ENFOQUE HOLÍSTICO DEL RIA Y NECESARIA COHESIÓN CON OTRAS NORMAS A ACTUALIZAR, ENTRE OTRAS, RESPONSABILIDAD CIVIL Y DERECHO DE CONSUMO[*]

Carmen Muñoz García[**]

Resumen

Estudiar el Reglamento de Inteligencia Artificial de manera aislada no es suficiente. Para cumplir con los objetivos de la UE de innovar y garantizar derechos, es necesario abordarlo desde un enfoque holístico y en conexión con las normas y valores de la Unión. El objetivo ambivalente fijado por el Reglamento no resulta extraño: esta tecnología, con un potencial indescriptible, avanza de manera ilimitada y, puesto que afecta a todos los ámbitos y sectores en los que participamos, a nuestra forma de trabajar y de relacionarnos, es imprescindible establecer límites éticos y jurídicos a su desarrollo e implementación. Su regulación no basta para proteger a las personas, sino que forma parte de un marco regulatorio integral y complejo que pretende poner de manifiesto que la tecnología está al servicio de las personas, de ahí que principios como la transparencia y la intervención humana hayan calado de manera directa en el texto y obliguen a conectar esta materia con el acervo de la UE.

Palabras clave

Reglamento de Inteligencia Artificial, obligaciones de transparencia, agentes de IA, Derecho del consumo, responsabilidad por los daños causados por productos defectuosos.

---

[*] Este trabajo se enmarca en la Cátedra Internacional de «IA Generativa: retos y riesgos", financiada por la Unión Europea-Next Generation EU; en el Proyecto cAIre OdiseIA for GOOGLE.org «AI Governance» (subproyecto #1.1 "Mapping of AI governance – recommendations and regulation" (en julio de 2024, en el contexto de este proyecto, la autora propuso el «enfoque holístico del RIA» y su cohesión con el acervo de la UE); en el Instituto de Derecho Europeo e Integración Regional de la UCM (IDEIR) y en el Grupo de Investigación UCM 971680, «Derecho de Daños. Derecho de la contratación».

[**] Profesora Titular de Derecho Civil de la Universidad Complutense de Madrid, acreditada a Catedrática. Codirectora de la Sección de IA del Spanish Hub, European Law Institute (ELI). Instituto de Derecho Europeo e Integración Regional (IDEIR). Codirectora del Diploma de Alta Especialización en IA y su impacto en el Derecho en la Escuela de Práctica Jurídica UCM. Miembro de OdiseIA. carmenmunoz@der.ucm.es

Abstract

Studying the Artificial Intelligence Regulation in isolation is not enough. In order to meet the EU's objectives of innovation and guaranteed rights, it is essential to approach it holistically and in connection with EU norms and values. The ambivalent objective set by the Regulation is not strange: this technology, with its indescribable potential, is advancing in an unlimited way and, since it affects all areas and sectors in which we participate, the way we work and relate to each other, it is essential to establish ethical and legal limits to its development and implementation. Its regulation is not enough to protect people, but forms part of a comprehensive and complex regulatory framework that aims to show that technology is at the service of people, which is why principles such as transparency and human intervention have permeated directly into the text and make it necessary to connect this matter with the EU acquis.

Key words

Artificial Intelligence Act, transparency obligations, agents AI, Consumer Law, Liability for defective products.

SUMARIO: I. Codificando el desarrollo de la IA, y, además, respetar el acervo de la UE. 1. Codificando el desarrollo de la IA. 2. Regular la IA sí, pero respetando de Derecho vigente de la UE. II. Los sujetos protegidos y la autonomía de la voluntad en la regulación de la IA. 1. La persona en el centro de la regulación de la IA. A. El ser humano en el «estándar jurídico» del RIA. B. La importancia de los sistemas destinados a interactuar con persona física. C. Cuando la IA causa daño a tercero ajeno a la tecnología. La responsabilidad extracontractual. 2. La transparencia: aspecto clave para conformar la autonomía de la voluntad. A. Transparencia para evitar los fallos y la desinformación creada por IA. B. Obligaciones de transparencia en el Reglamento de IA. C. En concreto, los sistemas destinados a interactuar con personas físicas. 3. Agentes que simulan comportamientos humanos y comprometen la autonomía de la voluntad ¿completamente autónomos? A. ¿Qué es un agente de IA? ¿Qué nivel de autonomía puede llegar a alcanzar? B. Los agentes de IA en la contratación con consumidores. III. Reflexiones finales. IV. Bibliografía.

I. CODIFICANDO EL DESARROLLO DE LA IA, Y, ADEMÁS, RESPETAR EL ACERVO DE LA UE

1. CODIFICANDO EL DESARROLLO DE LA IA

Ante el imparable desarrollo e implementación de la inteligencia artificial en nuestras vidas, que proporciona innumerables beneficios en cualquiera de los ámbitos en los que se desarrolla, no hay mucho que objetar si se respetan las normas jurídicas y los valores de la Unión, si se salvaguardan el derecho a la informa-

ción veraz y a la autonomía de la voluntad de las personas, junto con la libre competencia de las empresas en el marco de la economía de mercado. Se trata de evitar abusos y situaciones de poder que, mediante esta herramienta, minoren o anulen derechos de las personas y de quienes operan en el libre mercado.

Ahora bien, en el mejor de los casos, la potencialidad de estas herramientas que amplifican exponencialmente las capacidades de empresas, poderes públicos y privados, e incluso de individuos con grandes habilidades, pueden poner en riesgo derechos fundamentales, democracias o Estados de Derecho, ya que pueden derivar en prácticas ilícitas o antijurídicas con influencia en la información y en la toma de decisiones de personas y colectivos.

De ahí que, por la Unión Europea se considerase imprescindible establecer normas armonizadas en materia de inteligencia artificial que posibilitasen el desarrollo y uso de la herramienta de manera ética, segura y confiable. El largo proceso regulatorio llevado a efecto en la Unión, con arduas negociaciones políticas y fundados desencuentros, dio lugar a la aprobación del Reglamento de Inteligencia Artificial en junio de 2024 (en adelante, RIA o LIA –AI Act– por su traducción y abreviatura del inglés)(1), aunque el proceso regulatorio en la materia, ni de lejos, ha culminado. De primeras, además de implementar el marco de gobernanza en la Unión, serán necesarias normas propias en los distintos Estados miembros que habrán de establecer, entre otros aspectos, quienes son las autoridades nacionales competentes en la materia, y también, cuál es el régimen de sanciones por las infracciones del RIA, en los términos que fija el artículo 99 del Reglamento(2).

Para ayudar en la adecuada adaptación e implementación en los términos pretendidos por el Reglamento de la Unión, se irán sumando Directrices de la Comisión Europea (CE) sobre la aplicación práctica del RIA, que previstas en el artículo 96 del texto(3), no son vinculantes y habrán de irse actualizando según sea necesa-

---

(1) Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, 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.° 300/2008, (UE) n.°167/2013, (UE) n.° 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. Es el Reglamento de Inteligencia Artificial (RIA o AI Act, por sus siglas en inglés), publicado en el DOUE de 12 de julio de 2024.

(2) España, ha sido el primero de los 27 Estados miembros en tomar la iniciativa y presentar el régimen sancionador y de gobernanza previsto en el RIA. El Anteproyecto de *Ley para el buen uso y la gobernanza de la inteligencia artificial* para implementar en nuestro Derecho interno el Reglamento 2024/1689 fue aprobado el 18 de marzo de 2025 y ha sido remitido por el Gobierno para su tramitación. Consta de 4 capítulos: (I) disposiciones generales: objeto, definiciones y ámbito subjetivo; (II) gobernanza y supervisión: autoridad notificante, autoridades de vigilancia del mercado y gobernanza de los espacios controlados de prueba; (III) prohibiciones y excepciones: además de las prácticas prohibidas del artículo 5 del RIA, se incluye un largo listado de excepciones vía autorización de uso de «identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de garantía del cumplimiento del Derecho»; (IV) clasificación de las infracciones y sanciones. También, publicidad de las sanciones, procedimiento, actuaciones previas, medidas de carácter provisional, prescripción, a lo que se suma la reparación del daño e indemnización, y la posibilidad de que la autoridad de vigilancia del mercado competente establezca multas coercitivas.

(3) El artículo 96 del RIA encomienda a la CE la elaboración de directrices sobre la aplicación práctica del Reglamento (la definición de sistemas de IA, prácticas prohibidas, obligaciones y requisitos de los sistemas de IA de alto riesgo u obligaciones de transparencia, entre otras). Por lo pronto, el 4 y 6 de febrero de 2025, casi coincidiendo con la obligada aplicación de las disposiciones generales del RIA (arts. 1 a 4), y también de la prohibición de determinados sistemas de IA (art. 5 del RIA), que son de aplicación desde el 2 de febrero, la Comisión ha publicado, en cumplimiento del

rio. Además, las resoluciones del TJUE serán fundamentales para garantizar que se produzca una integración, interpretación y aplicación óptimas del Reglamento en los Estados miembros. Se trata de incorporar y desarrollar en el mercado aquellos «productos de IA» que deberán ser seguros y estar garantizados. No olvidemos que estas herramientas deberán encajar en el marco de las obligaciones y requisitos exigidos por la norma, respetando al mismo tiempo los derechos fundamentales contemplados en el Reglamento [4].

Antes de continuar con esta relevante labor armonizadora en materia de IA, conviene advertir que la sola capacidad de la inteligencia artificial para emular la conducta humana y de manipular la información y la autonomía de los individuos, puede inducir a comportamientos en las personas que nos convierte, a los usuarios medios y posiblemente también a la «persona física normalmente informada y razonablemente atenta y perspicaz» [5], en sujetos vulnerables frente a quienes producen y/o desarrollan estos tipos de *software*. Lo que exige de actuaciones éticas y un gran rigor técnico, administrativo y de vigilancia poscomercialización durante toda la vida del sistema de IA.

Al fin y a la postre, «la combinación de la inteligencia artificial con las nuevas tecnologías digitales es el germen de la economía de la vigilancia caracterizada por agravar el riesgo de control de los usuarios. En este contexto se desarrollan disrup-

---

citado artículo 96 del RIA, dos Directrices sucesivas: (i) Una primera sobre prácticas prohibidas: *Communication from the Commission - Commission Guidelines on prohibited artificial intelligence practices established byRegulation (EU) 2024/1689 (AI Act)*, y anexo; y (ii) una segunda sobre la definición técnico-jurídica de sistemas de IA para facilitar la aplicación del Reglamento: *Communication from the Commission - Commission Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act)*. Vistos respectivamente en: https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act; https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application. También conviene traer a colación que el 22 de abril de 2025 la CE ha abierto consulta pública para que por las partes interesadas se propongan directrices prácticas que ayuden a implementar de forma clara y accesible las normas sobre los modelos de IA de uso general. Se espera que, para el 2 de agosto de 2025, fecha en la que la normativa sobre los modelos será de aplicación, estén publicadas estas directrices, y, además, los Códigos de buenas prácticas. Visto en: https://digital-strategy.ec.europa.eu/en/news/commission-seeks-input-clarify-rules-generalpurpose-ai-models

(4) Dada la clara potencialidad y escalabilidad de la IA en el mercado interior de la Unión, el considerando 3 del RIA *advierte* que es incuestionable: *evitar la fragmentación normativa en materia de protección de derechos fundamentales*, y, además, *evitar que falte la seguridad jurídica* en los operadores. Para ello, es *imprescindible*: «garantizar un nivel elevado y coherente de protección en toda la Unión» y así alcanzar una IA fiable, *y, además*, «evitar las divergencias» que supongan límites a la innovación, la libre circulación, el despliegue y la adopción en el mercado de la Unión «de los sistemas de IA y los productos y servicios conexos», ¿Cómo? Mediante normas con obligaciones y requisitos uniformes para los operadores, y también, mediante la protección del interés general y de los derechos de las personas.

(5) Cuando el Reglamento europeo de IA se refiere a determinados sistemas de inteligencia artificial destinados a interactuar con personas físicas o a generar contenidos, y *tras admitir que pueden plantear riesgos específicos de suplantación o engaño –«con independencia de si cumplen las condiciones para ser considerados como de alto riesgo o no»–*, fija que el uso de estos sistemas debe estar sujeto, además de las obligaciones y requisitos exigidos –según el sistema de que se trate–, *a obligaciones de transparencia específicas*, y por tanto, es preciso comunicar a la persona física que interactúa con un sistema de IA, a excepción de que esto resulte evidente «desde el punto de vista de una persona física normalmente informada y razonablemente atenta y perspicaz, teniendo en cuenta las circunstancias y el contexto de utilización» (cdo. 132 y art. 50, apart. 1 del RIA).

tivas tácticas empresariales de comercialización que originan riesgos de abuso en particular de los consumidores. El desarrollo tecnológico potencia las capacidades del empresario de anticiparse, personalizar e influir de forma determinante en sus decisiones de consumo. Esta aptitud genera una nueva situación que requiere que el derecho examine y ajuste en el contenido de sus nomas de tutela»(6). Todo un punto de partida para entender cómo la IA es capaz de generar desequilibrios que se ven potenciados por sus capacidades/habilidades de esta herramienta en un mercado muy competitivo.

Se trata de una herramienta transformadora muy poderosa(7), enormemente beneficiosa para las personas y la sociedad, que avanza de forma *imparable e ilimitada* y para la que es necesario establecer *límites legales y éticos* sin ambages. De ahí que, desde las distintas fuerzas políticas, gobiernos y organizaciones públicas y privadas, y desde todas las áreas del Derecho se estén planteando análisis, debates, propuestas, acuerdos y acciones legislativas y de gobernanza que permitan avanzar en innovación sin menoscabar derechos. Si bien, las distintas posturas adoptadas por China, EE. UU o Europa, muestran intensidad regulatoria diferente y enfoques dispares según los posicionamientos políticos y la rivalidad tecnológica, principalmente entre los dos primeros, y que va, desde un control absoluto por parte del Gobierno hasta un completo reconocimiento autorregulatorio dirigido a empresas, organizaciones e instituciones.

A nivel europeo hemos puesto de manifiesto, desde el comienzo, que el proceso regulatorio propio ha culminado con *un primer hito normativo*, el Reglamento de Inteligencia Artificial. Se trata de la primera regulación que establece normas armonizadas para el mercado interior de la IA, que tiene por fin promover la adopción de una IA centrada en el ser humano(8) y fiable, y, además, es un hito a nivel mundial. Su impacto global se debe, básicamente: además de por el inapelable «efecto Bruselas»(9), porque la norma europea, que pretende el desarrollo e imple-

---

(6) Así lo expresa Ana Felicitas Muñoz Pérez «Economía de la vigilancia y manipulación mediante la IA. Libertad cognitiva y protección del consumidor», *La Ley mercantil*, núm. 121, febrero 2025.

(7) Los grandes cambios ya se están viendo en el ámbito de la salud, su capacidad para revolucionar el campo de la medicina es infinito y proliferan los ejemplos en los que esta tecnología se muestra con todo su potencial: diagnóstico temprano de enfermedades, monitorización de pacientes, genómica personalizada, registros de salud, asistencia domiciliaria o cirugía asistida por IA.

(8) Como ya señalara el ilustre jurista Antonio Hernández Gil, «Conceptos jurídicos fundamentales», T. 1, *Colección Obras completas,* ed. Espasa Calpe, Madrid, 1987, pp. 20 y ss., el *ius privatum*, se integra el *ius civile*, y en este, la persona es la base y el centro del Derecho. En concreto Jesús Delgado Echeverría, «I. La persona física», en Lacruz Berdejo, Sancho Rebullida, Luna Serrano, Delgado Echevarría, Rivero Hernández y Rams Albesa, *Elementos de Derecho civil. I. Parte General, vol. segundo. Personas*, ed. Dykinson, Madrid, 2000, p. 1 y ss., fija, sin ambages, que son personas, en primer lugar, y por antonomasia, todos los seres humanos y expresa que, como exige la Declaración universal de los derechos humanos (ONU, 1948), artículo 6: «todo ser humano tiene derecho, en todas partes, al reconocimiento de su personalidad jurídica», en definitiva, como incide el autor, a ser reconocido «como persona ante el Derecho», *as a person before the law*. Lo que no obsta para que, por una ficción del Derecho, más allá de los seres humanos, se reconozca la cualidad de personas a las personas jurídicas.

