

EL REGLAMENTO EUROPEO DE INTELIGENCIA ARTIFICIAL: ¿GARANTÍA DEL MERCADO O DEFENSA DE LOS DERECHOS?¹

*The European Artificial Intelligence Act:
Market Guarantee or Defense of Rights?*

MIGUEL ÁNGEL PRESNO LINERA

Universidad de Oviedo

presno@uniovi.es

Cómo citar/Citation

Presno Linera, M. Á. (2025).

El Reglamento europeo de inteligencia artificial:
¿garantía del mercado o defensa de los derechos?
Revista Española de Derecho Constitucional, 135, 13-49.
doi: <https://doi.org/10.18042/cepc/redc.135.01>

Resumen

Una vez aprobado el Reglamento europeo de inteligencia artificial, y con parte de sus exigencias ya en vigor, resulta oportuno analizar tanto su proceso de aprobación como el resultado final de una iniciativa compleja en la que se trató de compatibilizar el establecimiento de un marco jurídico uniforme para el desarrollo, la introducción en el mercado, la puesta en servicio y la utilización de sistemas de inteligencia artificial con un elevado nivel de protección de la salud, la seguridad y los derechos consagrados en la Carta de los Derechos Fundamentales de la Unión Europea, incluidos la democracia, el Estado de derecho y la protección del medio ambiente. En las siguientes líneas se expondrá el proceso de aprobación y se estudiará de manera crítica en qué medida se ha cumplido en la práctica este ambicioso programa de garantía simultánea del mercado y de los derechos.

¹ Este es uno de los resultados del Proyecto de investigación PID2022-136548NB-I00/FEDER, UE «Los retos de la inteligencia artificial para el Estado social y democrático de Derecho», financiado por el Ministerio de Ciencia e Innovación en la Convocatoria Proyectos de Generación de Conocimiento 2022. Agradezco sus sugerencias y comentarios a quienes evaluaron la primera versión de este texto.

Palabras clave

Unión Europea; inteligencia artificial; Reglamento europeo de inteligencia artificial; mercado interior; derechos fundamentales.

Abstract

Once the AI Act has been passed and some of its requirements are already in force, we should analyze both its approval process and its final outcome. It was a complex initiative which tried to reconcile the establishment of a uniform legal framework for the development, placing on the market, the putting into service and use of artificial intelligence systems with a high level of protection of health, safety and the rights enshrined in the European Charter of Fundamental Rights, including democracy, the rule of law and the protection of the environment. The following lines will describe the approval process and critically examine the extent to which this ambitious goal of simultaneously protecting the market and fundamental rights has been achieved in practice.

Keywords

European Union; artificial intelligence; *Artificial Intelligence Act*; internal market; fundamental rights.

SUMARIO

I. PRESENTACIÓN. II. EL PROCESO DE CONSTRUCCIÓN JURÍDICA DEL MODELO EUROPEO DE GOBERNANZA DE LA INTELIGENCIA ARTIFICIAL: 1. Actuaciones previas a la tramitación del Reglamento. 2. El proceso de aprobación del Reglamento. 3. El posproceso legislativo: las revisiones periódicas y los poderes concedidos a la Comisión para aprobar actos delegados y directrices. III. EL REGLAMENTO DE INTELIGENCIA ARTIFICIAL COMO UNA SUMA DE TECNOLOGÍAS REGULADAS Y CONSTITUTIVAS. IV. EL CONCEPTO DE INTELIGENCIA ARTIFICIAL EN EL REGLAMENTO Y EN LAS DIRECTRICES DE LA COMISIÓN EUROPEA. V. UNA REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL MÁS ORIENTADA AL MERCADO QUE A LA GARANTÍA DE LOS DERECHOS FUNDAMENTALES: 1. La exclusión de determinados sistemas de inteligencia artificial de la sujeción al Reglamento. 2. La prohibición, un tanto pacata, de ciertas prácticas de inteligencia artificial. 3. Un enfoque de la gobernanza de la inteligencia artificial basado en la aceptación del riesgo industrial. 4. La confianza en normas armonizadas y especificaciones comunes todavía poco desarrolladas. 5. La evaluación de la conformidad de los sistemas de alto riesgo confiada, esencialmente, a los propios proveedores. 6. Las insuficiencias de la evaluación del impacto de los sistemas de alto riesgo en los derechos fundamentales. 7. La posibilidad de exportar sistemas de inteligencia artificial prohibidos en la Unión Europea por ser lesivos de derechos fundamentales. 8. Los derechos, más bien retóricos, para presentar una reclamación y para recibir una explicación. 9. La ausencia de unos principios obligatorios comunes a todos los sistemas de inteligencia artificial. 10. La excesiva dilación de la entrada en vigor de algunas exigencias relativas a sistemas ya introducidos en el mercado o puestos en servicio. VI. A MODO DE BREVE CONCLUSIÓN. BIBLIOGRAFÍA.

