo y la Comisión se ha pronunciado recientemente sobre los parámetros analizados en este epígrafe en la *Declaración Europea sobre los Derechos y Principios Digitales* para la Unión Europea, de 26 de enero de 2022⁴. Si bien el texto tiene carácter declarativo, aborda, entre otras cuestiones, en el capítulo III, las “interacciones con algoritmos y sistemas de inteligencia artificial”, en la que se compromete políticamente, en aras a garantizar la libertad de elección, a la promoción y aplicación de los siguientes principios:

4 <https://digital-strategy.ec.europa.eu/en/library/declaration-european-digital-rights-and-principles#Declaration>

- velar por la transparencia en el uso de los algoritmos y la inteligencia artificial y porque en su interacción con ellos las personas estén empoderadas e informadas;
- asegurar que los sistemas algorítmicos se basen en conjuntos de datos adecuados para evitar la discriminación ilegal y permitir la supervisión humana de los resultados que afectan a las personas;
- garantizar que las tecnologías, como los algoritmos y la inteligencia artificial, no se utilicen para predeterminedar las decisiones de las personas en ámbitos como, por ejemplo, la salud, la educación, el empleo y la vida privada;
- proporcionar salvaguardias que garanticen que la inteligencia artificial y los sistemas digitales son seguros y se utilizan con pleno respeto de los derechos fundamentales de las personas”.

La finalidad última, en este contexto, debería ser evitar que la IA y los algoritmos decidan “sobre fines y objetivos futuros del desarrollo de la vida, al margen de las elecciones de las personas concernidas por esos diseños” y en ocasiones con sustitución plena de las decisiones humanas por la IA (Pérez Luño, 2021, p. 30).

En este contexto, siguiendo a Martínez de Pisón, la Filosofía ya ha dado serios pasos en aras de avanzar, justificar y establecer un código ético en el marco, por ejemplo con la denominada roboética, que “vendría a ser una ética aplicada cuyo objeto es el ámbito científico de la robótica” (2017, p.70) , que giraría en torno a tres bloques: el estudio de los efectos de los productos tecnológicos en el marco social, el análisis de las reglas de conducta en el desarrollo de la IA y la polémica sobre la decisiones autónomas por parte de las máquinas. No obstante, según el citado autor, la roboética sería insuficiente para tratar los problemas que para la IA presenta, siendo necesario como mínimo abordar, entre otras, las siguientes categoría éticas y jurídicas afectadas por la nueva realidad tecnológica: identidad, libertad, igualdad y responsabilidad.

En este contexto, como señalaba Ortega y Gasset en su *Meditación sobre la Técnica*, uno de los riesgos respecto a la técnica es la posible pérdida de la denominada “conciencia técnica”, individual y colectiva, y sus condiciones, incluidas las morales o éticas en que se está desarrollando. Se presenta, por tanto, de forma previa y paralela a la elaboración de una normativa sobre IA, la necesidad de un análisis desde una perspectiva ética de la IA que algunos autores incluyen dentro de la denominada “Ética Digital” cuestionando si no resultaría más útil y operativo regresar al debate inicial de cómo debe realizarse esta regulación, preventiva, por un lado, técnica, por otro. Esta necesidad no resulta extraña pues, como señala De Asís Roig, los avances no siempre han ido acompañados del “progreso moral” (2014, p. 18). La confianza tanto individual como colectiva dependerá en gran medida de cómo se aborde el debate ético, de la transparencia y comunicación de las reglas pretécnicas y de la determinación de los agentes que deben participar en el mismo.

3 El nuevo panorama de la regulación de la IA en la Unión Europea: los derechos fundamentales como punto de partida

Son muchos los sectores en los que tendrá repercusión la inteligencia artificial entre los cuales está el ámbito jurídico y sus normas. El mayor o menor acierto en esta regulación influirá de forma decisiva en la manera de abordar el avance tecnológico y sus consecuencias.