(9) El denominado «efecto Bruselas» responde a la idea de que a nivel regulatorio Europa es pionera en los muchos temas que legisla. El término fue acuñado por Anu Bradford, profesora estadounidense de derecho del comercio internacional. Realizando una analogía con el *Efecto California*, Bradford habla de «efecto Bruselas» para describir la manera en que la UE se ha convertido en una potencia reguladora mundial, con un marco normativo que ha acabado afectando indirectamente

mentación ética, segura y fiable de la IA, se ha convertido en un modelo o referente regulatorio para otros países (Canadá, Australia o Brasil, entre otros). Y aunque Europa, tecnológicamente, no sigue ni de cerca a países como China o EE. UU, probablemente, la reglamentación se traduce en una ventaja competitiva por segura y fiable. La seguridad jurídica para el adecuado desarrollo de los mercados también es una baza relevante en el comercio internacional(10), tanto para productores y desarrolladores como para generar confianza en los usuarios.

Así las cosas, como veremos más adelante, la inteligencia artificial está y debe estar al servicio de las personas y del bien común, y la persona debe ser y estar en el núcleo de esta regulación. Es más, todas las iniciativas legislativas y los acuerdos internacionales van en esa línea.

Además del Reglamento europeo de IA, cabe referirse al que puede considerarse el primer acuerdo internacional sobre la materia, el *Convenio Marco sobre Inteligencia Artificial y Derechos Humanos, Democracia y Estado de Derecho*(11), que negociado por casi sesenta países pretende ser el primer tratado internacional jurídicamente vinculante sobre la inteligencia artificial, fijando principios inalienables como la transparencia y la intervención humana en beneficio de las personas. Además, en el preámbulo se reconoce el valor y el potencial de la IA en beneficio humano, por lo que se pretende alcanzar la unidad en el tratamiento de esta tecnología y la cooperación con terceros Estados que comparten los mismos valores. Se trata de promover y proteger los derechos humanos, el bienestar individual y social, el desarrollo sostenible, la igualdad de género, y el empoderamiento de todas las mujeres y niñas, así como otros objetivos e intereses importantes, potenciando el progreso y la innovación. Deja claro que la IA ofrece oportunidades sin precedentes para la vida de las personas y que en el desarrollo y uso de cualquier sistema de

---

a sectores y mercados de todo el planeta. Sus reglas estrictas, sus mayores estándares de seguridad y calidad, su amplio y riguroso ordenamiento, no son obstáculo para que otras economías a la vanguardia, decidan seguir operando en el Mercado Único Europeo para prestar bienes o servicios a cerca de 500 millones de usuarios de la UE. Visto en Anu BRADFORD, *The Brussels Effect: How the European Union Rules the World*, Oxford University, 2020.

(10) Carmen Muñoz García, C. (2023) *Regulación de la inteligencia artificial en Europa. Incidencia en los regímenes jurídicos de protección de datos y de responsabilidad por productos*, 1ª edición, Tirant lo Blanch, Valencia, pp. 28-29.

(11) El *Convenio Marco del Consejo de Europa sobre Inteligencia Artificial y Recursos Humanos Derechos, democracia y Estado de Derecho* (CETS – No. 225), aprobado en mayo de 2024, y publicado para firmas de adhesión el 05.09.2024, nace con el propósito de tener alcance mundial, impulsando la adhesión, no sólo de los Estados miembros (integrado en la actualidad por 46 Estados, entre otros, Canadá, EE. UU, México y Japón), sino también de los no miembros. Primero, el Comité de Ministros del Consejo del Consejo de Europa, adoptó el *Convenio Marco del Consejo de Europa sobre Inteligencia Artificial y derechos humanos, democracia y Estado de derecho* que, negociado por 46 Estados miembros del Consejo de Europa, la Unión Europea y 11 Estados no miembros (Argentina, Australia, Canadá, Costa Rica, Estados Unidos de América, Israel, Japón, México, Perú, la Santa Sede y Uruguay), y en el que han contribuido observadores del sector privado, la sociedad civil y el mundo académico, fue abierto a firma durante la conferencia de los ministros de Justicia celebrada en Vilna (Lituania) el 5 de septiembre de 2024. Es, como señala la noticia de prensa del Consejo: *el primer tratado internacional legalmente vinculante destinado a garantizar que el uso de los sistemas de inteligencia artificial es totalmente coherente con los derechos humanos, la democracia y el Estado de derecho*. Convenio disponible en: https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatynum=225 . La noticia, disponible en: https://www.coe.int/es/web/portal/-/council-of-europe-opens-first-ever-global-treaty-on-ai-for-signature. Acceso al texto https://rm.coe.int/1680afae3c.

IA deben promoverse y protegerse los derechos humanos, la democracia y el Estado de Derecho. Deja en manos de los Estados miembros cómo implementar las obligaciones que deberán abordar las administraciones públicas y las empresas.

Desde otros planos, la OCDE incorpora cinco principios para impulsar un desarrollo y uso de la IA confiable[12], o el Consejo de Europa[13], que ha proporcionado en este breve período de tiempo principios que pretenden servir de base a nivel global y cuyo objetivo es la protección de los derechos humanos. Este último, el *Convenio Marco del Consejo de Europa sobre Inteligencia Artificial,* fija un desiderátum de profunda relevancia al afirmar que dado que los sistemas de IA –entiendo que también los modelos–, pueden socavar la dignidad humana y la autonomía individual, los derechos humanos, la democracia y el Estado de Derecho[14] es preciso asentar principios que se adapten a los nuevos desafíos de la IA. Sirvan de ejemplo, principalmente, los artículos 16 (gestión de riesgos e impacto)[15], 17

(12) En mayo de 2024, la Organización para la Cooperación y Desarrollo Económicos (OCDE) actualizó la Recomendación que sobre IA había adoptado inicialmente en mayo de 2019 y que inicialmente fue acogida por el G20 y la UE, y, además, entre otros, EE. UU y Japón. Si ya en 2019 se puso en valor el respeto a los derechos humanos y los valores democráticos, a la inclusión, la diversidad, la equidad, la innovación y el bienestar, también se fijaron como principios los siguientes: (i) crecimiento inclusivo, desarrollo sostenible y bienestar; (ii) respeto al Estado de Derecho, los derechos humanos, valores democráticos, también la equidad y la privacidad; (iii) transparencia y explicabilidad; (iv) robustez, seguridad de la IA y protección; y (v) rendición de cuentas. Con esos antecedentes, en noviembre de 2023 se revisa y adopta una definición de sistemas de IA, alineada con la que sería la definición que incorpora el Reglamento europeo de IA, y finalmente, en 2024, el Consejo de la OCDE procede a la actualización de los principios de 2019 para adaptarlos a los grandes cambios de esta tecnología. Con la adaptación proporciona mayor precisión respecto a los principios y proporciona aclaraciones para adoptar una IA más segura y comprometida con las personas. Sobre la Recomendación y su actualización: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

(13) El *Convenio Marco del Consejo de Europa sobre Inteligencia Artificial y Derechos Humanos, Democracia y Estado de Derecho* impulsa promover y proteger los derechos humanos, el bienestar individual y social, el desarrollo sostenible, la igualdad de género, y el empoderamiento de todas las mujeres y niñas, así como otros objetivos e intereses importantes, potenciando el progreso y la innovación. Deja claro que la IA ofrece oportunidades sin precedentes para la vida de las personas y que en desarrollo y uso de cualquier sistema de IA deben promoverse y protegerse los derechos humanos, la democracia y el Estado de Derecho https://rm.coe.int/1680afae3c

(14) Mención especial merece el análisis comparativo entre el RIA y el Convenio Marco del Consejo de Europa realizado por Dignum, V. (Umeå universitet, Suecia, académica, experta en Ciencias computacionales y directora del *AI Policy Lab*), sobre «How Europe is shaping AI for human rights» en *AI Policy Lab*, (5.09.2024). Visto en https://aipolicylab.se/2024/09/05/how-europe-is-shaping-ai-for-human-rights/

(15) En el capítulo V del Convenio Marco sobre evaluación y mitigación de riesgos e impactos adversos, un único precepto, el artículo 16 sobre Marco de gestión de riesgos e impactos fija, en 4 apartados lo siguiente:

(1) Cada Parte, teniendo en cuenta los principios establecidos en el Capítulo III, adoptará o mantendrá medidas para la identificación, evaluación, prevención y mitigación de los riesgos planteados por los sistemas de inteligencia artificial considerando los impactos reales y potenciales sobre los derechos humanos, la democracia y el Estado. de derecho.

(2) Dichas medidas serán graduadas y diferenciadas, según corresponda, y tendrán en cuenta los siguientes criterios: a) el contexto y el uso previsto de los sistemas de inteligencia artificial, en particular en lo que respecta a los riesgos para los derechos humanos, la democracia y el Estado de Derecho; b) la gravedad y probabilidad de posibles impactos; c) las perspectivas de las partes interesadas relevantes, en particular las personas cuyos derechos puedan verse afectados; d) la aplicación a lo largo de las actividades dentro del ciclo de vida del sistema de inteligencia artificial; e) incluir el seguimiento de riesgos e impactos adversos para los derechos humanos, la democracia y el

(sobre no discriminación)[16] y 18 (derechos de las personas con discapacidad y de los niños)[17] del Convenio Marco del Consejo de Europa.

La iniciativas internacionales para alcanzar una IA fiable y centrada en el ser humano no han cesado, a la de la OCDE y el *Convenio Marco sobre la inteligencia artificial* del Consejo de Europa, le siguen entre otros, el *Pacto para el Futuro y sus anexos*, en concreto, el anexo I sobre el *Pacto Digital Global* de la Asamblea de Naciones Unidas[18], la *Cumbre de Competencia* del G7[19], el *white paper* sobre

---

estado de derecho; f) incluir documentación de los riesgos, los impactos reales y potenciales y el enfoque de gestión de riesgos; y g) exigir, cuando proceda, pruebas de los sistemas de inteligencia artificial antes de ponerlos a disposición para su primer uso y cuando se modifiquen significativamente.

(3) Cada Parte adoptará o mantendrá medidas que busquen garantizar que se aborden adecuadamente los impactos adversos de los sistemas de inteligencia artificial sobre los derechos humanos, la democracia y el estado de derecho. Dichos impactos adversos y las medidas para abordarlos deben documentarse e informar las medidas de gestión de riesgos pertinentes descritas en el párrafo 2.

(4) Cada Parte evaluará la necesidad de una moratoria o prohibición u otras medidas apropiadas con respecto a ciertos usos de los sistemas de inteligencia artificial cuando considere que dichos usos son incompatibles con el respeto de los derechos humanos, el funcionamiento de la democracia o el Estado de derecho.

(16) En el capítulo VI, sobre aplicación del convenio, 6 preceptos, el primero de ello, el artículo 17, fija la no discriminación en cuanto a la aplicación de este marco jurídico.

(17) Dentro del capítulo VI, el artículo 18 fija los derechos de las personas con discapacidad y los de los niños, señalando al efecto que: Cada Parte, de conformidad con su derecho interno y las obligaciones internacionales aplicables tendrá debidamente en cuenta las necesidades y vulnerabilidades específicas en relación con el respeto de los derechos de las personas con discapacidad y de los niños.

(18) El 22 de septiembre de 2024, por la Asamblea General de Naciones Unidas, se aprobó la resolución sobre el *Pacto para el Futuro y sus anexos* para abordar las profundas transformaciones que están operando en este siglo y que es preciso impulsar mayor compromiso con la cooperación internacional a efectos de reafirmar y potenciar los tres pilares de Naciones Unidas (desarrollo, paz y seguridad y derechos humanos). En concreto, en el Anexo I, sobre el *Pacto Digital Global*, y dada la universalización que cabe esperar de una resolución consensuada en el seno del órgano principal de la ONU, se admite que, parte de que las tecnologías digitales están transformando el mundo y que sus beneficios para las personas y la sociedad son inmensos, pero también, lo son los riesgos. Por lo que es precisa la cooperación internacional para eliminar todas las brechas digitales y determinar y mitigar riesgos, además de garantizar la supervisión humana de la tecnología. Entre los objetivos, se fija como imprescindible mejorar la gobernanza internacional de la IA en beneficio de la humanidad. Además, se adoptan compromisos, de aquí a 2030, para alcanzar los numerosos y loables objetivos, entre los que figuran los siguientes: 1. *Eliminar todas las brechas digitales y acelerar los progresos en todos los Objetivos de Desarrollo Sostenible* (ODS); 2. *Ampliar la inclusión en la economía digital y sus beneficios para todos*; 3. *Fomentar un espacio digital inclusivo, abierto y seguro que respete, proteja y promueva los derechos humanos*, para lo que se emplaza, con urgencia, a que se adopten medidas con este fin (apartado 36, c); 4. *Promover enfoques de la gobernanza de datos que sean responsables, equitativos e interoperables para proteger los derechos humanos*, para lo que es preciso alcanzar normas eficaces de protección de los datos personales y de la privacidad (apartado 37); 5. *Mejorar la gobernanza internacional de la inteligencia artificial en beneficio de la humanidad*, para lo que se potencia adoptar un enfoque equilibrado, inclusivo y basado en los riesgos, con la participación de todos los países e interesados. Se trata de aprovechar los beneficios de la IA y mitigar sus riesgos respetando plenamente el derecho internacional, incluido el de los derechos humanos, así como otros marcos como la Recomendación sobre la Ética de la IA de la ONU para la Educación, la Ciencia y la Cultura. El reconocimiento del potencial de la IA para lograr avances en todos los ODS será también crucial y el papel decisivo de la ONU para una gobernanza internacional es incuestionable (apartados 50 y ss.) (p. 43).

(19) La autoridad italiana de competencia (AGCM) ha organizado una Cumbre de Competencia del G7 que, celebrada en Roma los días 3 y 4 de octubre de 2024, se ha centrado en la competencia y los desafíos regulatorios ante una inteligencia artificial imparable y generalizada, tanto en el desarrollo como en el despliegue de las herramientas. Principalmente, por lo que se refiere a los modelos y algoritmos de IA generativa. No faltaron los debates sobre los mercados digitales y los enfoques preventivos (ex ante) y correctivos (ex post) para abordar y hacer frente a los muchos desafíos que se

Governance in the Age of Generative AI: A 360° Approach for Resilient Policy and Regulation del Foro Económico Mundial[20], y recientemente, la celebración de la Cumbre *AI Action Summit*, que, celebrada en Paris los días 10 y 11 de febrero, ha sido objeto de importantes aportaciones y conclusiones[21]. Por lo pronto, constituye el mayor encuentro celebrado hasta la fecha con el denominador común de una IA ética, justa y accesible. Y aunque la Declaración final sobre *Inteligencia Artificial inclusiva y sostenible para las personas y el planeta* no ha sido suscrita por Estados Unidos y el Reino Unido, sí lo ha sido por más de sesenta países, entre ellos China. Parece que poner en valor la IA al servicio del ser humano, como hace el Reglamento de la Unión sigue siendo el eje central de esta Cumbre. Para lo que no falta una pretensión legítima y fundada: es preciso acrecentar y fortalecer la cooperación internacional para promover la coordinación en la gobernanza internacional para llevar a máximos los beneficios de la IA y minimizar los riesgos de la misma[22].