I. PRESENTACIÓN

En su conocido libro *Digital Empires. The Global Battle to Regulate Technology*, Anu Bradford (2023: 6-13) explica que, en la actualidad, son tres las potencias digitales dominantes—Estados Unidos, China y la Unión Europea—que metafóricamente pueden denominarse «imperios digitales». Estos imperios son las principales potencias tecnológicas, económicas y reguladoras y operan con la ambición y la capacidad de moldear el orden digital global hacia sus intereses y valores. Cada uno de estos territorios, prosigue Bradford, ha desarrollado un modelo de gobernanza distinto para sus economías digitales globales, coherente con sus compromisos ideológicos. Podría considerarse que estos tres modelos representan tres variedades de capitalismo digital basadas

en diferentes teorías sobre la relación entre los mercados, el Estado y los derechos individuales y colectivos: Estados Unidos ha sido pionero de un modelo impulsado, en gran medida, por el mercado; China, de un modelo dirigido por el Estado, y la Unión Europea (UE), de un modelo que, en teoría, tiene como motor los derechos fundamentales². Pero no estamos ante modelos puros: en Estados Unidos no solo impera el mercado, en China el Estado no lo controla todo y en la Unión Europea los derechos conviven y, en no pocas ocasiones, ceden ante el mercado.

En las siguientes páginas nos centraremos en el análisis de algunos de los contenidos más relevantes del modelo adoptado por la UE, que ha tenido su momento culminante el 12 de julio de 2024, cuando se publicó en el *Diario Oficial de la Unión Europea* el Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) 300/2008, (UE) 167/2013, (UE) 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; en adelante, RIA).

Ese fue el punto de llegada de un proceso de debate legislativo, institucional, social, cultural, tecnológico y económico que se remonta a varios años atrás y que partió del convencimiento, como se dice en el primer párrafo del *Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza* de la Comisión Europea, de 19 de febrero de 2020, de que la inteligencia artificial (en lo sucesivo, IA)

[...] se está desarrollando rápido. Cambiará nuestras vidas, pues mejorará la atención sanitaria (por ejemplo, incrementando la precisión de los diagnósticos y permitiendo una mejor prevención de las enfermedades), aumentará la eficiencia de la agricultura, contribuirá a la mitigación del cambio climático y a la correspondiente adaptación, mejorará la eficiencia de los sistemas de producción a través de un mantenimiento predictivo, aumentará la seguridad de los europeos y nos aportará otros muchos cambios que de momento solo podemos intuir. Al mismo tiempo, la IA conlleva una serie de riesgos potenciales, como

² Sobre el papel de la Unión Europea en una posible «Constitución del algoritmo», Balaguer Callejón (2022: 183-185); para una perspectiva metodológica, Bustos Gisbert (2024: 146-178); sobre el proceso de aprobación del Reglamento, Presno Linera y Meuwese (2024: 131-161); para una perspectiva comparada, Pérez-Ugues (2024: 129-156); para un balance desde la perspectiva del derecho administrativo, Mir Puigpelat (2024: 63-75); un enjuiciamiento muy crítico en Jakubowska *et al.* (2024).

la opacidad en la toma de decisiones, la discriminación de género o de otro tipo, la intromisión en nuestras vidas privadas o su uso con fines delictivos³.

Desde luego, es indudable que este documento es un buen ejemplo del propósito de los libros blancos de la Comisión⁴, pues contribuyó al debate, especialmente con las empresas tecnológicas, el Parlamento Europeo y el Consejo, en aras de alcanzar un acuerdo sobre la regulación del uso de los sistemas de IA.

El enorme potencial de la IA al que se refiere el *Libro Blanco* ya se había detectado con anterioridad, así como la conciencia de que había que dar respuestas jurídicas a preguntas como las que formuló el Grupo Europeo sobre Ética de la Ciencia y las Nuevas Tecnologías en su *Declaración sobre Inteligencia artificial, robótica y sistemas autónomos*, de 9 de marzo de 2018: ¿cómo podemos construir un mundo con IA y dispositivos autónomos interconectados que sea seguro y cómo podemos estimar los riesgos involucrados?, ¿quién es responsable de resultados no deseados y en qué sentido es responsable?, ¿cómo se deben rediseñar nuestras instituciones y leyes para que estén al servicio del bienestar de las personas y la sociedad, y para hacer de la sociedad un lugar seguro ante la aplicación de estas tecnologías?, ¿cómo evitar que, a través del aprendizaje automático, los datos masivos y las ciencias del comportamiento se manipulen las arquitecturas de toma de decisiones según fines comerciales o políticos? En suma, ¿cómo se puede prevenir que estas poderosas tecnologías sean utilizadas como herramientas para socavar sistemas democráticos y como mecanismos de dominación?⁵ Con el RIA se ha tratado de ofrecer algunas respuestas a estas preguntas.

II. EL PROCESO DE CONSTRUCCIÓN JURÍDICA DEL MODELO EUROPEO DE GOBERNANZA DE LA INTELIGENCIA ARTIFICIAL

1. ACTUACIONES PREVIAS A LA TRAMITACIÓN DEL REGLAMENTO

Ya se ha apuntado que el modelo europeo de gobernanza de la IA tiene, en teoría, como motor la garantía de los derechos fundamentales, pero en un contexto de mercado, y es, sobre todo, en este contexto, en el que se ha venido

³ <https://is.gd/hGVAue> (consulta: 21-4-2025).

⁴ Sobre los *Libros Blancos*, Senden (2004: 126-128).