En la actualidad ya existen normas que regulan algunos aspectos de la IA, como son las relativas al tratamiento de la información personal en estos sistemas ya señaladas en epígrafes anteriores.

No obstante, en aras de dotar de una mayor seguridad jurídica y evitar la dispersión normativa en el ámbito de la Unión Europea, la regulación de la IA está siendo tratada a nivel comunitario, siendo numerosa la actividad de los órganos de la UE en esta materia. Ya en octubre de 2017, el Consejo Europeo apuntó la necesidad urgente de abordar las nuevas tendencias tecnológicas, incluida la IA, instando a la Comisión a proponer un planteamiento sobre la misma. El objetivo, como se señala en la *Comunicación de la Comisión Europea de 25 de abril de 2018, sobre Inteligencia artificial para Europa*, es garantizar el establecimiento de un marco ético y jurídico, basado en los valores de la Unión y en su carta de Derechos Fundamentales. Posteriormente, el *Libro blanco sobre IA*, de 19 de febrero de 2020, formuló alternativas políticas para promover y regular esta tecnología y abordar sus riesgos, destacando la posición de la Unión Europea en torno a una serie de valores compartidos que la posicionan en un marco de liderazgo condicionado por promoción del uso ético de la inteligencia artificial. En estos momentos son ya varios los textos, con mayor o menor fuerza jurídica, en el ámbito de la Unión destinados a intentar conjugar la aplicación y uso de la IA con la protección de los derechos⁵. Entre los más recientes, pueden citarse la Declaración Europea sobre los Derechos y Principios Digitales para la Unión Europea⁶ y la Propuesta de Reglamento del parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia artificial) y se modifican determinados actos legislativos de la Unión. Éstos se encuentran estrechamente vinculados a otras normas que regulan esta cuestión en diversos ámbitos sectoriales como el citado RGPD, de 21 de abril de 2021, la Directiva sobre protección de datos en el ámbito penal o la normativa sobre igualdad.

De todos ellos, la propuesta de Reglamento constituye el texto más ambicioso. Parte de un enfoque basado en el riesgo, con el objetivo de cubrir todos los sistemas de IA de alto riesgo, con posibilidad de que los proveedores de IA que no sean de alto riesgo sigan un código de conducta; riesgos que han de “calcularse teniendo en cuenta su repercusión para los derechos y la seguridad” –punto 3.1 de la exposición de motivos-. Pretende ser una respuesta coordinada de la Unión a las implicaciones éticas y jurídicas de la IA, incluyendo entre sus objetivos específicos “garantizar que los sistemas de IA introducidos y usados en el mercado de la UE sean seguros y respeten la legislación vigente en materia de derechos fundamentales y valores de la UE”, prohibiendo específicamente determinadas prácticas en materia de IA por posible colisión directa con los derechos fundamentales. Los derechos humanos constituyen, al menos de forma teórica, el marco ético, jurídico y político de referencia en la materia. En este contexto, se considera que la protección de los mismos permitirá garantizar el posible desequilibrio de los ciudadanos frente a esta tecnología, pero también permitirá la creación de una legislación con una base sólida que evite la tan temida inseguridad jurídica por parte de las empresas y organismos afectados.

Con este objetivo se propone una definición jurídica de IA que pueda resistir el paso del tiempo, introduciendo una definición de “sistema de inteligencia artificial” entendido como “el

5 Para una relación temporal de todos ellos, Vid. Punto 1.1 de la exposición de motivos de la propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de inteligencia artificial) y se modifican determinados actos legislativos de la Unión. <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:52021PC0206>.