2. REGULAR LA IA SÍ, PERO RESPETANDO EL DERECHO VIGENTE DE LA UE

El desarrollo de la inteligencia artificial se ha convertido en la fuerza impulsora de grandes cambios normativos que se están produciendo a nivel global, desde tratados y acuerdos internacionales que se ocupan de principios y directrices que

---

presentan. La Comisión Europea puso en valor su apuesta regulatoria en esto temas tanto vía Reglamento 2022/1925 sobre mercados disputables y equitativos en el sector digital (Reglamento de Mercados Digitales), como el Reglamento 2024/1689 por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican otros Reglamentos y Directivas (Reglamento de Inteligencia Artificial. Disponible en https://mail.google.com/mail/u/0/#inbox/FMfcgzQXJZsNGvsjZlKTtJqrqFHLKTFC?compose=new

(20) El Foro Económico Mundial, reunido el 8 de octubre de 2024, ha elaborado un *White Paper* sobre *Governance in the Age of Generative AI: A 360° Approach for Resilient Policy and Regulation*, que dirigido principalmente a gobiernos y reguladores, pretende proporcionar políticas de gobernanza de IA generativa que mejoren la coordinación entre los distintos reguladores, promoviendo la colaboración intersectorial e interdisciplinar, asegurando entre otros aspectos: el intercambio de datos, sin menoscabar la privacidad, la transparencia y trazabilidad de los modelos de IA, así como la rendición de cuentas. No faltan aspectos y retos de interés común, que ya mencionados en este capítulo: la privacidad y la protección de datos, los derechos de consumidores, los derechos de autor y afines, o la competencia. La cooperación internacional para adoptar un enfoque armonizado de la respuesta regulatoria será más eficiente si se aborda al nivel más global posible. Disponible en https://www.weforum.org/publications/governance-in-the-age-of-generative-ai/

(21) Una síntesis sobre aspectos relevantes de la Cumbre, la encontramos de la mano de Enrique Benítez Palma, «Diez notas sobre la AI Action Summit de París», *LAB*, Observatorio Sector Público e Inteligencia Artificial, 9 de marzo de 2025. Visto en https://www.ospia.org/o-lab/diez-notas-sobre-la-ai-action-summit-de-paris

(22) Benítez…, op. cit. ant., pone de relieve las siguientes prioridades fijadas en la Cumbre de Paris: (i) Promover la accesibilidad de la IA para reducir las brechas digitales; (ii) Garantizar que la IA sea abierta, inclusiva, transparente, ética, segura y digna de confianza, teniendo en cuenta los marcos internacionales para todos; (iii) Hacer que prospere la innovación en IA facilitando las condiciones para su desarrollo y evitando la concentración del mercado impulsando la recuperación y el desarrollo industriales; (iv) Fomentar un despliegue de la IA que modele positivamente el futuro del trabajo y los mercados laborales y ofrezca oportunidades de crecimiento sostenible; (v) Hacer que la IA sea sostenible para las personas y el planeta; (vi) Reforzar la cooperación internacional para promover la coordinación en la gobernanza internacional.

fomenten un uso ético, seguro y confiable de la herramienta, hasta profundas reformas regulatorias que, iniciadas en la Unión Europea, se han constituido en todo un referente normativo a nivel mundial.

Con todo, la regulación de la inteligencia artificial no se agota en un conjunto armonizado de normas que proporcionan un enfoque normativo de los sistemas de IA desde una perspectiva basada en el riesgo y, en consecuencia, fija obligaciones y requisitos que deben cumplir estas herramientas en función del riesgo que generan[(23)]. Tampoco se limita a la creación de adecuados marcos de gobernanza –por parte de la Unión y de los Estados miembros–, ni a un régimen sancionador adecuado y eficiente. Es más, el Reglamento «debido a su carácter horizontal», debe ser plenamente coherente con las normas de la Unión en todos los sectores donde ya se utilice la IA, o en aquellos donde se vaya a utilizar en el futuro [(24)].

Afloran así, y por ello merecen una mención especial, otras normativas que, conexas a la inteligencia artificial, se ven impactadas por el desarrollo y uso de esta tecnología (Carta de los Derechos Fundamentales de la UE, protección de datos, derechos de los consumidores, propiedad intelectual, o responsabilidad civil), lo que me lleva proponer que el análisis y estudio del Reglamento se haga desde una «perspectiva holística»[(25)]. Así es, no basta con entender la IA y su regulación como una suma de sistemas y de modelos de IA, ni considerar cada sistema de manera aislada para fijar obligaciones y requisitos según el riesgo que genera. En ningún caso se pueden establecer tratamientos estancos para cada uno de los sistemas según el riesgo que generan. Se trata de una tecnología que está en constante evolución y su estudio debe abordarse teniendo en cuenta que la herramienta y su marco normativo deben implementarse con una arquitectura de gobernanza a dos (UE y EE.MM.), teniendo en cuenta el carácter horizontal de la norma, que, *de*

(23) El artículo 3, apartado 2 del Reglamento, define el riesgo como «la combinación de la probabilidad de que se produzca un perjuicio y la gravedad de dicho perjuicio». La mera contingencia o probabilidad de que el perjuicio se dé constituye *per se,* razón suficiente para, mediante normas imperativas, exigir que cualquier sistema de IA, también los modelos que operan en el mercado de la Unión, lo sean con todas las garantías que exige el reglamento para salvaguardar y garantizar los derechos protegidos.

(24) Así se expresa la Comisión Europea en la exposición de motivos, apartado 1.2 de la propuesta de Reglamento de IA que fue publicada el 21 de abril de 2021.

(25) La pretendida regulación integral de la IA, y su estudio, debe abordarse desde «una perspectiva holística», en coherencia con los valores, principios y legislación vigente de la Unión, Con esta expresión pretendo poner de relieve que el estudio y análisis del Reglamento de IA debe hacerse desde este enfoque holístico. No cabe *entender la IA, y su regulación, como una suma de sistemas y de modelos de IA* sujetos a un régimen de obligaciones y requisitos, e inmersos en un sistema óptimo de gobernanza en función del mayor o menor riesgo que representan para las personas. Debe tenerse en cuenta el carácter horizontal de la norma y la pretensión de la UE de crear un ecosistema integral en cohesión con valores, principios y normas europeas (derechos fundamentales, protección de datos, de consumidores, de propiedad intelectual o responsabilidad civil). Tal «perspectiva holística» permite implementar la regulación de los sistemas y modelos de la IA garantizando la seguridad jurídica que proporciona el ordenamiento jurídico de la Unión, y protegiendo los derechos de las personas. El enfoque holístico fue propuesto por la autora de esta publicación en julio de 2024, en el marco del Subproyecto 1.1. «Mapping of AI governance – recommendations and regulation AI Governance (Project #1)», cAIre Research Project OdisedIA FOR GOOGLE.org. Un extracto del mismo, y vinculado a las personas y colectivos vulnerables fue publicado en *Blog OdiseIA*, 6 de enero 2025 bajo el título: «¿Protege el Reglamento de IA a los vulnerables? Análisis desde una perspectiva holística y necesaria cohesión con otras normas y derechos consolidados» Visto en https://www.odiseia.org/post/protege-el-reglamento-de-ia-a-los-vulnerables.

facto, alcanza a ser completamente integral y abarcativa. Para ello, deberán tenerse en cuenta los siguientes aspectos:

1. Los límites éticos de la UE, y los adoptados en los sucesivos acuerdos internacionales;
2. Además, deberá ponerse en coherencia con la legislación sectorial, europea y nacional, que habrá de adaptarse, no sólo a las propiedades de autoaprendizaje y autonomía de esta tecnología, también, a los aspectos de complejidad, opacidad, e incluso, de imprevisibilidad.

En definitiva, se trata de que el ilimitado e imparable desarrollo de la IA, con todos los beneficios que ello supone para los individuos y la sociedad, se desarrolle con las mínimas restricciones técnicas[(26)] y las máximas garantías éticas y jurídicas para las personas. En última instancia, se trata de garantizar que los sistemas de IA sean seguros y respeten los derechos fundamentales, valores y normas de la UE[(27)].

Así, la incidencia de la inteligencia artificial en nuestro ordenamiento, tanto público como privado es inmensa y los retos que se plantean también, por lo que es preciso, como hemos señalado con anterioridad, establecer límites éticos y normativos efectivos y adecuados. Afrontar esta nueva era es tarea de todos, y para ello, es oportuno adentrarnos en algunos de los ámbitos del Derecho privado que se ven afectados por el desarrollo e implementación de la IA y que dan lugar a adaptar el marco regulatorio europeo en materias previamente reguladas y consolidadas.

Grosso modo, el uso de estas herramientas, junto al calado de la regulación vía Reglamento europeo de IA se expande a nuestro día a día (salud, educación, contratación laboral, acceso a prestaciones sociales, identidad digital, transportes públicos o privados, préstamos hipotecarios o el uso de plataformas digitales de cualquier tipo), y también a todos los ámbitos sin excepción, por lo que cualquier área vinculada al uso de estas herramientas y con impacto jurídico ha de ser: (i) o bien adaptada o reformulada para adecuarse al impacto que produce en derechos fundamentales, protección de datos, derechos de consumidores o responsabilidad civil; (ii) o bien regulada si no lo ha sido antes: control judicial o discriminación de los algoritmos. Y por descontado, teniendo en cuenta la especial consideración hacia las personas o colectivos vulnerables[(28)].

---

(26) La exposición de motivos antedicha, en el apartado 1.1. expresa el propósito de la Unión de alcanzar la excelencia y el liderazgo tecnológico, y, además, de innovar.

(27) También, en el mismo apartado citado anteriormente, se fija el propósito de alcanzar la excelencia y la confianza mediante una IA fiable, ética y segura en la UE.

(28) Sobre las personas y colectivos vulnerables merecen especial mención las publicaciones de: Lorenzo Cotino Hueso (2024) «Vulnerabilidad y colectivos vulnerables en el Derecho: ¿quiénes son y cómo se definen?» en el marco del Subproyecto 2.1. «Digital Futures Initiative (Project #2)», cAIre Research Project OdiseIA FOR GOOGLE.org. *Revista Catalana de Dret Public (RCDP blog)*, 18 de diciembre de 2024. Visto en https://eapc-rcdp.blog.gencat.cat/2024/12/18/vulnerabilidad-y-colectivos-vulnerables-en-el-derecho-quienes-son-y-como-se-definen-lorenzo-cotino/; Begoña González Otero «AI for Good: La idea de la vulnerabilidad humana en tela de juicio» en el marco del Subproyecto 1.4. «AI for Good (Project #1)», cAIre Research Project OdiseIA FOR GOOGLE.org. *Blog OdiseIA*, 6 de septiembre 2024. Visto en https://www.odiseia.org/post/ai-for-good-la-idea-de-la-vulnerabilidad-humana-en-tela-de-juicio-1, o el ya citado de Carmen Muñoz García, «¿Protege el Reglamento de IA a los vulnerables? Análisis desde una perspectiva holística y necesaria cohesión con otras normas y derechos consolidados AI Governance» op. cit.

Del impacto positivo de la inteligencia artificial da cumplida muestra la incidencia de la herramienta en ámbitos como el de la salud (aunque los ejemplos son innumerables), sirviendo este contexto para poner en valor la trascendencia de esta herramienta en uno de nuestros bienes y derechos más preciado. Su potencial transformador, que hace posible la implementación de una medicina preventiva, predictiva, personalizada, y de mayor alcance y precisión, mejora la atención sanitaria y propicia innumerables avances en investigación que hasta muy recientemente eran impensables[29]. De la relevancia de la IA en esta materia da cumplida muestra que el premio nobel de química de 2024, D. HASSABIS, fue merecedor del mayor reconocimiento científico internacional por el desarrollo de *AlphaFold*, una herramienta de IA capaz de describir la estructura de los 200 millones de proteínas conocidas, y que puso de relieve que «normalmente, el hecho de predecir la estructura de una proteína conlleva todo el trabajo de un estudiante de doctorado que dispone de medios sofisticados. Esto significa que descubrir hasta 200 millones de proteínas necesitaría cinco veces 200 millones de años»[30], todo un avance que no tiene parangón con lo que sería el progreso humano sin la ayuda de la IA.

Ahora bien, los riesgos son enormes y los retos también, entre estos: mantener la protección de datos y la privacidad en general, potenciar un verdadero consentimiento informado, la transparencia y explicabilidad cuando se participa en cualquier cesión de información o cuando se realizan pruebas médicas o se participa de prácticas automatizadas que clasifican, o excluyen, a pacientes para tratamientos propios. Y lo mismo cabe decir respecto a las decisiones automatizadas que perpetúan errores del pasado, como es la denegación de un préstamo anterior[31], o la privación del carnet de conducir años atrás. No faltan los riesgos derivados de las alucinaciones[32] que propician respuestas que

(29) Cabe poner de relieve que Demis HASSABIS, fundador de *DeepMind* (filial de Google) y Nobel de Química 2024 por su uso de la IA en concreto por el desarrollo de *AlphaFold*, afirma que, en la actualidad, cerca de dos millones y medio de científicos utilizan esta aplicación, lo que, de primeras, puede ayudar a desarrollar fármacos para curar enfermedades «especialmente aquellas que afectan los países más pobres y en que no invierten las grandes farmacéuticas». Visto en https://www.elconfidencial.com/tecnologia/2025-02-10/demis-hassabis-google-inteligencia-artificial-ia-paris-macron_4061065/ Sobre *AlphaFold* https://blog.google/technology/ai/google-deepmind-isomorphic-alphafold-3-ai-model/#life-molecules

(30) Demis Hassabis, op. cit. ant.

(31) Ante las decisiones totalmente automatizadas que producen efectos legales o significativos en el interesado, ver la sentencia del TJUE de 27 de febrero de 2025, asunto C-203/22 (*Dun & Bradstreet Austria*) que incide, como ya lo hiciere la STJUE de 7 de diciembre de 2023, asunto C-634/21 (*Schufa Holding*), en la relevancia de la debida aplicación del artículo 22 del RGPD. Según este, el afectado (i) por decisiones totalmente automatizadas, y (ii) que le producen efectos negativos, tiene derecho a ser informado de esta adopción, de una manera clara y suficientemente concisa. De esta manera, podrá verificar de la exactitud de los datos e impugnar cualquier decisión con conocimiento de causa.

(32) La primera acepción del verbo alucinar, según el Diccionario de la Real Academia Española, es «ofuscar, seducir o engañar haciendo que se tome una cosa por otra». A esta acepción conviene añadir que se trata de una acción involuntaria. Este significado resulta particularmente pertinente para describir un fenómeno cada vez más frecuente en el ámbito de la inteligencia artificial, especialmente en relación con los modelos de lenguaje natural (NLP). Dichos sistemas, diseñados para generar texto de manera autónoma, pueden producir resultados que, aun presentándose como coherentes y verosímiles, contienen afirmaciones fácticamente incorrectas, inexactas o carentes de respaldo empírico. Este fenómeno, comúnmente denominado «alucinación» en el ámbito técnico,

inducen a la toma de decisiones equivocadas, o la creación de los *deepfakes* que lesionan derechos personalísimos como el derecho al honor, a la intimidad o a la propia imagen.

Sin duda, estas situaciones requieren de respuestas jurídicas para las que muy probablemente no es suficiente el marco regulatorio existente hasta la fecha, entendiendo este en su conjunto, incluido el Reglamento de IA y también las normas conexas. A lo anterior, se suma un reto aún más relevante si cabe, teniendo en cuenta que, cuando se causan daños por el desarrollo y uso de la IA, debe ser posible para el perjudicado exigir, en condiciones de viabilidad, la responsabilidad civil, tanto contractual como extracontractual. Así debe ser si se trata de generar confianza en los usuarios ya que estos deben poder ser reparados.

Con todo, y volvemos al ámbito de la salud en el que todos los ciudadanos participamos, la integración de grandes modelos de lenguaje (LLM) en la atención sanitaria suscita gran interés, y ello, a pesar de la actual calidad de los resultados que no son todo lo precisos que sería deseable. Su capacidad de respuesta y de ofrecer soluciones con el entrenamiento previo de enormes volúmenes de datos, imágenes, textos, artículos, con las fuentes de internet, y además, con retroalimentación y supervisión humana, permiten que, los grandes modelos de lenguaje (LLM) potencien las capacidades de procesamiento del lenguaje natural (PNL) y generen información en segundos que podría escapar al ojo humano y a su conocimiento, y que además, podría suponer semanas, meses o años de análisis de datos y estudios. Los primeros datos apuntan a que los denominados modelos de inteligencia artificial generativa, como son los *chatbots* conversacionales, propicien un avance en la medicina a la que contribuirá la integración clínica, el asesoramiento, el diagnóstico, la prevención y el tratamiento personalizado. Con todo, el problema sigue siendo las alucinaciones o la información sesgada[33].