⁵ https://www.bioeticayderecho.ub.edu/archivos/pdf/EGE_inteligencia-artificial.pdf (consulta: 21-04-2025); sobre las «nuevas leyes de la robótica», Pasquale (2024), sobre el impacto de la IA en los derechos fundamentales, Presno Linera (2022).

articulando en los últimos ocho años; así, en su reunión de 19 de octubre de 2017, el Consejo Europeo concluyó que, para construir con éxito una Europa digital, la UE necesita, en particular:

[...] concienciarse de la urgencia de hacer frente a las nuevas tendencias, lo que comprende cuestiones como la inteligencia artificial y las tecnologías de cadena de bloques, garantizando al mismo tiempo un elevado nivel de protección de los datos, así como los derechos digitales y las normas éticas. El Consejo Europeo ruega a la Comisión que, a principios de 2018, proponga un planteamiento europeo respecto de la inteligencia artificial y le pide que presente las iniciativas necesarias para reforzar las condiciones marco con el fin de que la UE pueda buscar nuevos mercados gracias a innovaciones radicales basadas en el riesgo y reafirmar el liderazgo de su industria⁶.

Transcurrido poco más de un año, el 7 de diciembre de 2018, la Comisión Europea presentó la comunicación titulada «Plan Coordinado sobre la Inteligencia Artificial», junto con el Plan Coordinado sobre el Desarrollo y Uso de la Inteligencia Artificial «Made in Europe»-2018, preparado por los Estados miembros (como parte del Grupo sobre la Digitalización de la Industria Europea y la Inteligencia Artificial), Noruega, Suiza y la Comisión⁷. Entonces se ofreció un primer concepto de IA que irá cambiando a lo largo de este proceso: «[...] el término inteligencia artificial se aplica a los sistemas que manifiestan un comportamiento inteligente, pues son capaces de analizar su entorno y pasar a la acción —con cierto grado de autonomía— con el fin de alcanzar objetivos específicos». Se apunta aquí a la idea de «cierto grado de autonomía», que será esencial para la conceptualización de los sistemas de IA.

La preocupación por la garantía de los derechos y, al tiempo, la presencia europea en un mercado tecnológico cada vez más competitivo se exteriorizó en el Consejo de la Unión Europea de 11 de febrero de 2019, donde se destacó la importancia de garantizar el pleno respeto de los derechos de los ciudadanos europeos mediante la aplicación de directrices éticas para el desarrollo y el uso de la inteligencia artificial dentro de la Unión Europea y a nivel mundial, haciendo de la ética de la inteligencia artificial una ventaja competitiva para la industria europea⁸. Entonces todavía se apelaba a la ética, y no al derecho, como herramienta idónea para garantizar las libertades ciudadanas y, al tiempo, permitir el desarrollo tecnológico europeo.

⁶ <https://is.gd/t7ZvCE>, p. 7 (consulta: 21-4-2025).

⁷ <https://is.gd/ZbBnhz> (consulta: 21-4-2025).

⁸ <https://is.gd/Xu5d5e>, p. 8 (consulta: 21-4-2025).

Poco más de un año después, el 19 de febrero de 2020, la Comisión publicó el ya citado *Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza*⁹, donde se afirmó que «la Comisión respalda un enfoque basado en la regulación y en la inversión, que tiene el doble objetivo de promover la adopción de la inteligencia artificial y de abordar los riesgos vinculados a determinados usos de esta nueva tecnología. La finalidad del presente *Libro Blanco* es formular alternativas políticas para alcanzar estos objetivos [...].».

En octubre de 2020, el Parlamento Europeo aprobó diversas resoluciones en materia de IA en el ámbito de la ética¹⁰, la responsabilidad civil¹¹ y los derechos de propiedad intelectual¹², a las que siguieron, ya en 2021, resoluciones sobre el uso de la IA en los sectores educativo, cultural y audiovisual¹³ y en materia penal¹⁴.

2. EL PROCESO DE APROBACIÓN DEL REGLAMENTO

Antes de esta última resolución ya se había publicado el documento con el que formalmente se abrió el procedimiento de aprobación del RIA: el 21 de abril de 2021 se conoció la Propuesta de Reglamento del Parlamento Europeo y del Consejo elaborada por la Comisión¹⁵, donde la IA se definió entonces (art. 3.1) como «el *software* que se desarrolla empleando una o varias de las técnicas y estrategias que figuran en el anexo I y que puede, para un conjunto

⁹ <https://is.gd/hGVAue> (consulta: 21-4-2025).

¹⁰ Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas, 2020/2012(INL) (21-4-2025).

¹¹ Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre un régimen de responsabilidad civil en materia de inteligencia artificial, 2020/2014(INL) (consulta: 21-4-2025).

¹² Resolución del Parlamento Europeo, de 20 de octubre de 2020, sobre los derechos de propiedad intelectual para el desarrollo de las tecnologías relativas a la inteligencia artificial, 2020/2015(INI) (consulta: 21-4-2025).

¹³ Resolución del Parlamento Europeo, de 19 de mayo de 2021, sobre la inteligencia artificial en la educación, la cultura y el sector audiovisual, <https://is.gd/NJfvLw> (consulta: 21-4-2025).

¹⁴ Resolución del Parlamento Europeo, de 6 de octubre de 2021, sobre la inteligencia artificial en el Derecho penal y su utilización por las autoridades policiales y judiciales en materia penal, <https://is.gd/NJfvLw> (consulta: 21-4-2025).

¹⁵ <https://is.gd/MzpXrO> (consulta: 21-4-2025).

determinado de objetivos definidos por seres humanos, generar información de salida como contenidos, predicciones, recomendaciones o decisiones que influyan en los entornos con los que interactúa».