6 <https://digital-strategy.ec.europa.eu/en/library/declaration-european-digital-rights-and-principles#Declaration>

software que se desarrolla empleando una o varias de las técnicas y estrategias que figuran en el anexo I 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”. Las técnicas y estrategias de aprendizaje de IA que figuran en el mencionado anexo I son las siguientes: “estrategias de aprendizaje automático, incluidos el aprendizaje supervisado, el no supervisado y el realizado por refuerzo, que emplean una variedad de métodos, entre ellos el aprendizaje profundo. Estrategias basadas en la lógica y en el conocimiento, especialmente la representación del conocimiento, la programación (lógica) inductiva, las bases de conocimiento, los motores de interferencia y deducción, los sistemas expertos y de razonamiento (simbólico). Estrategias estadísticas, estimación bayesiana, métodos de búsqueda y optimización”.

La consideración del riesgo parece que va a ser determinante para establecer los requisitos legales obligatorios, apostando a priori únicamente por su aplicación en aquellos casos en los que el riesgo se considere alto. No cabe olvidar que la necesidad de gestión del riesgo para los derechos y libertades forma parte ya del concepto de responsabilidad proactiva que constituyó una de las novedades más relevantes del RGPD –artículo 24-. El principio de responsabilidad proactiva supone un cambio de paradigma que exige pasar de un sistema de protección reactivo frente al incumplimiento, a un modelo preventivo y proactivo (Lorenzo Cabrera, 2018, p. 123). Este principio requiere que el responsable del tratamiento realice un análisis de los datos que trata y de las demás circunstancias del tratamiento para después adoptar las medidas reales y eficaces que garanticen que el mismo cumple los requisitos legales. Para ello deberá tener en cuenta “la naturaleza, el ámbito, el contexto y los fines del tratamiento, así como los riesgos de diversa probabilidad y gravedad para los derechos y libertades de las personas físicas”.

En este sentido, la UE parece optar por un marco regulatorio que únicamente se aplique a los sistemas de IA de alto riesgo, con la opción de que los sistemas de IA de menor riesgo sigan o se sometan a un código de conducta. El debate sobre los grados de intervención reguladora no es, sin embargo, novedoso. Así, ya hace tiempo parte de la doctrina se pregunta si no resultaría más útil y operativo regresar al debate inicial de cómo debe realizarse esta regulación, preventiva, por un lado, técnica, por otro. Si bien, no pocos autores apuestan por la idoneidad de los denominados “sistemas de autorregulación”, señalando a los códigos de conducta como mecanismo idóneo para abordar los problemas, para otros, esta fórmula no es más que un reflejo de la ruptura del viejo modelo ilustrado y de la crisis de uno de sus rasgos: la tendencia a la desregularización, lo que implicaría la deslegalización y a contractualización de fenómenos sociales, que se traduce en “una retirada de la normatividad jurídica y en una mutación profunda de función y alcance” (De Julios Campuzano, 2018, p. 68). Advierten estos últimos que las posibles consecuencias negativas en el plano jurídico de la adopción de esta postura parecen demasiado amplias y que desdibuja los límites entre lo prohibido y lo permitido. Como indica De Julios-Campuzano “al difuminarse el límite entre seguridad y previsibilidad del Derecho y, con él, la libertad y responsabilidad del ciudadano, cambiaría igualmente la seguridad precaria, preventiva, manipulable y contextual en relación a determinados bienes jurídicos” (2018. 74).

En estos términos, la Unión Europea en su nueva propuesta reguladora de IA parece apostar por un intento de conciliación de las visiones expuestas, optando, por un lado, en la futura regulación de la IA por la elaboración de un Derecho de prevención de riesgos con

