

ENCUESTA: INTELIGENCIA ARTIFICIAL Y DERECHOS FUNDAMENTALES

Teoría y Realidad Constitucional (nº 55, 1º semestre de 2025)

Acceso a la encuesta completa en

<https://revistas.uned.es/index.php/TRC/issue/view/1977>

(versión relativa a las respuestas de diciembre 2024 de Lorenzo Cotino Hueso, Catedrático de Derecho Constitucional Universidad de Valencia)

Contenido

- 1. ¿Qué consideraciones generales puede hacernos sobre la incidencia de la inteligencia artificial en los derechos fundamentales reconocidos en la Constitución? ¿Hay algún derecho en particular o grupo de derechos en los que, a su juicio, esa incidencia tenga especial impacto?** 2
- 2. De forma más concreta ¿qué reflexiones puede hacernos sobre las consecuencias de la inteligencia artificial en los derechos fundamentales implicados en los procesos democráticos y, por tanto, en esos mismos procesos?** 4
- 3. Desde el punto de vista de la dogmática de los derechos fundamentales ¿Cree que la inteligencia artificial puede llegar a producir algún cambio en la forma en la que concebimos hoy los derechos fundamentales (su objeto, contenido, estructura...)?** 7
- 4. ¿Considera preciso incorporar a la Constitución algún nuevo derecho fundamental para poder hacer frente a los riesgos que comporta la inteligencia artificial?** 9
- 5. En un contexto de creciente presencia de los sistemas de inteligencia artificial ¿sería conveniente o necesario prever nuevos instrumentos de garantía de los derechos fundamentales?** 12
- 6. Desde la perspectiva de los derechos fundamentales ¿qué opinión le merecen el Reglamento europeo de inteligencia artificial y el Convenio Marco del Consejo de Europa sobre inteligencia artificial y derechos humanos, democracia y Estado de derecho?** 15

1. ¿Qué consideraciones generales puede hacernos sobre la incidencia de la inteligencia artificial en los derechos fundamentales reconocidos en la Constitución? ¿Hay algún derecho en particular o grupo de derechos en los que, a su juicio, esa incidencia tenga especial impacto?

Creo que no es nada exagerado afirmar que la inteligencia artificial (IA) supone un punto de inflexión en la historia de la humanidad, aunque nos falta suficiente perspectiva en este momento. No se trata simplemente de una evolución tecnológica más, sino de una auténtica revolución estructural que redefine las bases sobre las que descansa la sociedad, la economía, la política y, cómo no, los derechos fundamentales. La Constitución fue concebida en un mundo analógico, un mundo donde los riesgos eran previsibles, los actores estaban claramente identificados y los problemas podían abordarse dentro de un marco jurídico relativamente estable. Sin embargo, la IA ha desdibujado esos límites.

Con la llegada de Internet, los derechos más afectados parecían ser aquellos relacionados con la esfera personal. Con la IA el escenario es más complejo y de más amplio espectro. Y el fuerte impacto queda repartido en -prácticamente- todos los derechos fundamentales con proyección individual y, sobre todo, social y colectiva. Ahora bien, antes de un breve recorrido por el impacto -negativo- que puede tener la IA en los derechos, es asumir necesario -también para el Derecho- que la IA puede ser muy beneficiosa. Esta dimensión positiva y favorable rara vez se subraya desde una perspectiva jurídica constitucional. Se trata de asentar una premisa jurídica: reconocer el deber constitucional de progreso tecnológico como un principio jurídico fundamental, como por ejemplo Barrilao ha afirmado en general y hay que insistir ahora en particular. Así, cabe apreciar la innovación tecnológica y en particular la IA no como un fin en sí misma, sino un medio para mejorar las condiciones de vida de las personas y garantizar sus derechos fundamentales así como en su caso el Estado democrático y social de Derecho. Este deber puede traducirse en un deber de facilitar las condiciones para el uso de IA, con proyecciones específicas para el sector público. Es más, al igual que ocurrió con el acceso a Internet, no es descartable que en un futuro próximo surja la exigencia de un “derecho a la IA”, entendido como un derecho a beneficiarse de sus aplicaciones en ámbitos como la educación, la salud, la sostenibilidad o la justicia. Por ello, el Derecho también debe proteger el desarrollo de la IA y fomentar la innovación, al tiempo que la orienta y contiene.

Asentada la anterior premisa jurídica y pese a las ventajas de la IA, el jurista no puede ignorar que también presenta un lado más oscuro por el riesgo e impacto que implica para derechos fundamentales. Más allá de los riesgos generales que presenta la digitalización, internet o el software en general, hay que tener especialmente presente el plus de peligrosidad que implica la IA por su capacidad de aprendizaje, su autonomía y su adaptabilidad a contextos

cambiantes. Los humanos que diseñan la IA no son perfectos y pueden crear sistemas con errores y sesgos y, obviamente, también pueden actuar con malas intenciones. Además, la IA y los sistemas autónomos con autoaprendizaje adquieren cierta independencia de sus creadores. Así las cosas, la IA, los sistemas automatizados, el reconocimiento facial y tecnologías afines suponen *disparar y con fuego racheado* a los derechos fundamentales. El Reglamento de IA de la UE en su Considerando 48 menciona hasta veintiún derechos fundamentales que pueden verse afectados por sistemas de alto riesgo, además de derechos específicos de menores.