A propósito de esta propuesta, que remitía la definición de IA a un anexo y que no mencionaba el aprendizaje automático, el 6 de diciembre de 2022, el Consejo de la Unión Europea hizo pública su Orientación general de 25 de noviembre¹⁶, donde señaló que «para garantizar que la definición de los sistemas de IA proporcione criterios suficientemente claros para distinguirlos de otros sistemas de *software* más clásicos, el texto transaccional restringe la definición del artículo 3, apartado 1, a los sistemas desarrollados a través de estrategias de aprendizaje automático y estrategias basadas en la lógica y el conocimiento», es decir, se introduce entonces el criterio del aprendizaje automático como una de las características de los sistemas de IA; también se amplían los sistemas que se pretende prohibir y se introducen cambios en los considerados de «alto riesgo».

A continuación, cabe mencionar las enmiendas formuladas por el Parlamento Europeo al texto de la Comisión y aprobadas el 14 de junio de 2023¹⁷, que incorporaron una nueva definición de sistema de IA («un sistema basado en máquinas diseñado para funcionar con diversos niveles de autonomía y capaz, para objetivos explícitos o implícitos, de generar información de salida —como predicciones, recomendaciones o decisiones— que influya en entornos reales o virtuales»), y también de lo que se entendió por un modelo fundacional —«un modelo de sistema de IA entrenado con un gran volumen de datos, diseñado para producir información de salida de carácter general y capaz de adaptarse a una amplia variedad de tareas diferentes»—, al tiempo que, entre otras cosas, se ampliaron los sistemas de IA prohibidos; en suma, propuestas que iban bastante más allá de lo planteado por la Comisión.

En diciembre de 2023 tuvieron lugar los trílogos entre el Parlamento, la Comisión y el Consejo¹⁸ para limar las diferencias en cuestiones tan relevantes como la prohibición, o no, del uso de los sistemas de identificación biométrica remota en tiempo real en espacios de acceso público o el alcance de la regulación de los entonces denominados «modelos fundacionales», que pasaron a llamarse modelos de IA de uso general. En lo que a las prohibiciones se refiere, el resultado fue rebajar las exigencias del Parlamento en materia de garantía de los derechos y, parece, en aras de «mejorar el funcionamiento del mercado interior», pero también es verdad que este punto de llegada es más garantista

¹⁶ <https://is.gd/OHMIf3> (consulta: 21-4-2025).

¹⁷ <https://is.gd/RPOMnU> (consulta: 21-4-2025).

¹⁸ <https://is.gd/HEtdqH> (fecha de consulta: 21-4-2025)

en algunos puntos (la supervisión humana de los sistemas, la evaluación de impacto relativa a los derechos fundamentales para los sistemas de alto riesgo, la inclusión de los modelos de uso general...) que el punto de partida presentado en su día por la Comisión.

Finalmente, se aprobó la Resolución legislativa del Parlamento Europeo, de 13 de marzo de 2024, sobre la propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión, y, como ya se ha dicho, se publicó en el *Diario Oficial de la UE* el 12 de julio de 2024¹⁹.

3. EL POSPROCESO LEGISLATIVO: LAS REVISIONES PERIÓDICAS Y LOS PODERES CONCEDIDOS A LA COMISIÓN PARA APROBAR ACTOS DELEGADOS Y DIRECTRICES

Hemos visto que antes del inicio en sentido estricto del procedimiento de aprobación del RIA la Comisión Europea ya tuvo un papel relevante a través, sobre todo, del *Libro Blanco*; ese protagonismo lo va a tener también, una vez aprobado el Reglamento, en lo que Linda Senden (2004: 120) denomina «posproceso legislativo europeo», que, en parte, se llevará a cabo a través de actos de *soft law* como las directrices a las que remite expresamente el RIA, aunque también por medio de revisiones del propio Reglamento y con la aprobación de actos delegados previstos de manera explícita en él.

No cabe duda de que el legislador europeo es plenamente consciente de la necesidad de adaptación que va a precisar una norma que regula los sistemas de IA y por ello ya prevé una serie de evaluaciones y revisiones periódicas «habida cuenta de los rápidos avances tecnológicos y de los conocimientos técnicos necesarios para la aplicación efectiva del presente Reglamento» (considerando 174), y establece (art. 96) que la Comisión elaborará una serie de directrices —criterios interpretativos y ejemplos— para facilitar la aplicación práctica del RIA y, en particular, sobre los requisitos para cumplir por los sistemas de alto riesgo, las prácticas prohibidas y la definición de los sistemas de IA del art. 3.1. A este respecto, cabe mencionar la publicación de las directrices sobre prácticas prohibidas el 4 de febrero de 2025 y las relativas a la

¹⁹ Pueden verse los comentarios al RIA coordinados y/o editados por Cotino Hueso y Simón Castellanos (2024), Barrio Andrés (2024) y Huergo Lora y Díaz González (2025); sobre el proceso de aprobación del Reglamento, Presno Linera y Meuwese (2024: 131-161).

definición de sistemas de IA dos días después. Como es sabido, dichas directrices no son vinculantes y las interpretaciones autorizadas del RIA se reservan al Tribunal de Justicia de la Unión Europea²⁰.

A la misma Comisión se le otorgan (art. 97) los poderes para aprobar actos delegados —actos no legislativos de alcance general adoptados para completar o modificar, teniendo en cuenta el progreso científico o técnico, determinados elementos no esenciales de un acto legislativo— mencionados previamente en numerosos artículos.