el objetivo de anticipar la dimensión temporal del futuro y garantizar una vida digna de las futuras generaciones. La necesidad deriva de que, en el contexto actual, la técnica ya no es un instrumento para el hombre, sino que pasa a plano principal. De ahí, la importancia de la aplicación del denominado “principio de precaución”, esto es, la adopción de normas preventivas orientadas a minimizar la potencialidad de las posibles amenazas destructivas para la humanidad. Ciertamente que, frente a este principio, no pocos autores defenderían la superioridad del derecho a la libertad científica, como rasgo esencial del Estado de Derecho. Sin embargo, el legislador comunitario entiende que, con esta propuesta, no se trata de frenar el avance técnico sino de establecer un marco unitario con estándares mínimos que buscan la protección de derechos y bienes jurídicamente esenciales para proteger al ser humano de forma individual y colectiva. Por este motivo, combina este modelo con el de autorregulación en el desarrollo o utilización de aplicaciones de IA no consideradas de alto riesgo, a través de códigos de conducta cuya finalidad parece ser que los sistemas de IA que no sean de alto riesgo “cumplan de manera voluntaria los requisitos que son obligatorios para los sistemas de IA de alto riesgo”.

Partiendo de un enfoque basado en el riesgo, la propuesta de Reglamento distingue entre riesgo inaceptable, riesgo alto y riesgo bajo o mínimo, considerando incluidos dentro de los primeros los sistemas de IA los contrarios a los valores de la UE por violación de los derechos fundamentales y estableciendo su prohibición; éstas “engloban aquellas prácticas que tienen un gran potencial para manipular a las personas mediante técnicas subliminales que trasciendan o que aprovechan las vulnerabilidades de grupos vulnerables concretos, como los menores o las personas con discapacidad, para alterar de manera sustancial su comportamiento de un modo que es probable que provoque perjuicios físicos o psicológicos a ellos o a otras personas (...). La propuesta prohíbe igualmente que las autoridades públicas realicen calificación social basada en IA con fines generales. Por último, también se prohíbe, salvo excepciones limitadas, el uso de identificación biométrica remota en tiempo real en espacios de acceso público con fines de aplicación de la ley” –punto 5.2.2 exposición de motivos y artículo 5- Respecto a los sistemas de riesgo alto se incluye un listado no cerrado con el objetivo de completarlo de conformidad con los avances técnicos y compatibilizarlo con el necesario respeto y garantía de los derechos.

De nuevo, los requisitos de transparencia y comunicación de información a los usuarios y la vigilancia y supervisión humanas, analizados en el epígrafe anterior, aparecen como ejes de la nueva propuesta de regulación en un intento de garantizar una información adecuada y correcta a los usuarios respecto a los sistemas de alto riesgo, incluyendo en todo caso una herramienta de interfaz humano-máquina que permita la vigilancia humana –artículos 13 y 14 de la propuesta de Reglamento sobre IA-.

Si bien esta perspectiva resulta en parte novedosa, la Comisión Europea, desde hace ya algunos años está trabajando en la regulación de algunos sectores que pueden ser determinantes en su creación y desarrollo. Ésta abarca desde los temas de protección de datos personales, ya analizados, o la libre circulación de los no personales (Reglamento (UE) 2018/1807 del Parlamento Europeo y del Consejo, de 14 de noviembre de 2018 al marco normativo de la ciberseguridad⁷. En la actualidad, entre las seis prioridades de la Comisión Europea para el 2019-2024, con el objetivo de configurar el futuro digital de Europa, figuran la IA, la estrategia Europa de datos y la estrategia industrial europea, con un enfoque que se apoya en tres pilares:

⁷ Vid. <https://www.consilium.europa.eu/es/policies/cybersecurity/>

la tecnología al servicio de las personas, una estrategia económica digital justa, pero también competitiva, y una sociedad abierta, democrática y sostenible.

La Unión Europea deberá además abordar cómo afectará la IA a la normativa que ya está en vigor, como por ejemplo las normas destinadas a garantizar la igualdad, y si la misma será capaz de dar una respuesta efectiva a aplicación de estos sistemas. Así, la dificultad no solo residirá en intentar hacer frente a un marco común sino abordar un estudio de la legislación, tanto en el ámbito comunitario como estatal, para garantizar la correcta aplicación de la normativa existente. Como señala el citado libro blanco, el marco normativo se enfrentará a los retos de la posible opacidad de la IA, al aprendizaje automático, los temas de responsabilidad y cambios en el concepto de seguridad -*Informe de la Comisión al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica*, de 19 de febrero de 2020-

