

Gobernanza algorítmica: claves ético-jurídicas en tiempos de automatización

Algorithmic Governance: Ethical-Legal Keys in Times of Automation

Dra. (c) Ginette Isabel Rodrigo Romero

ORCID: [0009-0006-3695-086X](https://orcid.org/0009-0006-3695-086X)

Doctoranda en Derecho Internacional Público de la Universidad de Sevilla, España.

RESUMEN: En las últimas décadas, el avance de la inteligencia artificial (IA) ha generado una transformación profunda en la sociedad, afectando directamente los marcos normativos y los derechos fundamentales. La incorporación de sistemas automatizados en sectores sensibles como la salud o el empleo evidencia la necesidad de establecer marcos éticos y jurídicos que orienten su regulación y garanticen su desarrollo alineado con la dignidad humana. Este trabajo ofrece una aproximación al funcionamiento de los sistemas de IA y analiza críticamente los principios fundamentales que deben guiar su regulación, con énfasis en el enfoque propuesto por la Unión Europea. Se examinan especialmente el principio de precaución y el de transparencia algorítmica, como ejes normativos para enfrentar los riesgos estructurales y fortalecer los mecanismos de rendición de cuentas. La propuesta central radica en impulsar una gobernanza algorítmica inclusiva y responsable, que supere los principios abstractos y se traduzca en obligaciones jurídicas claras, orientadas a la protección de los derechos fundamentales y la legitimidad social de la IA.

Palabras clave: ética algorítmica, derechos fundamentales, sesgo discriminatorio, principio de precaución, transparencia, gobernanza digital.

ABSTRACT: In recent decades, the advancement of artificial intelligence (AI) has profoundly transformed society, directly impacting legal frameworks and fundamental rights. The integration of automated systems into sensitive sectors such as healthcare and employment underscores the urgent need to establish ethical and legal frameworks to guide AI regulation and ensure its development aligns with human dignity. This paper offers an overview of how AI systems operate and offers a critical analysis of the core principles that should guide their regulation, with a particular focus on the approach adopted by the European Union. Special attention is given to the precautionary principle and the principle of algorithmic transparency, which serve as normative tools to address structural risks and ensure accountability. The central proposal is to promote inclusive and responsible algorithmic governance—one that goes beyond abstract principles and is translated into clear legal obligations aimed on protecting fundamental rights and ensuring the social legitimacy of artificial intelligence.

Keywords: Algorithms ethics, Fundamental rights, Discriminatory bias, Precautionary principle, Transparency, Digital governance.

SUMARIO: 1. INTRODUCCIÓN. 2. EL CONTEXTO DE TRANSFORMACIÓN DIGITAL IMPERANTE. 3. LA INTELIGENCIA ARTIFICIAL EN CLAVE ÉTICA: DILEMAS Y TENSIONES. 4. PRINCIPIOS ÉTICOS COMO FUNDAMENTOS DE REGULACIÓN RESPONSABLE. 4.1. El principio de precaución. 4.2. El principio de transparencia. 5. HACIA UN MARCO JURÍDICO GLOBAL VINCULANTE Y UNA GOBERNANZA LEGÍTIMA. CONCLUSIONES. REFERENCIAS BIBLIOGRÁFICAS.

1. INTRODUCCIÓN

El desarrollo acelerado de tecnologías disruptivas como la inteligencia artificial (IA) en la actualidad plantea desafíos inéditos para los marcos jurídicos, éticos y políticos. La capacidad de los sistemas automatizados para intervenir en procesos de toma de decisiones que afectan la vida cotidiana, como por ejemplo el acceso a servicios públicos, el análisis predictivo de gustos y preferencias y la selección de personal en el ámbito laboral, exige una reflexión crítica sobre los límites, las garantías y los principios que deben regir su funcionamiento. Esta transformación normativa debe orientarse a asegurar la protección de los derechos fundamentales y a prevenir la consolidación de sesgos estructurales a través de la tecnología.

En ese sentido, este trabajo parte de la siguiente interrogante: ¿es la integración de preceptos éticos indispensable en los marcos jurídicos vinculantes para abordar los retos jurídicos que plantea la utilización de sistemas algorítmicos en la sociedad actual para así garantizar la protección efectiva de derechos fundamentales? Para responder esta pregunta, en primer lugar, es esencial analizar los fundamentos técnicos del funcionamiento de la IA y su impacto en los derechos fundamentales, así como el estudio e identificación de los marcos éticos y de gobernanza que orientan su regulación, especialmente de acuerdo con la experiencia regulatoria de la Unión Europea.

Como segunda cuestión, se examina el rol trascendental de principios éticos como el de precaución y el de transparencia algorítmica, para finalmente proponer lineamientos para una regulación orientada a la rendición de cuentas, la supervisión democrática y la protección efectiva frente a prácticas algorítmicas sesgadas y discriminatorias. La metodología adoptada es de carácter cualitativo, principalmente bibliográfica con enfoque jurídico y documental. El punto de partida es el análisis de fuentes normativas emitidas por los organismos internacionales adscritos a la Organización de las Naciones Unidas y acuerdos y resoluciones emitidos por la Unión Europea, así como literatura especializada en ética de la tecnología y derechos humanos.

El presente trabajo se inscribe en el marco de la gobernanza digital y pretende contribuir al debate académico sobre cómo construir una IA fiable, inclusiva y jurídicamente

responsable. Se parte del convencimiento de que el derecho, como instrumento regulador de las relaciones sociales, no puede mantenerse neutral ante las transformaciones tecnológicas, y debe asumir un rol proactivo en la construcción de un futuro digital justo.

2. EL CONTEXTO DE TRANSFORMACIÓN DIGITAL IMPERANTE

Prima facie, no resulta difícil entender la relación simbiótica entre los datos y máquinas, particularmente al hablar de sistemas de inteligencia artificial. Dentro de sus rasgos esenciales, se han reconocido la capacidad de aprendizaje, autonomía y razonamiento. En principio, el término IA fue acuñado por primera vez por John McCarthy en 1956 para referirse a la ciencia y la ingeniería de crear máquinas inteligentes, especialmente programas de computación inteligente (Russel & Norvig, 2004). Pero al día de hoy, sin embargo, ¿a qué nos referimos específicamente cuando hablamos de un sistema de inteligencia artificial?

La Real Academia Española (2024) define a inteligencia artificial como «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». Siguiendo esa línea, la IA se materializa a través de algoritmos, entendidos como códigos informáticos programados para ser capaces de dar respuestas a formulaciones o problemas.

