(versión previa a pruebas de imprenta, difundida para usos docentes)

Desentrañando el Reglamento de inteligencia artificial de la Unión Europea: una guía esencial

en AA.VV. Actualidad Legislativa y Jurisprudencial: El derecho ante el reto de las tecnologías disruptivas, Gobierno de Aragón, Zaragoza, 2024 (en prensa).

Lorenzo Cotino Hueso[1] www.cotino.es
Catedrático de Derecho Constitucional
Universidad de Valencia. Valgrai

Sumario

1. A modo de presentación; 2. La primera regulación global de la inteligencia artificial: Un marco pionero y sus elementos clave; 3. Qué IA se considera inaceptable y prohibida en la Unión Europea; 4. El modelo de riesgos: sólo se quiere regular la IA más peligrosa, de "alto riesgo"; 5. El "corazón" del Reglamento: ¿Qué obligaciones deben cumplir los sistemas de alto riesgo?; 6.¿Qué debemos saber los humanos cuando interactuamos con determinados sistemas de IA? (artículo 50); 7. Obligaciones respecto de la modelos de IA general como "Chatgpt"; 8. Innovación bajo control: Sandboxes y otras medidas para fomentar la IA; 9. Gobernanza de la IA en Europa: ¿Quiénes supervisan el cumplimiento del Reglamento en España y en la Unión?; 10.¿Cómo se supervisan los sistemas de IA después de su comercialización; 11. Otros aspectos del Reglamento: derechos, sanciones y aplicación gradual, códigos de conducta y delegaciones a la Comisión; 12. Para acabar, también en 2024 un Convenio de inteligencia artificial del Consejo de Europa;

1. A modo de presentación

El Reglamento de Inteligencia Artificial de la Unión Europea (RIA)[2] es un documento pionero, extraordinariamente complejo y detallado, diseñado para establecer un marco legal general de la inteligencia artificial en la Unión Europea,

[1] cotino@uv.es. OdiseIA. El presente estudio es resultado de investigación de los siguientes proyectos: MICINN Proyecto "Derechos y garantías públicas frente a las decisiones automatizadas y el sesgo y discriminación algorítmicas" 2023-2025 (PID2022-136439OB-I00) financiado por MCIN/AEI/10.13039/501100011033/; Proyecto "Algoritmic law" (Prometeo/2021/009, 2021-24 Generalitat Valenciana); "Algorithmic Decisions and the Law: Opening the Black Box" (TED2021-131472A-I00) y "Transición digital de las Administraciones públicas e inteligencia artificial" (TED2021-132191B-I00) del Plan de Recuperación, Transformación y Resiliencia. Estancia Generalitat Valenciana CIAEST/2022/1, Convenio de Derechos Digitales-SEDIA Ámbito 5 (2023/C046/00228673) y Ámbito 6. (2023/C046/00229475).

[2] Su nombre finalmente es 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 (Reglamento de Inteligencia Artificial), DOUE núm. 1689, de 12 de julio de 2024, DOUE-L-2024-81079. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

con pretensiones de influencia mundial. Para dar una idea de su extensión, es comparable a cada tomo del El Señor de los Anillos, que cuales ronda las 100.000 palabras, como el RIA.

Ni más ni menos que se pretende en estas páginas hacer una síntesis del mismo que permita dar al menos una respuesta esencial a cuestiones como las que siguen: ¿Se está regulando la IA en el mundo? ¿Qué caracteriza al Reglamento de Inteligencia Artificial de la Unión Europea como pionero a nivel mundial? ¿Cuáles son los sistemas de IA que se consideran inaceptables y prohibidos bajo el RIA? ¿Qué obligaciones deben cumplir? ¿Qué obligaciones añade el RIA para que sepamos que interactuamos con algunos sistemas de IA? ¿Cómo se regulan modelos de IA de propósito general, como ChatGPT? ¿Qué papel juegan los sandboxes regulatorios en el fomento de la innovación en IA bajo el RIA? ¿Quiénes supervisarán y controlarán que se cumple el RIA en España y en la Unión Europea? ¿Cómo se controlará que un sistema siga siendo seguro después de su comercialización? ¿Hay sanciones por incumplimiento? ¿Qué papel jugará el Convenio de inteligencia artificial del Consejo de Europa junto al Reglamento?

En la actualidad, la inteligencia artificial (en adelante, IA) se ha consolidado como una de las tecnologías más disruptivas e influyentes de nuestro tiempo, por su ingente capacidad de rápida transformación de sectores productivos, desde la medicina hasta la seguridad, y, también, entre muchas otras revoluciones, por su posible integración en objetos, productos o servicios digitales. Al menos en la Unión Europea (en adelante, UE) se percibe desde hace años la necesidad de un marco regulatorio general robusto y coherente que acompañe y encauce el desarrollo de esta tecnología. La Unión Europea, para bien o para mal, ha sido la pionera al establecer esta regulación general tanto para sí misma como para intentar influir en el resto del mundo con lo que se ha llamado el *efecto Bruselas,* algo que en cierta medida consiguió con la aprobación del Reglamento General de Protección de Datos (en adelante, RGPD).

Los pasos para abordar la IA en la UE se iniciaron hace más de siete años con intervención de Comisión, Parlamento y Consejo. Se busca una IA "made in Europe" caracterizada por estar diseñada éticamente para el respeto de los derechos y los principios democráticos.[3] El RIA ha seguido un largo y costoso procedimiento legislativo[4], del que cabe ahora destacar especialmente la

[3] Un análisis exhaustivo de los pasos y políticas de la UE en la materia hasta 2019, Cotino Hueso, L., "Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho" en *Revista Catalana de Derecho Público* nº 58 (junio 2019). http://revistes.eapc.gencat.cat/index.php/rcdp/issue/view/n58 http://dx.doi.org/10.2436/rcdp.i58.2019.3303

[4] Puede seguirse el mismo en los siguientes enlaces de referencia. Parlamento Europeo: https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?reference=2021/0106(COD)&l=en Y *Legislative Train Schedule*, también del Parlamento: https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-regulation-on-artificial-intelligence

propuesta de la Comisión Europea de 2021[5] y la posición del Consejo de la Unión de diciembre de 2022.[6] También son muy relevantes las enmiendas del Parlamento en junio de 2023[7]. El acuerdo político fue de 8 de diciembre de 2023 y desde entonces se ha ido concretando el texto y sus traducciones en las diversas fases de su adopción final hasta su publicación en el DOUE.

La propuesta de la Comisión Europea pretendía, de un lado, garantizar que los sistemas de IA utilizados en la UE e introducidos en el mercado europeo sean seguros y respeten los derechos de los ciudadanos; del otro, estimular la inversión y la innovación en el ámbito de la IA en Europa.

El texto definitivo del Reglamento de IA de la UE pretende armonizar las normas sobre inteligencia artificial, y lo hace con un enfoque basado en el riesgo, estableciendo un marco de obligaciones y requisitos distintos en función del nivel de riesgo de la tecnología de IA aplicable y su uso concreto. La técnica del enfoque armonizador exige, por ejemplo, unas obligaciones reforzadas para los sistemas de inteligencia artificial calificados como de alto riesgo; lo que deberá concretarse con la elaboración de estándares técnicos y el establecimiento de controles ad hoc para su aplicación.

Aunque el RIA entró en vigor en 2024, ciertamente cuenta con una aplicación y obligatoriedad muy escalonada, que llega incluso a los seis años. En cualquier caso, es posible prever un enorme impacto en el mercado y la sociedad. Pronto seremos testigos de las prohibiciones de tecnologías y usos concretos, de la implantación y actividad de las oficinas y autoridades europeas y nacionales de supervisión de la IA. Ya en febrero de 2024 se creó la nueva Oficina de la IA. Veremos la incoación y resolución de procedimientos sancionadores, de la aprobación de estándares técnicos y de la actividad de los organismos de certificación, de la proliferación de sistemas de gestión de riegos específicos, de la aparición de códigos de conducta voluntarios sectoriales, entre muchas otras cuestiones que serán reflejo de la norma que es objeto de estudio en el presente trabajo.

En modo alguno resulta sencillo hacer una síntesis de 106 mil palabras de una norma extraordinariamente compleja.[8] Valga este intento, eso sí, reenviando al lector interesado a los más de 35 estudios que analizan el RIA de modo sistemático en el *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea*, Aranzadi, 2024 que con Simón he coordinado.

[5] https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex%3A52021PC0206
[6] https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/es/pdf
[7] https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_ES.html
[8] Una visión general del RIA, Presno Linera, M. Á: "La propuesta de "Ley de Inteligencia Artificial" europea", en *Revista de las Cortes Generales*, nº 116, 2023, págs. 81-133. De este autor, de interés, *Derechos fundamentales e inteligencia artificial*, Marcial Pons, Madrid, 2023.