Los algoritmos generalmente dependen de un consumo masivo de datos, y la gestión de estos datos suele centrarse en la privacidad y protección de los mismos. Resulta de interés la transparencia del uso de esta tecnología, qué datos han servido para su entrenamiento, validación y concreto caso de uso, el control y en su caso consentimiento informado sobre los mismos o el cumplimiento de principios esenciales como la minimización y calidad de los datos. De igual forma, la intimidad se ve muy afectada por sistemas de reconocimiento facial, vigilancia automatizada y el análisis predictivo de comportamientos individuales. El uso de estas tecnologías por parte de empresas y administraciones públicas puede generar escenarios de vigilancia masiva incompatibles con un Estado de derecho. Los *deepfakes* y los sistemas de manipulación de contenido también impactan el honor y la imagen propia. Como se mencionará, además de los derechos de la personalidad, quedan en riesgo la democracia y la confianza en la información. La IA nos categoriza y clasifica, lo cual tiene potenciales efectos en innumerables casos de uso.

Más allá de los derechos de la personalidad, uno de los derechos más afectados es, sin duda, el derecho a la igualdad y a la no discriminación; como he podido analizar con profundidad, los algoritmos de IA, aunque se presenten como herramientas neutrales y objetivas, no están exentos de sesgos y errores. Estos sesgos fácilmente se proyectan en individuos o en colectivos generando tratamientos diferentes que pueden no estar suficientemente justificados. Es más, hay un especial peligro de que los sistemas IA incidan en tradicionales colectivos especialmente prohibidos de discriminar. O lo que es también negativo, es fácil que los efectos de la IA se proyecten en nuevos colectivos que generan los algoritmos, que no coinciden con los tradicionales, pero sí que pueden implicar una especial vulnerabilidad. Estos sesgos o errores de los algoritmos pueden surgir, de un lado, de la falta de calidad y gobernanza los datos utilizados para entrenar o implementar los modelos. Del otro lado, pueden surgir por los problemas de robustez, exactitud y fiabilidad del sistema autónomo de IA. Como se ha señalado, el sistema autónomo que de IA que aprende y es adaptable, bien puede adoptar derivas o sesgos. Estos tratamientos diferenciados pueden tener consecuencias que son especialmente graves cuando se utilizan en ámbitos como el empleo, la educación, la concesión de créditos bancarios, las prestaciones, los contratos y ayudas sociales o la

determinación de penas. Las discriminaciones algorítmicas, ya sean conscientes o no, pueden perpetuar y amplificar desigualdades estructurales preexistentes.

En cualquier caso, como ahora se expone, la IA impacta colectiva y supraindividualmente en nuestros derechos y en la democracia.

2. De forma más concreta ¿qué reflexiones puede hacernos sobre las consecuencias de la inteligencia artificial en los derechos fundamentales implicados en los procesos democráticos y, por tanto, en esos mismos procesos?

Internet y las redes sociales, especialmente a través de la integración de la IA en las principales plataformas digitales, han alterado de forma decisiva el ecosistema informativo y las dinámicas de participación política. Desde el fenómeno del *clickbait* y las burbujas de filtro hasta la generación automática de contenidos orientados a maximizar la polarización, la IA incide directamente en la calidad y la legitimidad de los procesos democráticos. De manera innegable, redes sociales con sus potentes algoritmos de recomendación, han amplificado la fragmentación y la polarización de la sociedad y del debate público, como hace más de veinte años vaticinara Sunstein. Ello pone en riesgo significativo a las democracias consolidadas, mientras que, curiosamente, no parece afectar del mismo modo a regímenes autoritarios o a las *democraduras* o “democracias iliberales”.

Somos espectadores de la progresiva sustitución de los medios de comunicación tradicionales, sujetos a deberes éticos y jurídicos profesionales y cierta por plataformas que parecen repeler la regulación. Las plataformas en general operan bajo lógicas comerciales centradas en la monetización y esto además se está trasladando a los medios tradicionales y nuevos. Así, el *clickbait* busca generar el mayor número posible de visualizaciones y clics, lo que va ligado a titulares sensacionalistas o contenidos emocionales que generan reacciones extremas en las audiencias. Este proceso viene atado a algoritmos de recomendación capaces de aislar al usuario en su propia burbuja ideológica como expuso Parisier, reforzando prejuicios y limitando la exposición a perspectivas críticas o divergentes. El resultado es la exacerbación de la polarización, así como la llamada democracia sentimental.

En relación con la desinformación, los efectos de la IA se han hecho cada vez más notorios desde 2016, año en el que se popularizó el debate sobre la injerencia extranjera en las elecciones estadounidenses y luego el Brexit. Aunque todavía no se ha probado con absoluta fiabilidad que estas campañas de desinformación —orquestradas o no mediante IA— hayan sido determinantes en algún resultado electoral concreto, resulta innegable que la IA -y la IA generativa- facilita la capacidad de generar y de difundir selectivamente material engañoso. Por una parte, las redes neuronales generativas permiten crear con facilidad contenido falso (*deepfakes*, voces sintéticas, textos simulados) que complica las tareas de verificación y sembrar la duda sobre cualquier material audiovisual. Además esta situación contribuye a un escepticismo generalizado,

facilitando que se pongan en duda las informaciones que sí que son veraces y, al mismo tiempo, se incrementa la duda pública sobre todo aquello que se publica, con el consiguiente desgaste de la confianza en las instituciones.

Además, la IA permite una personalización masiva y detallada de la propaganda política, así como la microsegmentación de audiencias y la difusión selectiva de mensajes engañosos.

En Estados Unidos, la amplia interpretación de la libertad de expresión garantizada por la Primera Enmienda, especialmente en el contexto de internet, limita significativamente la capacidad de las autoridades para regular e imponer obligaciones a las plataformas.