Por todo lo anterior, resulta conveniente comenzar justificando que, aunque hay ciertos déficits normativos en el Reglamento, es necesario visibilizar la forma en que la inteligencia artificial está afectando a la mayoría de los marcos regulatorios existentes, por lo que veremos qué, desde el análisis del RIA, son precisas «adaptaciones o reformulaciones»[34] de normas consolidadas que, debido al enorme impacto de la inteligencia artificial, necesitan ser revisadas, como

---

tiene su origen en la arquitectura probabilística de dichos modelos, la cual se basa en la predicción estadística de secuencias lingüísticas a partir de grandes volúmenes de datos, sin que exista una comprensión semántica ni una validación ontológica del contenido generado. Sirva como ejemplo el siguiente: un ciudadano noruego solicitó a ChatGPT que le proporcionara información sobre él mismo. El *output* que generó resultó ser una historia falsa que lo describía como alguien condenado a veintiún años de prisión por el asesinato de dos de sus hijos e intento de asesinato del tercero. La información era sobre él, pero junto a datos reales y verificables había alucinaciones que le causaban graves daños y perjuicios. Este caso se ha dado a conocer en: https://noyb.eu/en/ai-hallucinations-chatgpt-created-fake-child-murderer

(33) Bright Huo, Amy Boyle, Nana Marfo, et al, «Large Language Models for Chatbot Health Advice Studies A Systematic Review». *PubMed*, 4 febrero 2025. Visto en https://pubmed.ncbi.nlm.nih.gov/39903463/

(34) En una primera publicación sobre el impacto de la IA en la responsabilidad civil: Carmen Muñoz García, «Adaptar o reformular la Directiva 85/374 sobre responsabilidad por daños causados por productos defectuosos a la Inteligencia Artificial». *Diario La Ley. Ciberderecho*, 28 de febrero de 2022. Reeditado como «Adaptar… Últimas novedades y un cambio de rumbo». *Revista Crítica de Derecho Inmobiliario*, vol. 793, noviembre-diciembre 2022, pp. 2886-2908

son las del Derecho de consumo y la responsabilidad civil. Y ello, además, teniendo en cuenta que, conforme al considerando 9 del RIA, este debe también entenderse en cohesión con el Reglamento sobre «seguridad de los productos», al que complementa[35].

Al fin y a la postre, como se indica en el considerando 1 del RIA, se trata de mejorar el funcionamiento del mercado interior de la UE, por lo que, si el propósito de la ley es determinar cómo debe ser la IA que se desarrolle, comercialice y use en el mercado de la Unión, su regulación uniforme e imperativa pretende que esté centrada en el ser humano. Para ello, como hemos dicho con anterioridad, debe ser segura y fiable, y garantizar un nivel óptimo de protección de la salud, la seguridad y los derechos fundamentales consagrados en la Carta de los Derechos Fundamentales de la Unión Europea (en lo sucesivo, «Carta»), incluidos la democracia, el Estado de Derecho y la protección del medio ambiente.

Puede decirse que, en materia de IA, Europa busca reforzar la confianza en una tecnología que resulte competitiva y beneficiosa para las personas y las empresas. Dicho esto, no será fácil para la UE encontrar el equilibrio entre intereses tan contrapuestos (empresariales y personales), aunque tampoco lo ha sido en otros contextos previos a la publicación del Reglamento de IA, como el de consumidores, la responsabilidad civil o la protección de datos. Así, si tenemos en cuenta los intereses contrapuestos en otras normativas nucleares conexas de la Unión, vemos que, en el caso de que la regulación propia de la IA no proporcione respuestas claras sobre un conflicto de intereses, habrán de aplicarse las normas vigentes que ayuden a resolver las situaciones de riesgo derivadas del uso de herramientas inteligentes. Sirvan como ejemplo, entre otros: (i) el RGPD, que contrapone al interesado frente al responsable o encargado del tratamiento, (ii) la normativa de consumidores, que enfrenta al consumidor frente al empresario o profesional, (iii) o como la Directiva sobre responsabilidad por producto defectuoso enfrenta, inicialmente, al consumidor frente al productor (fabricante o importador), aunque en la nueva Directiva 2024/2853 sobre responsabilidad por daños causados por productos defectuosos[36], se protege a la persona física frente al «operador económico». Por lo tanto, no es de extrañar que el Reglamento de IA contraponga, en líneas parecidas, aunque no iguales, a la persona física frente al operador de sistemas de inteligencia artificial. Y ello a sabiendas de la prevalencia de todos los derechos y valores consolidados en el Derecho de la Unión.

---

(35) El citado considerando 9 expresa: «Las normas armonizadas que se establecen en el presente Reglamento deben aplicarse en todos los sectores y, en consonancia con el nuevo marco legislativo, deben entenderse sin perjuicio del Derecho vigente de la Unión, en particular en materia de protección de datos, protección de los consumidores, derechos fundamentales, empleo, protección de los trabajadores y seguridad de los productos, al que complementa el presente Reglamento. En consecuencia, permanecen inalterados y siguen siendo plenamente aplicables todos los derechos y vías de recurso que el citado Derecho de la Unión otorga a los consumidores y demás personas que puedan verse afectados negativamente por los sistemas de IA, también en lo que respecta a la reparación de los posibles daños de conformidad».

(36) Directiva 2024/2853 del Parlamento Europeo y del Consejo, de 23 de octubre de 2024, sobre responsabilidad por los daños causados por productos defectuosos y por la que se deroga la Directiva 85/374/CEE del Consejo, de 23 de octubre de 2024, publicada en el DOUE el 18 de noviembre de 2024 (en adelante, Directiva 2024/2853 o Directiva sobre responsabilidad por los daños causados por productos defectuosos).

II. LOS SUJETOS PROTEGIDOS Y LA AUTONOMÍA DE LA VOLUNTAD EN LA REGULACIÓN DE LA IA

1. LA PERSONA EN EL CENTRO DE LA REGULACIÓN DE LA IA

A. El ser humano en el «estándar jurídico» del RIA

El Reglamento de IA reconoce que estamos ante una regulación de esta herramienta, completamente disruptiva y con capacidades ilimitadas, para el adecuado funcionamiento del mercado interior de la UE. También se constata que la finalidad de regular esta materia y de proporcionar un régimen riguroso e imperativo, con obligaciones y requisitos ambiciosos, es establecer un marco uniforme que garantice que esta tecnología opera de acuerdo con los valores y derechos de la UE en beneficio del ser humano, y que la haga fiable y garantista. No es una ley para la persona física, pero sí crea el marco jurídico óptimo para protegerla de forma eficiente. Y se ha pretendido normativizar para promover la adopción de una IA centrada en el ser humano, lo que no es baladí a lo largo de una regulación exhaustiva y de carácter esencialmente imperativo.

En un contexto en el que aumentan los automatismos y la generación de contenidos al margen de actos voluntarios, la defensa del ser humano y su desarrollo en un contexto social óptimo se ha convertido en objetivo esencial. Para ello, la autonomía de la voluntad, que se vincula con la autodeterminación de las personas debe ser el eje de todo el Derecho privado, como poder de disposición al tiempo de dar forma a la esfera jurídica individual. Se trata de autorregular los intereses propios mediante reglas ajustadas a Derecho y asumidas por cada persona individualmente, quien podrá crear, modificar o extinguir relaciones jurídicas, y disponer de sus derechos a voluntad, con las limitaciones propias de las normas imperativas.

No hay duda, la posición del Reglamento es precisa, y aunque no declara expresamente quién es el sujeto protegido, el marco regulatorio riguroso lo es para proteger a la persona física, una vez que las normas se establecen para garantizar que el desarrollo, la introducción en el mercado, la puesta en servicio y la utilización de sistemas de IA admitidos en el mercado de la Unión, debe abordarse conforme a valores y derechos en beneficio del ser humano. Se trata de «promover la adopción de una inteligencia artificial (IA) garantizando al mismo tiempo un elevado nivel de protección de la salud, la seguridad y los derechos fundamentales consagrados en la Carta de los Derechos Fundamentales de la Unión Europea (en lo sucesivo, «Carta»), incluidos la democracia, el Estado de Derecho y la protección del medio ambiente»[(37)]. Y añade poco después que, dado el alcance y las repercusiones de la IA en la sociedad, y en cualquiera de los contextos en los que nos desarrollamos, la IA debe cumplir como requisito previo a su inserción en el mercado interior de la UE, que sea «una tecnología centrada en el ser humano. Además, debe ser una herramienta para las personas y tener por objetivo último aumentar el bienestar humano»[(38)].

(37) Así se expresa en el primero de los considerandos del RIA.
(38) Sin atisbos, así lo expresa en el considerando 6 del RIA.

Es más, la clasificación de los sistemas de IA atendiendo al riesgo que generan en: 1. «de riesgo inaceptable»; 2. «de alto riesgo»; 3. «de riesgo limitado o medio»; y 4. «de riesgo bajo o mínimo», lo es teniendo en cuenta que este se considera en función de que suponga una amenaza o peligro para la salud, la seguridad, los derechos fundamentales, la democracia, el Estado de Derecho y la protección del medio ambiente.

El marco regulatorio nos confirma esta idea a lo largo de sus considerandos y artículos. Además de los antedichos en cuanto a una regulación: «en beneficio del ser humano» y «debe ser una herramienta para las personas» (cdos. 1 y 6 del RIA, respectivamente), se incide en que no cabe eludir su cumplimiento y se trata de «garantizar la protección efectiva de las personas físicas ubicadas en la Unión» (cdo.22); fija los deberes de orientación e información cuando estemos ante sistemas de alto riesgo «incluyan mecanismos destinados a orientar e informar a las personas físicas» (cdo. 73, y en parecidos términos el cdo. 93); mientras que, a los proveedores de modelos de IA de uso general, les dispensará de determinadas obligaciones al tiempo de suministro de un producto o un servicio siempre que «los derechos de las personas físicas no se vean afectados» (cdo. 97). También es de destacar lo que se refiere respecto a los sistemas concretos de IA destinados a interactuar con personas físicas o a generar contenidos «es preciso comunicar a las personas físicas que están interactuando con un sistema de IA, excepto cuando resulte evidente desde el punto de vista de una persona física normalmente informada y razonablemente atenta y perspicaz, teniendo en cuenta las circunstancias y el contexto de utilización» (cdo. 132).

El Reglamento es aún más tajante al ocuparse de los abundantes riesgos de esta tecnología que, *con la clara pretensión de emular la inteligencia humana*, también hace posible manipular la voluntad de las personas o desarrollar o usar sistemas con riesgos indebidos, inadecuados e incluso lesivos para el ser humano:

a) En cuanto a los *sistemas de IA prohibidos*[39] se fija en el artículo 5 que se prohíbe esta tecnología cuando utilice «técnicas subliminales que transciendan la conciencia de una persona o de técnicas deliberadamente manipuladoras o engañosas» (apartado 1, letra a); «que explote alguna de las vulnerabilidades de una persona física» (apartado 1, letra b); que realice «evaluaciones de riesgos de personas físicas con el fin de valorar o predecir el riesgo de que una persona física cometa un delito basándose únicamente en la elaboración del perfil de una persona física o en la evaluación de los rasgos y características de su personalidad» (apartado 1, letra d); «que creen o amplíen bases de datos de reconocimiento facial» (apartado 1, letra e); que ayuden a colegir «las emociones de una persona física en los lugares de trabajo y en los centros educativos, excepto… por motivos médicos o de seguridad» (apartado 1, letra f);

(39) Desde el pasado 2 de febrero de 2025 son de aplicación tanto el capítulo I del RIA (disposiciones generales), como el capítulo II (sistemas prohibidos). Con todo, el 4 de febrero de 2025, la Comisión Europea ha presentado las *Directrices sobre prácticas prohibidas*. Todo un hito para ayudar a esclarecer el alcance de todas estas prohibiciones que, aunque inicialmente se presentan como absolutas, parecen no serlo tanto. Sobre estas, Lorenzo Cotino Hueso (2025) «¿Cuándo "no es no"? Criterios para definir los sistemas de inteligencia artificial prohibidos en la Unión Europea» *Revista General de Derecho Administrativo,* 2025, nº (en prensa).

b) Por lo que se refiere a los *sistemas de alto riesgo*, por un lado, el artículo 6, apartado 3 del RIA señala que «no obstante lo dispuesto en el apartado 2, un sistema de IA a que se refiere el anexo III no se considerará de alto riesgo cuando no plantee un riesgo importante de causar un perjuicio a la salud, la seguridad o los derechos fundamentales de las personas físicas, también al no influir sustancialmente en el resultado de la toma de decisiones»; por otro lado, para estos sistemas, en el artículo 14 se exige como requisito esencial el de la supervisión humana; además, los responsables del despliegue de estos sistemas cuando tomen decisiones, o ayuden a tomarlas: «informarán a las personas físicas de que están expuestas a la utilización de los sistemas de IA de alto riesgo» (art. 26, apartado 11);
c) También el artículo 50 del RIA, en relación con otros sistemas, *los de riesgo medio*, se exigen obligaciones de transparencia de los proveedores y responsables del despliegue. Estos «garantizarán que los sistemas de IA destinados a interactuar directamente con personas físicas se diseñen y desarrollen de forma que las personas físicas de que se trate estén informadas de que están interactuando con un sistema de IA, excepto cuando resulte evidente desde el punto de vista de una persona física razonablemente informada, atenta y perspicaz, teniendo en cuenta las circunstancias y el contexto de utilización», en la misma línea que lo anticipado en el considerando 132 de la norma.

Todo lo anterior sirve para avalar que la tecnología está al servicio de las personas, y que el ser humano es la razón de ser de esta norma, y aunque no faltan argumentos que llevan a mantener dicha afirmación, la premisa inicial del Reglamento sirve de base a este propósito: se trata de que el sistema no suponga una amenaza o riesgo para los derechos fundamentales de la persona física, la salud y la seguridad.

B. La importancia de los sistemas destinados a interactuar con persona física

En conexión con el recién mencionado artículo 50 de RIA, y dado que en materia de IA se exige que medie el requisito de «transparencia», es esencial entender que los proveedores de estas herramientas deben asumir esta obligación para propiciar un desarrollo y uso ético, seguro y confiable de la IA. La transparencia constituye, desde hace años, una exigencia en la contratación con los consumidores (parte débil en la relación contractual). Y es ahora, a través del Reglamento de IA, que se ha incorporado, como principio y como norma[(40)], para proteger a los usuarios frente a la opacidad, imprevisibilidad y autonomía de estas herramientas. Y es así, no sólo para los sistemas que suponen un riesgo alto para los derechos fundamentales, la salud o la seguridad, también se exige para los sistemas destinados a interactuar con personas físicas, y que, por la propia naturaleza de esta tecnología, conllevan riesgos de *suplantación, engaño, falsedad o manipulación*.

(40) Así lo veremos con ocasión de este requisito de «transparencia» del que nos ocupamos en apartado diferente.

En concreto, es en el ámbito de la denominada «inteligencia artificial generativa» (ejemplo tipo de los modelos de IA de uso general)[41], en el que las impactantes y singulares características de generalidad y flexibilidad en la creación de contenidos de esta herramienta, propiciarán los malos usos a los que hemos aludido.

C. Cuando la IA causa daño a tercero ajeno a la tecnología. La responsabilidad extracontractual

En consonancia con lo anterior, es innegable que cualquier operador de inteligencia artificial puede producir daños a terceros ajenos a la herramienta mediante el desarrollo o uso de la IA. Su despliegue y utilización no agota sus efectos respecto a quienes se sirven de la inteligencia artificial, sino que sus capacidades para generar contenido, e incluso para la toma de decisiones, pueden afectar a terceros que no participan de esta tecnología. Por ejemplo, los servicios públicos relativos a prestaciones sociales, asistencia sanitaria, o búsqueda de empleo que operan con IA, y que pueden dejar fuera del sistema a quienes, cumpliendo los requisitos, queden excluidos. ¿A qué puede deberse esa exclusión no justificada? Puede ser por defectos del sistema desde el diseño (que opera por sí mismo o integrado en un producto), a sesgos en el entrenamiento o a un factor genérico que el sistema discrimina, también por falta de actualizaciones, o por un uso indebido o inadecuado de los datos por parte de quien los proporciona al sistema. No obstante, si no se corrigen los errores del sistema, estos se perpetúan y agravan.