2. La primera regulación global de la inteligencia artificial: un marco pionero y sus elementos clave

La aceleración de la regulación mundial de la IA al tiempo del Reglamento de la Unión Europea

El RIA en su tramitación y aprobación final ha estimulado el interés global en la regulación de la IA.[9] Por ejemplo, el 16 de junio de 2022, Canadá presentó la *Digital Charter Implementation Act* de 2022[10] como parte de la *Artificial Intelligence and Data Act (AIDA)*. [11] En diciembre de 2022, Brasil[12] inició la tramitación del proyecto de ley n.º 2338 en el Senado para regular la IA, mientras que el Reino Unido adoptó un enfoque más flexible en la regulación tras el *Brexit*. [13]

El surgimiento de tecnologías como ChatGPT a finales de 2022 impulsó aún más estos desarrollos. Por ejemplo, China implementó las *"Medidas provisionales para la gestión de servicios de IA generativa"* el 13 de julio de 2023. [14] En septiembre de 2023, Canadá dio a conocer su *"Código de Conducta Voluntario para el Desarrollo y Gestión Responsable de Sistemas Generativos Avanzados de IA"*. Desde el G7, el *"Proceso de Hiroshima"* resultó en el *"Código Internacional de Conducta para el Desarrollo de Sistemas Avanzados de IA"* el 30 de octubre de 2023. [15] Ese mismo día, Estados Unidos adoptó la significativa *Orden ejecutiva sobre el desarrollo y la utilización seguros y fiables de la inteligencia artificial*, también el 30 de octubre de 2023. [16] De manera más

---

[9] Una visión general de las propuestas de regulación mundial, muy reciente, Unesco, *Consultation paper on AI regulation. Emerging Approaches Across the World*, 16 de agosto de 2024, Unesco. https://unesdoc.unesco.org/ark:/48223/pf0000390979
De 2023, IALAB, Propuestas de regulación y recomendaciones de inteligencia artificial en el mundo, IALAB-Thomson Reuters, Buenos Aires, https://ialab.com.ar/wp-content/uploads/2023/08/Propuestas-de-regulacion-y-recomendaciones-de-IA-en-el-mundo-1.pdf

[10] https://www.parl.ca/legisinfo/en/bill/44-1/c-27

[11] El texto en https://www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading El estado de tramitación en https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act y https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document

[12] *Projeto de lei nº 2338 do Senado, que regulamenta a IA*, texto inicial en https://legis.senado.leg.br/sdleg-getter/documento?dm=9347622&ts=1702407086098&disposition=inline&_gl=1*1ifop8*_ga*MTg0Njk3ODg3MS4xNzA1Mzk4MDU2*_ga_CW3ZH25XMK*MTcwNTM5ODA1Ni4xLjAuMTcwNTM5ODA1Ni4wLjAuMA_ Procedimiento en https://www25.senado.leg.br/web/atividade/materias/-/materia/157233

[13] Así, destaca Gobierno de Reino Unido: *A pro-innovation approach to AI regulation*, Department for Science, Innovation and Technology and Office for Artificial Intelligence, 29 de marzo 2023, actualización 3 de agosto. https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper

[14] http://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm

[15] https://www.mofa.go.jp/files/100573473.pdf

[16] *Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*. https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/ Cabe seguir también la nota informativa https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/ e

específica, California aprobó el *Artificial Intelligence Accountability Act*[17] el 3 de enero de 2024, mientras que Illinois cuenta con su *Artificial Intelligence Video Interview Act (820 ILCS 42/)*[18]. El 17 de mayo de 2024 el gobernador de Colorado firmó la ley de protección del consumidor para las interacciones con la inteligencia artificial, que entrará en vigor del 1 de febrero de 2026, se trata de la primera legislación integral en los Estados Unidos dirigida a los "sistemas de inteligencia artificial de alto riesgo".[19] En Iberoamérica, existen diversas iniciativas legislativas en curso. [20] Como se señala al final de este estudio, 2024 ha traído no sólo el RIA final, sino un Convenio del Consejo de Europa.

Una visión general del Reglamento: fundamentos y disposiciones generales

El RIA tiene como objetivo regular el uso de la inteligencia artificial dentro de la Unión Europea, con la finalidad de mitigar los riesgos inherentes a su aplicación. El ámbito de aplicación del RIA sigue la estela del Reglamento de protección de datos (RGPD) buscando no sólo el efecto Bruselas de influencia, sino la extraterritorialidad en su aplicación. Así, conforme al artículo 2, abarca tanto a los proveedores de sistemas de IA que se comercialicen o pongan en servicio en la UE, independientemente de su origen, como a aquellos cuya salida o resultados se utilicen en el territorio de la Unión. Asimismo, incluye a los implementadores, entendidos como los operadores de estos sistemas, excluyendo a los usuarios finales o afectados. El RIA también es aplicable a los proveedores de estos sistemas, a sus representantes autorizados, importadores y distribuidores, asegurando así un marco normativo amplio y comprensivo para la regulación de la inteligencia artificial en el contexto europeo.

El RIA excluye de su aplicación a las autoridades públicas de terceros países y a las organizaciones internacionales que empleen sistemas de IA en el marco de la cooperación policial o judicial con la UE o sus Estados miembros. De igual modo, no se aplica a sistemas destinados exclusivamente a fines militares, seguridad nacional o actividades de investigación y desarrollo científico.

El artículo 3 del RIA define un sistema de inteligencia artificial como aquel que, operando con distintos niveles de autonomía, puede adaptarse después de su despliegue. Este tipo de sistemas, con objetivos explícitos o implícitos, procesan la entrada de datos para generar resultados como predicciones, contenido, recomendaciones o decisiones que influyen en entornos físicos o

[17] *California Senate Bill 896*, https://legiscan.com/CA/text/SB896/id/2868456

[18] https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=4015&ChapterID=68#:~:text=An%20employer%20may%20not%20use,use%20of%20artificial%20intelligence%20analysis

[19] *SB24-205, Consumer Protections for Artificial Intelligence. Concerning consumer protections in interactions with artificial intelligence systems*. https://leg.colorado.gov/bills/sb24-205

[20] Existen propuestas normativas de legisladores particulares y otras vinculadas a estrategias nacionales en la materia como en Perú, Chile, Colombia, por mencionar algunos.

virtuales. El RIA establece una jerarquía de riesgos basada en el uso de la IA y asigna obligaciones proporcionales a los riesgos identificados.

Gran parte del articulado del RIA se centra en la regulación de los sistemas de alto riesgo, que se clasifican en dos grandes categorías: finalidades de alto riesgo y productos de alto riesgo. En la primera categoría, el RIA enumera y describe un conjunto de sistemas cuya finalidad o uso se considera de alto riesgo, como los sistemas de identificación biométrica, protección de infraestructuras críticas, selección y promoción de personal, control fronterizo, o los empleados por las Fuerzas y Cuerpos de Seguridad del Estado o la Administración de Justicia. La Comisión Europea puede actualizar esta lista mediante actos delegados. Estos sistemas no serán considerados de alto riesgo en casos específicos, tales como cuando el sistema realice una tarea procedimental limitada, mejore el resultado de una actividad humana previamente completada, o detecte patrones de toma de decisiones sin reemplazar o influir en la evaluación humana sin una revisión adecuada.

Por otro lado, existen productos ya regulados por normativas armonizadas de la UE, sujetos a evaluaciones de conformidad. Entre ellos se encuentran dispositivos médicos, ascensores, juguetes o maquinaria. Un sistema de IA que forme parte de estos productos, o que constituya un componente de seguridad de ellos, estará sujeto a la normativa armonizada correspondiente y será considerado de alto riesgo bajo el RIA.

El RIA impone a los sistemas de alto riesgo una serie de requisitos obligatorios que deben demostrarse mediante pruebas de conformidad. Estos requisitos incluyen la gestión de riesgos, la gestión de datos, la documentación técnica, la transparencia hacia el implementador, la precisión, la robustez ante fallos, la ciberseguridad, la trazabilidad y la supervisión humana.

Además de los sistemas de alto riesgo, el RIA también regula los sistemas de inteligencia artificial de propósito general, que pueden ser utilizados para una variedad de finalidades, tanto directamente como integrados en otros sistemas de IA. El RIA establece obligaciones específicas para los proveedores de estos sistemas, siendo más estrictas cuando dichos sistemas se consideran de riesgo sistémico.

El RIA define diversas entidades de supervisión. Entre ellas, los organismos de evaluación de conformidad, encargados de realizar las pruebas de conformidad, deben estar habilitados por una autoridad notificante. Esta última notificará a la Comisión Europea sobre los organismos habilitados, documentando sus procedimientos. Las autoridades de supervisión de mercado serán responsables de monitorizar el correcto funcionamiento de los sistemas de IA de alto riesgo ya en el mercado, identificando riesgos, incidentes u otras situaciones que requieran medidas correctivas. Los productos de IA considerados de alto riesgo estarán supervisados por la autoridad designada en la normativa correspondiente.

En el caso de sistemas de identificación biométrica utilizados por las Fuerzas y Cuerpos de Seguridad del Estado, en el ámbito migratorio o en la

Administración de Justicia, la supervisión recaerá sobre las autoridades de protección de datos competentes, como la Agencia Española de Protección de Datos en algunos casos.