En cualquier caso quedaba en manos de la sociedad civil y de las propias empresas y plataformas la posibilidad de combatir la desinformación y la manipulación. Sin embargo, las soluciones de mercado resultan en gran medida ineficaces si los gigantes tecnológicos determinan que sus beneficios continúan ligados a algoritmos que promueven la polarización. La situación empeora con el nuevo Gobierno de Trump y Elon Musk, quien, con su plataforma X ideologizada y partidista, complementa un legislativo y judicial que probablemente no cambiarán esta dinámica. Así, ni va a haber soluciones normativas ni desde las plataformas y el mercado. De ahí que se anuncie la desaparición casi completa de secciones de verificación de datos y moderación de contenidos. La situación es muy distinta en la UE, donde sin duda que hay una actitud más intervencionista y reguladora. Con el fin de preservar la integridad de los procesos democráticos y de proteger derechos fundamentales (libertad de expresión e información, protección de datos, infancia, etc.), la UE ha desarrollado varias normas que obligan a realizar muchas acciones a las plataformas así como a controlar sus algoritmos y sistemas de IA que puedan incidir de manera significativa en la formación de la opinión pública. Destaca la *Digital Services Act* (DSA), que introduce obligaciones específicas de transparencia y responsabilidad para grandes plataformas, exigiéndoles la implementación de mecanismos de gestión de riesgos sistémicos y auditorías periódicas. Se trata de un enfoque bastante acertado que no pasa por un control ni dirección de contenidos público ni una evaluación de su constitucionalidad por los gobiernos, que es lo que hay que evitar a toda costa. El Centro de Transparencia Algorítmica de Sevilla para toda la UE va a jugar un papel esencial.

Junto a la DSA, el emergente RIA califica como sistemas de “alto riesgo” aquellos diseñados para influir en procesos electorales o moldear el comportamiento de los votantes. Aunque habrá que ver cómo se implementa y qué autoridades se hacen cargo de su cumplimiento efectivo, supone que estos sistemas de IA cumplan importantes obligaciones. También el RIA regula algunas obligaciones interesantes de informar sobre la existencia de deepfakes así como de contenidos generados por IA (art. 50). Otro hito de la UE es el Reglamento europeo sobre transparencia y segmentación de la publicidad política, que dispone requisitos muy estrictos en la publicidad segmentada. Entre

otras cosas, prohíbe la utilización de datos sensibles para fines de microsegmentación política y exige la trazabilidad de la publicidad en repositorios públicos. Asimismo, el acervo europeo en materia de protección de datos personales, con el RGPD, puede actuar como un dique frente a la creación de perfiles indiscriminados e hiperpersonalización de la propaganda electoral. Nuestro TC ha prestado atención a esta materia y, al menos en teoría, ha establecido importantes garantías para la regulación de estas cuestiones (STC 76/2019, de 22 de mayo de 2019). Lamentablemente, la dejadez está siendo muy importante, por ejemplo, respecto de la asignación de atribuciones y competencias a órganos como la Junta Electoral, como el mismo TS ha señalado expresamente.

Sin embargo, estas medidas regulatorias enfrentan importantes problemas e incertidumbres de cumplimiento dentro de la UE y, sobre todo, desafíos geoestratégicos internacionales, dado que el alcance de las principales plataformas excede el territorio de la Unión. Hacer cumplir estas normativas puede convertirse en una cuestión de relaciones internacionales y un pulso político de la UE con EEUU.

Irónicamente es la propia democracia la que parece dar cobertura a estos contenidos y limita las posibilidades de actuar frente a ellos. Es posible que en los próximos años las democracias liberales deban reflexionar si los contenidos e informaciones estructuralmente y pretendidamente inveraces gozan de alguna protección constitucional cuando están dirigidos, precisamente, a minar el propio sistema constitucional.

Aunque soy relativamente escéptico de la respuesta educativa y cultural frente a estos fenómenos sentimentales, una de las escasas medidas que quedan a las democracias liberales y así lo ha recogido la misma Carta de Derechos digitales, es la necesidad de una mayor alfabetización mediática y digital de la ciudadanía que permita reconocer patrones de desinformación. Una sociedad que en su caso pueda elegir entre las plataformas y servicios que no inviertan de forma sincera en moderación de contenidos, etiquetado confiable de información y diseño de algoritmos que prioricen la fiabilidad por encima del sensacionalismo.

Y como *guinda* de este peligroso *pastel*, cabe al menos mencionar otro peligro significativo que la IA introduce en los procesos electorales: su capacidad para amplificar los riesgos de ciberseguridad en los sistemas electorales. La IA facilita ataques más sofisticados y personalizados, como el *phishing* dirigido a funcionarios electorales o ciudadanos clave, el despliegue de malware indetectable y la saturación de sistemas críticos mediante ataques distribuidos de denegación de servicio (DDoS). Este cóctel tóxico de vulnerabilidades técnicas y narrativas manipuladoras puede erosionar gravemente la confianza pública en el proceso electoral y deslegitimar gobiernos democráticamente electos, incluso cuando no existan pruebas concretas de manipulación efectiva. Y las tecnologías cuánticas que están por llegar pueden disparar los peligros de

integridad de procesos electorales que cada vez más tienen su soporte en sistemas informáticos.

3. Desde el punto de vista de la dogmática de los derechos fundamentales ¿Cree que la inteligencia artificial puede llegar a producir algún cambio en la forma en la que concebimos hoy los derechos fundamentales (su objeto, contenido, estructura...)?