Ahora bien, conviene precisar que el estudio de los sistemas de IA no se limita exclusivamente a los algoritmos, sino que hay algo superior a ellos; a razón de que también existen otros eslabones que forman parte de esta cadena digital. La profesora Gutiérrez García (2024) sostiene que el algoritmo se implementa en un código de programación, siendo aquel solo una parte de este, y que funciona en conjunto con librerías de datos, a los cuales dicho código se adapta y ejecuta una solución tecnológica en forma de *software* o *software* y *hardware*.

Por lo tanto, los sistemas algorítmicos están vinculados con grandes volúmenes de datos o *big data*. Este es definido como el conjunto de nuevas tecnologías asociado con grandes cantidades de datos; cuya utilidad radica en el reconocimiento y explotación de los mismos como un activo para la competitividad económica, la mejora de procesos, prestación de servicios públicos y privados (Hernández Peña, 2022) entre otras muchas aplicaciones que definen la sociedad digital de este siglo.

En este contexto, la cualidad fundamental del *big data* es que el análisis y procesamiento de grandes volúmenes de información permite identificar patrones y aspectos que serían difíciles de detectar cuando se trabaja con conjuntos de datos más reducidos o en menor escala (Delgado Martín, 2020). Así, la relación entre *big data* y algoritmos queda establecida; no obstante, es igualmente crucial detenerse a examinar la natura-

leza de estos últimos. Como se mencionó anteriormente, los algoritmos no actúan de forma aislada, sino que operan conjuntamente con diversos métodos, instrucciones y datos.

Domínguez Álvarez señala que, dentro del universo de algoritmos, algunos son deterministas, cuando sus reglas están completamente definidas, mientras que los no deterministas se caracterizan por la ausencia de reglas plenamente establecidas y cierta autonomía en el procesamiento de la información, lo que les permite generar nuevos resultados (2024, p. 53). Sin embargo, es fundamental diferenciar también que en los sistemas de inteligencia artificial existe un campo perteneciente a los algoritmos que hacen uso de un enfoque de *machine learning* (aprendizaje automático), y dentro de este, los que aplican el *deep learning* (aprendizaje profundo).

El *machine learning*, a su vez, comprende distintos tipos de aprendizaje automático, como el aprendizaje supervisado, el aprendizaje por refuerzo, el aprendizaje no supervisado y el aprendizaje semi-supervisado, cuya principal diferencia radica en el etiquetado de datos (Palma Ortigosa, 2022). De este modo, mientras que los sistemas de aprendizaje supervisado utilizan datos previamente etiquetados, en el aprendizaje no supervisado el algoritmo descubre nuevos patrones en los datos sin una clasificación previa.

Por su parte, los sistemas de *deep learning* se distinguen por su capacidad de aprendizaje sin entrenamiento previo, procesando información en forma de imágenes, sonido y video mediante redes neuronales artificiales organizadas en capas. Esto les otorga una notable precisión en la resolución de problemas complejos (Hernández Peña, 2022). Cabe destacar que algunas aplicaciones de *deep learning* ya forman parte de la vida cotidiana, como los asistentes virtuales *Siri*, *Alexa* o *Google Assistant*.

Ciertamente, en los Estados democráticos de Derecho, la irrupción y despliegue de la IA, en especial respecto a los sistemas de *machine learning*, podría afectar el objeto de algunos derechos fundamentales (Presno Linera, 2022). Así pues, es esencial subrayar que no es baladí el dilema de la singularidad tecnológica frente a las consecuencias jurídicas derivadas de la actuación de sistemas autónomos de *machine learning* sin supervisión humana.

3. LA INTELIGENCIA ARTIFICIAL EN CLAVE ÉTICA: DILEMAS Y TENSIONES

Aunque a principios de este siglo se asumía que los sistemas tecnológicos operaban bajo un principio de neutralidad—entendiéndose como ecuanímenes y objetivos—esta idea ha sido progresivamente cuestionada, usualmente por los resultados generados por algoritmos no deterministas. En consecuencia, la confianza inicialmente depositada en la inteligencia artificial ha erosionado, al evidenciarse la presencia de sesgos hereda-

dos de los desarrolladores de estos sistemas, muchas veces de forma involuntaria, pero que, en última instancia, pueden resultar perjudiciales para la ciudadanía.

Adicionalmente la autonomía de la que gozan algunos sistemas de IA nos lleva a la apremiante necesidad de colocarnos en una perspectiva ética y a cuestionar, como bien apunta Muñoz Vela (2021) sobre el objetivo principal y actual de la IA. En ese sentido, ¿la inteligencia artificial se configura como un medio para satisfacer necesidades, resolver problemas o mejorar la vida del ser humano? ¿o su finalidad se orienta a replicar con mayor fidelidad y profundidad la inteligencia humana en sistemas de información y máquinas?

Para decidir entre ambas cuestiones, es necesario aterrizar en el ámbito ético. La ética suele definirse como el conjunto de principios morales que orientan la conducta humana en diversos ámbitos de la sociedad e históricamente ha sido estrechamente vinculada con la política y el derecho. Empero, una perspectiva que integre estas tres disciplinas conduciría a "una construcción racional de la situación del hombre en la vida social teniendo en cuenta todos los elementos morales, políticos y jurídicos enraizados en un antropocentrismo laico, que es el paradigma de la modernidad" (Peces-Barba Martínez, 1995, p. 56).

Aunque esta visión trilateral es tentadora, la moralidad y los principios que sustentan la ética, para el profesor Fernández (1990), se distinguen entre moral personal y moral social. Tal argumento puede ser corroborado, en el sentido que la sociedad efectivamente alimenta el contenido moral y aprueba o rechaza ciertos comportamientos individuales, otorgándoles una calificación moral. No obstante, el fundamento de los principios éticos emana no solo de la moralidad, sino de los derechos fundamentales, específicamente, de la dignidad del ser humano.

Dicho esto, no hay que perder de vista la dimensión y trascendencia de la moral personal. Pese a que un rasgo social sea la sujeción al derecho y a la moral como los dos «sistemas normativos más relevantes en el ámbito de regulación de la vida social» (Garzón Valdés, 1998, p. 19), la autonomía cobra vital relevancia, pues la moral social puede colisionar con las aspiraciones personales de cada individuo.