Finalmente, para favorecer la gobernanza del RIA, se establecen varios órganos, incluyendo la Oficina Europea de la IA y el Consejo Europeo de Inteligencia Artificial, además de dos órganos consultivos: el grupo científico de expertos y el foro consultivo. Finalmente el RIA consta de 113 artículos repartidos en trece capítulos, más trece anexos. En el Capítulo I ("Disposiciones generales"), se establecen las bases del RIA, incluyendo el objeto, ámbito de aplicación, definiciones clave, y alfabetización en IA. En el Capítulo II ("Prácticas prohibidas de IA"), se especifican las prácticas de inteligencia artificial que están completamente prohibidas dentro de la UE, asegurando la prohibición de usos peligrosos o éticamente inaceptables de la IA. El Capítulo III ("Sistemas de IA de alto riesgo") es uno de los más extensos, dividido en cinco secciones. Este capítulo abarca desde la clasificación de los sistemas de IA como de alto riesgo, los requisitos y obligaciones de los proveedores e implementadores, hasta las normas de evaluación de conformidad y la cooperación con las autoridades competentes. En el Capítulo IV ("Obligaciones de transparencia para los proveedores e implementadores de determinados sistemas de IA"), se establecen las responsabilidades de transparencia para asegurar que la información relevante sobre los sistemas de IA esté disponible para los usuarios. El Capítulo V ("Modelos de IA de uso general") regula los modelos de IA de propósito general, abarcando su clasificación, obligaciones de los proveedores, y los códigos de buenas prácticas. El Capítulo VI ("Medidas de apoyo a la innovación") fomenta la innovación en IA mediante la creación de sandboxes regulatorios y otras medidas que facilitan las pruebas en condiciones reales, especialmente para pymes y startups. El Capítulo VII ("Gobernanza") se divide en dos secciones: la primera aborda la gobernanza a nivel de la Unión, incluyendo la creación de la Oficina de IA y del Comité Europeo de Inteligencia Artificial, mientras que la segunda trata sobre las autoridades nacionales competentes. En el Capítulo VIII ("Base de datos de la UE para sistemas de IA de alto riesgo"), se crea una base de datos centralizada para los sistemas de IA de alto riesgo, garantizando transparencia y trazabilidad. El Capítulo IX ("Seguimiento poscomercialización, intercambio de información y vigilancia del mercado") detalla las obligaciones de seguimiento de los sistemas de IA tras su comercialización, el intercambio de información sobre incidentes graves, y la vigilancia del mercado. En el Capítulo X ("Códigos de conducta y directrices"), se fomenta la creación de códigos de conducta para la aplicación voluntaria de requisitos específicos, y se establecen directrices por parte de la Comisión. El Capítulo XI ("Delegación de poderes y procedimiento de comité") regula la delegación de poderes a la Comisión y el procedimiento de comités para la toma de decisiones relacionadas con la IA. El Capítulo XII ("Sanciones") detalla las sanciones aplicables por incumplimiento, incluyendo multas administrativas para instituciones, órganos, organismos de la Unión y proveedores de modelos de IA

de uso general. Finalmente, en el Capítulo XIII ("Disposiciones finales"), se incluyen modificaciones a otros reglamentos y directivas de la UE, además de disposiciones sobre la evaluación y revisión del reglamento.

Los trece anexos del RIA detallan la legislación de armonización aplicable, listan las infracciones penales relevantes y describen los requisitos técnicos y procedimientos necesarios para la conformidad de los sistemas de IA, incluyendo los de alto riesgo. Además, proporcionan pautas para la transparencia, registros y documentación técnica, y especifican los criterios para la designación de modelos de IA con riesgo sistémico.[21]

Son más de cincuenta las definiciones que incluye el RIA, por destacar alguna, alineándose con la definición de la OCDE,[22] el RIA define al sistema de inteligencia artificial como aquel que opera con diferentes niveles de autonomía, capaz de adaptarse tras su despliegue y de inferir, a partir de los datos que recibe, cómo generar resultados tales como predicciones, contenido, recomendaciones o decisiones que puedan influir en entornos físicos o virtuales. La figura del "Proveedor" se refiere a la persona física o jurídica, o entidad pública, que desarrolla o para quien se desarrolla un sistema de IA, y lo pone en

[21] En el Anexo I se presenta una lista de la legislación de armonización de la Unión, estableciendo el marco legal que interacciona con las disposiciones del Reglamento de IA. El Anexo II contiene una lista detallada de las infracciones penales a las que se refiere el artículo 5, apartado 1, párrafo primero, letra h), inciso iii), especificando las conductas delictivas que están prohibidas cuando se emplea inteligencia artificial. El Anexo III identifica los sistemas de IA de alto riesgo según lo establecido en el artículo 6, apartado 2, detallando aquellos sistemas que requieren un control y vigilancia más estrictos debido a los riesgos que implican. En el Anexo IV se describen los requisitos de la documentación técnica que deben cumplir los sistemas de IA de alto riesgo, de acuerdo con el artículo 11, apartado 1. Esta documentación es esencial para demostrar el cumplimiento de las normas de seguridad y calidad. El Anexo V proporciona el formato y contenido de la Declaración UE de conformidad, necesaria para que los sistemas de IA puedan ser comercializados dentro del mercado único europeo. El Anexo VI detalla el procedimiento de evaluación de la conformidad basado en el control interno, un proceso que los proveedores deben seguir para asegurar que sus sistemas cumplen con los requisitos establecidos. El Anexo VIII incluye información relacionada con la evaluación del sistema de gestión de la calidad y la documentación técnica. Este anexo se divide en varias secciones, cada una especificando la información que los proveedores y desplegadores de sistemas de IA de alto riesgo deben presentar al momento del registro, en conformidad con el artículo 49. El Anexo IX describe la información que debe presentarse en el registro de los sistemas de IA de alto riesgo enumerados en el anexo III, especialmente en lo que respecta a los ensayos en condiciones reales, de conformidad con el artículo 60. El Anexo X enumera los actos legislativos de la Unión relativos a los sistemas informáticos de gran magnitud en el espacio de libertad, seguridad y justicia, proporcionando el contexto normativo más amplio en el que se sitúan las regulaciones de IA. El Anexo XI se refiere a la documentación técnica exigida para los proveedores de modelos de IA de uso general, dividiéndose en dos secciones que especifican tanto la información básica que todos los proveedores deben proporcionar, como la información adicional requerida para aquellos modelos con riesgo sistémico. El Anexo XII detalla los requisitos de transparencia que deben cumplir los proveedores de modelos de IA de uso general, especialmente en lo que respecta a la documentación técnica que deben facilitar a los proveedores posteriores que integren el modelo en su sistema de IA. Finalmente, el Anexo XIII establece los criterios para la designación de modelos de IA de uso general con riesgo sistémico, como se describe en el artículo 51, proporcionando un marco para la clasificación y gestión de estos sistemas de mayor riesgo.

[22] https://www.oecd-ilibrary.org/science-and-technology/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_623da898-en

servicio o lo comercializa bajo su nombre o marca, ya sea con fines lucrativos o no. Por otro lado, se define al "Implementador" como la persona física o jurídica, pública o privada, bajo cuya autoridad se utiliza el sistema. Cabe destacar que el RIA excluye a los implementadores que sean personas físicas utilizando sistemas de IA en un contexto puramente personal y no profesional.

3. Qué IA se considera inaceptable y prohibida en la Unión Europea

El artículo 5 (Capítulo II) es uno de los más importantes. Ahí, el RIA prohíbe una serie de sistemas de inteligencia artificial que se consideran particularmente peligrosos debido a su potencial para causar daños físicos o psicológicos, o para afectar negativamente los derechos fundamentales de las personas. Entre estos se incluyen aquellos sistemas que utilizan técnicas subliminales con la intención de alterar el comportamiento de una persona, de manera que puedan provocarle daños a ella o a terceros.

Asimismo, se prohíben los sistemas que se aprovechan de las vulnerabilidades específicas de ciertos grupos, como aquellos definidos por su edad, discapacidad o situación social o económica, y que de esta forma podrían distorsionar su comportamiento, generando posibles daños. Los sistemas que categoricen a personas a partir de datos biométricos para inferir información sensible también están prohibidos, ya que estos procesos pueden conducir a la toma de decisiones discriminatorias.

Además, se consideran inaceptables los sistemas de inteligencia artificial diseñados para crear perfiles de personas basados en su comportamiento, si esto conduce a la elaboración de un "baremo social" que podría resultar en un trato desproporcionadamente desfavorable para ciertas personas o grupos, especialmente en contextos distintos a aquellos en los que se recopilaron los datos. Otro tipo de sistemas prohibidos son aquellos que pretenden evaluar o predecir la probabilidad de que una persona cometa un delito, basándose únicamente en perfiles o características personales. Esta prohibición admite una excepción: cuando estos sistemas se utilicen para apoyar la evaluación humana en la determinación de la implicación de una persona en actividades delictivas.

También están prohibidos los sistemas que buscan crear o expandir bases de datos de reconocimiento facial mediante la recopilación indiscriminada de imágenes faciales de internet o de grabaciones de sistemas de videovigilancia. Por último, se prohíbe la utilización de sistemas que infieran emociones de personas físicas en el entorno laboral o en instituciones educativas, salvo que dichos sistemas se empleen con fines médicos o de seguridad.

El uso de la identificación biométrica en tiempo real en lugares públicos por parte de fuerzas de seguridad o autoridades públicas ha sido desde el inicio, uno de los caballos de batalla, más discutidos política y socialmente que ha supuesto una regulación final muy compleja que, además, debe superponerse con el

RGPD.[23] El artículo 5 establece una prohibición general, con algunas excepciones para casos específicos como la búsqueda de víctimas de delitos, la prevención de amenazas graves, la prevención de ataques terroristas, y la persecución de crímenes como el terrorismo o la trata de seres humanos. Para la autorización de estos sistemas se requerirá una evaluación previa de la escala y probabilidad del daño potencial, además de la obtención de una autorización judicial o administrativa, y se deberán imponer restricciones temporales, geográficas y personales. Fuera de estas excepciones, el uso de identificación biométrica en tiempo real sigue estando sujeto al Reglamento General de Protección de Datos (RGPD) 2016/679 y a la Directiva 2016/680 sobre el tratamiento de datos personales por parte de autoridades.

Debe subrayarse la grave dificultad que entraña la interpretación concreta de algunos de los sistemas prohibidos, y sobre todo hay que tener en cuenta que este artículo cinco de prohibiciones y sistemas inaceptables debe leerse conjuntamente con la descripción de los sistemas de alto riesgo, especialmente establecida en el anexo III. Y es que debe señalarse que muchos de los sistemas que no están prohibidos por excepcionarse en este artículo 5 pasan a ser de alto riesgo. Asimismo, en ocasiones no es fácil distinguir cuando el sistema es prohibido o es de alto riesgo, como puede suceder con los relacionados con el cumplimiento de la ley.