Son muchas, por tanto, las características de las tecnologías pueden hacer difícil comprobar el cumplimiento de la protección de los derechos fundamentales e impedir su eficacia. El reto de determinar los destinatarios de la norma y su alcance geográfico será otra de las dimensiones que deberá abordarse. Más allá incluso de esta dimensión futura o amplitud de la regulación, como indica Martínez de Pisón, las discusiones en la ética y en la teoría de los derechos humanos, en función de la evolución, podrían extenderse hacia el debate de si los seres no humanos pueden tener derechos si se le atribuye la condición de agentes morales. Así, en este marco, únicamente cabe partir del hecho indubitado de que el desarrollo de la tecnología está alterando la naturaleza y la relación del hombre con la misma. Como han advertido autores como De Julios-Campuzano, el Derecho de la Modernidad que supeditaba su respuesta a la constatación de un conflicto, no resulta aplicable siendo necesaria la “incorporación de una dimensión futura”. La necesidad de un nuevo Derecho ante la aparición de nuevas formas de amenaza para los derechos fundamentales, derivadas de las nuevas tecnologías, parece constatarse. En este sentido, como indica Frosini es necesario un cambio de perspectiva en relación con los derechos humanos, puesto que de los derechos comprendidos en un catálogo cerrado se ha pasado a una concepción abierta y progresiva de éstos, cuya finalidad es la de su adecuación a las nuevas necesidades del hombre como “*creador del mundo tecnológico*” (Frosini, 1996, p.95).

4 Algunas consideraciones finales

La articulación jurídica de la relación entre IA y derechos fundamentales se presenta compleja en un mundo tecnológicamente cambiante en el que Derecho trata de adaptar estos avances, en la mayoría de los supuestos, a categorías ya existentes que pone en riesgo la seguridad jurídica y que no alcanzan, en ocasiones, a solucionar los problemas que la IA artificial presenta para el Derecho y para los derechos. Teniendo en cuenta el estado de los avances tecnológicos relacionados con la IA, los derechos más afectados por su utilización en la actualidad son el derecho a la vida privada y el derecho a la igualdad y no discriminación. En este sentido, el imparable progreso tecnológico ha facilitado la fabricación, diseño y utilización de máquinas dedicadas a los procesos de decisión de perfiles individuales y al análisis e incluso resolución de gran variedad de relaciones públicas y privadas, sometiendo a la privacidad a unos riesgos significativos y distintos a los tradicionales para la eficacia del derecho fundamental (Lucas Murillo

de la Cueva, 2022). Posiblemente, estos riesgos obligarán a cambiar el modelo de protección de datos, que ya ha comenzado en el ámbito de la UE con el RGPD a través de la introducción de ciertos principios como el de responsabilidad proactiva, la privacidad desde el diseño o por defecto. No obstante, la IA presenta también un riesgo directo o indirecto para otros derechos que pueden verse comprometidos en la deriva de un uso torticero de la tecnología. Entre los derechos que pueden verse afectados por un mal uso o creación de IA, cabe señalar la libertad de expresión, el derecho a la información, la libertad de reunión, la libertad de pensamiento, el derecho a una tutela judicial efectiva y a un juicio justo, el derecho a la educación e incluso el derecho a la vida y a la salud.

En este contexto, el derecho a la protección de datos personales, como libertad fundamental en sí misma y como derecho instrumental ordenado a la protección de otros derechos fundamentales frente al uso de la tecnología, cumplirá una doble misión. Por una parte, garantizar a los ciudadanos el control de sus datos de carácter personal, permitiéndole conocer qué datos, quién y para qué los obtiene y utiliza. Por otra, como mecanismo de protección de otras libertades. No obstante, en ambos casos, se enfrentará a nuevos desafíos derivados del análisis de grandes cantidades de datos y su conexión, incluso de conjuntos de datos que, en sí mismos, no contienen datos personales. En este contexto, se presenta como esencial fortalecer el principio de transparencia para poder conocer quién diseña esas categorías, quién decide su significado y quién decide bajo qué circunstancias esas categorías sean decisivas.