La Unión Europea, algunos años atrás, ya ha emitido pronunciamientos con la publicación de las Directrices éticas para una IA fiable, elaboradas por un Grupo de expertos de Alto Nivel sobre IA designados por la Comisión Europea. Para asegurar la fiabilidad de la IA, tres componentes han sido identificados, los cuales deben satisfacerse a lo largo del ciclo de vida del sistema:

La IA debe ser lícita, de modo que se que garantice el respeto de todas las leyes y reglamentos aplicables; ii) también ha de ser ética, es decir, asegurar el cumplimiento de los principios y valores éticos; y iii) debe ser robusta, tanto desde el punto de vista técnico como social, puesto que los sistemas de IA, incluso si las intenciones son buenas, pueden provocar daños accidentales. (Grupo de expertos de alto nivel sobre IA, 2019, p. 6)

El documento no se pronuncia respecto al elemento de licitud, pero sí que lo hace respecto al componente ético. Al respecto, el Grupo de expertos (2019) también enarbola principios éticos imperativos sustentados en los derechos fundamentales como derechos morales y legales: i) el respeto a la autonomía humana; ii) prevención del daño; iii) equidad y iv) explicabilidad. Su correcta aplicación e intervención en los sistemas de IA se presenta con miras a potenciar el bienestar individual y colectivo. En otras palabras, los mecanismos de IA deben ayudar a la sociedad y al sistema democrático en base a un diseño centrado en las personas.

Sin duda, el culmen de estas Directrices éticas es el fomento de una IA que lejos de perjudicar a los seres humanos, proteja la dignidad humana e impulse el desarrollo de la tecnología en armonía con los derechos fundamentales, como el empleo, la educación, la seguridad y la privacidad. En ese sentido y de forma adicional, también argumenta que la autonomía humana es igualmente respaldada cuando se requiere que los sistemas de IA actúen como facilitadores de una sociedad democrática, próspera y equitativa, apoyando la acción humana y además contribuyendo la supervisión humana (Grupo de expertos de alto nivel sobre IA, 2019).

Los principios éticos descritos, en buena medida, facilitan una toma de postura frente a las interrogantes planteadas respecto al objetivo de la IA. Así, desde la perspectiva jurídica europea, pero también global como veremos posteriormente, se propicia la explotación de las capacidades de la IA para mejorar la calidad de vida de las personas, impulsar la competitividad de las empresas en el mercado global, la sostenibilidad ambiental y la mejora en servicios públicos.

Al día de hoy, debemos reconocer que actitudes tecnofóbicas no preservarán la humanidad de muchas facetas de nuestra existencia (Campione & Pietropaoli, 2024). Lejos de la intencionalidad de que los sistemas algorítmicos repliquen o reemplacen la inteligencia humana, la propuesta que defendemos es la promoción de tecnologías emergentes, bajo necesarias e imperantes reglas éticas que eviten colisión con derechos; para lo cual se requiere la supervisión por parte de los poderes públicos. Por aditamento, también subrayamos la necesidad de regulación trasnacional que unifique criterios y principios éticos algorítmicos.

A nivel global, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) ha aportado el instrumento normativo Recomendación sobre la ética de la Inteligencia Artificial, publicado en 2022. Este instrumento, elaborado mediante un enfoque mundial, pretende dar una orientación responsable a las tecnologías de la IA en armonía con el derecho internacional, la dignidad humana y el desarrollo. La Recomendación ha sido y continúa siendo trascendental para articular valores y principios éticos, así como su repercusión en los sistemas de IA. Uno de sus objetivos, versa sobre la aplicación voluntaria de los Estados miembro de sus disposiciones y medidas, en particular las medidas legislativas (UNESCO, 2022).

Las instituciones de la Unión Europea también han hecho lo propio, al aprobar diversos informes y documentos como el de *Inteligencia Artificial en la era digital*, en el que se remarca la firme voluntad de la UE de liderar un marco regulador en materia de IA, para lo cual, se asume la necesidad de definir el enfoque y así actuar como referente normativo a escala mundial (Parlamento Europeo, 2022). Esta estrategia reguladora general de la UE, como afirma el catedrático de la Universidad de Sevilla, Dr. Llano Alonso (2024), no solo pretende que Europa sea competitiva, sino que "se generen las condiciones necesarias para establecer un ecosistema de confianza para los ciudadanos, la sociedad en su conjunto y las empresas en relación con las nuevas tecnologías que impulsan la revolución 4.0" (p. 130).

En definitiva, lo que plantea el marco regulador de la UE es que la IA "pueda desplegar todos sus beneficios y se eviten los daños que sin duda pueda generar, ya sea por un diseño defectuoso, por un uso malicioso, o simplemente por la pérdida de control sobre sus resultados" (Peguera Poch, 2020, p. 53). Antes bien, el carácter ético que evocamos está unido a la capacidad libre y consciente del ser humano para distinguir que está bien y que está mal y, en el caso de los algoritmos, *a priori*, a la necesidad de límites en su programación y diseño (Salazar García, 2022).

En esa línea, las instituciones de la UE han enfrentado también la exigencia de adaptación de la legislación europea en materia de IA frente a sus riesgos derivados. En ese derrotero de armonización de normas, el Parlamento mediante Resolución 2020/20142/(INL) consideró que la acción legislativa estaba justificada para evitar la fragmentación de la normativa, pero también puso especial hincapié en que los principios éticos comunes solo son eficaces cuando están también asentados en Derecho y se establece quiénes son los responsables de garantizar, evaluar y supervisar la conformidad de la IA y tecnologías conexas de alto riesgo (Parlamento Europeo, 2020).

El resultado es harto conocido: el Reglamento (UE) 2024/1689 de 13 de junio de 2024, conocido como Reglamento de Inteligencia Artificial, primera norma jurídica de IA del mundo. Así pues, se ha consolidado el liderazgo europeo en materia de inteligencia artificial, pese a que su aplicación se circunscribe a sus Estados miembro. Sin embargo, es claro que esta norma influirá en otros países que también opten por regular de forma específica la inteligencia artificial en sus ordenamientos jurídicos internos.

4. PRINCIPIOS ÉTICOS COMO FUNDAMENTOS DE REGULACIÓN RESPONSABLE

A pesar de la evidente utilidad de la Recomendación de la UNESCO como instrumento internacional y otras iniciativas europeas como la del Libro Blanco sobre la inteligencia artificial que "ofrece alternativas políticas para facilitar un desarrollo de la inteligencia artificial seguro y fiable en Europa, que respete plenamente los valores y los derechos de los ciudadanos de la UE" (Comisión Europea, 2020, p. 3); sí que deviene en im-

prescindible el deseable y necesario equilibrio entre la regulación, la normalización, la competitividad, el consenso internacional y el compromiso de la industria en relación a los sistemas de IA (Muñoz Vela, 2021). La consecuencia es evidente, pues solo de ese modo se evitarán conflictos en la aplicación de sistemas de IA con marcos normativos restrictivos, o del otro lado, proteccionistas, ambos con el mismo resultado perjudicial para los ciudadanos.