4. El modelo de riesgos: sólo se quiere regular la IA más peligrosa, de "alto riesgo"

El modelo de riesgos para regular la inteligencia artificial

Ante la IA, es fundamental adoptar técnicas de responsabilidad "proactiva", "demostrada" o "accountability" desde el diseño, siguiendo el principio jurídico de "más vale prevenir que curar"[24]. Inspirándose en las leyes de la robótica de Asimov, esto implica la integración del cumplimiento ético y legal directamente en el "Código", conforme a los términos expuestos por Lessig. Este enfoque, conocido como compliance o cumplimiento normativo de origen anglosajón, comienza a ser reconocido en el Derecho continental europeo, aplicándose en áreas como el medio ambiente o los estudios de impacto de género en

[23] Sobre el tema, mis estudios, "Reconocimiento facial automatizado y sistemas de identificación biométrica bajo la regulación superpuesta de inteligencia artificial y protección de datos", en Balaguer, F. y Cotino, L. (2023): *Derecho público de la inteligencia artificial*, F. Jiménez Abad-Marcial Pons, acceso , más breve, "Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos. Mejor regular bien que prohibir mal", en *El Cronista del Estado Social*, IUSTEL, *monográfico Inteligencia artificial*, nº 100, septiembre-octubre 2022, pp. 68-79. acceso

[24] Sobre la aplicación del principio de responsabilidad proactiva y el diseño para el ámbito de la IA y el big data, Hernández Peña, J. C. (2022): *El marco jurídico de la inteligencia artificial. Principios, procedimientos y estructuras de gobernanza*, Aranzadi, Cízur, págs. 119-172; Martínez Martínez, R.: "Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo", *Revista catalana de dret públic*, nº 58, 2019, págs. 64-81. También, mi estudio, "Nuevo paradigma en la garantías de los derechos fundamentales y una nueva protección de datos frente al impacto social y colectivo de la inteligencia artificial", en Cotino Hueso, Lorenzo (editor), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thompson-Reuters Aranzadi, FIADI, Cizur, 2022, págs. 69-105.

disposiciones generales, [25] y extendiéndose para incluir la discriminación algorítmica pública. Este modelo se generaliza en la UE en diversos ámbitos digitales[26] y en la legislación europea sobre comercialización de productos potencialmente peligrosos.

El modelo de responsabilidad proactiva se ha proyectado claramente en la UE como el "Ethics & Rule of law by design X-by design", insignia de la "IA confiable" "Made in Europe", [27] que fundamenta las políticas y normativas de IA en la UE. [28] Además, el alto grupo de expertos de la UE adoptó un modelo de ética en el diseño con su exhaustiva lista de evaluación, [29] su modelo ALTAI de evaluación o FUTURE-AI para salud. [30] El Libro Blanco de la IA de 2020 clarificó el modelo de riesgos, que se concretaría un año después en el RIA. Este enfoque supone un aumento de obligaciones y garantías en proporción al impacto o riesgo del sistema de IA. Es notable destacar las contribuciones de Mantelero[31] y Simón en España[32] en el ámbito de la protección de datos y posteriormente para los sistemas de IA.

El Capítulo III inicialmente regula la "Clasificación de sistemas IA de alto riesgo". Es de interés contextualizar para una mejor comprensión, que el RIA ha encajado la regulación de la IA en el ámbito de la seguridad y garantía de los productos, las normas de armonización y el modelo del llamado "nuevo marco legislativo". Se trata del marco por el que se establecen unas bases comunes para la comercialización, evaluación y vigilancia de productos en la Unión Europea[33]. Todo sea dicho, es un modelo con el que los juristas en general no

---

[25] Impuestos a nivel estatal por artículos 22.2 para leyes y 24.1.b para reglamentos, Ley 50/1997, de 27 de noviembre, del Gobierno; artículo 19 Ley Orgánica 3/2007 para disposiciones de carácter general y los planes de especial relevancia económica y artículos 26. 3 f) Ley 40/2015. Existe abundante normativa autonómica.

[26] Barrio Andrés, M.: *El cumplimiento basado en el riesgo o risk-based compliance, pieza cardinal del nuevo Derecho digital europeo*, Real Instituto Elcano, ARI 34/2023, 20 abril, 2023, https://media.realinstitutoelcano.org/wp-content/uploads/2023/04/el-cumplimiento-basado-en-el-riesgo-o-risk-based-compliance-pieza-cardinal-del-nuevo-derecho-digital-europeo-real-instituto-elcano.pdf

[27] Comisión Europea, *Directrices éticas para una IA fiable*, 2019, https://data.europa.eu/doi/10.2759/14078 versión 2018, págs. 19, 21, 29.

[28] Un análisis exhaustivo en mi estudio "Ética en el diseño… cit.

[29] Comisión Europea, *Directrices éticas… cit.*, en especial Capítulo III y listado, págs. 33-41.

[30] Portal ALTAI https://futurium.ec.europa.eu/en/european-ai-alliance/pages/welcome-altai-portal

[31] Respecto del ámbito de los datos Mantelero, A.: "Toward a New Approach to Data Protection in the Big Data Era", en Gasser U. y otros (dir.): *Internet Monitor 2014: Reflections on the Digital World.* Cambridge (MA): Berkman Center for Internet and Society at Harvard University, págs. 84 y ss. Para la IA, Mantelero, A.: *Beyond Data. Human Rights, Ethical and Social Impact Assessment*, Springer, Information Technology and Law Series IT&LAW 36, 2022, https://link.springer.com/book/10.1007/978-94-6265-531-7

[32] Simón Castellano, P.: *La evaluación de impacto algorítmico en los derechos fundamentales*, Aranzadi, Cízur, 2023. También, "Taxonomía de las garantías jurídicas en el empleo de los sistemas de inteligencia artificial", *Revista de Derecho Político*, nº 117, 2023, págs. 153-196.

[33] Las tres textos legales que conforman el Nuevo Marco Legislativo son: el Reglamento (CE) n.º 765/2008 del Parlamento Europeo y del Consejo por el que se establecen los requisitos de acreditación y vigilancia del mercado de los productos; la Decisión n.º 768/2008/CE del Parlamento Europeo y del Consejo sobre un marco común para la comercialización de los productos y; el Reglamento (UE) 2019/1020 del Parlamento Europeo y del Consejo relativo a la vigilancia del mercado y la conformidad de los productos.

estamos muy familiarizados. Ello influye también en la definición de lo que es un sistema de alto riesgo.

Identificación de sistemas de IA de alto riesgo: ¿Cuáles son y por qué?

Así, en primer término, estos sistemas incluyen aquellos que se implementan como productos o componentes de seguridad ya cubiertos por la legislación armonizada de la Unión Europea, lo que implica que están sujetos a evaluaciones de conformidad realizadas por terceros, como es el caso de dispositivos médicos, ascensores, aviones o maquinaria. Las normas armonizadas aplicables a estos sistemas están enumeradas en el Anexo II del RIA. Sin poder entrar ahora en muchos detalles que serían necesarios, hay que señalar que la clave está en delimitar qué productos peligrosos de la Unión Europea se considera que lo son tanto como para que esta evaluación de conformidad debe ser realizada por un tercero (un llamado "organismo, notificado"), pues por lo general, se suelen establecer auto evaluaciones por los propios proveedores.

En segundo lugar, son de alto riesgo los sistemas IA utilizados para "influir sustancialmente en el resultado de la toma de decisiones" (art. 6.3 RIA) para hasta 25 finalidades enumeradas en el Anexo III. Así, el Anexo III del RIA especifica una serie de sistemas que se clasifican en diversas categorías de alto riesgo, tales como: sistemas de identificación biométrica que operan sin la participación activa del sujeto, infraestructuras críticas como el tráfico, la electricidad o el agua, así como la gestión del acceso a la educación o la planificación del desarrollo académico en el ámbito educativo y de formación profesional.

Asimismo, se incluyen sistemas relacionados con la selección de personal y la gestión de relaciones laborales, la gestión del acceso a servicios esenciales tanto públicos como privados, y actividades vinculadas a las fuerzas y cuerpos de seguridad, como la valoración de pruebas o sospechosos. También se consideran de alto riesgo los sistemas empleados en migración, asilo y control de fronteras, como polígrafos o la valoración de solicitudes, y aquellos utilizados en la administración de justicia y procesos democráticos.

Es importante destacar que la Comisión Europea posee una capacidad para añadir o retirar sistemas de estas categorías, adaptándose a las circunstancias cambiantes, así como para ir elaborando criterios y determinaciones que permitan concretar cuando un sistema es de alto riesgo.

No obstante, hay situaciones en las que estos sistemas de IA no se considerarán de alto riesgo. Por ejemplo, si el sistema realiza una tarea procedimental limitada, mejora el resultado de una actividad humana previamente completada o detecta patrones de toma de decisiones o desviaciones sin pretender reemplazar o influir en la evaluación humana sin una revisión adecuada. En tales casos, el proveedor debe documentar el proceso que llevó a considerar que su sistema no es de alto riesgo y debe proceder a su registro en una base de datos específica.

5. El "corazón" del Reglamento: ¿Qué obligaciones deben cumplir los sistemas de alto riesgo?

Obligaciones esenciales que deben cumplir los sistemas de IA parar ser seguros

El RIA implica esencialmente que los sistemas de IA de alto riesgo (artículo 6, anexo I y III) deben cumplir ciertos requisitos. "La clasificación de un sistema de IA como "de alto riesgo" debe limitarse a aquellos sistemas de IA que tengan un efecto perjudicial importante en la salud, la seguridad y los derechos fundamentales de las personas de la Unión" (Cons. 46). Como se señaló al inicio, el Considerando 48 menciona veintiún derechos fundamentales afectados por estos sistemas, además de derechos específicos de menores.