Finalmente, y en el penúltimo precepto (el 112) del RIA, se articula un procedimiento de evaluación y revisión que va desde el año a partir de la entrada en vigor para la lista del anexo III y la lista de prácticas prohibidas del art. 5 a la exigencia de que, antes del 2 de agosto de 2028 y de la misma fecha de 2029, y posteriormente cada cuatro años, se revisen, en el primer caso, la necesidad de ampliar los ámbitos enumerados en el anexo III o de añadir nuevos ámbitos, la necesidad de modificar la lista de sistemas de IA que requieren medidas de transparencia adicionales con arreglo al art. 50, y la necesidad de mejorar la eficacia del sistema de supervisión y gobernanza; en el segundo caso (a más tardar, el 2 de agosto de 2029), la Comisión presentará al Parlamento Europeo y al Consejo un informe sobre la evaluación y revisión del Reglamento. En función de sus conclusiones, dicho informe irá acompañado, en su caso, de una propuesta de modificación del presente Reglamento. A más tardar, el 2 de agosto de 2031, la Comisión evaluará la ejecución del Reglamento e informará al respecto al Parlamento Europeo, al Consejo y al Comité Económico y Social Europeo, teniendo en cuenta sus primeros años de aplicación. En función de sus conclusiones, dicho informe irá acompañado, en su caso, de una propuesta de modificación. Este proceso de evaluación y revisión del RIA expresamente previsto parece bien justificado y mantendrá el papel colegislador del Parlamento y el Consejo.

En cuanto a los ya citados actos delegados, previstos para determinadas exigencias de los sistemas de alto riesgo y de los modelos de uso general, el art. 97 incluye mecanismos expresos de control, como la posibilidad de formular objeciones y, en su caso, revocar la delegación por parte del Parlamento y el Consejo. No obstante estas cautelas, es llamativo, y un tanto

²⁰ Para la elaboración de las Directrices la Oficina de Inteligencia Artificial de la Comisión puso en marcha un proceso de consulta invitando a las partes interesadas, incluidos los proveedores de sistemas de IA, las empresas, las autoridades nacionales, el mundo académico, las instituciones de investigación y la sociedad civil, a presentar sus aportaciones durante un período de cuatro semanas que concluyó el 11 de diciembre de 2024.

preocupante, el poder que se concede a la Comisión para algo tan relevante como modificar la lista del anexo III, que es el relativo a los sistemas de alto riesgo del art. 6.2.

Por lo que respecta a las directrices, no cabe obviar que, como ya se ha dicho más arriba, están previstas no solo para todos los requisitos de los sistemas de alto riesgo, sino también para la delimitación de las responsabilidades a lo largo de la cadena de valor de la IA (art. 25) y, lo que quizás sea más importante, tanto para las prácticas prohibidas en el art. 5 como para la aplicación de la definición de sistema de IA del art. 3.1.m, cuya pronta aprobación hemos mencionado con anterioridad.

Que el recurso a este instrumento se haya realizado de manera expresa por el legislador europeo parece que deja a las directrices de la Comisión al resguardo de eventuales cuestionamientos de los Estados miembros y de las otras instituciones de la UE, singularmente del Parlamento, que en ocasiones anteriores se ha quejado de que los instrumentos de *soft law* europeos vengan siendo utilizados como alternativas a la legislación, con lo que, como señala Fajardo del Castillo (2024: 203-204), no solo se estaría menoscabando el papel asignado a aquella institución en el procedimiento legislativo, sino que también se estarían eludiendo las garantías democráticas previstas en los tratados para su desarrollo. El recurso al *soft law* como alternativa al legislador también pondría en cuestión la justificación democrática o la legitimidad de ese instrumento (Peters, 2011: 39); piénsese que uno de los apartados del RIA que mayor debate suscitó y en el que se apreciaron las diferencias de posición del Parlamento respecto al proyecto presentado por la Comisión fue, precisamente, el de las prácticas prohibidas.

III. EL REGLAMENTO DE INTELIGENCIA ARTIFICIAL COMO UNA SUMA DE TECNOLOGÍAS REGULADAS Y CONSTITUTIVAS

En el considerando inicial del RIA se proclama:

El objetivo del presente Reglamento es mejorar el funcionamiento del mercado interior mediante el establecimiento de un marco jurídico uniforme, en particular para el desarrollo, la introducción en el mercado, la puesta en servicio y la utilización de sistemas de inteligencia artificial [...] a fin de promover la adopción de una inteligencia artificial centrada en el ser humano y fiable, 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, incluidos la democracia, el Estado

de Derecho y la protección del medio ambiente, proteger frente a los efectos perjudiciales de los sistemas de IA en la Unión, así como brindar apoyo a la innovación. El presente Reglamento garantiza la libre circulación transfronteriza de mercancías y servicios basados en la IA, con lo que impide que los Estados miembros impongan restricciones al desarrollo, la comercialización y la utilización de sistemas de IA, a menos que el presente Reglamento lo autorice expresamente.

Así pues, el RIA cumpliría, en primer lugar, y tal y como se proclama en las anteriores palabras, una función reguladora de la tecnología —entendida como instrumento técnico dotado de un componente material (Hildebrandt, 2015: 59)— de los sistemas de IA y lo haría al servicio de los objetivos que ahí se mencionan, vinculados, en pocas palabras, al desarrollo del mercado y la libre circulación de mercancías y servicios y a la garantía de los derechos fundamentales. En este sentido, este Reglamento no ofrecería novedades estructurales respecto, por ejemplo, al Reglamento (UE) 2016/679 del Parlamento y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (Reglamento General de protección de datos, RGPD).