Como hemos puesto de manifiesto en otra publicación precedente la protección de las personas físicas en el Reglamento, debe extenderse «al margen de cualquier relación contractual, con independencia de que la persona, para obtener bienes o servicios, opere con IA en el mercado y se vea afectada de manera indirecta por los efectos negativos que esta genere o propicie, ya sea por los automatismos, la creación de contenidos, o por las decisiones contrarias a Derecho que adopte». Y es así «porque habrá ocasiones en las que la persona física, sin ser parte de una relación jurídica contractual, sin operar en el mercado para la obtención de un bien o servicio, o sea, sin haber realizado actuaciones propias en el mercado, esté sometido a las acciones u omisiones de la IA a través de actuaciones de terceros»[42]. Sin duda, serán muchos los supuestos que encajarán en esta posibilidad ya que los riesgos de esta tecnología serán incontables.

En este caso, *la responsabilidad civil extracontractual será el cauce para obtener la reparación de los daños*, si bien, y a falta de una normativa expresa que contemple estos supuestos singulares que son los perjuicios causados por inteligencia artificial, no es desdeñable exigir la reparación por otras vías contempladas en el acervo comunitario, al margen de las sanciones por infracción del marco

(41) Sobre estos, ver Carmen Muñoz García «Modelos de IA de uso general y sistemas de IA de riesgo limitado y mínimo», en Barrios Andrés, M, (coord.), *El Reglamento Europeo de Inteligencia Artificial*, Editorial Tirant lo Blanch, Valencia, 2024, pp. 87-109.

(42) Carmen Muñoz García, «La persona física como sujeto de derechos en el Reglamento Europeo de IA», artículo remitido para su evaluación y publicación a la *RCD,* 2025.

normativo contemplado en el RIA[43]. Así, para obtener la reparación cuando medien daños:

— Por el momento, rige la Directiva 85/374, *sobre la responsabilidad por los daños causados por productos defectuosos*, aún en vigor[44]. Esta, continuará siendo de aplicación con igual régimen de responsabilidad objetiva por daños causados por productos defectuosos, salvo excepciones, pero con toda la dificultad que ello entraña en cuanto al alcance del concepto «producto», la dificultad probatoria del defecto del producto y la relación causal entre el defecto y el daño.
— Si bien, será la Directiva 2024/2853, de 23 de octubre de 2024, *sobre responsabilidad por los daños causados por productos defectuosos* (DRPD)[45], y que deroga la precedente, la que por el momento tendrá que amparar los daños que se produzcan por la IA, a falta de una regulación singular para esta tecnología. Esta Directiva, que entró en vigor a los veinte días de su publicación, y será aplicable a los productos introducidos en el mercado o puestos en servicio después del 9 de diciembre de 2026,

(43) Es clave garantizar la gobernanza y eficacia del Reglamento vía imposición de «sanciones y otras medidas de ejecución», como se denomina en el considerando 168 del RIA, que se configura en los siguientes artículos del texto: (1) en el artículo 99 se establece un régimen sancionador general dirigido a los operadores por falta de cumplimiento del Reglamento, cuya competencia es de los EE. MM.; (ii) en el artículo 100 se establece un régimen sancionador específico para la Administración de la UE (instituciones, órganos y organismos de la UE), atribuyendo su competencia al Supervisor Europeo de Protección de Datos; y (iii) un régimen sancionador específico para los proveedores de los denominados «modelos de IA de uso general», que se contiene en el artículo 101 del RIA, cuya competencia se atribuye a la CE. Ver al respecto Joaquín Delgado Martín «Capítulo VI. Régimen sancionador», en Barrios Andrés, M. (dir.), *El Reglamento Europeo de Inteligencia Artificial*, Editorial Tirant lo Blanch, Valencia, 2024, p. 201 y ss.

(44) La Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, *relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos*, mostró durante años su adaptabilidad a los nuevos tiempos, pero entraña profundas dificultades para su aplicación a productos de la era digital y la economía circular. También ha dejado patente, a lo largo de estos años de vigencia, que el problema mayor para el perjudicado ha sido el de la carga de la prueba del defecto del producto causante del daño y de la relación causal. Estas dificultades se han mostrado en toda su extensión con: (i) los productos sanitarios y farmacéuticos, (ii) la complejidad, opacidad y autonomía de las nuevas tecnologías, y (iii) en el contexto de la economía circular, cuando median empresas que modifican producto.

(45) Si bien, en fecha de 28 de septiembre de 2022 la Comisión Europea presentó dos propuestas para adaptar la Directiva 85/374/CEE del Consejo, de 25 de julio de 1985, *relativa a la aproximación de las disposiciones legales, reglamentarias y administrativas de los Estados Miembros en materia de responsabilidad por los daños causados por productos defectuosos*, a las nuevas tecnologías y a la economía circular, sólo una de ellas ha culminado el proceso legislativo con éxito. Se trata de la Directiva, 2024/2853, de 23 de octubre de 2024, *sobre responsabilidad por los daños causados por productos defectuosos* y por la que se deroga la Directiva 85/374/CEE del Consejo, publicada el 18 de noviembre en el DOUE, fija en el primero de los considerandos, con absoluta firmeza que: «establece normas comunes en materia de responsabilidad por los daños causados por productos defectuosos, con el fin de eliminar las divergencias entre los ordenamientos jurídicos de los Estados miembros que puedan falsear la competencia y afectar a la circulación de mercancías dentro del mercado interior. Una mayor armonización de las normas comunes sobre responsabilidad por los daños causados por productos defectuosos establecidas en dicha Directiva contribuiría aún más a la consecución de estos objetivos, al tiempo que supondrá un mayor grado de protección de la salud o la propiedad de los consumidores y de otras personas físicas».

forma parte de un marco normativo complejo, integral, que pretende establecer normas armonizadas para abordar los riesgos derivados del desarrollo y uso de la IA. Se impone un sistema de responsabilidad objetiva, a pesar de contener excepciones, e incluye en el cuerpo normativo lo que no cubría la precedente Directiva de 1985: (i) son también producto los «archivos de fabricación digital» y los «programas informáticos»; (ii) su aplicación se dirige a proteger «al consumidor y otras personas físicas»; (iii) cabe exigir la responsabilidad a cualquier «operador económico». A pesar de todo, podremos encontrarnos supuestos de difícil cobertura a través de esta norma. Por ejemplo, ¿será posible determinar si el fallo de un dispositivo médico equipado con IA y que se usa para analizar imágenes médicas o el de un algoritmo que gestiona el frenado en un coche autónomo se debe a un defecto del producto o a una recomendación incorrecta? ¿Se trata de un fallo del *software* que no se integra bien en el producto?

— Así las cosas, no es ocioso recordar que, al tiempo de proponerse la Directiva precedente, en fecha de 28 de septiembre de 2022, se presentó una propuesta complementaria con la DRPD, y que establecía un régimen propio para los daños causados por la IA. Esta proposición, que fue denostada por muchos[(46)], no logró el consenso entre los colegisladores europeos, y finalmente, ha sido retirada por la Comisión en marzo de 2025. No es extraño, ni respondía a los objetivos iniciales de la Comisión Europea desde el Libro Blanco sobre la IA, de 19 de febrero de 2020, ni a la clara pretensión del PE desde su resolución de 20 de octubre de 2020 que pretendían armonizar la regulación de la IA, y abordar los problemas del desarrollo y uso de la IA de manera uniforme para todo el mercado de la UE. La propuesta de Directiva finalmente retirada por la Comisión era un marco de mínimos, que no atendía al riesgo de los sistemas[(47)], y que proclamaba una responsabilidad basada en la culpa, alejada de los modelos de responsabilidad objetiva, que dificultaban cualquier pretendida reclamación por daños.

— Con todo, nos queda una pregunta más: ¿y si es el Estado el que, sirviéndose de la IA, contraviene el Derecho de la Unión y causa daños a personas ajenas al desarrollo y uso de esta? El denominado «principio de res-

(46) Entre otros, Carmen Muñoz García tanto en la publicación sobre la «Directiva 2024/2853 sobre responsabilidad por los daños causados por productos defectuosos. Contexto y armonización de máximos para proteger, ahora sí, a consumidores y "a otras personas físicas"» *La Ley Unión Europea*, número 132, enero 2025, como en la monografía sobre *Regulación de la inteligencia artificial en Europa. Incidencia en los regímenes jurídicos de protección de datos y de responsabilidad por producto*, op. cit., p. 184 y ss., en la que se emite una crítica a la armonización de mínimos y a la situación de indefensión en la que colocaba a cualquier perjudicado en cuanto básicamente establecía un régimen de responsabilidad basado en la conducta culposa del operador y con fases probatorias muy complejas. También Máximo J. Pérez García, «La responsabilidad por los daños causados por los productos defectuosos y la Directiva (UE) 2024/2853, de 23 de octubre», *Blog Facultad de Derecho*, UAM, 14 enero 2025. Visto en https://www.blog.fder.uam.es/responsabilidad-por-los-danos-causados-por-los-productos-defectuosos-y-la-directiva-ue-2024-2853-de-23-de-octubre/

(47) La idea de la responsabilidad atendiendo al riesgo también late en los Principios de Derecho Europeo de responsabilidad Civil (*European Group on Tort Law*, 2005). En el art. 5:101 se establece que: «La persona que lleva a cabo una actividad *anormalmente peligrosa* responde por el daño característico del riesgo que tal actividad comporta y que resulta de ella».

ponsabilidad extracontractual del Estado», ampliamente tratado por autores de Derecho Administrativo y Derecho civil, permite esta vía para reclamar por daños causados por la IA cuando es el Estado, a través de cualquiera de sus órganos e instancias, quien causa daño a terceros[48]. Al respecto, y a falta de norma expresa, desde el TJUE se ha venido configurando un régimen de responsabilidad del Estado por daños a los particulares cuando la Administración Pública, por acciones u omisiones, infringe el DUE y no cumple su obligación de «adoptar todas las medidas generales o particulares apropiadas para asegurar el cumplimiento de las obligaciones que les incumben en virtud del Derecho comunitario. Entre esas obligaciones se encuentra la de eliminar las consecuencias ilícitas de una violación del Derecho comunitario», y que encuentra su fundamento en el artículo 4, apartado 3, segundo párrafo del Tratado de la Unión Europea[49]. En España, será la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público (LRJSP), la que más allá de la regla general de responsabilidad del Estado por cualquier lesión que sufran los particulares, como consecuencia del funcionamiento «normal o anormal» de los servicios públicos fije una responsabilidad objetiva, en la que cabe la excepción de los casos de fuerza mayor o de daños que el particular tenga el deber jurídico de soportar (art. 32, apartado 1 LRJSP)[50]. Esta denominada «responsabilidad patrimonial de las Administraciones Públicas» en nuestro Derecho, introduce *ex novo,* entre otros supuestos, la responsabilidad del Estado si el perjuicio deriva de la aplicación de una norma contraria al Derecho de la Unión Europea, conforme al apartado 5 del artículo 32 LRJSP. Pues bien, el Reglamento europeo de IA forma ya parte de este marco normativo de la Unión, al igual que lo es su enfoque holístico, lo que obliga a considerar su aplicación en un marco integral en cohesión con otras normas y principios vigentes. Con esto, ¿qué requisitos adicionales exige el apartado 5 del citado artículo 32 para aplicar la responsabilidad del apartado 3? Se trata de los siguientes: a) que la norma jurídica europea incumplida confiera derechos a los particulares; b) que el incum-

(48) Sobre la materia: Carmen Muñoz García «Unificación del principio de responsabilidad extracontractual del Estado por daños a los particulares. Delimitando el requisito de violación suficientemente caracterizada», en Aguilera Morales. M. (dir.), *Tribunal de justicia de la Unión Europea, Justicia civil y Derechos fundamentales*, en Aranzadi Thomson Reuters, 2020. También ver, entre otros: Edorta Cobreros Mendazona «La pertenencia a la Unión Europea y su repercusión en la responsabilidad patrimonial», *Revista de Administración Pública*, 2016, Nº 200, pp. 315-339; Gabriel Domenech Pascual «Repensar la responsabilidad patrimonial del Estado por normas contrarias a Derecho», *Indret,* núm. 4, 2022, pp. 168-228.

(49) El citado artículo 4, apartado 2, segundo párrafo, se expresa en los siguientes términos: «Los Estados miembros adoptarán todas las medidas generales o particulares apropiadas para asegurar el cumplimiento de las obligaciones derivadas de los Tratados o resultantes de los actos de las instituciones de la Unión».

(50) En el artículo 32, apartado 1 de la LRJSP se proclama entre los principios de responsabilidad la regla general siguiente: «Los particulares tendrán derecho a ser indemnizados por las Administraciones Públicas correspondientes, de toda lesión que sufran en cualquiera de sus bienes y derechos, siempre que la lesión sea consecuencia del funcionamiento normal o anormal de los servicios públicos salvo en los casos de fuerza mayor o de daños que el particular tenga el deber jurídico de soportar de acuerdo con la Ley», si bien, dicha responsabilidad objetiva no lo será tanto, entre otras razones porque los tribunales exigen un mínimo de culpa que diluye tal objetivación.

plimiento esté suficientemente caracterizado; y c) que exista nexo causal entre el incumplimiento de la obligación impuesta a la Administración responsable por el Derecho de la Unión Europea y el daño sufrido por los particulares[51].

Todo lo anterior no excluye la aplicación del derecho interno de cada Estado miembro, como sería, por ejemplo, la aplicación del régimen común en materia de responsabilidad civil por hecho propio o ajeno o por los bienes que nos pertenecen, según el caso.

2. LA TRANSPARENCIA: ASPECTO CLAVE PARA CONFORMAR LA AUTONOMÍA DE LA VOLUNTAD

A. Transparencia para evitar los fallos y la desinformación creada por IA

Que los principales *chatbots* conversacionales (ChatGPT, Copilot, Gemini) muestran importantes problemas de fiabilidad no es algo nuevo. Generan respuestas y crean contenidos, pero incluyen errores fácticos, jurídicos o científicos, y, además, citan fuentes imprecisas, alteradas, e incluso, inventadas.

La IA ética y responsable es clave para alcanzar su máximo desarrollo, y para ello, desde todos los poderes y sectores habrá de abogarse por una mayor transparencia por parte de productores y desarrolladores de la IA, reguladores, e incluso, medios de difusión. También estos últimos proporcionan, a través de la red, informaciones que sirven de retroalimentación a la IA, por lo que la veracidad o no de estas será clave al tiempo de que cualquier *chatbot* ofrezca información lo más veraz posible, que adopten medidas de mejora y que se evite la información errónea.

En este contexto, y a sabiendas de que esta tecnología puede llegar a operar con elevados niveles de autoaprendizaje y de autonomía, y que puede inferir de la información de entrada, por sí mismo, contenidos, recomendaciones o decisiones, también predicciones[52] –sin que sea preceptivo conocer cómo ha llegado a esos resultados de salida–, será esencial que los usuarios, cuando interactúan con la IA, estén informados. Se trata de generar confianza, y para ello son buenas, entre otras medidas, las siguientes:

1. Conocer si los desarrolladores adoptan estándares que propician buenas prácticas acorde con la normativa europea, incluidas pruebas previas a la implementación;

(51) Para mayor detalle ver Carmen Muñoz García «Unificación del principio de responsabilidad…», op. cit. p. 305 y ss.

(52) Si en el artículo 3.1 del RIA, el primero de los conceptos que proporciona la norma es sobre qué debe entenderse por «sistema de IA» a los efectos del Reglamento: «un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales». Pues bien, aunque de la definición se extraen las principales características, será la capacidad de inferencia, referida al proceso de obtención de resultados de salida como predicciones, contenidos o decisiones, que influirán en entornos físicos o virtuales, la que marque la gran diferencia.

2. Cómo se procesan los datos y que la información no siempre es veraz;
3. Cómo opera la herramienta para garantizar una IA segura y accesible;
4. También, cómo desarrollar y mejorar destrezas para obtener mejores resultados,
5. Y siempre, estar en condiciones de rendir cuentas. De ahí que las medidas de transparencia siempre se puedan probar.