De este modo, existe una apremiante obligación, como punto de partida, de delimitar principios éticos que sirvan de hilo conductor en el diseño y funcionamiento de diversos tipos y sistemas de IA. En este punto, es menester diferenciar entre dos tipos de IA: en primer lugar, la IA débil, que se circunscribe a sistemas capaces de realizar tareas específicas, carentes de autonomía y sin capacidad de actuación de forma racional (Muñoz Vela, 2021), como lo son los mecanismos de traducción de idiomas y el reconocimiento de imágenes. La IA fuerte, en cambio, entendida como la IA del futuro, se caracteriza porque el comportamiento 'inteligente' de la máquina estará mucho más desarrollado en tantas competencias como lo podría estar una persona, es decir, que gozaría de cierta autonomía (Salazar García, 2022).

Ya IA NOW (2018), instituto de investigación interdisciplinario de la Universidad de New York abocado a comprender las implicancias sociales de las tecnologías de IA, anticipaba que el desarrollo de sistemas de IA socialmente justos y equitativos requerirá algo más que un lenguaje ético, por bien intencionado que sea. De ahí que el deber inicial de identificar, atenuar y no exacerbar sesgos discriminatorios es una de las piedras angulares en la conformación del marco ético jurídico de los sistemas de IA. De esta manera, en el desarrollo de este trabajo nos ocuparemos de dos principios base, a decir, el principio de precaución, y el de transparencia.

4.1. El principio de precaución

Las primeras apariciones del principio de precaución en ordenamientos jurídicos nacionales se remontan a Suecia a través de la *Swedish Environmental Protection Act* de 1969 y *Vorsorgeprinzip*, principio precursor del principio *per se*, durante el mismo periodo (Sunstein, 2003). Con el paso de los años, es factible identificar la presencia del principio de precaución en normas de derecho internacional, cuyo objeto, como no puede ser otro, está unido al concepto de riesgo y gestión ante posibles daños.

En efecto, la definición general otorgada a este principio viene de la Declaración de Río de 1992, que destaca que cuando haya peligro de daño grave o irreversible, la falta de certeza científica absoluta no deberá utilizarse como razón para postergar la adopción de medidas eficaces en función de los costos para impedir la degradación del medio ambiente. Si bien es cierto que este principio ha sido especialmente aplicado en materia sanitaria y ambiental, ello no obsta a subsumirlo en el marco ético regulador de la IA.

Ahora bien, Europa ha mostrado mayor apertura a aplicar el principio frente a riesgos específicos, en comparación con Estados Unidos (Wiener & Rogers, 2002). La Comunicación de la Comisión UE del 02 de febrero de 2000, de forma certera esboza que su objetivo se orienta a informar a todas las partes interesadas, "el modo en que la Comisión aplica o pretende aplicar el principio de precaución cuando se ve obligada a tomar decisiones relativas al control de riesgo", aunque también asevera que el documento "no pretende poner punto final al debate, sino contribuir a la reflexión en curso tanto a escala comunitaria o internacional" (Comisión Europea, 2000, p. 8). De esta manera, la UE evita caer en proteccionismos innecesarios, a través de la interacción con los criterios de otros principios como los de proporcionalidad, no discriminación y coherencia.

En esa línea, el principio también ha sido respaldado por la UE con su incorporación al derecho originario. El Artículo 191 del Tratado de Funcionamiento de la Unión Europea (TFUE), estipula que la política de la UE en el ámbito de medio ambiente busca un nivel de protección elevado. Conviene precisar paralelamente, que la dimensión del principio de precaución "va más allá de las problemáticas asociadas a los riesgos a corto o medio plazo, puesto que se refiere también a cuestiones a largo plazo e incluso ligadas al bienestar de las generaciones futuras" (Comisión Europea, 2000, p. 7).

Roberto Andorno (2004) enumera una serie de condiciones para su puesta en práctica: i) situación de incertidumbre acerca del riesgo; ii) evaluación científica del riesgo, iii) perspectiva de un daño grave o irreversible; iv) proporcionalidad de las medidas; v) transparencia de las medidas y, vi) inversión de la carga de la prueba. En materia de sistemas de inteligencia artificial, se requiere también la evaluación de tipos de riesgo. De acuerdo con ello, una correcta aplicación del principio se traduce en la inexistencia de resultados adversos por parte de los algoritmos.

No le falta razón a Rafael de Asís cuando afirma que "hay que ser conscientes de que los sesgos existen con y sin inteligencia artificial. Es más, gracias a ella puede ser posible identificar estos sesgos y eliminarlos" (De Asís, 2024, p. 31). En efecto, una correcta práctica precautoria puede dar resultados favorables respecto a uno de los principales riesgos en el uso de IA, que es el de aparición y/o perpetuación de sesgos discriminatorios, a través de la evaluación científica de riesgos y otras políticas anticipatorias. Por ahora, no obstante, algunos sistemas de IA crean nuevas barreras de accesibilidad digital que afectan a colectivos vulnerables como el de las personas con discapacidad (Álvarez García, 2023).

Así pues, las discriminaciones producidas por sesgos pueden tener diferentes orígenes, ya que pueden provenir de los datos, así como también de los desarrolladores que, de forma inconsciente, o no, los trasladan a los algoritmos. En ambos casos, los algoritmos reflejan los sesgos de la sociedad; por lo que insistimos que evaluaciones de impacto en los derechos fundamentales previas al despliegue de sistemas de IA con riesgo alto, pueden ser una herramienta útil ante el potencial riesgo de vulneraciones de derechos a

la igualdad, a la salud e integridad física y psíquica, el acceso al empleo, a la educación y a la administración de justicia sólo por mencionar algunos.

Sin duda, otro riesgo latente en el uso de sistemas de IA dimana de la afectación de derechos fundamentales que garantizan la dimensión privada de las personas, es decir, la protección de datos, la privacidad y la propia imagen. Las Administraciones públicas no han sido ajenas a la revolución digital; puesto que también hacen uso de algoritmos para optimizar servicios. Por un lado, están los *chatbots* o asistentes virtuales que requieren datos personales de los ciudadanos para la tramitación de procedimientos administrativos o consultas adicionales, y por otro, los sistemas de video vigilancia en espacios públicos que hacen uso de reconocimiento facial y que además puede recopilar información biométrica.

Ciertamente, una manifestación del principio precautorio se traduce en las (auto) evaluaciones de impacto algorítmico. Estas involucran, como afirma Simón Castellano, "a los principales actores vinculados con la iniciativa, diseño, desarrollo, comercialización y/o implementación de los sistemas de IA" (2023, p. 28) y su objetivo se traslada a la minimización de los efectos negativos de los algoritmos tanto en los derechos fundamentales como al riesgo latente de discriminación, sea por motivos de género, de estatus migratorio, discapacidad, etc.