En caso de sistemas de alto riesgo, se requiere cumplir con un conjunto de obligaciones preventivas que, de verificarse, permiten una evaluación de conformidad. Ello permite que el sistema IA de alto riesgo se incluya en un registro y pueda ponerse en el mercado bajo el marcado "CE" de la UE. Así, en la Sección dos se regulan los requisitos para los sistemas de IA de alto riesgo. Los sistemas de inteligencia artificial clasificados como de alto riesgo deben cumplir con una serie de exigencias específicas que garantizan su seguridad y la protección de los derechos fundamentales. En primer lugar, se requiere la implementación de un sistema de gestión de riesgos que tenga en cuenta, especialmente, los riesgos que estos sistemas puedan presentar para la salud, la seguridad y los derechos fundamentales en función de su propósito.

Además, es imprescindible establecer una gobernanza adecuada y una gestión rigurosa de los datos utilizados para el entrenamiento y prueba de los sistemas. use datos de alta calidad (art. 10 y 16 RIA). Se parte de que los datos deben "estar libres de errores y ser completos en vista de la finalidad prevista del sistema" (Considerando 67). Si los datos de los que se alimenta la IA son basura, sus resultados necesariamente serán basura. El proveedor de un sistema de alto riesgo ha de generar documentación técnica que demuestra el cumplimiento de los requisitos y los elementos técnicos del sistema, para ponerla a disposición de las autoridades (art. 11). La documentación debe incluir un contenido mínimo especificado, que puede ser modificado por la Comisión cuando se considere necesario.

También se ha de diseñar la IA de alto riesgo para que genere registros y logs de funcionamiento (art. 12). Se trata de los sistemas de IA más peligrosos estén localizados y que se puedan vigilar, así como trazar exactamente cómo han funcionado, así como verificar que cumplen las obligaciones preventivas. Asimismo, al menos para los sistemas peligrosos, se hace preciso diseñar el sistema algorítmico de manera que pueda ser auditable, esto es, creando registros entendibles para que se pueda investigar si funcionó bien en todo momento. Y obviamente el sistema de IA ha de estar bien diseñado y comprobarse que funciona bien en todo momento. También hay obligaciones de "transparencia y comunicación de información a los responsables del

despliegue" (art. 13), esto es, que los sistemas tengan buenos "manuales de instrucciones" y las empresas o administraciones -los implementadores- que los utilicen los conozcan y puedan también dar suficiente información a la ciudadanía, que son los afectados últimos por la IA. Así, se debe proporcionar información detallada a los implementadores sobre las capacidades del sistema, los requisitos de equipamiento, el ámbito de aplicación, el nivel de precisión, y las condiciones de uso que puedan implicar riesgos, así como los mecanismos de supervisión humana disponibles.

Las medidas proactivas pasan también en general por las obligaciones de supervisión humana (art. 14). Es fundamental que estos sistemas permitan la supervisión humana durante su uso, con el fin de minimizar cualquier riesgo residual para la salud, la seguridad y los derechos fundamentales, incluso después de que se hayan implementado medidas de mitigación. Los implementadores deben poder monitorizar los sistemas y entender las decisiones o resultados que estos generan. En casos de identificación biométrica remota, se requiere que la salida del sistema sea verificada por al menos una persona física, y posiblemente dos, para garantizar su fiabilidad.

Finalmente, los sistemas de IA evitarán los sesgos y la discriminación si están diseñados con "solidez" técnica, cumpliendo obligaciones para prevenir o minimizar "errores, fallos, incoherencias, situaciones inesperadas" (Considerando 75) y la eliminación o reducción de resultados sesgados una vez comercializados (art. 15. 3º).- Estos sistemas deben ser diseñados con tolerancia a errores o inconsistencias que puedan surgir en su interacción con el entorno, especialmente cuando interactúan con personas u otros sistemas. Además, deben incorporar medidas de ciberseguridad -que no sean *hackeables*-. adecuadas y proporcionadas a las circunstancias específicas del sistema, especialmente para proteger contra la manipulación de los datos de entrenamiento.

Todas estas exigencias serán desarrolladas además por todo un armazón de normas técnicas y de armonización, que serán acompañadas de la normalización técnica de la IA que está actualmente en fase de desarrollo. Para los sistemas que no son de alto riesgo, todas estas obligaciones serán los referentes de las buenas prácticas a seguir voluntariamente en códigos, sellos o certificaciones a las que se sumen (Considerandos 165-166) los proveedores públicos o privados de esos sistemas de IA.

Responsabilidades de proveedores, implementadores y otros actores clave

En el marco del RIA, se establecen obligaciones claras para las diferentes partes que intervienen en la puesta en marcha, comercialización y funcionamiento de los sistemas de IA de alto riesgo. Los proveedores asumen la mayor parte de estas responsabilidades, debiendo cumplir estrictamente con los requisitos establecidos y superar las pruebas de conformidad. Además, están obligados a mantener un sistema de gestión de calidad bien documentado y actualizado, que incluya la preservación de todos los registros pertinentes

relacionados con el sistema. Asimismo, deben colaborar activamente con las autoridades, lo cual implica registrar el sistema, demostrar su conformidad con los requisitos reglamentarios cuando sea necesario, y notificar cualquier incumplimiento o riesgo identificado, así como las medidas correctivas adoptadas en consecuencia.

Los implementadores fueron llamados "usuarios" en el periodo de elaboración, pero que nada tienen que ver con personas afectadas por una IA. Pues bien, los implementadores de sistemas IA de alto riesgo, especialmente si se trata de organismos públicos o empresas que prestan servicios públicos, tienen la obligación adicional de realizar una evaluación de impacto sobre los derechos fundamentales relacionada con el sistema de IA. Por su parte, los importadores y distribuidores están encargados de conservar la documentación técnica y asegurarse de que los productos hayan pasado por la evaluación de conformidad requerida y cuenten con el marcado CE. Además, deben garantizar que exista un representante autorizado del proveedor dentro de la Unión Europea y colaborar con las autoridades cuando sea necesario.

Los implementadores también tienen responsabilidades clave en la supervisión humana de los sistemas, debiendo monitorizar su funcionamiento, guardar registros detallados y cooperar con las autoridades competentes. Cualquier modificación sustancial de un sistema de alto riesgo existente o de un sistema de inteligencia artificial de propósito general, que lo convierta en un sistema de alto riesgo, se considera como la creación de un nuevo sistema de alto riesgo, lo cual genera las mismas obligaciones mencionadas anteriormente.

La evaluación de conformidad que se obtiene si se cumplen estándares y normas y el registro en la base de datos de sistemas de alto riesgo

En la sección 5 de este capítulo III, se abordan las normas armonizadas, las evaluaciones de conformidad, los certificados y el registro de los sistemas de IA. Si un sistema de IA de alto riesgo o un sistema de IA de propósito general cumple con las normas técnicas armonizadas publicadas en el Diario Oficial de la Unión Europea, se presumirá que cumple con los requisitos del RIA. En ausencia de estas normas, la Comisión tiene la facultad de aprobar especificaciones comunes que actuarán como sustituto de las normas técnicas hasta que estas estén disponibles.

En la mayoría de los casos, la conformidad de los sistemas de alto riesgo y de propósito general se puede demostrar a través de un procedimiento de evaluación interna. No obstante, para sistemas de identificación biométrica regulados por el RIA, si no se han aplicado normas armonizadas o solo parcialmente, se requerirá la intervención de un organismo notificado.[34]

[34] Sobre este tema, de especial complejidad, Palma Ortigosa, A.: "¿Quién es quién en el Reglamento Europeo de Inteligencia Artificial? Las autoridades notificantes y los organismos notificados", *Actualidad Jurídica Iberoamericana*, nº 21, 2024, monográfico, pp. 598-617 https://revista-aji.com/numero-21/

Los sistemas que superen la evaluación de conformidad recibirán un certificado emitido por los organismos notificados, con una validez máxima de cuatro años para los sistemas de IA del Anexo II y de cinco años para los sistemas del Anexo III, pudiendo ser renovado si se superan las pruebas pertinentes. Estos sistemas también deberán llevar el marcado CE, y los incluidos en el Anexo III deberán registrarse en un registro europeo de sistemas de alto riesgo.

Finalmente, se prevé que una autoridad de supervisión de mercado puede autorizar, en casos excepcionales y justificados, y por un tiempo limitado, la operación de un sistema de alto riesgo que no haya completado las pruebas de conformidad, cuando esto sea necesario para proteger la vida de las personas, el medio ambiente o infraestructuras clave. Sin embargo, estas pruebas deberán realizarse lo antes posible, y si no se superan, el sistema deberá ser retirado.

Antes de que un sistema de alto riesgo sea puesto en servicio o comercializado, el proveedor está obligado a registrarlo en la base de datos de la UE. Asimismo, los implementadores del sector público deberán registrarse antes de usar el sistema. No obstante, no están obligados al registro los sistemas de alto riesgo destinados al orden público, control de fronteras, inmigración o asilo, ni sus implementadores.

El artículo 71, Capítulo VII del RIA se encarga de regular la creación y el mantenimiento de una base de datos a nivel de la UE, destinada a registrar los sistemas de inteligencia artificial de alto riesgo, tal como se definen en el anexo III. Esta base de datos será gestionada por la Comisión Europea en coordinación con los Estados miembros. Además, se requiere que los implementadores del sector público que utilicen estos sistemas de IA proporcionen información adicional específica durante el proceso de registro.