Esta función reguladora de los sistemas de IA se concretaría, por mencionar algunas partes relevantes del RIA, en el ámbito de aplicación (art. 2), en las prácticas prohibidas (art. 5) o en los requisitos que se imponen a los sistemas de alto riesgo (Sección 2 del capítulo III). En estos preceptos, la IA se convierte en el objeto de la regulación y ello implica la introducción en un texto jurídico de conceptos y expresiones propios, en este caso, de la ciencia de la computación y de la misma IA, algo que, como es bien sabido, ya viene sucediendo con las normas que disciplinan ámbitos como la protección ambiental o alimentaria, lo que puede implicar, como es obvio, una mayor dificultad de comprensión de las normas, a pesar de que incluyan una suerte de glosario o apartado de definiciones, como hace el propio RIA en el art. 3. Aquí la tecnología ocupa una posición «exterior» al derecho (San Martín Segura, 2023a: 207) en cuanto objeto de este al que se somete a una regulación imperativa.

Pero, en segundo lugar, el RIA, además, incluye numerosas normas en las que la tecnología, la IA, ocupa una posición «interior» al derecho (San Martín Segura, 2023a: 209), en el sentido de que los sistemas de IA o bien actúan como un apoyo o suplemento a su aplicación, pero sin determinarla, o bien se insertan como auténticos decisores jurídicos. Ambas funcionalidades cuentan con conocidos antecedentes en el derecho público y, a este respecto,

en España podemos mencionar, como ejemplo de la primera, el Registro electrónico general previsto en el art. 16²¹ de la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas; como muestra de la segunda, la actuación administrativa automatizada contemplada en el art. 41 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público²².

Y en el RIA encontramos ejemplo de la funcionalidad de apoyo en el apdo. 8.a del anexo III cuando menciona los «sistemas de IA destinados a ser utilizados por una autoridad judicial, o en su nombre, para ayudar a una autoridad judicial en la investigación e interpretación de hechos y de la ley, así como en la garantía del cumplimiento del Derecho a un conjunto concreto de hechos, o a ser utilizados de forma similar en una resolución alternativa de litigio»; una muestra de la funcionalidad decisoria aparece en el apdo. 4.b de ese mismo anexo: «[...] sistemas de IA destinados a ser utilizados para tomar decisiones que afecten a las condiciones de las relaciones de índole laboral o a la promoción o rescisión de relaciones contractuales de índole laboral, para la asignación de tareas a partir de comportamientos individuales o rasgos o características personales o para supervisar y evaluar el rendimiento y el comportamiento de las personas en el marco de dichas relaciones».

Parece mantenerse una diferencia sustancial entre la función de apoyo o ayuda y la función decisoria de estos sistemas que, en todo caso, al ser sistemas

²¹ En su apdo. 3 dispone que el registro electrónico de cada Administración u organismo garantizará la constancia, en cada asiento que se practique, de un número, epígrafe expresivo de su naturaleza, fecha y hora de su presentación, identificación del interesado, órgano administrativo remitente, si procede, y persona u órgano administrativo al que se envía, y, en su caso, referencia al contenido del documento que se registra. Para ello, se emitirá automáticamente un recibo consistente en una copia autenticada del documento de que se trate, incluyendo la fecha y hora de presentación y el número de entrada de registro, así como un recibo acreditativo de otros documentos que, en su caso, lo acompañen, para garantizar la integridad y el no repudio de estos.

²² «1. Se entiende por actuación administrativa automatizada cualquier acto o actuación realizada íntegramente a través de medios electrónicos por una Administración Pública en el marco de un procedimiento administrativo y en la que no haya intervenido de forma directa un empleado público.
2. En caso de actuación administrativa automatizada deberá establecerse previamente el órgano u órganos competentes, según los casos, para la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente. Asimismo, se indicará el órgano que debe ser considerado responsable a efectos de impugnación».

de alto riesgo, estarán sometidos a supervisión humana por mandato del art. 14 RIA para «prevenir o reducir al mínimo los riesgos para la salud, la seguridad o los derechos fundamentales que pueden surgir cuando se utiliza un sistema de IA de alto riesgo conforme a su finalidad prevista o cuando se le da un uso indebido razonablemente previsible»²³.

No obstante, la diferencia entre ayuda y decisión y la propia eficacia de la supervisión humana resultan cuando menos discutibles si tenemos en cuenta la constatada existencia de un sesgo de automatización, que admite el propio RIA cuando requiere que las personas físicas a quienes se encomienda la supervisión humana deban ser conscientes de la posible tendencia a confiar automáticamente o en exceso en los resultados de salida generados por un sistema de IA de alto riesgo («sesgo de automatización»)²⁴, en particular con aquellos

²³ El recurso a la supervisión humana como garantía de control no es ni mucho menos algo novedoso: explica Leta Jones que, desde hace más de cincuenta años, con las primeras regulaciones frente al desarrollo informático, se han impuesto restricciones al tratamiento totalmente automatizado para las tecnologías computacionales, a través de la figura del *human in the loop* y de otros mecanismos de gobernanza análogos (2017: 220 y ss.); sobre la experiencia que puede aportar a estos efectos el «control humano significativo» previsto en el derecho internacional humanitario, Obregón Fernández y Lazcoz Moratinos (2021: 8-20); más en general, sobre los cuatro modelos de interacción en la relación entre humanos y máquinas dentro del proceso de toma de decisiones (*Human in the Loop*, *Human in the Loop for Exceptions*, *Human on the Loop* y *Human out the Loop*), véase la detallada exposición de Llano Alonso (2024: 135-155); sobre la supervisión humana en el RIA, Lazcoz Moratinos (2024: 681-699).