Los derechos fundamentales se han concebido históricamente como derechos subjetivos para proteger a los individuos frente al poder y en las últimas décadas también frente a privados. Sin embargo, en general en los entornos digitales y especialmente con la IA se requiere apreciar que sus efectos impactan especialmente de manera colectiva, a comunidades e incluso a la estructura misma de las sociedades democráticas. En consecuencia, es preciso superar una visión tradicional de derecho subjetivo de los derechos fundamentales. Esta concepción individualista heredada del Derecho romano se centra en la tutela de un interés subjetivo del titular frente a interferencias externas. Sin embargo, esta perspectiva resulta insuficiente porque los sistemas de IA tienen consecuencias estructurales en grupos y comunidades. Por ejemplo, los algoritmos en los servicios públicos inteligentes que gestionan la sanidad, el transporte o la educación, pueden servir para la toma de decisiones discrecionales y colectivas políticas y administrativas. Como señalan autores como Hoffmann-Riem y Mantelero, es necesario adoptar un enfoque que reconozca los efectos colectivos y que permita evaluar jurídicamente el daño social, una categoría análoga al daño medioambiental, difícil de aprehender desde una visión tradicional subjetiva. El uso masivo de datos y la capacidad de los algoritmos para crear perfiles y predecir comportamientos colectivos requieren mecanismos de protección que vayan más allá del caso concreto y aborden los efectos sociales del uso de IA. Y no se trata sólo del uso de IA por el poder público. El impacto colectivo de la IA no se limita al ámbito público; su uso en el sector privado también tiene efectos significativos en los derechos fundamentales en áreas como el consumo, el trabajo, la educación y muchos otros sectores. Con este enfoque destacan especialmente los análisis de gran calado en esta dirección de Hoffmann-Riem y especialmente Mantelero, de Tullio con apoyo de Vedder, Rodotà, Peña Gangadharan, Crawford, Faleiros, Luers, Meier, Perlich y Thorp. Al momento de una valoración jurídica y constitucional de los riesgos, daños y restricciones, es necesario realizar un “recalibrado”, como sugirió Citron en relación al debido proceso. Es imprescindible evaluar no solo casos individuales, sino también el impacto masivo de las operaciones continuadas de algoritmos, sistemas de IA o el manejo extensivo de datos. Como se dirá, ello tiene especial incidencia en los mecanismos de garantía.

Como he sostenido hace años, la respuesta desde los derechos fundamentales exige en parte volver a los principios, esto es a la dignidad que está en la base de los derechos y con ella a la dimensión objetiva de los derechos fundamentales. Este enfoque concibe los derechos fundamentales no solo como

garantías individuales, sino como principios estructurales que orientan todo el ordenamiento jurídico. En el ámbito de la IA, esta dimensión objetiva justifica la intervención del legislador para garantizar que las tecnologías se desarrollen y utilicen de forma compatible con los derechos fundamentales. También permite integrar objetivos sociales, como la sostenibilidad o la equidad, en la interpretación y aplicación de los derechos. La dignidad ha de seguir el rumbo principio rector de cualquier desarrollo tecnológico. Son muy numerosas las proclamaciones sobre la dignidad y los derechos fundamentales como elementos fundamentales en la ética de la IA y su marco jurídico, pero más allá de la retórica habitual, considero que acudir a la dignidad y la dimensión objetiva de los derechos es jurídicamente útil:

- la dignidad actúa como reactivo para establecer límites jurídicos claros. El concepto permite determinar cuándo una tecnología cruza líneas éticas y jurídicas.

- hermenéuticamente se puede acudir a la dignidad cuando es preciso actualizar el contenido y garantías de los derechos ya existentes frente a la IA. Es más, si es preciso, la dignidad también sirve como palanca para reconocer nuevos derechos.

- la dignidad y la dimensión objetiva son apoyos para impulsar nuevas regulaciones o jurisprudencia para maximizar la eficacia de los derechos y garantías efectivas ante la IA.

- sirve para justificar la intervención pública en todas las ramas del Derecho, incluyendo el Derecho de consumo, privado o el de la competencia, para facilitar que integren entre sus objetivos la realización de los derechos.

- proporciona cobertura jurídica a la cláusula de progreso tecnológico, así como a los objetivos sociales del uso de la IA, como la sostenibilidad y la IA para el bien común.

- permite sostener nuevas respuestas y mecanismos de tutela efectivos, en su caso ,la creación de nuevas instituciones y organismos específicos.

- puede ser el apoyo para garantizar los derechos de las futuras generaciones o de la humanidad en general frente a riesgos de la IA.

Además de lo anterior, desde la dogmática de los derechos fundamentales se necesita insistir en la garantía de la reserva de ley en derechos fundamentales y en su caso vinculada al principio de legalidad del sector público que utiliza IA. No sólo son precisas regulaciones generales, sino que es imprescindible que estas normas cumplan con estándares elevados de calidad legislativa. Como he tenido ocasión de examinar con detalle, no son pocos los tribunales europeos de Países Bajos, Alemania, Francia, Eslovaquia o de Gales que invalidan el uso de sistemas automatizados o con IA por la falta de una regulación legal de calidad. Las normativas que regulan el uso tecnologías de alto impacto como la IA deben ser claras, precisas y previsibles, no sólo para habilitar su uso que impacte en derechos, sino que también deben establecerse garantías específicas por la propia norma legal. Además, se deben evitar las habilitaciones en blanco o remisión a normativas técnicas o reglamentarias que no cubren los estándares

exigibles. Cabe recordar que el RIA no es norma que sirva como regulación legal que sirva para legitimar los usos de alto riesgo que regula, sino que se requiere una regulación estatal o de la UE que cubra los estándares exigibles. Y por cuanto a la situación en España, baste decir que en diversos países se ha declarado inconstitucional en ocasiones una normativa que contaba con garantías que ni de lejos tenemos aquí respecto de los tratamientos masivos de datos que se llevan a cabo por la AEAT, la TGSS, CNMC o inspección de trabajo, entre otras.

La falta de una regulación democrática de calidad nos priva de poder usar legítimamente estas tecnologías en áreas como la seguridad, fraude, salud, educación, laboral, personalización de servicios públicos, marketing y un largo etcétera. Estas carencias pueden también llevar a anular actuaciones públicas por violaciones de derechos fundamentales.