6.¿Qué debemos saber los humanos cuando interactuamos con determinados sistemas de IA? (artículo 50)

El artículo 50 supone un Capítulo, sobre Obligaciones de transparencia de ciertos sistemas de IA y genera una tipología especial de sistemas de inteligencia artificial con un régimen jurídico específico. Se establecen una serie de requisitos esenciales que los proveedores e implementadores de sistemas de inteligencia artificial deben cumplir para garantizar la transparencia en el uso de estos sistemas, especialmente cuando interactúan directamente con personas. Los proveedores tienen la responsabilidad de asegurarse de que las personas físicas sepan que están interactuando con un sistema de IA, salvo que dicha interacción sea evidentemente obvia, como podría ser el caso en ciertos contextos donde la intervención de IA es claramente identificable. Existen excepciones para este requisito, particularmente en casos relacionados con la persecución de delitos, donde revelar la presencia de un sistema de IA podría comprometer las investigaciones.

En cuanto a los sistemas de IA que generan contenido sintético, como audio o video, los proveedores deben garantizar que exista un mecanismo

técnico que permita corroborar que dicho contenido ha sido creado o manipulado artificialmente. Esta medida busca prevenir el engaño o la confusión en la percepción del contenido por parte del público.

Además, los implementadores de sistemas de categorización biométrica o de reconocimiento de emociones tienen la obligación de informar a las personas sobre el uso de dichos sistemas en su interacción. No obstante, se prevén excepciones en contextos de persecución del crimen, donde la divulgación de esta información podría interferir con la eficacia de las investigaciones.

Por último, en relación con los sistemas de IA que producen imágenes o sonidos que simulan ser reales, como en el caso de los llamados "deep fakes", los implementadores deben advertir sobre la naturaleza artificial de este contenido, con algunas excepciones, como en la persecución de delitos o en el caso de obras que sean claramente creativas, satíricas o ficticias.

7. Obligaciones respecto de la modelos de IA general como "Chatgpt"

El RIA en su fase final y tras el surgimiento de tecnologías como ChatGPT y similares, introdujo regulaciones específicas para los sistemas de IA de propósito general, imponiéndoles una serie de obligaciones generales que se intensifican para aquellos considerados de riesgo sistémico. Estas regulaciones están bajo el control, dirección e impulso de la nueva Oficina de IA de la Unión Europea, que será responsable de establecer códigos de conducta para precisar estas obligaciones. Cabe apuntar que el modelo más o menos regulatorio e incisivo fue uno de los grandes caballos de batalla en la negociación final en la que especialmente Alemania Francia e Italia se oponían a otros países que pretendían una regulación más taxativa, llevándose a un modelo híbrido.

El Capítulo 1 establece "Normas de clasificación de los modelos de IA de propósito general". Así´, el RIA distingue entre modelos de IA de propósito general y aquellos clasificados como de propósito general con riesgo sistémico. Estos últimos son definidos como aquellos que, o bien poseen capacidades de alto impacto, o bien han sido designados como tales por la Comisión, o como resultado de una alerta emitida por el Grupo Científico de Expertos Independientes. El Capítulo 2 regula "Obligaciones de los proveedores de modelos de IA de propósito general". Todos los proveedores de sistemas de inteligencia artificial de propósito general están sujetos a una serie de obligaciones, entre las que se incluyen: la elaboración de documentación técnica adecuada, el diseño de una política de cumplimiento con la Directiva sobre derechos de autor, y la creación de un resumen suficientemente detallado sobre el contenido utilizado durante el entrenamiento del sistema de IA.

Además de las obligaciones generales aplicables a todos los sistemas de IA de propósito general, el Capítulo 3 dispone "Obligaciones de los proveedores de modelos de IA de propósito general con riesgo sistémico". Los proveedores de aquellos clasificados como de riesgo sistémico deberán llevar a cabo evaluaciones de los modelos con el fin de identificar posibles riesgos sistémicos, tomar medidas para mitigar dichos riesgos en su caso, documentar y notificar

incidentes graves, y garantizar un nivel adecuado de protección de la ciberseguridad de estos sistemas.

8. Innovación bajo control: Sandboxes y otras medidas para fomentar la IA

Este Capítulo de Medidas de apoyo a la innovación, esencialmente sobre sandboxes, establece las directrices para implementar medidas de apoyo a la innovación, centrándose en los sandboxes regulatorios. Según el RIA, las autoridades nacionales deben crear en dos años al menos un sandbox regulatorio destinado a desarrollar, entrenar, probar y validar sistemas de inteligencia artificial bajo su orientación, supervisión y apoyo. Cualquier proveedor, actual o potencial, que cumpla con los criterios establecidos y sea seleccionado, podrá participar en estos sandboxes.

La Comisión Europea, a través de actos de ejecución, definirá las características comunes que deberán incluir todos los sandboxes implementados por las autoridades nacionales. Estas características abarcarán aspectos como los procedimientos de planificación y ejecución, incluyendo la elegibilidad, el proceso de solicitud, selección, participación, seguimiento y salida.

La participación en un sandbox estará limitada temporalmente, dependiendo de la escala del proyecto. Basado en un plan específico, si los participantes cumplen con las directrices establecidas por las autoridades, estarán exentos de sanciones administrativas por infracciones legales relativas al sistema supervisado en el sandbox. Sin embargo, esta exención no limita los poderes de supervisión de las autoridades ni exime a los participantes de responsabilidad por cualquier daño causado durante su participación.

Dentro del sandbox, y bajo ciertas condiciones, se permitirá el tratamiento de datos personales recolectados para otros fines, siempre que sean esenciales para el desarrollo de sistemas que persigan un interés público sustancial, como en las áreas de salud, medio ambiente, sostenibilidad energética, movilidad, calidad del servicio público y seguridad de infraestructuras críticas.

Las autoridades nacionales deberán acordar con los participantes las condiciones para realizar pruebas en entornos reales, garantizando siempre la seguridad, la salud y el respeto a los derechos fundamentales. El RIA establece términos y limitaciones, que incluyen, entre otros, la aprobación previa y el registro del plan de pruebas del participante, el consentimiento de los afectados y la reversibilidad de las decisiones tomadas durante las pruebas.

Al finalizar su participación, los participantes recibirán una certificación de las actividades realizadas y un informe final, el cual podrá ser valorado por organismos notificados durante la evaluación de conformidad. Si los participantes lo consienten, estos informes estarán disponibles para la Comisión Europea y el Consejo Europeo de Inteligencia Artificial, y, en algunos casos, podrían publicarse en la plataforma de información única.

Además, se han establecido medidas específicas para PYMEs, incluidas las start-ups, tales como formación adaptada sobre el RIA, canales de comunicación específicos, apoyo en el cumplimiento del RIA y la publicación de información y mejores prácticas. Cabe destacar que las microempresas no estarán obligadas a mantener un sistema de gestión de calidad, aunque sí deberán gestionar los riesgos asociados.

9. Gobernanza de la IA en Europa: ¿Quiénes supervisan el cumplimiento del Reglamento en España y en la Unión?

Es importante señalar que en el RIA se introduce la creación de diversos organismos clave para la gobernanza de la inteligencia artificial en la Unión Europea. Entre ellos, se encuentra la Oficina Europea de la IA, [35] establecida el 24 de enero de 2024, cuya función principal es centralizar el desarrollo y fortalecimiento de las capacidades de la Unión en materia de IA. Asimismo, el RIA regula la conformación del Consejo Europeo de Inteligencia Artificial y la creación de dos órganos consultivos: un grupo científico de expertos y un foro consultivo.

El Capítulo VI del RIA se enfoca en la gobernanza, detallando las competencias asignadas a la Oficina Europea de la IA y al Consejo Europeo de Inteligencia Artificial. Este Consejo, compuesto por un representante de cada Estado miembro y con la participación del Supervisor Europeo de Protección de Datos como observador, tendrá como tareas principales la coordinación entre autoridades competentes, la recopilación de conocimientos técnicos y la promoción de mejores prácticas a nivel europeo. Además, contribuirá a la armonización de pruebas de conformidad y la supervisión de sandboxes y pruebas en condiciones reales, en colaboración con organismos expertos en diversas áreas, como servicios digitales, financieros, ciberseguridad, y criptomonedas.

El Consejo también asesorará a la Comisión Europea en la definición de directrices para la implementación del RIA, en la preparación de actos de ejecución o delegados, y en asuntos internacionales relacionados con la IA. Asimismo, ofrecerá recomendaciones sobre especificaciones, estándares y guías, y apoyará a las autoridades nacionales en la supervisión del mercado de IA.

Por último, los órganos consultivos, como el foro consultivo y el grupo científico de expertos independientes, complementarán las actividades del Consejo proporcionando asesoramiento técnico y científico en la ejecución de las disposiciones del RIA.

Ya respecto de la gobernanza en cada Estado miembro, se designarán autoridades de supervisión del mercado (art. 59 del RIA) encargadas de monitorizar el desempeño de los sistemas de IA, incluidos aquellos considerados

[35] *Decisión de la Comisión de 24.1.2024 por la que se crea la Oficina Europea de Inteligencia Artificial*, Bruselas, 24.1.2024C(2024) 390 final, https://ec.europa.eu/newsroom/dae/redirection/document/101625

de alto riesgo, identificando riesgos emergentes, incidentes y situaciones que puedan requerir intervención. Además, los organismos de evaluación de conformidad, responsables de realizar las pruebas correspondientes, deberán estar habilitados por una autoridad notificante. Cada Estado miembro tiene la facultad de establecer nuevas instituciones o asignar estas competencias a autoridades existentes, como la Agencia Española de Protección de Datos o, para el uso de algoritmos en el ámbito judicial (Anexo III.8º), el Consejo General del Poder Judicial. [36] España ha sido pionera en la UE al crear la Autoridad Española de Supervisión de la IA (AESIA). [37] Estas autoridades deben actuar de manera independiente, si bien es cuestionable que en el caso español de la AESIA su regulación permita contar con este requisito.[38]

En el mecanismo de gobernanza nacional hay que ubicar a las autoridades notificantes y organismos notificados, regulados en la Sección IV del Capítulo III. El RIA establece que cada Estado miembro debe designar una autoridad notificante. Esta autoridad es responsable de establecer y aplicar los procedimientos necesarios para evaluar, designar, notificar y supervisar a los organismos de evaluación de conformidad. La autoridad notificante tiene la opción de delegar la evaluación y supervisión en una organización de acreditación, lo que permite garantizar que los organismos notificados puedan verificar la conformidad de los sistemas de IA de alto riesgo con el RIA.