En este escenario y nuevamente en el ámbito europeo, el Reglamento General de Protección de Datos (RGPD) hace frente a estos riesgos a través de medidas y principios, dentro de los cuales se puede inferir que está el de precaución. Aunque no lo regula de manera directa, el artículo 24, sobre la responsabilidad del encargado del tratamiento, dispone que, teniendo en cuenta los riesgos de diversa probabilidad y gravedad para los derechos y libertades de las personas físicas, el responsable del tratamiento aplicará técnicas y organizativas apropiadas a fin de garantizar y poder demostrar que el tratamiento es conforme al RGPD.

Ergo, se aprecia una alineación con la lógica precautoria en la anticipación de riesgos frente una posible colisión en el tratamiento de datos con la privacidad o seguridad de los ciudadanos. Si bien es cierto que el principio de precaución presenta cuestionamientos relativos a su costosa aplicación o la limitación al desarrollo tecnológico ante riesgos especulativos (Sunstein, 2009); ello no impide que su aplicación sea provechosa, máxime cuando está supeditada a requisitos de proporcionalidad y también a la interacción con otro principio ético crucial como lo es indudablemente, el principio de transparencia, que abordamos a continuación.

4.2. El principio de transparencia

El principio de transparencia es usado en diferentes ámbitos del derecho, administrativo, internacional, constitucional, pero también es susceptible de ser aglutinado en otras ciencias como la economía y la política. De acuerdo con ello, Cotino Hueso (2022)

esgrime que la transparencia "es, pues, una cualidad instrumental en beneficio de unos valores, normas o finalidades" (p. 27), y que en el ámbito de la IA "tiene un contenido objetivo variable, además de que el ordenamiento jurídico la reconoce como una obligación para distintos sujetos" (p. 28).

En esa línea, en la era de automatización de procesos y de análisis de datos generada por la relación simbiótica *machine learning–big data*, se torna indispensable abordar una nueva vertiente del principio: la transparencia algorítmica. Esta no solo constituye un imperativo ético y técnico, sino que además, su adjetivación normativa permite delinearla como un estándar jurídico que exige conocer qué datos se utilizan, cómo se utilizan, quiénes lo utilizan, para qué los utilizan y cómo se llega a partir de los datos a tomar las decisiones que afectan a la esfera vital de quien reclama la transparencia (Sangüesa, 2018).

Por otro lado, y volviendo al principio rector *per se*, la Recomendación de la UNESCO sobre el uso de IA, establece que la transparencia es necesaria para que los regímenes nacionales e internacionales pertinentes en materia de responsabilidad funcionen eficazmente, y que tiene por objetivo:

Proporcionar información adecuada a los respectivos destinatarios para permitir su comprensión y fomentar la confianza. En el caso específico de los sistemas de IA, la transparencia puede permitir a las personas comprender cómo se implementa cada etapa en un sistema de IA, en función del contexto y la sensibilidad del sistema en cuestión. (UNESCO, 2022, p. 22)

Consecuentemente, el núcleo central de la aplicación de este principio ético no solo implica el conocimiento de las decisiones tomadas por los mecanismos algorítmicos, sino también del contexto de prácticas sociales y políticas que los rodean (Valle Escolano, 2022). De este modo, el principio de transparencia nos sumerge en la comunicación entre los poderes públicos y la ciudadanía en garantía del derecho de acceso a la información algorítmica, —estrechamente vinculado con la explicabilidad— a lo largo del ciclo de vida de la IA. Además, se erige como un mecanismo de control de sesgos e individualización de responsables en eventuales procedimientos sancionadores.

En resumidas cuentas, el desconocimiento o la falta de comprensión por parte de los responsables humanos sobre los resultados generados por la IA pueden revelar deficiencias en la aplicación del principio de transparencia. La opacidad algorítmica, en particular, puede ser muy compleja y revela también un desafío significativo, puesto que los resultados obtenidos mediante algoritmos de *machine learning* pueden insertarse bajo una lógica de 'caja negra'. Como afirma Araya Paz (2021) esta circunstancia evidencia la inminente dificultad humana para comprender plenamente el proceso de toma de decisiones de estos sistemas.

En el marco del Reglamento de Inteligencia Artificial de la UE se disponen obligaciones de transparencia para los proveedores de modelos de IA de uso general a lo largo de

la cadena de valor de la IA que incluye elaborar documentación, mantenerla actualizada y facilitar información sobre el modelo de IA para su uso a favor de proveedores posteriores. Del mismo modo, esta documentación debe ponerse a disposición, previa solicitud, de la Oficina de IA y de las autoridades nacionales competentes (Reglamento UE, 2024/1689, considerando 101).

En el caso de sistemas de alto riesgo, el Reglamento estipula que se diseñarán y desarrollarán de un modo que se garantice que funcionan con un nivel de transparencia suficiente para que los responsables del despliegue interpreten y usen correctamente sus resultados de salida; y que se garantizará un tipo y un nivel de transparencia adecuados para el proveedor y el responsable del despliegue (Reglamento UE, 2024/1689, artículo 13). Queda plenamente delimitado entonces, que este principio se dimensiona y concreta en el acceso a datos e información, mediante un equilibrado balance entre estos y el impacto del algoritmo en el resultado de decisiones.

También la opacidad puede estar presente en los algoritmos usados por las Administraciones públicas por diversos motivos. Cerrillo i Martínez (2020) argumenta que, un motivo de limitación del acceso a la información puede justificarse en pos de proteger bienes o derechos como la propiedad intelectual, pero también puede ser por causas técnicas u organizativas. Floridi (2024) en cambio, enfatiza que la falta de transparencia puede deberse a la maleabilidad de los algoritmos, en el sentido que pueden reprogramarse de forma continua y dinámica, lo cual puede recaer en una excesiva difuminación respecto a su utilidad y las oportunidades que brinda.

Uno de los casos más relevantes sobre falta de transparencia se asocia a los procedimientos de contratación utilizados por algunas empresas. En 2015, el sistema de IA utilizado por *Amazon* para mecanizar el proceso de selección de talento reveló que el sistema no era neutral; puesto que prefería contratar hombres que mujeres. Expuesto el sesgo discriminatorio, la empresa ya no podía confiar en él, por lo que el proyecto finalmente fue disuelto (Dastin, 2018). Cabe destacar que Amazon no reveló mayor información sobre su algoritmo reclutador, en clara vulneración del derecho a la igualdad y no discriminación.

Por otro lado, el sesgo de automatización es otro de los claros enemigos de la transparencia. La excesiva confianza depositada en los algoritmos irrumpe en la esfera reguladora. Los Principios de Asilomar sobre IA (2017) ya identificaban que los sistemas de IA no son infalibles, y que además de fallar, pueden causar perjuicios en el proceso de análisis de datos de entrada, el proceso y los resultados de salida. Otro modo de empleo de la transparencia algorítmica, en el ejemplo anterior, también incluye identificar las causas y el momento del error, para así superar el fallo y evitar su repetición en el futuro.