Ahora bien, con todo, no será suficiente con entender cuanto antecede, sino que será necesario que los usuarios comprendan de manera pormenorizada los fallos de la herramienta para operar con criterio y adoptar decisiones verdaderamente queridas por el sujeto que interactúa con esta tecnología.

B. Obligaciones de transparencia en el Reglamento de IA

Con la finalidad de alcanzar la máxima eficiencia del RIA, además de establecerse obligaciones para los operadores y requisitos para los distintos sistemas de IA según el enfoque normativo basado en el riesgo, se imponen obligaciones de transparencia con distinta intensidad.

Esta exigencia contemplada en la norma no es nueva en este marco regulatorio, que desde el inicio, en la propuesta de la regulación de la IA por la Comisión Europea (21 abril de 2021), contiene obligaciones mínimas de transparencia para determinados sistemas –que no sean prohibidos ni de alto riesgo–, y en concreto decía: «cuando se utilizan robots conversacionales o ultra falsificaciones»[(53)] por el peligro que supone esta herramienta para los derechos fundamentales y la seguridad que el legislador europeo busca proteger.

La transparencia aparece como requisito y obligación vinculada a la trazabilidad, si bien, una y otra van referidas a aspectos diferentes. Mientras que la transparencia, término más genérico, implica el deber de información que debe proporcionar el operador entendido en toda su extensión[(54)], la trazabilidad, se integra en esta, como posibilidad de identificación del histórico del producto, como así veremos en las siguientes líneas.

Con carácter dispositivo, el considerando 27 del Reglamento pone en valor las Directrices éticas para una IA fiable de 2019, elaboradas por el Grupo independiente de expertos de alto nivel sobre IA creado por la Comisión. Estas directrices no son vinculantes, aunque sí se recomiendan, y hacen hincapié, entre otros principios, en la transparencia. Y esta, aunque no es vinculante, se define en el citado considerando en los siguientes términos:

---

(53) Así se expresa en la Exposición de motivos, apartado 2, de la propuesta de Reglamento *por el que se establecen normas armonizadas en materia de inteligencia artificial* (Ley de Inteligencia Artificial), de 21 de abril de 2021. Sin duda, serán las ultra falsificaciones o «deepfakes», una de las actividades más frecuentes y reprobables en cuanto se produce una manipulación que altera la realidad con un fin, en su mayoría ilícito. Si a ello le sumamos que mediante estas herramientas se crean contenidos creíbles, de fácil difusión y con coste cero, el daño está claro. Basta como ejemplo lo ocurrido en Almendralejo (Badajoz), en 2023, cuando quince menores, superpusieron con IA las caras de compañeras sobre falsos desnudos.

(54) El término «operador» proporcionado por el artículo 3, apartado 8 del RIA, comprende a todos los sujetos que enumera en apartado precedentes el mencionado artículo 3: *proveedor, fabricante del producto, responsable del despliegue, representante autorizado, importador o distribuidor*. Y que se definen en los apartados 3 a 7 de dicho artículo.

> «Por "transparencia" se entiende que los sistemas de IA se desarrollan y utilizan de un modo que permita una trazabilidad y explicabilidad adecuadas, y que, al mismo tiempo, haga que las personas sean conscientes de que se comunican o interactúan con un sistema de IA e informe debidamente a los responsables del despliegue acerca de las capacidades y limitaciones de dicho sistema de IA y a las personas afectadas acerca de sus derechos»

Y lo hace en el sentido de que la herramienta se desarrolle y utilice de un modo que permita una trazabilidad y explicabilidad óptimas, y que, al mismo tiempo, permita que las personas sepan que interactúan con un sistema de IA, con todos los riesgos que ello conlleva. Permítanme el atrevimiento, pero esta música me suena a la misma letra que ya conocemos en relación con el alcance del «control de transparencia» respecto a elementos esenciales de un contrato, previsto como excepcional en la Directiva 93/13 *sobre las cláusulas abusivas en los contratos celebrados con consumidores*(55). Ese ir más allá de la transparencia formal para lograr la transparencia material (extensión, alcance y completo conocimiento del impacto de la relación en la que se participa)(56) lleva a un conocimiento de lo que libremente se decide como expresión de la autonomía de la voluntad, y a una defensa a ultranza de quienes se adhieren a una relación jurídica.

Por otro lado, y al margen del carácter dispositivo de la transparencia y de los otros principios incluidos en las Directrices éticas de 2019(57), es necesario destacar que la obligación de transparencia también está recogida en el marco normativo del Reglamento. Este deber, se alinea con la pretensión inicial de la Comisión Europea al tiempo de publicar la propuesta de Reglamento en abril de 2021. El objetivo es imponer la obligación de transparencia teniendo en cuenta el potencial riesgo que conllevan determinados sistemas y así garantizar la implementación de una IA confiable y segura. Y así ocurre tanto para los sistemas de alto riesgo (artículo 13 del RIA) como para los sistemas de IA destinados a interactuar directamente con personas físicas (artículo 50 del RIA). La transparencia, sin ambages y con carácter vinculante, se materializa en los siguientes supuestos:

---

(55) Es esencial, en la protección de los consumidores, la tutela frente a cláusulas abusivas en base a la falta de negociación individualizada que se alinea con las cláusulas predispuestas que dinamizan el un mercado cada vez más globalizado. Pues bien, mientras en la mayoría de las ocasiones la negociación se centra en los elementos esenciales del contrato, y se desdibuja respecto a los aspectos no esenciales, para estos últimos, la Directiva 93/13, *sobre las cláusulas abusivas en los contratos celebrados con consumidores* prevé, de manera expresa, la declaración de abusividad cuando la cláusula ha sido predispuesta y contrariamente a la buena fe causa desequilibrio importante (entre los derechos y obligaciones) al consumidor, según el artículo 3, apartado 1 de dicha Directiva. Al margen de esta regla general, el artículo 4, apartado 2 de la Directiva 93/13, más allá de la declaración de abusividad respecto a las cláusulas accesorias del contrato, permite declarar abusivas cláusulas que versan sobre el objeto principal del contrato, «siempre que dichas cláusulas se redacten de manera clara y comprensible».

(56) De este tema, profusa doctrina y jurisprudencia ampliamente desarrollada y consolidada por el Tribunal de Justicia de la Unión Europea con ocasión, principalmente, de la reiterada interpretación y aplicación de la Directiva 93/13 sobre cláusulas abusivas. Sin duda, como ya he dicho en otras ocasiones, el tema de la abusividad en la relación contractual con consumidores se ha convertido en uno de los de mayor proyección jurídica en los tribunales nacionales y en el de Luxemburgo

(57) Del considerando 27 del RIA.

— En el cap. III, dedicado a los «Sistemas de IA de alto riesgo»[58]. Por un lado, la sección 2, en "*Requisitos de los sistemas de IA de alto riesgo*" (arts. 8 a 15), incluye algunos de los principios aludidos en el considerando 27 del RIA. En concreto, el artículo 13, apartado 1 del RIA prevé y exige a «proveedores o proveedores potenciales de sistemas de IA de alto riesgo»: transparencia suficiente para que los responsables del despliegue interpreten y usen correctamente los resultados de salida (*outputs*). Además, en la sección 3, en el artículo 26, entre las obligaciones de los responsables del despliegue de los sistemas de IA de alto riesgo a que se refiere el anexo III[59] que tomen decisiones o ayuden a tomar decisiones relacionadas con personas físicas, figura la siguiente: «informarán a las personas físicas de que están expuestas a la utilización de los sistemas de IA de alto riesgo».
— Tras los primeros capítulos del Reglamento (el I sobre disposiciones generales, el II sobre prácticas de IA prohibidas y el capítulo III acerca de los sistemas de IA de alto riesgo), en el capítulo IV, en un único artículo, el 50 del RIA, se ocupa de las «Obligaciones de transparencia de los proveedores y responsables del despliegue de determinados sistemas de IA». En él se impone –con carácter vinculante–, el deber de transparencia para determinados sistemas de IA[60]. Lo que implica que los proveedores y desarrolladores del despliegue de determinados sistemas deben garantizar la transparencia entendida esta en el sentido que nos proporciona el considerando 27 del RIA: «los sistemas de IA se desarrollan y utilizan de un modo que permita una trazabilidad y explicabilidad adecuadas».

---

(58) Dentro de este capítulo, la sección 1ª fija cuándo un sistema de IA lo es de alto riesgo. Así, el artículo 6 del RIA, primero de los artículos que contribuye a aclarar qué sistemas se enmarcan en esta calificación porque suponen un riesgo considerable para los derechos fundamentales, la salud, la seguridad, el Estado de Derecho, la democracia y el medioambiente, establece dos supuestos diferentes para catalogarlo como de alto riesgo, el primero vinculado al anexo I, y el segundo vinculado al anexo III. Por lo que se refiere al primero de los supuestos, el sistema habrá de reunir dos requisitos: 1. Que «el sistema de IA esté destinado a ser utilizado como componente de seguridad de un producto que entre en el ámbito de aplicación de los actos legislativos de armonización de la Unión enumerados en el anexo I, o que el propio sistema de IA sea uno de dichos productos, 2. Y que «el producto del que el sistema de IA sea componente de seguridad con arreglo a la letra a), o el propio sistema de IA como producto, deba someterse a una evaluación de la conformidad de terceros para su introducción en el mercado o puesta en servicio con arreglo a los actos legislativos de armonización de la Unión enumerados en el anexo I» (máquinas, juguetes, ascensores, productos sanitarios, vehículos, aeronaves,…). Por lo que se refiere al segundo de los supuestos contemplados, también se considerarán sistemas de alto riesgo aquellos expresamente contemplados en el anexo III del Reglamento (biometría, infraestructuras críticas, servicios privados o públicos esenciales, entre otros).

(59) De los trece anexos que se integran en el Reglamento de IA, es el anexo III el que fija cuáles son los «sistemas de IA de alto riesgo a que se refiere el artículo 6, apartado 2»: 1. Biometría; 2. Infraestructuras críticas; 3. Educación y formación profesional; 3. Empleo; 4.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. Sistemas de IA que garanticen el cumplimiento del Derecho; 7. Migración, asilo y gestión del control fronterizo; 8. Administración de justicia y procesos democráticos. Y todo ello en la medida en que su uso está permitido por el DUE o el nacional aplicable.

(60) Como ya hemos indicado, igual deber de información y transparencia se ha impuesto previamente para los «sistemas de alto riesgo» (artículo 13 del Reglamento).

C. **En concreto, los sistemas destinados interactuar con personas físicas**

El esfuerzo del legislador europeo por implementar la transparencia en IA en toda su extensión, no se agota en las generalidades de las directrices (considerando 27), ni a las especificidades de las normas expuestas en el apartado anterior (artículos 13 y 25 respecto a sistemas de alto riesgo, y artículo 50 para determinados sistemas de IA –ni son los prohibidos, ni son los de alto riesgo–). Se impone la necesidad de implementar la transparencia desde el diseño del modelo y del sistema, y durante la vida de la herramienta. Es más, aunque el Reglamento aboga a modo de *desiderátum* por incorporar la transparencia *como principio* en estas herramientas, parece que también *incorpora la transparencia en sentido formal, y, además, material*. Y es así en cuanto exige:

- que la documentación técnica que se elabore de los modelos y sistemas sea suficientemente clara y precisa y actualizada[(61)]: se trata de que la información sea comprensible y siga actualizándose durante toda la vida útil del sistema.
- que sea posible «la trazabilidad y explicabilidad»,
- y, además, «que las personas sean conscientes de que se comunican o interactúan con un sistema de IA e informe debidamente a los responsables del despliegue acerca de las capacidades y limitaciones de dicho sistema de IA y a las personas afectadas acerca de sus derechos» (tal y como se especifica en el considerando 27).

Si este único artículo del capítulo IV, precedente de la regulación específica de los modelos de IA de uso general (capítulo V), pretende abordar las obligaciones de transparencia de los proveedores y responsables del despliegue de determinados sistemas de IA para que no se diluya la protección de las personas que se pretende en el Reglamento en su conjunto, cabe plantearse dos cuestiones: (i) qué supuestos se contemplan –con sus excepciones correspondientes–, y (ii) cuál es la medida de transparencia exigida a los proveedores de estos otros sistemas de IA:

---

(61) Así se expresa en el considerando 71 para los sistemas de IA de alto riesgo: *resulta esencial disponer de información comprensible sobre el modo en que se han desarrollado y sobre su funcionamiento durante toda su vida útil. A tal fin, es preciso llevar registros y disponer de documentación técnica que contenga la información necesaria para evaluar si el sistema de IA de que se trate cumple los requisitos pertinentes y facilitar la vigilancia poscomercialización. Dicha información debe incluir las características generales, las capacidades y las limitaciones del sistema y los algoritmos, datos y procesos de entrenamiento, prueba y validación empleados, así como documentación sobre el sistema de gestión de riesgos pertinente, elaborada de manera clara y completa*. La documentación técnica debe mantenerse adecuadamente actualizada durante toda la vida útil del sistema de IA, También para los modelos de IA de uso general, según el considerando 101: *Los proveedores de modelos de IA de uso general tienen una función y una responsabilidad particulares a lo largo de la cadena de valor de la IA, ya que los modelos que suministran pueden constituir la base de diversos sistemas de etapas posteriores, que a menudo son suministrados por proveedores posteriores que necesitan entender bien los modelos y sus capacidades, tanto para permitir la integración de dichos modelos en sus productos como para cumplir sus obligaciones en virtud del presente Reglamento o de otros reglamentos. Por consiguiente, deben establecerse medidas de transparencia proporcionadas, lo que incluye elaborar documentación y mantenerla actualizada y facilitar información sobre el modelo de IA de uso general para su uso por parte de los proveedores posteriores*. En ambos casos, en línea con los artículos 11 y 53 del RIA, respectivamente.

1) principalmente cuando se trata de sistemas de IA o modelos destinados *a interactuar con personas físicas*, y que por su propia naturaleza *pueden conllevar riesgos de suplantación o engaño, falsedad o manipulación, entre otros*. En este caso, deben diseñarse y desarrollarse de forma que las personas físicas sepan que interactúan con IA (art. 50, apartado 1), con dos excepciones[62]. En esta línea, el considerando 101 exige medidas de transparencia proporcionadas, lo que incluye elaborar documentación y mantenerla actualizada;

2) cuando se trate de sistemas de IA, «entre los que se incluyen los sistemas de IA de uso general[63] que *generen contenido sintético* de audio, imagen, video o texto». Este otro supuesto merece especial atención en cuanto que la herramienta tiene *capacidad de crear contenido artificial que puede inducir a error* a cualquier usuario (puede emular la voz de un familiar o replicar la imagen de un rival electoral, generar noticias falsas o crear el video de un producto engañoso). En este caso, el proveedor está obligado a velar porque los resultados de salida del sistema de IA estén marcados en un formato legible por máquina y que sea posible detectar que han sido generados o manipulados de manera artificial (art. 50, apartado 2)[64];

3) cuando se trate de «un sistema de *reconocimiento de emociones o de un sistema de categorización biométrica*», en este caso, los responsables del despliegue deberán informar del funcionamiento de la herramienta a las personas físicas afectadas, y, además, tratarán sus datos personales conforme a la normativa específica de protección de datos[**65**], según los casos (art. 50, apartado 3)[66];

4) no es casualidad que, a mayor abundamiento, y parece que significativamente vinculado al sistema contemplado en el apartado 2 del artículo 50, se especifique qué, cuando estamos ante «un sistema de IA que genere o manipule imágenes o contenido de *audio o video que constituyan una suplantación*», los responsables del despliegue, deberán hacer público que estos contenidos o imáge-

---

(62) Se exceptúa dos supuestos: *excepto cuando resulte evidente desde el punto de vista de una persona física razonablemente informada, atenta y perspicaz, teniendo en cuenta las circunstancias y el contexto de utilización. Esta obligación no se aplicará a los sistemas de IA autorizados por ley para detectar, prevenir, investigar o enjuiciar delitos, con sujeción a las garantías adecuadas para los derechos y libertades de terceros, salvo que estos sistemas estén a disposición del público para denunciar un delito penal*.

(63) Estos otros «sistemas de IA de uso general» según el artículo 3, apartado 66, se definen como: «Un sistema de IA basado en un modelo de IA de uso general y que puede servir para diversos fines, tanto para su uso directo como para su integración en otros sistemas de IA».