Los organismos de evaluación de conformidad que deseen ser "notificados", esto es, ser autorizados a poder realizar evaluaciones de conformidad) deben obtener una certificación que acredite el cumplimiento de los requisitos establecidos por un organismo nacional de acreditación. Luego, deben presentar su solicitud y la documentación correspondiente a la autoridad notificante, que a su vez notificará a la Comisión Europea una vez que se hayan cumplido los requisitos necesarios.

[36] Para el caso de sistemas de identificación biométrica utilizados por Fuerzas y Cuerpos de Seguridad del Estado, en migración o en Administración de Justicia, las autoridades de supervisión serán, as autoridades de protección de datos competente en esos ámbitos, pudiendo ser en algunos supuestos la AEPD.

[37] La creación legal y habilitación fue por la Ley 28/2022, de 21 de diciembre, en concreto, su D.A 7ª reguló la creación. Asimismo, según Ley 28/2022, la AESIA está adscrita orgánica al Ministerio de Asuntos Económicos y Transformación Digital, a través de su Secretaria-o SEDIA.El Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto.

[38] Se creó como un organismo público bajo el artículo 91 de la Ley 40/2015 y no como una agencia estatal independiente de acuerdo con los artículos 108 bis y siguientes de la Ley 40/2015. En sus estatutos regulados por el Real Decreto 729/2023, de 22 de agosto, el artículo 8 menciona como "Principios de actuación de la Agencia" la "Autonomía" y la "Independencia técnica" y señala que "la Agencia actuará con plena autonomía". No obstante, estas afirmaciones parecen insuficientes, especialmente dada una evidente falta de independencia orgánica que afecta a la independencia funcional. Incluso se establece explícitamente la "dependencia" de la Dirección de un órgano político como el Secretario de Estado, quien es el Presidente y desempeña algunas funciones materiales (art. 23.1º).

10.¿Cómo se supervisan los sistemas de IA después de su comercialización

Resulta de interés el Capítulo VIII que regula la Monitorización post-comercialización, intercambio de información y vigilancia del mercado. Así, los proveedores de sistemas de alto riesgo deberán implementar un sistema de monitorización post-comercialización que sea adecuado a los riesgos identificados y a la tecnología de IA utilizada. Este sistema recogerá los datos necesarios para asegurar el cumplimiento de los requisitos establecidos para los sistemas de alto riesgo en el Capítulos III,. La estructura del sistema deberá estar documentada en un plan que se ajuste al esquema propuesto por la Comisión.

Respecto del Intercambio de información sobre incidentes graves, los proveedores estarán obligados a notificar cualquier incidente a las autoridades de vigilancia del mercado del Estado miembro donde se haya producido. A su vez, estas autoridades informarán del incidente a las autoridades encargadas de la protección de los derechos fundamentales.

Asimismo se dispone que las autoridades de supervisión del mercado reportarán a la Comisión las actividades realizadas en relación con el RIA. Para los sistemas de alto riesgo regulados por normativa armonizada, la autoridad supervisora del mercado será la indicada por dicha normativa, salvo que el Estado miembro justifique la designación de otra autoridad. En el caso de los sistemas de alto riesgo utilizados por instituciones financieras, el supervisor financiero será la autoridad competente, a menos que el Estado miembro designe otra autoridad justificada.

Para los sistemas de alto riesgo destinados a la identificación biométrica en contextos policiales, migratorios, de aplicación de la ley y justicia, las autoridades de vigilancia del mercado serán las correspondientes autoridades de protección de datos, como la Agencia Española de Protección de Datos en algunos casos.

Las instituciones de la Unión que estén bajo el ámbito del RIA serán supervisadas por el Supervisor Europeo de Protección de Datos. Los Estados miembros deberán facilitar la coordinación entre sus autoridades nacionales de supervisión del mercado, incluidas aquellas designadas en virtud del RIA y las que supervisen la aplicación de normativa armonizada. La Oficina Europea de IA prestará asistencia en estas tareas de coordinación.

Si es necesario, las autoridades de supervisión del mercado podrán acceder a los datos utilizados para el entrenamiento, validación y testeo de los sistemas a través de APIs, así como al código fuente, si fuera necesario, para evaluar la conformidad del sistema de alto riesgo con los requisitos del RIA o si otros métodos han sido insuficientes para verificar dicho cumplimiento.

Estas autoridades tendrán competencia y poderes para asegurar que las pruebas en condiciones reales se realicen de acuerdo con el RIA. Podrán otorgar o denegar permisos para realizar dichas pruebas, y, en caso de incidentes graves, exigir modificaciones al proveedor, suspender o finalizar las pruebas.

Las autoridades nacionales que supervisen o impongan obligaciones relacionadas con la protección de derechos fundamentales, cuando ejerzan sus

competencias respecto al uso de sistemas de alto riesgo, podrán acceder a la documentación generada conforme a este RIA. Será necesario notificar a la Comisión cuáles son esas autoridades. Si la documentación resulta insuficiente, podrán solicitar a la autoridad de supervisión del mercado que organice pruebas específicas.

Si una autoridad de supervisión del mercado considera que un sistema presenta riesgos, podrá evaluarlo y, si descubre que no cumple con el RIA, requerir al proveedor que tome medidas correctivas. Si el incumplimiento es transnacional, la autoridad informará a la Comisión y a otros Estados miembros. Si el proveedor no toma las acciones correctivas necesarias, la autoridad podrá tomar medidas para prohibir o restringir el sistema de alto riesgo. El RIA detalla cómo se aplican estas medidas, especialmente en casos transnacionales, y establece procedimientos para resolver desacuerdos entre Estados miembros.

Las autoridades también podrán exigir medidas correctivas o incluso retirar del mercado sistemas de alto riesgo que, aunque cumplan con los requisitos, puedan presentar riesgos o no demuestren adecuadamente su conformidad.

Por cuanto a la confidencialidad, en virtud de este capítulo, toda persona física o jurídica involucrada en la aplicación del RIA deberá garantizar la confidencialidad de los datos obtenidos como resultado de dichas actividades. La información intercambiada entre las autoridades nacionales y la Comisión no se revelará sin previa consulta con la autoridad que la originó, especialmente cuando se trate de sistemas utilizados en actividades de orden público, control fronterizo, inmigración o asilo. El intercambio de datos operativos sobre estas actividades no es obligatorio. Los proveedores tendrán la responsabilidad de custodiar la documentación de estos sistemas y de proporcionar copias a los agentes autorizados de las autoridades de supervisión.

11. Otros aspectos del Reglamento: derechos, sanciones y aplicación gradual, códigos de conducta y delegaciones a la Comisión

¿Reconoce el Reglamento derechos a los afectados por la IA?

En relación con el reconocimiento de nuevos derechos dentro del RIA, es fundamental destacar que este se enfoca principalmente en los fabricantes (proveedores) y en quienes implementan productos de IA. Sin embargo, ha recibido críticas por pasar por alto a las personas afectadas por estos sistemas de IA, quienes ni siquiera son mencionadas en el texto. [39] Respecto a la incorporación de derechos, las enmiendas 628 y 629 de junio de 2023 propusieron incluir ciertos derechos, culminando en el reconocimiento del derecho a presentar una reclamación ante una autoridad de vigilancia del mercado, así como el derecho a obtener explicaciones sobre decisiones individuales, como se indica en los artículos 85-87. Es relevante señalar que el

[39] Hasta junio de 2023 no se incluyó entre las definiciones la de “persona afectada”, junto con diversos derechos en las enmiendas del Parlamento (Enmienda 174, art. 3. 1º, 8 bis). Las “personas afectadas” finalmente no se definen pero sí que forman parte del “alcance” del artículo 2.1.g) RIA. “personas afectadas que estén establecidas en la Unión.”

texto final aprobado atenúa la propuesta inicial del Parlamento. Como menciona López-Tarruella, estas disposiciones se mantienen más en el ámbito de las buenas intenciones que en medidas concretas. [40]

Así, el derecho a presentar una reclamación, a diferencia del derecho establecido en el artículo 77 del RGPD, se articula en el RIA como un derecho a presentar una solicitud ante la autoridad de vigilancia del mercado, la cual podría tomar en consideración para acciones o investigaciones futuras. No obstante, según el autor mencionado, estas solicitudes podrían verse entorpecidas por la compleja distribución de competencias entre las distintas autoridades de vigilancia del mercado, derivada del artículo 74 del RIA. En cuanto al derecho a recibir explicaciones sobre decisiones individuales, este se restringe a ciertos sistemas de alto riesgo, y en el caso de personas físicas, la obtención de dichas explicaciones deberá seguir los procedimientos establecidos en el artículo 22 del RGPD.

Códigos de conducta: Estándares voluntarios para sistemas no regulados

Cómo se ha señalado los sistemas que no están regulados por el RIA de inteligencia artificial (esto es los que no estén prohibidos, no sean de alto riesgo, no sean de inteligencia artificial general o que tengan las especiales obligaciones de transparencia del artículo 50, no contarán con ninguna obligación por el RIA. No obstante, claro está, que tendrán que cumplir cualquier obligación de cualquier otra normativa, como pueda ser el reglamento de protección de datos, por ejemplo.