Finalmente, me permito señalar que no hay que abusar de acudir al derecho de protección de datos como “comodín” frente a la falta de regulación de nuevas tecnologías. Además, aplicar el régimen de protección de datos a la IA conlleva dificultades estructurales. Es difícil mantener la finalidad de uso de los datos (art. 5.1.b RGPD), dado que continuamente los algoritmos descubren nuevas correlaciones, haciendo impredecibles los usos futuros de los datos. Igualmente es difícil conseguir el consentimiento o una regulación legal para unas finalidades específicas ante finalidades cambiantes. El principio de minimización de datos (art. 5.1.c RGPD) contrasta con la necesidad general de grandes volúmenes de datos en el entrenamiento o validación de sistemas de IA.

4. ¿Considera preciso incorporar a la Constitución algún nuevo derecho fundamental para poder hacer frente a los riesgos que comporta la inteligencia artificial?

Se trata de una cuestión que genera un debate interesante. En los últimos tiempos se afirman “derechos de cuarta o última generación” o nuevos “derechos digitales” con relación a la necesidad de adaptar los derechos fundamentales a los nuevos entornos digitales o en particular por el impacto de la IA. No cabe duda en que en el contenido y facultades de los derechos fundamentales ya existentes hay que introducir garantías de mayor transparencia, explicabilidad y trazabilidad de todo sistema de IA que tenga un impacto significativo en los colectivos, las personas y sus derechos. Especialmente ello es así en el caso de uso público de IA. También hay que profundizar y dotar de mayores garantías a la no discriminación frente a algoritmos, así como a la privacidad y a las libertades informativas y la participación democrática.

A primera vista, el reconocimiento de nuevos derechos puede tener una gran fuerza simbólica y pedagógica, una importante función de concienciación y de generación de un marco institucional más firme. Puede servir como catalizador de nuevas regulaciones o de un mayor escrutinio y atención por parte de los tribunales. Dada la magnitud de la transformación que implica la IA no hay

que descartar el reconocimiento constitucional de nuevos derechos y garantías. Pero en cualquier caso hay que evitar declaraciones vacías que, en la práctica, no aportan soluciones. En todo caso, antes de optar por una reforma constitucional -extraordinariamente difícil en España- hay que potenciar la adaptación de la jurisprudencia constitucional y hay que poner el foco en el desarrollo legal (art. 81 CE) y regulación del ejercicio (art. 53 CE) de estos derechos. Se ha de tener en cuenta normativa y jurisprudencialmente la obligación de hacer efectivos los derechos en razón de su dimensión objetiva (arts. 9.2º, 10. 1 y 53 CE). Ello incluye también la necesidad de dotar de presupuestos de eficacia: recursos presupuestarios, autoridades supervisoras con competencias reales, sanciones disuasorias y vías procesales claras para la impugnación de decisiones automatizadas. Hay que evitar regulaciones grandilocuentes de derechos, pero sin implementación ni efectos prácticos. Que no vuelva a suceder lo que pasó con los “derechos digitales” de la Ley Orgánica 3/2018, que, ni gozaron de un valor simbólico -por la falta de rango constitucional- ni han contado con una regulación concreta y de desarrollo suficiente para tener un impacto tangible. Como decía un humorista hace años, “menos samba y más trabajar”. De lo contrario, estas estrategias pueden generar frustración y restar credibilidad al propio sistema de derechos.

La Carta de Derechos Digitales de España de 2021 ha resultado un referente y estímulo no sólo nacional, sino también en Iberoamérica y la UE, como la Declaración Europea sobre los Derechos y Principios Digitales para la Década Digital (2023/C 23/01), de 15 de diciembre de 2022 o la Carta Iberoamericana de Principios y Derechos en Entornos Digitales, de 25 de marzo de 2023. De la Carta española hay que destacar los derechos relacionados con la IA en la actuación administrativa (artículo XVIII.6º), pues incluye innovaciones y concreciones que deberían inspirar al legislador nacional. El apartado XXV sobre IA es también interesante, pues se proclaman una serie de mandatos “en el desarrollo y ciclo de vida de los sistemas”: “no discriminación”, “transparencia, auditabilidad, explicabilidad, trazabilidad, supervisión humana y gobernanza” y que deberá “garantizarse la accesibilidad, usabilidad y fiabilidad” (2º). También, en este apartado XXV se proyecta para la IA el artículo 22 RGPD, incluyendo la explicabilidad y el derecho a la revisión humana ex post, si bien, la Carta amplía su aplicación pues va más allá de las decisiones “únicamente” automatizadas (3º). También hay garantías ante el uso de la IA en los “Derechos en el ámbito laboral” (artículo XIX), la salud (artículo XXIII) y las neurotecnologías (artículo XV). Así pues, En todo caso, la Carta de derechos digitales asienta los elementos básicos de lo que podría ser una reforma constitucional, pero sigue faltándole cualquier rango normativo así como la necesaria concreción de una regulación más práctica.

Distintos instrumentos internacionales y nacionales sobre IA, no siempre se crean derechos subjetivos nuevos, sino que más bien se profundiza en la interpretación de derechos fundamentales ya reconocidos. El Convenio sobre Inteligencia Artificial del Consejo de Europa de 2024 (Convenio IA) que he tenido

ocasión de analizar en el primer estudio sobre el mismo, alude más de cuarenta veces los derechos humanos, pero se cuida de no establecer de manera expresa nuevos derechos específicos, sino que se limita a reafirmar y concretar los existentes, principalmente en lo referente a la igualdad y no discriminación (arts. 9 y 17), la privacidad y protección de datos (art. 10) o los derechos de personas con discapacidad y de menores (art. 18). Por su parte, el Reglamento de la Unión Europea de IA de 2024 (RIA) se dirige fundamentalmente a fabricantes y proveedores de sistemas y a quienes los implantan en sus empresas o entidades. Es por ello que el RIA olvida a los afectados o interesados y deja en un segundo plano el reconocimiento expreso que hace de algunos derechos de los destinatarios o afectados por la IA, pese a que se introducen algunos avances en materia de recursos, reclamaciones y explicabilidad y transparencia.