En estos momentos se está intentando abordar en la UE un marco jurídico genérico de la IA. No obstante, tal y como apuntan tanto la doctrina como los órganos implicados en la elaboración de ese marco, uno de los requisitos previos para la creación y puesta en marcha de dispositivos de IA es la necesidad de generar confianza individual y social. Dicho reto pasa por la necesidad previa de una base mínima de alfabetización digital de la ciudadanía y la aplicación del principio de transparencia, que debe ir acompañado de la supervisión humana de los sistemas en aras de proteger la autonomía humana y evitar los efectos adversos. Resulta esencial la creación de unas directrices éticas sólidas para una inteligencia artificial fiable, previo debate pretécnico sobre el avance tecnológico en cuestión y sobre los agentes implicados en los mismos. Este análisis pretécnico llega tarde a la implantación de la denominada IA débil, pero debería considerarse respecto a la implantación de mecanismos de IA general o fuerte.

Dentro del marco del Estado de Derecho la confianza en la técnica queda condicionada además a la forma en la que se realice su regulación y las garantías de las que se dote a la misma. Esta normativa debe tener su base en la idea central de dignidad humana y en la protección de los valores que fundamentan los derechos humanos. El marco jurídico regulador de la IA debe tener como punto de partida los valores de la Unión y la carta de Derechos Fundamentales. En este sentido, los derechos humanos constituyen el marco ético, jurídico y político de referencia en la materia.

La apuesta la Unión Europa por la elaboración de un derecho de prevención de riesgos con el objetivo de anticipar la dimensión temporal de futuro y garantizar una vida digna de las futuras generaciones parte del hecho de que la IA no es mala en sí misma, sino que el problema será definir sus finalidades, el uso de la misma y los nuevos sesgos que se derivan de su empleo. Desde esta perspectiva, la propuesta de la UE parece centrada en el concepto de riesgo y en encontrar un marco legislativo apropiado para los sistemas de riesgo alto, dejando en manos

de los códigos de conducta la adopción de medidas respecto a la inteligencia artificial con menor riesgo; dicha postura presenta algunos problemas básicos desde la teoría de los derechos fundamentales, pues cabe recordar que la autorregulación no cabe ni respecto a su contenido ni desarrollo. No resulta posible, por tanto, olvidar que los bajos desafíos técnicos de la inteligencia artificial pueden presentar grandes riesgos para los derechos y, en especial, para la garantía y eficacia de los mismos.

Resulta curioso, el contraste de esta propuesta con declaraciones previas sobre la necesidad de que los derechos constituyan el marco básico desde el cual afrontar los desafíos que presentan los problemas relacionados con la IA. En este sentido, algunos textos e informes ya han destacado la problemática que para los derechos fundamentales puede implicar el enfoque de la norma basado en el análisis del riesgo por el impacto que estos sistemas pueden tener sobre los individuos, sobre la sociedad y sus derechos. Parecen necesarias, por tanto, ciertas revisiones o estudios de la propuesta, tendentes, como mínimo, a determinar su función y alcance y a dotar de una mayor seguridad y previsibilidad a la norma impidiendo que se desdibujen los límites entre lo permitido y lo prohibido, especialmente cuando se trata de una regulación que incide de forma directa en derechos y libertades fundamentales.

Referencias

BOSTROM, N.; *Superinteligencia. Caminos, peligros, estrategias*, Teell, 3ª edición, 2016.

CHURNIN, S.: *Inteligencia artificial: retos éticos y jurídicos, y la influencia de los derechos humanos*, Servicio de Publicaciones de la Facultad de Derecho de la Universidad Complutense de Madrid, Madrid, 2012.