(64) También se exceptúa de esta obligación de transparencia los supuestos en los que: *los sistemas de IA desempeñen una función de apoyo a la edición estándar o no alteren sustancialmente los datos de entrada facilitados por el responsable del despliegue o su semántica, o cuando estén autorizados por ley para detectar, prevenir, investigar o enjuiciar delitos*.

(65) En concreto hace una remisión expresa a los siguientes textos: Reglamento 2016/679, *relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos*, el Reglamento 2018/1725, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones, órganos y organismos de la Unión, y a la libre circulación de esos datos, y conforme a la Directiva 2016/680, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por parte de las autoridades competentes para fines de prevención, investigación, detección o enjuiciamiento de infracciones penales o de ejecución de sanciones penales, y a la libre circulación de dichos datos.

(66) Se exceptúan: los sistemas de IA utilizados para la categorización biométrica y el reconocimiento de emociones que hayan sido autorizados por ley para detectar, prevenir e investigar delitos, con sujeción a las garantías adecuadas para los derechos y libertades de terceros y de conformidad con el Derecho de la Unión.

nes han sido generados o manipulados de manera artificial (art. 50, apartado 4). También para este artificio que constituye una suplantación por quien no es su titular ni está autorizado hay una excepción, y aparte, una matización[(67)]. Respecto a esta última, hay que precisar que el contenido artificial es una crítica o sátira, y para esta, la transparencia exigida lo es con menor alcance en cuanto que bastará con «hacer pública la existencia de dicho contenido generado o manipulado artificialmente» (art. 50, apartado 4). A mi entender, esta exigencia guarda relación con la transparencia formal exigida en cualquier relación contractual cuando las dos partes en la misma están en una clara relación de superioridad y de inferioridad.

En este mismo apartado se contempla el supuesto no aludido en líneas precedentes, y es que el sistema de IA «genere o manipule texto que se publique con el fin de informar al público sobre asuntos de interés público», para este concreto supuesto el deber de transparencia obliga a los responsables a divulgar que el contenido se ha generado o manipulado artificialmente.

No obstante lo anterior, el artículo 50, en el párrafo 5 del RIA, establece para todos los supuestos contemplados en los apartados 1 a 4, el deber de proporcionar a todas las personas físicas afectadas, «de manera clara y distinguible», la información antedicha que «se ajustará a los requisitos de accesibilidad aplicables».

Con todo, y a tenor del párrafo 6 de dicho artículo, lo contenido en los apartados 1 a 4 no minorará las obligaciones y requisitos técnicos exigidos en el capítulo III (sistemas de alto riesgo), ni disminuirá los parámetros que sobre la obligación de transparencia existen tanto en el Derecho de la Unión como en el Derecho interno.

No falta el deber de la Oficina de IA, dependiente de la Comisión Europea, de fomentar y facilitar la implementación de códigos de buenas prácticas[(68)] que propicien la aplicación efectiva de las obligaciones que ayuden a la detección y el etiquetado de contenidos generados o manipulados de manera artificial.

3. AGENTES QUE SIMULAN COMPORTAMIENTOS HUMANOS Y COMPROMETEN LA AUTONOMÍA DE LA VOLUNTAD ¿COMPLETAMENTE AUTÓNOMOS?

A. ¿Qué es un agente de IA? ¿Qué nivel de autonomía puede llegar a alcanzar?

Tras la irrupción de ChatGPT en noviembre de 2022 estamos asistiendo a la innumerable proliferación de los grandes modelos de lenguaje (LLM), principalmente durante el año 2024. Ahora bien, será en 2025 cuando nos planteemos qué

(67) El deber de transparencia no se aplicará cuando *la ley autorice su uso para detectar, prevenir, investigar o enjuiciar delitos*. A continuación se admite un supuesto para el que la transparencia no se exige con igual intensidad que la prescrita en todos los supuestos del artículo 50, y es: *Cuando el contenido forme parte de una obra o programa manifiestamente creativos, satíricos, artísticos, de ficción o análogos*, en ese caso, la obligación de transparencia *se limitará a la obligación de hacer pública la existencia de dicho contenido generado o manipulado artificialmente de una manera adecuada que no dificulte la exhibición o el disfrute de la obra.*

(68) A ellos se refiere el artículo 56 del RIA, dentro del capítulo V dedicado a los modelos de IA de uso general.

niveles de autonomía pueden llegar a alcanzar estos modelos de IA creados para simular comportamientos humanos y realizar tareas específicas. De base:

— Mediante el *software* toma decisiones basadas en datos de preentrenamiento, tomados de internet, y, además, aprovechando el contexto que se le presenta en los *inputs*;
— Aunque no razonan interactúan con las personas, y también con otros agentes como si fuera un humano, en un continuo diálogo;
— Crea contenidos y es capaz de ofrecer respuestas multimodales. Ejecuta tareas y genera su propio autoaprendizaje, sin que en muchas ocasiones se puedan explicar sus decisiones;
— Con los *inputs* que les proporcionamos, identifican patrones y responde en todo caso adaptándose a lo que se le solicita, realizando tareas complejas de forma autónoma e independencia humana.

Estas herramientas de IA generativa o agentes autónomos son grandes modelos de lenguaje con capacidades en crecimiento exponencial que se desarrollan y emplean en todos los sectores. Gracias a sus amplias habilidades (como la automatización de tareas y procesos simples y complejos, la capacidad de entender y prevenir enfermedades, las plataformas educativas, asistenciales o sanitarias, o la creación de tratamientos farmacológicos personalizados o de experiencias individualizadas que proporcionan innumerables beneficios), podemos vislumbrar los múltiples riesgos e implicaciones éticas y jurídicas de la herramienta.

En línea con lo anterior, por lo pronto, no es ocioso tener en cuenta la siguiente advertencia: son muchas las tecnológicas que desarrollan agentes de IA para optimizar tareas y alcanzar objetivos, pero también, que crean, reproducen y emulan la inteligencia humana con gran precisión. Sus capacidades para emular son tal, que surgen importantes dificultades, entre otras: distinguir inteligencia humana de IA, los riesgos derivados de las alucinaciones de la herramienta y la manipulación para distorsionar el comportamiento de las personas con la discutible redefinición de la autonomía privada de las personas. También se acrecientan los riesgos respecto a la protección de datos y la privacidad, la identidad propia, y, además, respecto a aspectos esenciales del Derecho privado como son: el derecho de obligaciones, empezando por el consentimiento, continuando por las modalidades para dejar sin efecto un contrato, los incumplimientos o la responsabilidad civil. Estos aspectos, y otros muchos, están inmersos en una revolución fundamental que exige adaptaciones, reformulaciones, revisiones y seguimiento de todo cuanto surge en torno a estos[(69)].

No olvidemos que, cuando la IA generativa se incorpora al mercado para proporcionar bienes o servicios, la herramienta está sujeta a las mismas garantías que se exigen a cualquier otro producto o servicio. No podría ser de otra manera si tenemos en cuenta que los ámbitos en los que se desarrolla representan un alto riesgo para las personas. Así ocurre con los sistemas de alto riesgo presentes en el

(69) Sirva de ejemplo los avances en réplicas artificiales de personas humanas Al respecto, la síntesis realizada por Felipe Espinosa Wang, «La IA puede replicar tu personalidad con 85% de precisión» *DW,* 6 de enero de 2025, en relación con el estudio realizado por investigadores de Sanford y Google DeppMind, Visto en https://www.dw.com/es/tu-otro-yo-digital-la-ia-alcanza-precisi%C3%B3n-del-85-en-replicar-personalidades/a-71230370

sector público (salud, administración de justicia, sector educativo e investigación científica o prestaciones sociales[70]), así como en la industria y en sectores ampliamente regulados: la seguridad y el control de fronteras; el empleo, la seguridad vial, aérea y ferroviaria; las infraestructuras críticas; los seguros; o el sector financiero y bancario. Si bien es verdad que también puede darse en el sector privado (sanidad privada).

Para propiciar la óptima prestación de bienes o servicios las empresas, administraciones públicas u organizaciones públicas o privadas que incorporen en su actividad IA, deberán proveer un sistema de gestión de calidad, una evaluación de impacto y un análisis de riesgos. También sistemas de auditoría y mejora continua de algoritmos si fuese preciso, y, sobre todo, transparencia en su sentido más amplio. Además, deberán tenerse en cuenta directrices éticas o códigos de conducta (si fueran modelos de IA habrán de cumplirse las obligaciones fijadas en los arts. 51 a 55, según los casos, preferiblemente vía código de buenas prácticas, en los términos señalados en el art. 56). Con todo, la supervisión humana será un factor relevante, esencial, en aplicaciones de alto riesgo y en aquellas que deban guardar obligaciones de transparencia con independencia de la mayor o menor autonomía de la herramienta y del agente. Sólo de esta manera se salvaguardarán derechos fundamentales, la salud, la seguridad, la democracia, el Estado de Derecho y el medio ambiente, y, además, evitaremos que estos agentes perpetúen los errores que pudiese tener el sistema que no esté sujeto a una supervisión humana significativa.

Ahora bien, no será lo mismo un «asistente de IA» que está diseñado principalmente para ayudar a los usuarios con tareas específicas y cotidianas (*Siri, Alexa*, robot doméstico) que un «agente de IA» con propósitos más amplios y con capacidades para realizar una variedad de tareas complejas sin intervención humana (drones o coches autónomos, robots industriales o sistemas de diagnóstico de enfermedades o de catástrofes naturales). Estos pueden incluir el análisis de datos, identificar tendencias, aprender patrones, e incluso, tomar decisiones empresariales en tiempo real. La automatización de procesos

Sin duda, el desarrollo e implementación de agentes de IA en empresas, sector público y privado, deberán llevarse a cabo con todas las garantías, se trata de que estos agentes que emulan la inteligencia humana alcancen la máxima eficiencia allí donde se implementen y, además, que los usuarios comprenden las capacidades de la herramienta y los riesgos asociados a ella. Con todo, la empresa u organización incorpora al agente a su negocio y asume la herramienta como propia, y lo hace con una doble finalidad: ahorrar costes y dar un mejor servicio a sus clientes, lo que derivará en supuestos de responsabilidad complejos y en muchos casos de respuesta solidaria con el derecho del que respondió a exigir la repetición[71].

---

(70) En febrero de 2025 Google ha presentado un agente de IA para potenciar la investigación científica «AI co-scientist»: un multi agente de IA basado en su herramienta «Gemini» que, aunque es de propósito general, está pre entrencado para el desarrollo y optimización de investigaciones biomédicas. Sobre esta, Gottweis, Wei-Hung WENG, Alexander Daryin, et. al. *Towards an AI co-scientist*, Google, 18 de febrero de 2025. Visto en https://storage.googleapis.com/coscientist_paper/ai_coscientist.pdf

(71) Para estos supuestos nos remitimos a cuanto se ha dicho sobre responsabilidad civil en apartados precedentes.

B. **Los agentes de IA en la contratación con consumidores**

En este ámbito concreto partimos de dos premisas: una primera, que el Reglamento de IA es una norma dirigida a fijar los marcos regulatorios de esta herramienta que tiene como objetivo que, en el mercado interior de la UE, el desarrollo e implementación de la misma se haga en beneficio de los seres humanos. Y por otro, que siendo incuestionables los beneficios de la IA, los riesgos repercuten, de manera directa o indirecta, en las personas: la manipulación, el impacto en el empleo, los sesgos o la discriminación son solo un ejemplo de los peligros que pueden darse para los individuos y colectivos, y en todo tipo de sectores.

Ahora bien, lo anterior no excluye que nos centremos en la repercusión que estos agentes tienen en el ámbito del Derecho de consumo por ser los consumidores el mayor colectivo de sujetos afectados por esta tecnología. De base, es esencial determinar ante qué tipo de relación estamos cuando hay un consumidor, por regla general, persona física[72], que actúa con un propósito ajeno a una actividad empresarial o profesional y que opera en un contexto contractual, en una situación de debilidad. Por lo que el alcance de esta normativa lo es en el contexto de relaciones jurídicas contractuales, y sólo en estas. Esto no es baladí, se contrata a través de un agente de IA que no es un tercero, sino que está bajo el control de quien se sirve de él.

Empresas y profesionales se sirven de herramientas LLM (agentes de IA) para que interactúen con las personas y puedan realizar tareas cada vez más complejas. La variedad de modelos (ChatGPT, Copilot, Gemini 2.0, Flash 2.0…) dan muestra de las capacidades y múltiples usos que pueden llegar a desarrollar. Su eficiencia es alta pero no suficientemente certera como lo muestran las no pocas «alucinaciones» en las que incurren. Los indiciadores y análisis son constantes y muestran las

(72) La Directiva 2011/83 admite extender el concepto a de consumidor: «a las personas jurídicas o físicas que no sean «consumidores» en el sentido de la presente Directiva, como organizaciones no gubernamentales, empresas de reciente creación o pequeñas y medianas empresas», lo que se trasladó a los Derechos internos. En España se llevó a efecto mediante el artículo único de la Ley 3/2014, de 27 de marzo que modificaba el Real Decreto Legislativo 1/2007, de 16 de noviembre (*texto refundido de la Ley General para la Defensa de los Consumidores y Usuarios y otras leyes complementarias*, TRLGCU 1/2007) que, entre otras modificaciones relevantes, realizó una ampliación del concepto del consumidor. Así, tras dicha modificación, en el artículo 3 del TR, se proporciona una denominación más amplia frente al proporcionado por las normas de consumo de la Unión. Se define al consumidor en sentido negativo y como contraposición al concepto de empresario, incluyendo a: «las personas físicas que actúen con un propósito ajeno a su actividad comercial, empresarial, oficio o profesión. Son también consumidores a efectos de esta norma las personas jurídicas y las entidades sin personalidad jurídica que actúen sin ánimo de lucro en un ámbito ajeno a una actividad comercial o empresarial. 2. Asimismo, a los efectos de esta ley y sin perjuicio de la normativa sectorial que en cada caso resulte de aplicación, tienen la consideración de personas consumidoras vulnerables respecto de relaciones concretas de consumo, aquellas personas físicas que, de forma individual o colectiva, por sus características, necesidades o circunstancias personales, económicas, educativas o sociales, se encuentran, aunque sea territorial, sectorial o temporalmente, en una especial situación de subordinación, indefensión o desprotección que les impide el ejercicio de sus derechos como personas consumidoras en condiciones de igualdad». La consideración como consumidor de *las personas jurídicas y las entidades sin personalidad jurídica que actúen sin ánimo de lucro en un ámbito ajeno a una actividad comercial o empresarial*, tras Ley 3/2014, de 27 de marzo, por la que se modifica el TRLGDCU 1/2007, es sin duda, un claro ejemplo de cómo el concepto de consumidor que generalizan las normas europeas puede expandirse por los Estados miembros a otros sujetos, y cómo se expande el régimen tuitivo a ficciones admitidas en Derecho.

amplias capacidades, pero también los errores que tratan de ser minorados, de ahí que se generen versiones mejoradas.

A los efectos de una aproximación a los múltiples riesgos que propicia esta tecnología, profundamente disruptiva, propongo una clasificación de los riesgos específicos a los que se enfrenta el Derecho de consumo una vez que la inteligencia artificial puede afectar a la autonomía de la voluntad y aumentar el riesgo de vulnerabilidad de una persona, especialmente cuando opera como consumidor:

Aspecto	Impacto de la IA
Toma de decisiones	La IA puede influir en las decisiones al proporcionar recomendaciones basadas en datos y algoritmos
Privacidad y datos personales	Recopilación y análisis de datos por parte de la IA que puede comprometer la privacidad y autonomía
Libertad de elección	La IA puede limitar las opciones disponibles al favorecer información y sugerir opciones específicas
Responsabilidad y ética	Cuanto mayor sea la delegación de decisiones a la IA, más cuestiones se plantean sobre la responsabilidad y la ética
Manipulación y persuasión	La IA puede ser utilizada para influir en comportamientos y decisiones a través de técnicas persuasivas

* Carmen Muñoz. Imagen diseñada con ayuda de Copilot[(73)]

En este punto permítanme advertir que esto es sólo una pequeña aproximación de los muchos retos que en esta materia se nos avecinan. De ellos, habrá que ir dando cumplida muestra según avance la IA y conozcamos incidentes en los que se ponen en riesgo derechos de los consumidores y se compromete la autonomía y la libertad contractual.