Ahora bien, la reticencias al reconocimiento de nuevos derechos subjetivos no debe confundirse con un inmovilismo jurídico. En este sentido, tanto el Convenio IA como especialmente el Reglamento IA implican regulaciones en ocasiones muy precisas cuyo cumplimiento finalmente sirve para la garantía de los derechos de las personas afectadas por la IA.

Como excepción, si pueden considerarse como derechos específicos en el Convenio IA algunas obligaciones positivas de documentación y registro (art. 14 Convenio IA). También otras medidas para garantizar los “recursos” efectivos del artículo 14, así como el artículo 15 regula las “Garantías procesales”, que implican los deberes de supervisión humana y notificación del artículo 15. Cabe señalar que esta garantía no se limita a decisiones “basadas únicamente en el tratamiento automatizado”, sino que se extienden también a decisiones influenciadas significativamente por sistemas de IA. Asimismo se permite una reinterpretación del derecho a la tutela judicial efectiva y del propio derecho a un recurso eficaz (art. 13 del Convenio Europeo de Derechos Humanos), en la dirección de que los afectados puedan ejercer acciones para impugnar posibles vulneraciones derivadas de sistemas de IA. Estos mecanismos fortalecen las garantías procesales frente a decisiones automatizadas o fuertemente asistidas por algoritmos. Por consiguiente, hay un refuerzo de la efectividad de derechos ya existentes, que podría entenderse como una evolución natural de su contenido. En mayor medida se aprecia este fenómeno en el RGPD o en las obligaciones del RIA para los sistemas de alto riesgo. Sus previsiones sirven para concretar y adaptar garantías frente a la opacidad o las eventuales discriminaciones en el uso de algoritmos. Pese a no centrarse ni en los afectados ni en sus derechos, el RIA ha incluido el derecho a presentar una reclamación ante una autoridad de vigilancia del mercado y el derecho a explicación de decisiones tomadas individualmente en los artículos 85-87. Como señala López-Tarruella en el Tratado sobre el RIA que he podido coordinar, queda más bien en el ámbito de las buenas intenciones y queda en cierto modo a la sombra del artículo 22 RGPD, pero habrá que ver la interpretación y criterios que se desarrollan.

Dicho lo anterior, no hay que renunciar a reconocer en la Constitución nuevos derechos subjetivos o nuevas facultades de los derechos ya existentes o nuevas técnicas de garantía en favor de los afectados por sistemas de IA. Ante las dificultades para una reforma del texto constitucional, no hay que excluir la vía de la interpretación, bien por los avances internacionales que puedan darse, que sirven para definir el perfil exacto de los derechos de la Constitución (art. 10.2º CE). Asimismo, no hay que olvidar el papel de la dignidad humana y en su caso el libre desarrollo de la personalidad (art. 10. 1º CE). Se trata de pilares del ordenamiento político y fuente de garantías ante la IA que generan una tensión jurídica, normativa e interpretativa para exigir una adaptación constante de los derechos y sus mecanismos de protección ante las nuevas amenazas. De igual modo, no hay que olvidar todo el potencial interpretativo que aún da de sí el artículo 18. 4º CE, que no sólo sirve para dar cobertura a la protección de datos, sino a todo el ámbito de la privacidad reconocido en el artículo 18 CE.

Como muestra de lo señalado, el botón de los llamados neuroderechos. Chile es el país pionero, con una flamante reforma constitucional y una regulación legal, pero lo cierto es que suponen más ruido que nueces. Si bien, han estimulado una más interesante y concreta interpretación constitucional por la Corte. El apartado XXV de la Carta de derechos digitales sistematiza bien lo que podría ser un reconocimiento constitucional de neuroderechos, tema que De la Garza o Rollnert han analizado. El reconocimiento constitucional bien puede tener el aludido papel simbólico, pedagógico y político. Sin embargo, hoy día considero que no resulta necesario regular constitucionalmente la privacidad mental, la identidad personal, la libertad cognitiva y la integridad psicológica. Ello es así porque estos (neuro)derechos ya están reconocidos por la libertad ideológica (art. 16 CE), la integridad física y moral (art. 15 CE) o la protección de datos personales y el derecho a la privacidad (art. 18 CE). Por tanto, más que nuevos derechos, lo que se requiere es una adecuada adaptación normativa y jurisprudencial para interpretar estos derechos en el contexto de las neurotecnologías. Creo que más que el camino de Chile, parece más sólido seguir más Propuestas legislativas más concretas como la de México desde julio de 2024.

5. En un contexto de creciente presencia de los sistemas de inteligencia artificial ¿sería conveniente o necesario prever nuevos instrumentos de garantía de los derechos fundamentales?

Desde el punto de vista de las garantías, cabe remitir en parte a lo ya señalado respecto del reconocimiento de nuevos derechos; la creación de nuevos instrumentos de garantía no pasa forzosamente por la formulación de nuevos derechos, sino por la adaptación orgánica y procedimental de nuestro sistema jurídico. En muchos casos, basta con afinar y actualizar la legislación. En todo caso, creo que hay que abordar especialmente tres frentes con relación a las garantías.