A la postre, la naturaleza del principio tradicional de transparencia ha sido discutida en algunas ocasiones, puesto que algunos autores enfatizan que "no es un principio ético

en sí mismo, sino una condición pro-ética para permitir o perjudicar otras prácticas o principios éticos" (Turilli & Floridi, 2009, p. 111). Sin embargo, su trascendencia es innegable, al sumarle el conocimiento y la explicabilidad, entendiendo este último como un elemento más de la transparencia (Vestri, 2022).

5. HACIA UN MARCO JURÍDICO GLOBAL VINCULANTE Y UNA GOBERNANZA ALGORÍTMICA LEGÍTIMA

Ante las dos cuestiones planteadas respecto al fin último de la IA, nos hemos decantado por la primera. La inteligencia artificial se ha insertado en la sociedad para transformarla notablemente y por ello, debemos reconocer las enormes ventajas y oportunidades de desarrollo que ofrecen los algoritmos. Precisamente porque la IA ha llegado para permanecer y contribuir a la mejora de la calidad de vida humana como herramienta de apoyo, el derecho no puede mantenerse al margen de esta evolución digital.

La Unión Europea ha adoptado una visión antropocéntrica, en la que los sistemas digitales están al servicio de las personas. Esta perspectiva refleja un compromiso con la promoción de competencias digitales, así como la innovación, impulso de la economía global y la sostenibilidad ambiental. Para lograr estos objetivos, el primer paso ha sido el establecimiento de principios éticos que rigen la IA, dado que se trata de una tecnología diferencial que goza de competencias como la capacidad de aprender por sí misma, la generación de predicciones y la toma de decisiones (Belloto Gómez, 2022).

Cabalmente, el Reglamento de Inteligencia Artificial de la UE supone un impulso para que la ONU replique una normativa universal, en función a que conforme la tecnología va desarrollándose, los principios éticos requieren también evolucionar a obligaciones jurídicas. En ese orden de ideas, Megías Quirós (2024) apunta a que la posibilidad de alcanzar el acuerdo de regulación transnacional universal en el marco de las Naciones Unidas presenta algunas dificultades, puesto que "quienes tratan de promover iniciativas carecen de peso suficiente para vencer las reticencias visibles en el seno de la organización" (p. 64).

Sin circunloquios, y para responder la interrogante inicial, recalamos que la ética es indispensable en todo el ciclo de vida de los sistemas de IA, del mismo modo como lo ha sido a lo largo del tiempo en la sociedad contemporánea. El marco ético es condición imperante, pero no resulta suficiente si no viene acompañado de legislación que lo fortalezca. Ya el Informe del Alto Comisionado para los Derechos Humanos de las Naciones Unidas de 03 de agosto de 2018 sobre el derecho a la privacidad en la era digital exponía que este derecho "es fundamental para el goce y el ejercicio de los derechos humanos dentro y fuera de internet. Constituye uno de los fundamentos de la sociedad democrática y tiene un papel clave en la realización de una amplia gama de derechos humanos" (Organización de las Naciones Unidas, 2018, s/n).

Empero, las conclusiones del Informe únicamente remiten al marco internacional de derechos humanos como "base sólida para formular respuestas a los múltiples desafíos que se plantean en la era digital" (2018, s/n). Al día de hoy, los Informes y recomendaciones no vinculantes emitidos por la ONU, no se condicen estrictamente con el propósito de protección de los derechos humanos en una realidad algorítmica sesgada inevitable.

En efecto, las tecnologías emergentes de la era digital repercuten en la dimensión objetiva de los derechos, por lo que es de crucial envergadura, en la esfera jurídica internacional "revisar los marcos jurídicos que desarrollan y protegen los derechos de los ciudadanos y disciplinan el diseño de políticas públicas, poner fin a lagunas y fijar condiciones básicas que permitan la transformación digital ordenada y eficiente" (Barrio Andrés, 2022, p. 388).

El alcance limitado del cúmulo de informes y recomendaciones no vinculantes en cierta medida, denotan incuria por parte de las Naciones Unidas. El punto de llegada en materia ético-algorítmica, es la configuración y armonización de los preceptos de la Recomendación de la UNESCO para una IA fiable en un marco ético universal cuya vocación es su reconocimiento posterior en las legislaciones nacionales, sin perjuicio de la adopción de un instrumento normativo de derecho internacional obligatorio para todos los Estados miembro de la ONU.

Así pues, el marco regulatorio universal debe garantizar la explicabilidad de los posibles errores de los sistemas de IA. De acuerdo con el acápite de la Recomendación, denominado "Ámbitos de acción política", la UNESCO realza que el proceso de elaboración de políticas o medidas deben ser eficaces, como por ejemplo un mecanismo normativo que vele por todas las personas interesadas, que incluye a las empresas del sector privado, las instituciones universitarias y de investigación, y la sociedad civil. El objetivo es que todas ellas elaboren instrumentos de evaluación de impacto en los derechos humanos, el estado de derecho, la democracia y la ética (UNESCO, 2022, s/n).

Lo anteriormente mencionado, no es óbice para que los gobiernos, paralelamente, también adopten un marco regulador que establezca procedimientos para que las autoridades públicas, en particular, lleven a cabo evaluaciones del impacto ético de los sistemas de IA para atenuar riesgos, evitar consecuencias perjudiciales y facilitar la participación de los ciudadanos (UNESCO, 2022, s/n). A través de estos parámetros, es posible apreciar que la aplicación de los principios éticos anteriormente mencionados se condice con el principio de equidad y no discriminación en defensa de la dignidad del ser humano, con la rendición de cuentas como manifestación *ex nunc* del principio de transparencia.

A manera de colofón, el eje vertebrador de un sistema de gobernanza de la IA es la apertura hacia una aplicación homogénea de un marco jurídico global que delimite principios éticos que deberán cumplir todos los sistemas y tecnologías de IA. Esto en buena parte, requiere del consenso obtenido entre los operadores políticos y las instituciones privadas creadoras de sistemas de IA, pero sustancialmente, constituye una

garantía propia de los Estados de Derecho en defensa de la democracia y los derechos humanos.

CONCLUSIONES

La omnipresencia de los sistemas de inteligencia artificial es percibida en la observación de conductas, en la predicción de necesidades o preferencias y también en el proceso de sugerencias y toma de decisiones atribuido principalmente al *machine learning*. Como consecuencia directa, la vida cotidiana ya no solo gira en torno a vivencias físicas, sino también a vivencias digitales.