El capítulo IX sobre Códigos de conducta regula que la Oficina de la IA y los Estados Miembros impulsarán la creación de códigos de conducta destinados a fomentar la adopción voluntaria de algunos requisitos del RIA para sistemas de alto riesgo en aquellos sistemas de IA que no se consideran de alto riesgo. Además, promoverán la aplicación voluntaria de otros requisitos, como los relacionados con la sostenibilidad medioambiental o la accesibilidad. Estos códigos de conducta podrán ser elaborados por empresas, organizaciones de la sociedad civil y otros agentes relevantes. Se apuesta así por un desarrollo del mercado de las certificaciones respecto de la inteligencia artificial que no sea de alto riesgo.

Régimen sancionador: ¿cuáles son las consecuencias del incumplimiento?

Por cuanto al régimen sancionador, cabe señalar que será preciso un desarrollo y concreción por los estados miembros para que reúna los mínimos exigibles por el principio de legalidad, sancionadora y tipicidad. En todo caso, el RIA sí que establece unos parámetros básicos. Respecto de las sanciones, se ajustarán en función de las circunstancias y considerarán el tamaño del proveedor.

---

[40] El primer estudio sobre el tema López-Tarruella Martínez, A.: “Vias de recurso para los particulares en el reglamento de inteligencia artificial”, en Cotino Hueso, L. y Simón Castellanos, P. (coords.): *Tratado sobre el Reglamento … cit.*

El montante es importante, así, hasta 35 millones de euros, o el 7% del volumen anual de negocio para empresas (3% en el caso de PYMEs) por incumplir las prohibiciones del RIA. Hasta 15 millones de euros o el 3% del volumen anual de negocio (2% para PYMEs) por incumplir las obligaciones del Reglamento aplicables a proveedores, importadores, distribuidores e implementadores. Hasta 7.5 millones de euros o el 1% del volumen anual de negocio (1% para PYMEs) por suministrar información incorrecta a organismos notificados o autoridades nacionales competentes. Hasta 1.5 millones de euros para entidades de la Unión por incumplimiento de prohibiciones, o 750 mil euros por incumplimiento de obligaciones. Para los proveedores de modelos de IA de propósito general se contempla una sanción máxima de 15 millones de euros o el 3% del volumen anual de negocio.

Delegación de poderes a la Comisión para ir adaptando el Reglamento

En los próximos meses y años, la Comisión Europea desempeñará un papel esencial en el desarrollo y concreción de los criterios y normas establecidas en el RIA, ya sea de manera directa o a través de la Oficina de Inteligencia Artificial.

Uno de los mecanismos para llevar a cabo estas funciones es la delegación de poderes, regulada en el Capítulo XI del RIA. En virtud de esta delegación, la Comisión tiene la facultad de modificar la lista de tipos de sistemas que se consideran de alto riesgo en el Anexo III. Asegurar que la documentación requerida sea suficiente para garantizar la conformidad de los sistemas con el RIA, teniendo en cuenta los avances técnicos. Revisar y actualizar los Anexos IV (documentación técnica), VI (evaluación de conformidad mediante control interno) y VII (evaluación de conformidad a través del sistema de control de calidad y documentación) para adecuarlos al progreso tecnológico. Exigir que ciertos sistemas de alto riesgo, para los cuales el RIA permite la demostración de conformidad mediante control interno, sean sometidos a pruebas de conformidad por parte de una entidad notificada y, finalmente, modificar el Anexo V (contenido de la declaración de conformidad) para adaptarlo a los avances técnicos.

Disposiciones finales y la (lenta) aplicación del Reglamento

Este Capítulo final se centra en la modificación de otras disposiciones para alinearlas con el nuevo Reglamento. Además, establece un cronograma para la implementación del RIA. Así, aunque el RIA ya está en vigor tras 20 días desde su publicación, entrará en plena aplicación dos años después de su publicación en el DOUE. Las prohibiciones específicas relativas a los sistemas de IA comenzarán a regir a los seis meses de dicha publicación. A los doce meses después de la publicación, entrarán en vigor las disposiciones relativas a las autoridades notificantes, las autoridades nacionales de supervisión, los modelos de IA de propósito general, el marco de gobernanza del RIA, así como las sanciones establecidas. Finalmente, a los treinta y seis meses, se incorporará la

consideración de alto riesgo para los sistemas de IA regulados por legislación de armonización, tales como juguetes, ascensores o productos sanitarios. Debe notarse que las obligaciones sólo serán plenamente exigibles al sector público, ni más ni menos que después de seis años.

12. Para acabar, también en 2024 un Convenio de inteligencia artificial del Consejo de Europa

Se expuso al inicio el contexto actual de aceleración regulatoria mundial de la IA, y particularmente bajo la influencia del RIA de la Unión Europea. Pues en este contexto se sitúa la adopción del Convenio sobre inteligencia artificial, derechos humanos, democracia y Estado de Derecho del Consejo de Europa del Consejo de Europa,[41] finalmente aprobado el 17 de mayo de 2024 durante la 133ª Sesión del Comité de Ministros en Estrasburgo. Este Convenio, que será el tratado número 225 en la serie de tratados, [42] se abrirá a la firma el 5 de septiembre de 2024 durante la Conferencia de Ministros de Justicia en Vilnius. El documento es amplio, con diecisiete considerandos y treinta y seis artículos distribuidos en ocho capítulos. [43]

Este Convenio representa un esfuerzo considerable para armonizar y establecer un marco común mínimo en Europa, orientado hacia una visión global centrada en los derechos humanos, la democracia y el Estado de Derecho, en lugar de centrarse únicamente en la economía y el mercado. Europa aspira a posicionarse como un *faro luminoso* y una piedra angular en el ámbito normativo de la IA, no solo a nivel continental sino también global. Para los Estados que no son miembros de la UE, este Convenio de IA puede desempeñar un papel normativo y jurídico fundamental. En el caso de la UE y sus Estados miembros que ratifiquen el Convenio, este podría complementar, fortalecer y elevar la protección de los derechos. Es importante resaltar que, a diferencia del RIA, el Convenio de IA no se restringe únicamente a los sistemas de IA de alto riesgo. Además, para los Estados que ratifiquen el Convenio, este integra como "principios" en el ámbito jurídico y normativo aquello que hasta ahora eran solo principios éticos consensuados en torno a la IA. Estos principios jurídicos tienen un gran potencial en manos de los operadores legales para derivar normas y obligaciones concretas, siguiendo la tradición establecida en las últimas décadas con los principios de protección de datos. Además, el Convenio de IA otorga a las personas afectadas por la IA ciertos derechos y garantías, estableciendo mecanismos para la protección de estos derechos ante autoridades que deben

[41] Sobre el mismo, mi estudio The Council of Europe Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, forthcoming in CERIDAP Journal, nº 3/2024, acceso en https://ceridap.eu/fascicoli/ También en español en SSRN y una versión inicial en *Revista Administración & Cidadanía*, EGAP, 2024.

[42] https://rm.coe.int/1680afae3c

[43] Un Preámbulo con diecisiete Considerandos y treinta y seis artículos estructurados en ocho capítulos: Disposiciones generales (I); Obligaciones generales (II); Principios (III); Recursos (IV); Evaluación y Mitigación de Riesgos e Impactos adversos (V) y los relativos a Aplicación (VI); Mecanismo de seguimiento y cooperación (VII) y Cláusulas finales (VIII).

ser independientes. Un aspecto especialmente relevante del Convenio de IA es su enfoque basado en el riesgo, particularmente en el artículo 16, que requiere una evaluación y mitigación continua de los riesgos e impactos adversos de cualquier tipo de sistema de IA.

Por lo tanto, el Convenio de IA tiene el potencial de complementar el RIA de la UE. Más allá de esta perspectiva, el valor del Convenio va más allá. Si se me permite, *el Convenio aporta una dimensión lírica a la prosa técnica que representa el RIA*. Mientras que el RIA establece las bases y estructuras para un ecosistema de IA seguro y confiable, el Convenio se centra en su impacto en las personas y en la sociedad democrática. El RIA es meticuloso, detallado y preciso, trazando un camino claro a través de la complejidad técnica y jurídica, estableciendo estándares firmes y obligaciones concretas para los proveedores y usuarios de sistemas de IA. En contraste, el Convenio introduce una perspectiva más elevada, integrando normativamente los valores fundamentales, los principios éticos y los derechos humanos que deben guiar la evolución de la IA. El Convenio no solo tiene un valor simbólico y metajurídico, sino que es un instrumento normativo con capacidad de integración casi constitucional en los ordenamientos jurídicos de los Estados parte, y posee un gran potencial interpretativo. Por ello, el Convenio de IA supera a muchos de los instrumentos declarativos y de *soft law* que ya resultaban redundantes, ineficaces e incluso tediosos. Naturalmente, será esencial seguir de cerca su proceso final de aprobación, ratificación y la implementación de sus efectos.

Para concluir. El RIA representa un hito fundamental en la regulación de tecnologías emergentes a nivel mundial. Su enfoque basado en la gestión de riesgos establece un marco normativo que no solo busca mitigar los posibles peligros asociados con la inteligencia artificial, sino también fomentar la innovación bajo estrictas garantías. A lo largo del análisis se ha expuesto cómo el RIA, identifica la IA más peligrosa para prohibirla o bien para imponer obligaciones específicas para los sistemas de alto riesgo. En los próximos años se comprobará la capacidad de influencia de esta regulación en el mundo, su "efecto Bruselas". Este fenómeno ya se aprecia en algunos casos. También veremos si el RIA supone un lastre al desarrollo de la IA en la UE o nos hace líderes en IA ética en el diseño. Igualmente podremos apreciar en los próximos años el desarrollo del Convenio de IA del Consejo de Europa en 2024, más fácil de asumir para muchos países.