En primer lugar, en muchos supuestos no basta con una protección basada únicamente en intereses individuales, sino que debe atenderse a la dimensión supraindividual de los riesgos. La IA puede impactar, por ejemplo, sobre un amplio colectivo de consumidores o incluso repercutir en el debate público, la integridad de los procesos democráticos y la libertad de expresión y de formación de la voluntad política. En la Carta de Derechos Digitales propuse incluir un apartado sobre “Garantías de impacto social”, inspirado en parte por Mantelero. También la Recomendación sobre IA de la UNESCO subraya la importancia de un enfoque colectivo y la inclusión de mecanismos de participación social en el uso de sistemas de IA. En los derechos de últimas generaciones desde hace tiempo se va incorporando la protección de intereses colectivos y supraindividuales y se están reforzando las garantías y regulaciones para acciones colectivas efectivas del consumidor. En 2022, el Tribunal de Justicia de la Unión Europea ha interpretado favorablemente la normativa de protección de datos en este contexto. El artículo 15 del Convenio IA exige la identificación y evaluación de riesgos considerando todas las partes interesadas, incluidas aquellas cuyos derechos podrían verse afectados. La participación de la sociedad civil se menciona en el artículo 5 sobre “Integridad de los procesos democráticos y respeto del Estado de Derecho” y, especialmente, el artículo 20 sobre “consulta pública” dispone que “Cada Parte se esforzará por garantizar que las cuestiones importantes [...] sean, según proceda, objeto de debate público y de consulta a las múltiples partes interesadas a la luz, en particular, de las implicaciones sociales, económicas, jurídicas, éticas y medioambientales pertinentes.”

Una segunda vertiente de la necesaria reforma o refuerzo de las garantías se halla en la introducción para la IA del modelo del “más vale prevenir que curar”. Inspirándose en las leyes de la robótica de Asimov, esto implica la integración del cumplimiento ético y legal directamente en el “Código”, conforme a los términos expuestos por Lessig. El modelo preventivo, se ha consolidado en el RGPD como un elemento esencial: es el deber de realizar evaluaciones de impacto (artículo 35 RGPD), la exigencia de la protección de datos “desde el diseño y por defecto” (artículo 25 RGPD) o la obligación de registro interno de actividades (artículo 30 RGPD). Este modelo se generaliza en la UE en diversos ámbitos digitales como Barrio ha expuesto especialmente y en la legislación europea sobre comercialización de productos potencialmente peligrosos. Se trata la técnica del *compliance* o cumplimiento normativo de origen anglosajón, de integrar la ética y el cumplimiento legal directamente en el “código” o diseño de los sistemas. Este modelo se proyecta ahora, con matices, al ámbito de la IA. El alto grupo de expertos de la UE adoptó un modelo de ética en el diseño con su exhaustiva lista de evaluación, su modelo ALTAI de evaluación. En España, la Carta de Derechos Digitales establece generalmente el “principio de cumplimiento normativo desde el diseño” para todos los campos (1.4). Y para la Administración, “será necesaria una evaluación de impacto en los derechos digitales al diseñar algoritmos para decisiones automatizadas o semiautomatizadas” (XVIII.7º). En desarrollo de la Carta, el artículo 23.1º Ley

15/2022 afirma con cierta *suavidad* que “se promoverá la realización de evaluaciones de impacto que determinen el posible sesgo discriminatorio”. En el RIA destaca el artículo 9 del RIA que exige a los proveedores de sistemas de alto riesgo un sistema de gestión de riesgos que incorpore el impacto en los derechos fundamentales. Ello supone la identificación y evaluación de riesgos, considerando tanto el uso previsto como un mal uso razonablemente previsible. Las pruebas se basan en parámetros y umbrales de tolerancia definidos previamente. Además, tras la comercialización del producto se evalúan riesgos adicionales y se implementan medidas específicas de gestión de riesgos. Es un proceso de mejora continua a lo largo del ciclo de vida del sistema de IA de alto riesgo, que requiere revisiones y actualizaciones periódicas. Además, quien adquiere y utiliza el sistema está obligado a realizar un estudio de impacto sobre derechos fundamentales antes de su implementación (art. 27 RIA). Debe evaluarse si hay necesidad de implantar el sistema IA de alto riesgo por la entidad y, fundamentalmente, identificar y mitigar riesgos que tiene para los derechos fundamentales que sean aceptables. Todo esto debe documentarse adecuadamente.

En tercer lugar, se requieren autoridades o instancias dotadas de potestades de control, investigación y sanción. Así como en protección de datos se cuenta con autoridades independientes (como la Agencia Española de Protección de Datos), también se ha planteado la necesidad de crear, o en su caso atribuir funciones a órganos capaces de supervisar específicamente la IA. El RIA contempla la existencia de autoridades de supervisión del mercado a nivel nacional (artículo 59). También deben existir organismos de evaluación de conformidad, responsables de realizar las pruebas de conformidad y que deben estar habilitados por una autoridad notificante, que debe existir al menos una por país y puede ser la autoridad de supervisión del mercado. Asimismo, en febrero de 2024 se ha creado la Oficina Europea de la IA, que centraliza para la Comisión Europea lo relativo a la IA general. Cuanto más vinculada esté una autoridad al ámbito de los derechos fundamentales o a la protección del consumidor, mayor independencia se requiere, hasta el nivel de independencia de las instituciones de protección de datos (Conclusiones del Abogado General, de 14 de enero de 2021, Asunto C-718/18). Sin embargo, el carácter centralizado de la AESIA y su proximidad a no pocos derechos fundamentales llevan a pensar que habría de garantizarse un mayor estándar de independencia que actualmente no se cumple con su regulación. El Reglamento IA exige de modo general la independencia, y la refuerza y la asegura en algunos casos más vinculados con la identificación biométrica y los usos de seguridad pública y cumplimiento de la ley. El artículo 26 del Convenio IA subraya de manera clara que los mecanismos de supervisión deben “ejercer sus funciones de manera independiente e imparcial”, asegurándose de contar con recursos y competencias suficientes.

6. Desde la perspectiva de los derechos fundamentales ¿qué opinión le merecen el Reglamento europeo de inteligencia artificial y el Convenio Marco del Consejo de Europa sobre inteligencia artificial y derechos humanos, democracia y Estado de derecho?