Este giro copernicano ha revolucionado la sociedad del siglo XXI, al insertar nuevas formas de convivencia y también una nueva forma de ejercicio de derechos fundamentales. Así también, la incorporación de los sistemas de IA en las Administraciones públicas ha desentrañado variaciones en el quehacer jurídico debido a la automatización de procesos. Del otro lado de esta funcionalidad, algunos riesgos ya han sido detectados en la utilización de tecnologías emergentes en perjuicio de los ciudadanos o colectivos vulnerables. Estos riesgos ya no solo versan respecto a la afectación de los derechos del ámbito personal, como lo son el derecho a la privacidad, seguridad y libertad de expresión, sino que también vulneran principalmente el derecho a la igualdad y no discriminación.

La evolución de los algoritmos, desde el *machine learning* hasta el *deep learning* han potenciado el entorno digital y la toma de decisiones automatizadas. La aparente autonomía y la capacidad de influencia de estos sistemas exige un marco ético sólido, lo que implica el sometimiento de los sistemas de IA a directrices como la transparencia, la precaución, la supervisión, así como la explicabilidad. En ese orden de ideas, el horizonte prioritario a tener en cuenta en la era digital es que las máquinas no son perfectas, y por ello, se requiere la supervisión e intervención humana en mérito a preceptos éticos durante todo el ciclo de vida de la IA.

A la luz de tales consideraciones, el avance de los sistemas de IA demuestra que su impacto es social y político, pero también jurídico. El derecho en ese sentido, como ciencia reguladora de la acción humana libre, no puede hacer caso omiso a las transformaciones digitales bajo un pseudo principio de neutralidad. Los marcos éticos, en ese orden de ideas, requieren evolucionar en exigencias normativas con criterios axiológicos universales y paralelamente esta regulación debe, a futuro, trascender del ámbito nacional y convertirse en un esfuerzo global que garantice la responsabilidad algorítmica.

La implementación de principios como el de precaución, es vital para una correcta anticipación de riesgos en el uso de sistemas algorítmicos. Si bien no es posible negar que ciertas situaciones de incertidumbre ya han sido superadas por la inminente aparición de sesgos algorítmicos, la gestión y control de otros riesgos reafirma la necesidad

del principio. A partir de evaluaciones de impacto previas y posteriores realizadas por los gobiernos, se asume el reto de contrarrestar efectos adversos como mecanismo preventivo.

De otro lado, la transparencia algorítmica como bifurcación del principio de transparencia tradicional se sitúa como un requisito para la construcción de confianza en la inteligencia artificial. Este principio, sustancialmente, permite conocer los criterios de funcionamiento y toma de decisiones de los sistemas automatizados que lo convierte en un instrumento de control de arbitrariedades o discriminaciones algorítmicas. *A priori*, su aplicación promueve el fortalecimiento del derecho a la información y también es un resguardo para la conservación de valores esenciales de la sociedad democrática como la justicia, la dignidad humana y la igualdad.

De manera temprana, la Unión Europea ha recogido la perspectiva de uso de IA centrada en la fiabilidad, seguridad e inclusión, al reconocer que su efectividad dependía de su capacidad para armonizar estrategias normativas. En ese contexto, sus instituciones han asumido el reto de regular el funcionamiento de los sistemas algorítmicos, preliminarmente desde el Libro Blanco sobre IA y otros instrumentos emitidos por la Comisión Europea hasta el reciente Reglamento de Inteligencia Artificial. A través de estas normas, la UE promueve el enfoque basado en la protección de los derechos fundamentales, impulsando estándares que pueden seguir la lógica del efecto Bruselas y en consecuencia, ser replicados a nivel internacional. La trascendencia de esta normativa radica en el potencial para inspirar legislaciones a través del fomento del modelo de gobernanza algorítmica alineados con la responsabilidad de control humano sobre los sistemas inteligentes.

La capacidad innovadora de algunos sistemas, en especial los de *deep learning*, hace que los ciudadanos deban contar con información comprensible para entender los datos del sistema, así como los datos de salida como una expresión de transparencia. En mérito a ello, la regulación de la inteligencia artificial no puede limitarse a un mero marco técnico: debe consolidarse como una herramienta jurídica con vocación humanista, que garantice que el desarrollo algorítmico se subordine siempre al respeto de los derechos fundamentales y a los valores democráticos que nos definen como sociedad.

REFERENCIAS BIBLIOGRÁFICAS

- Álvarez, H. (2023). *La discriminación algorítmica por razón de discapacidad*, *Revista Internacional de Pensamiento Político*, 18, 17-41.
- Andorno, R. (2004). *Validez del principio de precaución como instrumento jurídico para la prevención y la gestión de riesgo*, en: C. M. Romeo Casabona (ed.), *Principio de precaución, biotecnología y derecho*, Granada: Editorial Comares S.L.

- Araya, C. (2021). *Transparencia algorítmica, ¿un problema normativo o tecnológico?*, *CUHSO (Temuco)*, 2(31), 306-334.
- Barrio Andrés, M. (2022). Garantía de los derechos en los entornos digitales, en L. Cotino Hueso (coord.), *La carta de los derechos digitales*, (pp. 363-393). Valencia: Tirant lo Blanch .
- Belloto, J. (2022). Explicabilidad e inteligencia artificial. Recomendaciones "aterizadas" al día a día de las empresas, en L. Cotino Hueso y J. Castellanos Claramunt (eds.), *Transparencia y explicabilidad de la inteligencia artificial*, (pp. 199-224). Valencia: Tirant lo Blanch.
- Campione, R. y Stefano P. (2024). *Los artefactos de la inteligencia jurídica: personas y máquinas*, Madrid: Editorial Dykinson S.L.
- Cerrillo, A. (2020). El impacto de la inteligencia artificial en las administraciones públicas: estado de la cuestión y una agenda, en A. Cerrillo i Martínez y M. Peguera Poch (coords.), *Retos jurídicos de la inteligencia artificial*, (pp.75-92). Cizur Menor (Navarra): Thomson Reuters Aranzadi.
- Comisión Europea. (2000). Comunicación de la Comisión sobre el recurso al principio de precaución (COM(2000) 1 final). [en línea] <<https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52000DC0001>>.
- Comisión Europea. (2020). *Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza* (COM(2020) 65 final). <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52020DC0065>.
- Conferencia de las Naciones Unidas sobre el Medio Ambiente y Desarrollo (1992). *Declaración de Río*.
- Cotino, L. (2022). Transparencia y explicabilidad de la inteligencia artificial y "compañía" (comunicación, interpretabilidad, inteligibilidad, auditabilidad, testabilidad, comprobabilidad, simulabilidad...) Para qué, para quién y cuánta, en L. Cotino Hueso y J. Castellanos Claramunt (eds.), *Transparencia y explicabilidad de la inteligencia artificial*, (25-70). Valencia: Tirant lo Blanch.
- Dastin, J. (2018). Amazon abandona un proyecto de IA para la contratación por su sesgo sexista. 14 de octubre *Reuters*. [en línea], <<https://www.reuters.com/article/world/amazon-abandona-un-proyecto-de-ia-para-la-contratacin-por-su-sesgo-sexista-idUSKC-N1M00M4/>>. [Consulta: 16/03/2025.]
- De Asís, R. (2024). Inteligencia artificial y derechos humanos, en F. Ibáñez López-Pozas (coord.), *Inteligencia artificial: los derechos humanos en el centro*, (pp. 17-32). Madrid: Editorial Dykinson S.L.
- Delgado, J. (2020). *Judicial-Tech, el proceso digital y la transformación tecnológica de la justicia. Obtención, tratamiento y protección de datos en la justicia*, Madrid: Wolters Kluwer España S.A.
- Domínguez, J. L. (2024). *Iusalgoritmia: las reglas del derecho y la inteligencia artificial*, Valencia: Tirant lo Blanch.