El apdo. 4 del art. 14 dispone que «el sistema de IA de alto riesgo se ofrecerá al responsable del despliegue de tal modo que las personas físicas a quienes se encomienda la supervisión humana puedan, según proceda y de manera proporcionada: a) entender adecuadamente las capacidades y limitaciones pertinentes del sistema de IA de alto riesgo y poder vigilar debidamente su funcionamiento, por ejemplo, con vistas a detectar y resolver anomalías, problemas de funcionamiento y comportamientos inesperados; b) [...]; c) interpretar correctamente los resultados de salida del sistema de IA de alto riesgo, teniendo en cuenta, por ejemplo, los métodos y herramientas de interpretación disponibles; d) decidir, en cualquier situación concreta, no utilizar el sistema de IA de alto riesgo o descartar, invalidar o revertir los resultados de salida que este genere; e) intervenir en el funcionamiento del sistema de IA de alto riesgo o interrumpir el sistema pulsando un botón de parada o mediante un procedimiento similar que permita que el sistema se detenga de forma segura».

²⁴ Una posible muestra de este sesgo en un sistema que ni siquiera es de IA la encontramos en *VioGén*, si, como se ha venido sosteniendo (González Álvarez *et al.*, 2018: 37), hasta en el 95 % de los casos se mantiene la puntuación de riesgo asignada

sistemas que se utilizan para aportar información o recomendaciones con el fin de que personas físicas adopten una decisión²⁵.

En este sentido, Ben Green (2022: 105681), al estudiar las políticas de supervisión humana en sus interacciones con los algoritmos, encontró que estas políticas adolecen de dos fallas significativas: primera, las evidencias sugieren que las personas no pueden proporcionar una adecuada supervisión; segunda, la suposición incorrecta de una supervisión humana efectiva legitima el uso de algoritmos defectuosos e irresponsables.

IV. EL CONCEPTO DE INTELIGENCIA ARTIFICIAL EN EL REGLAMENTO Y EN LAS DIRETRICES DE LA COMISIÓN EUROPEA

Como recuerda Asunción Gómez-Pérez (2023: 20), a finales de agosto de 1955, John McCarthy (del Dartmouth College), Marvin Minsky (de la Universidad de Harvard), Nathaniel Rochester (de IBM) y Claude Shannon (de los laboratorios Bell) propusieron que diez investigadores llevaran a cabo, durante un período de dos meses en el verano de 1956, un estudio sobre inteligencia artificial en el Dartmouth College. Plantearon demostrar la siguiente conjectura: «Cualquier aspecto del aprendizaje o cualquier otra característica de la inteligencia puede, en principio, estar tan precisamente descrita que se puede construir una máquina para simularla. Se intentará descubrir cómo hacer que las máquinas utilicen el lenguaje, formen abstracciones y conceptos, resuelvan

automáticamente por el algoritmo, a pesar de que las autoridades policiales pueden estimar que existe un riesgo superior al que predice el sistema; más ampliamente, Presno Linera (2023), San Martín Segura (2023b) y, especialmente, Martínez Garay (2024: 25-72).

²⁵ Además, el art. 14.4 exige que «el sistema de IA de alto riesgo se ofrecerá al responsable del despliegue de tal modo que las personas físicas a quienes se encomienda la supervisión humana puedan, según proceda y de manera proporcionada: a) entender adecuadamente las capacidades y limitaciones pertinentes del sistema de IA de alto riesgo y poder vigilar debidamente su funcionamiento, por ejemplo, con vistas a detectar y resolver anomalías, problemas de funcionamiento y comportamientos inesperados; b) [...]; c) interpretar correctamente los resultados de salida del sistema de IA de alto riesgo, teniendo en cuenta, por ejemplo, los métodos y herramientas de interpretación disponibles; d) decidir, en cualquier situación concreta, no utilizar el sistema de IA de alto riesgo o descartar, invalidar o revertir los resultados de salida que este genere; e) intervenir en el funcionamiento del sistema de IA de alto riesgo o interrumpir el sistema pulsando un botón de parada o mediante un procedimiento similar que permita que el sistema se detenga de forma segura».

tipos de problemas hasta ahora reservados a los humanos y se mejoren a sí mismas»²⁶.

A partir de este momento fundacional de la IA se han proporcionado numerosas definiciones —se han contabilizado hasta 55— desde las ópticas política, de investigación e industrial (Samoili *et al.*, 2020: 11), que, en general, aluden al desarrollo de sistemas que imitan o reproducen el pensamiento y obrar humanos, actuando racionalmente —en el sentido de hacer lo correcto en función de su conocimiento— e interactuando con el medio. La IA pretende sintetizar o reproducir los procesos cognitivos humanos, tales como la percepción, la creatividad, la comprensión, el lenguaje o el aprendizaje (Russel y Norvig, 2008: 1 y ss.). Para ello, la IA utiliza todas las herramientas a su alcance, entre las que destacan las proporcionadas por la computación, incluidos los algoritmos. No obstante, los sistemas de IA no usan cualquier algoritmo, sino, esencialmente, los que aprenden a base del procesamiento de datos.