Tanto el RIA como el Convenio IA representan avances positivos al reforzar el principio de la “IA centrada en la persona” y especialmente garantizar mecanismos preventivos de control y evaluación de riesgos. El RIA es más concreto y con un estándar de garantías más elevado que el convenio, que es muy flexible. El Convenio IA amplía el campo de acción, pone acento en la proyección sobre la democracia y el Estado de derecho, y establece principios vinculantes para todos los tipos de IA, sean o no de alto riesgo. Sin embargo, este andamiaje normativo en el viejo continente suscita algunas dudas respecto al lugar que ocupará la Unión Europea en la carrera por la innovación en IA. No pocos actores del mercado temen que la excesiva complejidad de las obligaciones técnicas y documentales desaliente la inversión en investigación y desarrollo de sistemas de IA en Europa. Este binomio puede proyectar a Europa como el referente mundial de la IA ética y responsable, sin embargo, también puede quedar afectada la innovación en razón de las obligaciones del RIA: requisitos de registro de sistemas, los procesos de evaluación de conformidad, el sometimiento a auditorías técnicas y, en definitiva, la densa burocracia que puede emerger.

Es muy positivo a mi juicio que el RIA prohíba algunos sistemas de IA por ser incompatibles con el Estado de Derecho y los derechos. Eso sí, podría haberlo hecho algo mejor y sin tantos espacios a la futura interpretación. En todo caso y especialmente, el RIA identifica unos sistemas de IA de “alto riesgo”, precisamente, por afectar a los derechos fundamentales y la seguridad de las personas. Por una parte, será de alto riesgo la integración de IA en productos ya peligrosos, sujetos a la normativa de la UE dentro del “nuevo marco legislativo” de la UE (Anexo I). Por otra parte, son de alto riesgo los sistemas IA utilizados para “influir sustancialmente en el resultado de la toma de decisiones” (art. 6.3 RIA) para hasta 25 finalidades enumeradas en el Anexo III agrupadas en: Biometría; 2. Infraestructuras críticas; 3. Educación y formación profesional; 4. Empleo, gestión de los trabajadores y acceso al autoempleo; 5. Acceso a servicios privados esenciales y a servicios y prestaciones públicos esenciales y disfrute de estos servicios y prestaciones; 6. Aplicación de la ley; 7. Migración, asilo y gestión del control fronterizo; 8. Administración de justicia y procesos democráticos.

Pues bien, siguiendo un modelo más técnico y de mercado, respecto de los sistemas de alto riesgo el RIA exige el marcado “CE”, diseñando procedimientos de evaluación de conformidad y estableciendo obligaciones específicas para fabricantes e implantadores. Los proveedores y usuarios de sistemas de alto riesgo están obligados a acometer evaluaciones de impacto, a mitigar los eventuales sesgos, a generar documentación técnica exhaustiva, a garantizar la supervisión humana y a llevar un registro de logs que posibilite la trazabilidad y

la explicación de las decisiones automatizadas. Asimismo obliga a que estos sistemas sean robustos, exactos y no *hackeables*. Todas estas exigencias serán desarrolladas además por todo un armazón de normas técnicas y de armonización, que serán acompañadas de la normalización técnica de la IA que está actualmente en fase de desarrollo. Para los sistemas que no son de alto riesgo, todas estas obligaciones serán los referentes de las buenas prácticas a seguir voluntariamente en códigos, sellos o certificaciones a las que se sumen los proveedores públicos o privados de esos sistemas de IA. Además, el RIA otorga a las autoridades de supervisión la capacidad de requerir documentación o solicitar modificaciones en aquellos sistemas que puedan poner en peligro la seguridad o los derechos individuales y colectivos. Se trata de un enfoque, que aunque provenga del Derecho privado y la seguridad de los productos, supone un avance notable y sitúa a la UE a la cabeza de la regulación mundial de la IA.

El Convenio IA sin duda alguna es más inconcreto y quizá tenga un valor más simbólico que jurídico y normativo. Muestra de ello es que Israel o Estados Unidos han firmado el Convenio de manera inmediata. El Convenio IA busca arraigar en el plano jurídico internacional los valores fundamentales que deben regir el desarrollo de la IA: el respeto a los derechos humanos, la democracia y el Estado de derecho. El Convenio IA contempla un abanico de garantías y principios que no se circunscriben a los llamados sistemas de alto riesgo. Ello puede tener un potencial interpretativo importante, pero muy posiblemente esta extensión de garantías a todos los sistemas ya quedará en sede de cada Estado o parte firmante. También es relevante que el Convenio resalta la importancia de la participación social, la consulta pública y la inclusión de todas las partes interesadas en la evaluación de los posibles impactos de la IA (artículo 20). Asimismo, expresa la necesidad de contar con autoridades independientes y efectivas encargadas de vigilar el cumplimiento de los principios consagrados (artículo 26). Quizá el mérito esencial del Convenio IA es que puede aplicarse también a países que no integran la Unión Europea e incluso de fuera de Europa, lo que potencia su vocación de instrumento internacional de referencia. Como he señalado en algún lugar, el Convenio “pone la lírica” allí donde el RIA “pone la prosa”. El RIA aporta la sistemática legal, la definición de obligaciones precisas, los procedimientos de evaluación y las sanciones para quienes vulneren la normativa. El Convenio IA, en cambio, traslada al plano vinculante los valores y principios fundamentales que deben guiar la actividad legislativa y la actuación administrativa y judicial en relación con la IA. Si se produce una sinergia acertada entre ambos —el RIA como fundamento técnico y el Convenio como brújula axiológica—, se logrará un modelo de gobernanza de la IA capaz de combinar la rigurosidad y la claridad normativa con la efectividad en la protección de los derechos humanos.