III. **CONSIDERACIONES FINALES**

Desde hace años, el objetivo principal de la Unión Europea viene siendo desarrollar un ecosistema de IA de excelencia en innovación y en protección de valores y derechos fundamentales para generar confianza en los mercados. Si en 2018 la Comisión Europea adoptó un plan coordinado con todos los EE.MM. para aunar inversión y concienciación en IA, en febrero de 2020 presentó, junto al *Libro Blanco de la IA*, la *Estrategia Europea de Datos* y el *Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet*

(73) Tras advertir a la herramienta de los múltiples riesgos que propician los *LLMs*, le he pedido que lleve a efecto una tabla que exprese todo lo que le he sistematizado. Es cierto que ha reproducido casi todas mis aportaciones, pero también que ha sido repetitivo en aspectos que no constituyen Derecho de consumo (derechos de autor…). No cabe duda de los *prompts* óptimos introducidos por la autora. En todo momento, eran concisos, claros y llevaban por núcleo central los riesgos para los consumidores.

de las cosas y la robótica[74], con el objetivo de configurar un ecosistema integral óptimo. Se trata de aprovechar e impulsar las oportunidades y beneficios de la tecnología en un mercado único de la Unión que se pretende seguro, ético y confiable.

Después llegaría la propuesta de Reglamento de IA (abril de 2021), finalmente aprobado como Reglamento 2024/1689, de 13 de junio, *de Inteligencia Artificial* (RIA), y otros desarrollos normativos que habían quedado fijados al tiempo de la Estrategia Europea de Datos. Pilares de esta última lo son, tanto el Reglamento 2022/868, de 30 de mayo, *de Gobernanza de Datos* (RGD, o DGA, por sus siglas en inglés), como el Reglamento 2023/2854, *sobre normas armonizadas para un acceso justo a los datos y su utilización* (Reglamento de Datos, o DA). El primero proporciona un marco para aumentar la confianza en el intercambio voluntario de datos en beneficio de las empresas y ciudadanos, mientras que el segundo pretende garantizar la equidad en el acceso y uso de los datos generados en la Unión en todos los sectores económicos, incluido el industrial, y facilitar el intercambio y la reutilización de datos. Además, los usuarios de dispositivos interconectados también participarán de los datos generados por estos. Como parte de esta estrategia, se sucederán otros hitos como son los sucesivos Espacios europeos de datos en sectores críticos, entre otros, el de la salud[75], las infraestructuras críticas o el sector financiero.

La relevancia del Reglamento de IA no solo deriva de ser la primera norma mundial que regula esta materia, sino que constituye todo un hito jurídico integral sobre esta tecnología, ya que además de promover la innovación segura, aborda los problemas derivados de su desarrollo, implementación y uso, y su impacto para derechos fundamentales, la salud o la seguridad de las personas. También por los desafíos que supone para los valores y principios éticos. Esto obliga, dado su impacto global, a abordar el estudio de la materia en conexión con todo el acervo de la Unión y con las normas de Derecho interno de cada Estado miembro. El progreso tecnológico es ilimitado, pero los riesgos deben minimizarse tanto como sea posible, de ahí la necesidad de imponer límites éticos y jurídicos.

En este ecosistema que proporciona innumerables beneficios para las personas, salvaguardar la autonomía privada y la libertad contractual, así como la privacidad y otros derechos fundamentales constituye uno de los mayores retos a los que nos enfrentamos. De ahí que, para conformar nuestra voluntad cuando interactuamos con IA, es imprescindible hacerlo siendo conscientes de que lo hacemos con esta herramienta. Por ello, la información y la transparencia se constituyen en elementos esenciales para su uso. Por otro lado, también es necesaria la protección cuando los organismos públicos o privados toman decisiones basadas exclusivamente en IA y el sistema nos priva de derechos, aunque no seamos usuarios directos de esta y otras tecnologías.

---

(74) Con este Informe de la Comisión al PE, al Consejo y al Comité Económico y Social Europeo, se pretendía crear un marco jurídico claro y garantista que proporcione protección a los *consumidores* –decía entonces–, y seguridad a las empresas. Vid. https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX:52020DC0064

(75) En esta materia se ha dado un gran paso al crear el Espacio Europeo de Datos de Salud (EEDS, EHDS por sus siglas en inglés), mediante el Reglamento (UE) 2025/327 del Parlamento Europeo y del Consejo, de 11 de febrero de 2025, relativo al Espacio Europeo de Datos de Salud, y por el que se modifican la Directiva 2011/24/UE y el Reglamento (UE) 2024/2847.

Por tanto, ante las características de esta tecnología disruptiva, es del todo necesario adaptar y reformular la mayoría de la normativa vigente en los ámbitos del Derecho privado y del Derecho público. La remisión a todas las materias en las que la IA tiene impacto requiere un marco seguro que se irá construyendo a medida que se implemente y desarrolle la tecnología. La casuística irá sucediéndose e iremos avanzando en las posibles respuestas jurídicas. En las primeras páginas anticipo que nos estamos adentrando en la regulación de la IA, pero para ayudar a una adaptación e implementación adecuadas en los términos pretendidos por el Reglamento de la Unión, la Comisión Europea (CE) seguirá publicando directrices sobre la aplicación práctica del RIA (art. 96 del RIA. También, el TJUE tendrá que abordar el fenómeno y su impacto en otras normas del DUE debidamente consolidadas.

De no producirse esta adaptación de la normativa vigente a la IA, que muestra amplias capacidades y habilidades en tiempo real, nos colocará en situaciones de vulnerabilidad y nos dejará a merced de quienes se sirven de la IA para desestabilizar vidas y valores vigentes y ampliamente consolidados en nuestro acervo comunitario. La fragilidad de la persona en este ecosistema queda de manifiesto desde que se producen réplicas artificiales de personas, lo que dificulta distinguir la inteligencia humana de la artificial. Las múltiples actuaciones de fraude, manipulación, suplantación o engaño de cualquier tipo, o las alucinaciones creadas, son solo algunos ejemplos.

Los problemas seguirán aumentando a medida que la herramienta adquiera mayores capacidades y habilidades para tomar decisiones de manera autónoma. Aparecen, y se imponen, los agentes de IA (para contratar un billete de avión, un seguro o para interpretar imágenes en radiología o proporcionar asistencia sanitaria). Es importante destacar que, además de proporcionarnos innumerables beneficios, dinamizar el mercado y mejorar nuestras vidas, podrán influir en nuestras decisiones, e incluso, suplir nuestra propia voluntad, lo que planteará cuestiones éticas y jurídicas de difícil respuesta y compleja solución. Por este motivo, es fundamental concienciar a la ciudadanía sobre estos riesgos, informarles y «alfabetizar» (cdo. 20 y art. 4 del RIA). Esto, se convierte en un deber urgente e insoslayable que, más allá de exigirse en empresas y organismos públicos y privados, debe implementarse a toda la ciudadanía, ya sea porque opere con la IA o porque sea un tercero ajeno al uso de la herramienta y que podría verse afectado por ella. Con todo, siempre nos quedará la responsabilidad civil, contractual y extracontractual.

IV. **BIBLIOGRAFÍA**

Moisés Barrio Andrés, *Manual de Derecho digital,* 2.ª edición, Editorial Tirant lo Blanch, Valencia, 2022.

Moisés Barrio Andrés, (dir.), *Legal Tech. La transformación digital de la abogacía*, 2.ª edición, La Ley, Madrid, 2023.

Enrique Benítez Palma, «Diez notas sobre la AI Action Summit de París», *LAB*, Observatorio Sector Público e Inteligencia Artificial, 9 de marzo de 2025. Visto en https://www.ospia.org/o-lab/diez-notas-sobre-la-ai-action-summit-de-paris

Anu Bradford, «Introduction: The Brussels Effect», *The Brussels Effect: How the European Union Rules the World*, Oxford Academic, 2020, ed. online, 19 diciembre 2019. Visto en https://academic.oup.com/book/36491/chapter-abstract/321181662?redirectedFrom=fulltext

Ángel Carrasco Perera, «Análisis de la nueva directiva de responsabilidad por daños causados por defectos de productos», *Web CESCO,* diciembre 2024. Visto en: https://centrodeestudiosdeconsumo.com/images/Analisis_de_la_nueva_Directiva_de_responsabilidad_por_danos.pdf

Edorta Cobreros Mendazona «La pertenencia a la Unión Europea y su repercusión en la responsabilidad patrimonial», *Revista de Administración Pública*, 2016, núm. 200, pp. 315-339.

Lorenzo Cotino Hueso «Vulnerabilidad y colectivos vulnerables en el Derecho: ¿quiénes son y cómo se definen?» en el marco del Subproyecto 2.1. «Digital Futures Initiative (Project #2)», *cAIre Research* FOR GOOGLE.org. *Revista Catalana de Dret Public* (*RCDP blog*), 18 de diciembre de 2024. Visto en https://eapc-rcdp.blog.gencat.cat/2024/12/18/vulnerabilidad-y-colectivos-vulnerables-en-el-derecho-quienes-son-y-como-se-definen-lorenzo-cotino/

Lorenzo Cotino Hueso «¿Cuándo "no es no"? Criterios para definir los sistemas de inteligencia artificial prohibidos en la Unión Europea», *Revista General de Derecho Administrativo,* 2025, (en prensa).

Lorenzo Cotino Hueso y Jorge Castellanos Claramunt J. (dir.), *Transparencia y explicabilidad de la inteligencia artificial,* Editorial Tirant lo Blanch, Valencia, 2023.

Jesús Delgado Echeverría, «I. La persona física», en Lacruz Berdejo, Sancho Rebullida, Luna Serrano, Delgado Echevarría, Rivero Hernández y Rams Albesa, *Elementos de Derecho civil. I. Parte General, vol. segundo. Personas*, ed. Dykinson, Madrid, 2000.

Joaquín Delgado Martín «Capítulo VI. Régimen sancionador», en Barrios Andrés, M, (dir.), *El Reglamento Europeo de Inteligencia Artificial*, Editorial Tirant lo Blanch, Valencia, 2024.

Virginia Dignum, «Responsible Artificial Intelligence -- from Principles to Practice» Cornell *University,* 2022. Visto en https://arxiv.org/abs/2205.10785v1

Virginia Dignum, «How Europe is shaping AI for human rights» en *AI Policy Lab*, (5.09.2024). Visto en https://aipolicylab.se/2024/09/05/how-europe-is-shaping-ai-for-human-rights/

Gabriel Domenech Pascual «Repensar la responsabilidad patrimonial del Estado por normas contrarias a Derecho», *Indret,* núm. 4, 2022, pp. 168-228. Visto en https://indret.com/wp-content/uploads/2022/10/1731.pdf

Begoña González Otero «AI for Good: La idea de la vulnerabilidad humana en tela de juicio» en el marco del Subproyecto 1.4. «AI for Good (Project #1) », *cAIre Research* Project FOR GOOGLE.org. *Blog OdiseIA*, 6 de septiembre de 2024. Visto en https://www.odiseia.org/post/ai-for-good-la-idea-de-la-vulnerabilidad-humana-en-tela-de-juicio-1

Bright Huo, Amy Boyle, Nana Marfo, *et al*., «Large Language Models for Chatbot Health Advice Studies A Systematic Review». *PubMed*, 4 febrero 2025. Visto en https://pubmed.ncbi.nlm.nih.gov/39903463/

Juraj Gottweis, Wei-Hung Weng, Alexander Daryin, et. al. *Towards an AI co-scientist*, Google, 18 de febrero de 2025. Visto en https://storage.googleapis.com/coscientist_paper/ai_coscientist.pdf

José María Martín Faba, «La inteligencia artificial en la nueva Directiva de responsabilidad por los daños causados por productos defectuosos, ¿realidad o expectativa?»,

Revista CESCO de Derecho de Consumo, núm. 53, enero-marzo 2025. Visto en https://ruidera.uclm.es/server/api/core/bitstreams/204518d1-86d2-47f7-aa6e-58aff76e7466/content

Carmen Muñoz Garcia «Unificación del principio de responsabilidad extracontractual del Estado por daños a los particulares. Delimitando el requisito de violación suficientemente caracterizada», en Aguilera Morales. M. (dir.), *Tribunal de justicia de la Unión Europea, Justicia civil y Derechos fundamentales*, en Aranzadi Thomson Reuters, 2020, pp. 305-338.

Carmen Muñoz Garcia, «Adaptar o reformular la Directiva 85/374 sobre responsabilidad por daños causados por productos defectuosos a la Inteligencia Artificial». *Diario La Ley. Ciberderecho*, 28 de febrero de 2022. Reeditado como «Adaptar… Últimas novedades y un cambio de rumbo». *Revista Crítica de Derecho Inmobiliario*, vol. 793, noviembre-diciembre 2022, pp. 2886-2908.

Carmen Muñoz Garcia, «¿ChatGPT en la universidad? ¿Complementar el aprendizaje y cambiar el modelo educativo?», *Diario La Ley*, núm. 10283, 10 de mayo de 2023.

Carmen Muñoz Garcia, *Regulación de la inteligencia artificial en Europa. Incidencia en los regímenes jurídicos de protección de datos y de responsabilidad por productos*. Editorial Tirant lo Blanch, Valencia, 2023.

Carmen Muñoz García «Modelos de IA de uso general y sistemas de IA de riesgo limitado y mínimo», en Barrios Andrés, M, (coord.), *El Reglamento Europeo de Inteligencia Artificial*, Editorial Tirant lo Blanch, Valencia, 2024, pp. 87-109.

Carmen Muñoz García «Artículo 55 del Reglamento de Inteligencia Artificial», en Barrios Andrés, M, (coord.), *Comentarios al Reglamento de Inteligencia Artificial*, Editorial La Ley, Madrid, 2024.

Carmen Muñoz García «¿Protege el Reglamento de IA a los vulnerables? Análisis desde una perspectiva holística y necesaria cohesión con otras normas y derechos consolidados AI Governance» Extracto del documento original de 1 de julio de 2024, en el marco del Subproyecto 1.1. «Mapping of AI governance – recommendations and regulation AI Governance (Project #1)», *cAIre Research* Project FOR GOOGLE.org. Publicado extracto documento original en *Blog OdiseIA*, 6 de enero 2025. Visto en https://www.odiseia.org/post/protege-el-reglamento-de-ia-a-los-vulnerables

Carmen Muñoz García «Directiva 2024/2853 sobre responsabilidad por los daños causados por productos defectuosos. Contexto y armonización de máximos para proteger, ahora sí, a consumidores y "a otras personas físicas"» *La Ley Unión Europea*, número 132, enero 2025.

Ana Felicitas Muñoz Pérez «Economía de la vigilancia y manipulación mediante la IA. Libertad cognitiva y protección del consumidor», *La Ley mercantil*, núm. 121, 2025, febrero.

Susana Navas Navarro, *ChatGPT y modelos fundacionales*, Editorial Reus, Madrid, 2023.

OCDE *Recommendation of the Council on Artificial Intelligence, Principles for responsible stewardship of trustworthy AI.* amended on May 2024, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449. Ver *Recommendation of the Council on Artificial Intelligence* (22 mayo 2019*) https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449* AI principles (3 mayo 2024) https://www.oecd.org/en/topics/ai-principles.html#:~:text=The%20OECD%20AI%20Principles%20promote,stand%20the%20test%20of%20time.

Máximo J. Pérez García, «La responsabilidad por los daños causados por los productos defectuosos y la Directiva (UE) 2024/2853, de 23 de octubre», *Blog Facultad de Derecho*, UAM, 14 enero 2025. Visto en https://www.blog.fder.uam.es/responsabilidad-por-

los-danos-causados-por-los-productos-defectuosos-y-la-directiva-ue-2024-2853-de-23-de-octubre/

Alan Mathison TURIN, «*Computing Machinery and Intelligence*», *Mind* LIX, Oxford Academy, octubre 1950, pp. 433-460. Visto en https://redirect.cs.umbc.edu/courses/471/papers/turing.pdf