- Fernández, E. (1990). *Estudios de ética jurídica*, Madrid: Editorial Debate.
- Floridi, L. (2024). *Ética de la inteligencia artificial. Principios, retos y oportunidades*, Barcelona: Herder Editorial S.L.
- Future of Life Institute. (2017). *Principios de Asilomar sobre IA*.
- Garzón, E. (1998). Derecho y moral, en R. Vázquez (comp.), *Derecho y moral*, (pp. 19-55). Barcelona: Editorial Gedisa S.A.
- Grupo de Expertos de Alto Nivel sobre IA. (2019). *Directrices éticas para una inteligencia artificial fiable*. Comisión Europea. <https://digital-strategy.ec.europa.eu/es/library/ethics-guidelines-trustworthy-ai>
- Gutiérrez, E. (2024). *Inteligencia artificial y derechos fundamentales. Hacia una convivencia en la era digital*, A Coruña: Editorial Colex S.L.
- Hernández, J. C. (2022). *El marco jurídico de la inteligencia artificial. Principios, procedimientos y estructuras de gobernanza*, Cizur Menor, Navarra: Thomson Reuters Aranzadi.
- IA NOW Institute (2018). *AI NOW Report 2018*. New York University.
- Llano, F. (2024). *Homo ex machina. Ética de la inteligencia artificial y derecho digital ante el horizonte de la singularidad tecnológica*, Valencia: Tirant lo Blanch.
- Megías, J. (2024). La necesidad de una regulación transnacional de la inteligencia artificial, en A. Ballesteros Soriano (coord.), *Inteligencia artificial y derechos: perspectivas iusfilosóficas*, (pp. 63-82). Madrid: Editorial Aranzadi S.A.U.
- Muñoz, J. (2021). *Cuestiones éticas de la inteligencia artificial y repercusiones jurídicas. De lo dispositivo a lo imperativo*, Pamplona: Thomson Reuters Aranzadi.
- Organización de las Naciones Unidas (2018). *Informe del Alto Comisionado para los Derechos Humanos sobre el derecho a la privacidad en la era digital (A/HRC/39/29)*.
- Palma, A. (2022). *Decisiones automatizadas y protección de datos. Especial atención a los sistemas de inteligencia artificial*, Madrid: Editorial Dykinson S.L.
- Parlamento Europeo (2020). Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre el marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL)), <https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_ES.html>.
- Parlamento Europeo (2022). Informe sobre la Inteligencia Artificial en la era digital. https://www.europarl.europa.eu/doceo/document/A-9-2022-0088_ES.html.
- Parlamento Europeo y Consejo de la Unión Europea. (2024), Reglamento (UE) 2024/1689 de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las

- Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial). <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689>.
- Peces-Barba, G. (1995). *Ética, poder y derecho. Reflexiones ante el fin de siglo*, Madrid: Centro de Estudios Constitucionales.
- Peguera, M. (2020). En búsqueda de un marco normativo para la inteligencia artificial, en A. Cerrillo i Martínez y M. Peguera Poch (coord.), *Retos Jurídicos de la inteligencia artificial*, (pp. 41-56). Cizur Menor (Navarra): Thomson Reuters Aranzadi.
- Presno, M. (2022). *Derechos fundamentales e inteligencia artificial*, Madrid: Marcial Pons.
- Real Academia Española. (2024). Definición de inteligencia artificial. <https://dle.rae.es/inteligencia>.
- Russel, S. y Norvig, P. (2004). *Inteligencia artificial: un enfoque moderno*, Madrid: Pearson Educación.
- Salazar, I. (2022). Retos actuales de la ética en la Inteligencia Artificial, en L. Cotino Hueso (dir.), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, (pp. 53-66). Cizur Menor, Navarra: Thomson Reuters Aranzadi.
- Sangüesa, R. (2018). Inteligencia artificial y transparencia algorítmica: "It's complicated", *BiD: textos universitaris de biblioteconomia i documentació* (41). <https://bid.ub.edu/es/41/sanguesa.htm>.
- Simón, P. (2023). Las evaluaciones de impacto algorítmico en los derechos fundamentales: hacia una efectiva minimización de riesgos, en L. Cotino Hueso y J. Castellanos Claramunt (eds.), *Algoritmos abiertos y que no discriminen en el sector público*, (pp. 27-56). Valencia: Tirant lo Blanch.
- Sunstein, R. (2003). Beyond the precautionary principle, *University of Pennsylvania Law Review*, 151(3), 1003-1058.
- Sunstein, R. (2009). *Leyes del miedo. Más allá del principio de precaución*, Madrid: Safekat S.L.
- Turilli, M. y Floridi, L. (2009). The ethics of information transparency, *Ethics and Information Technology*, 11(2), 105-112.
- UNESCO. (2022). *Recomendación sobre la ética de la inteligencia artificial*. París: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura.
- Valle, R. (2022). Transparencia en la inteligencia artificial y en el uso de los algoritmos: una visión de género, en L. Cotino Hueso y J. Castellanos Claramunt (eds.), *Transparencia y explicabilidad de la inteligencia artificial*, (pp. 85-107). Valencia: Tirant lo Blanch.
- Vestri, G. (2022). Denegación vs. derecho de acceso al código fuente en los sistemas algorítmicos, en L. Cotino Hueso y J. Castellanos Claramunt (eds.), *Transparencia y explicabilidad de la inteligencia artificial*, (pp. 111-133). Valencia: Tirant lo Blanch.
- Wiener, J. y Rogers M. (2002). Comparing precaution in the United States and Europe, *Journal of Risk Research*, 5, 217-249.