DE ASIS ROIG, R.; *Una mirada a la robótica desde los derechos humanos*, Instituto de Derecho Humanos “Bartolomé de las Casas”, Universidad Carlos III de Madrid, Madrid, 2014.

DE ASÍS ROIG, R.; *Derechos y tecnologías*, Dykinson, Madrid, 2022.

DE JULIOS-CAMPUZANO, A.; “Los derechos humanos en la sociedad del riesgo. Crisis del estado, justicia intergeneracional y medioambiente”, *Cuadernos Electrónicos de Filosofía del Derecho*, núm. 37, 2018.

DÍEZ-PICAZO, L.; *Derecho y masificación social. Tecnología y Derecho privado (Dos esbozos)*. Cuadernos Civitas, Madrid, 1987.

FROSINI, V.; “Los derechos humanos en la era tecnológica”, en PEREZ LUÑO, A. E. (Coord); *Derechos humanos y constitucionalismo en el tercer milenio*, Marcial Pons, Madrid, 1996.

GARRIGA DOMÍNGUEZ, A.; “La elaboración de perfiles y su impacto en los derechos fundamentales. Una primera aproximación a su regulación en el Reglamento de Protección de datos de la Unión Europea”, *Derechos y Libertades*, núm.38, época II, 2018.

LORENZO CABRERA, S; “Posición jurídica de los intervinientes en el tratamiento de datos personales. medidas de cumplimiento”, en AAVV.: *Protección de datos, responsabilidad activa y técnicas de garantía*, Editorial Reus, madrid, 2018,

LLANO ALONSO, F.H.; *Homo Excelsior. Los límites ético-jurídicos del transhumanismo*, Tirant lo Blanch, Valencia, 2018.

KAPLAN, J.; *Inteligencia artificial. Lo que todo el mundo debe saber*, Teell, 2017.

PECES-BARBA MARTÍNEZ, G.; “La libertad del hombre y el Genoma”, en *El Derecho ante el Proyecto Genoma*, Vol. I, Fundación BBV con la colaboración de Universidad de Deusto-Diputación Foral de Biskaia, Bilbao, 1994.

ORTEGA Y GASSET, J.; *Meditación de la Técnica*, Colección Austral núm. 1360, Espasa-Calpe, Madrid, 1965.

PECES-BARBA MARTÍNEZ, G.; *Ética, poder y Derecho. Reflexiones ante el fin de siglo*, Centro de Estudios Constitucionales, Madrid, 1995.

PÉREZ LUÑO, A.E.; “Vittorio Frosini y los nuevos derechos en la sociedad tecnológica”, *Informática e Diritto*, Le Monnier, Firenze, 1992.

PÉREZ LUÑO, A.E.; “El posthumanismo no es humanismo”, *Derechos y Libertades*, núm. 44, enero, 2021.

PINTO FONTANILLO, J.A.; *El Derecho ante los retos de la inteligencia artificial. Marco ético y jurídico*, Edisofer, 2020.

MARTÍNEZ DE PISÓN, J.: “Yo, Robot: de la biología a la singularidad. ¿Nuevas preguntas para la Filosofía del Derecho?”, *REDUR* 15, diciembre 2017.

MCCARTHY, J.; “What Is Artificial Intelligence”. Sección “Basic Questions”, 2007. Enlace: <http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html>

LUCAS MURILLO DE LA CUEVA, P.; “Aspectos de la actualidad del derecho fundamental a la protección de datos personales”, *La Ley Privacidad*, núm. 11, enero-marzo 2022.

TOMÁS MALLÉN, B.; “Transparencia y protección de datos: nuevos desafíos para la garantía europea de los derechos fundamentales”, en RALLO LOMBARTE, A. y GARCÍA MAHAMUT, R.: *Hacia un nuevo Derecho europeo de protección de datos*, Tirant lo Blanch, Valencia, 2015.