La dificultad de ofrecer una definición acabada de la IA se presenta también en el ámbito jurídico, como se puede comprobar leyendo las diferentes versiones que se han ido ofreciendo durante el proceso de aprobación del RIA, en cuyo texto vigente se ha conceptualizado como un sistema basado en máquinas y diseñado para funcionar con distintos niveles de autonomía, que puede mostrar capacidad de adaptación tras su despliegue y que, para objetivos explícitos o implícitos, infiere, a partir de las entradas que recibe, salidas tales como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales (art. 3). En esta definición cabe descomponer varios elementos, tal y como hacen las Directrices de la Comisión sobre la definición de sistemas de inteligencia artificial, de 6 de febrero de 2025, donde se habla de (1) un sistema basado en máquinas, (2) que esté diseñado para operar con diferentes niveles de autonomía, (3) que pueden exhibir adaptabilidad después del despliegue; (4) que, para objetivos explícitos o implícitos, (5) infiere, a partir de la entrada que recibe, cómo generar salidas (6) como predicciones, contenido, recomendaciones o decisiones (7) que pueden influir en entornos físicos o virtuales.

No todos estos elementos son igualmente importantes ni tienen que estar de forma necesaria presentes en todas las fases para definir un sistema como de IA; destacaremos dos que sí lo son: la inferencia y la autonomía, que, aunque están muy relacionadas, pueden diferenciarse.

²⁶ McCarthy *et al.* (1955). Pueden verse al respecto el libro de Mitchell (2024: 27 y ss.) y el discurso de ingreso de la profesora Asunción Gómez-Pérez en la Real Academia Española con el título «Inteligencia artificial y lengua española», disponible en: <https://is.gd/aY3kWG> (consulta: 21-4-2025).

Respecto de la primera, en el considerando 12 del RIA se puede leer que «[e]sta capacidad de inferir se refiere al proceso de obtención de los resultados, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en los entornos físicos y virtuales, y a la capacidad de los sistemas de IA para derivar modelos o algoritmos, o ambos, a partir de entradas o datos».

En palabras de las Directrices, la primera categoría de técnicas de IA mencionada en ese considerando es la de los «enfoques de aprendizaje automático que aprenden de los datos cómo alcanzar determinados objetivos». Esta categoría incluye una gran variedad de enfoques que permiten a un sistema «aprender» y que, como se especifica en las Directrices, son el aprendizaje supervisado²⁷, el no supervisado²⁸, el autosupervisado²⁹ y el aprendizaje por refuerzo³⁰. La segunda categoría de técnicas mencionadas en el considerando

²⁷ El sistema de IA aprende de las anotaciones (datos etiquetados), por lo que los datos de entrada se emparejan con la salida correcta. El sistema utiliza esas anotaciones para aprender una asignación de entradas a salidas y luego la generaliza a nuevos datos invisibles. Ejemplos de sistemas de IA basados en el aprendizaje supervisado son los sistemas de clasificación de imágenes entrenados con un conjunto de datos de imágenes, en los que cada imagen se etiqueta (por ejemplo, objetos como coches), los sistemas de diagnóstico de dispositivos médicos entrenados con imágenes médicas etiquetadas por expertos humanos y los sistemas de detección de fraudes que se entrena con datos de transacciones etiquetados.

²⁸ En este caso, el modelo se entrena con datos sin etiquetas ni salidas predefinidas. Mediante el uso de diferentes técnicas el sistema se entrena para encontrar patrones, estructuras o relaciones en los datos sin una orientación explícita sobre cuál debería ser el resultado. Así, por ejemplo, los sistemas de IA utilizan el aprendizaje no supervisado para agrupar compuestos químicos y predecir posibles nuevos tratamientos para enfermedades en función de sus similitudes con los fármacos existentes.

²⁹ En el aprendizaje autosupervisado el sistema de IA aprende utilizando los datos para crear sus propias etiquetas u objetivos. Un sistema de reconocimiento de imágenes que aprende a reconocer objetos mediante la predicción de los píxeles que faltan en una imagen es un ejemplo de un sistema de IA basado en el aprendizaje autosupervisado.

³⁰ Estos sistemas aprenden de los datos recopilados de su propia experiencia a través de una función de «recompensa». A diferencia de los sistemas de IA que aprenden de datos etiquetados (aprendizaje supervisado) o que aprenden de patrones (aprendizaje no supervisado), los sistemas de IA basados en el aprendizaje por refuerzo aprenden de la experiencia. Al sistema no se le dan etiquetas explícitas, sino que aprende por ensayo y error, refinando su estrategia en función de la retroalimentación que recibe del entorno. Un brazo robótico habilitado para IA que puede realizar tareas como agarrar objetos es un ejemplo de un sistema de IA basado en el aprendizaje por refuerzo. El aprendizaje por refuerzo también se puede utilizar, por ejemplo, para

12 del RIA son «enfoques basados en la lógica y el conocimiento que infieren a partir del conocimiento codificado o la representación simbólica de la tarea que se va a resolver»³¹.

Pues bien, es esta capacidad de inferencia de un sistema de IA (es decir, su capacidad para generar resultados como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales) la clave para lograr su autonomía, lo que, como se explica en las Directrices del 6 de febrero, consiste en que los sistemas de IA están diseñados para funcionar con «cierto grado de independencia de las acciones respecto a la intervención humana y de capacidades para funcionar sin intervención humana».

La referencia a «cierto grado de independencia de acción» en el considerando 12 del RIA excluye los sistemas diseñados para funcionar únicamente con la plena participación e intervención humana manual, que puede ser directa —por ejemplo, a través de controles manuales— o indirecta —por ejemplo, a través de controles automatizados basados en sistemas que permiten a los humanos delegar o supervisar las operaciones del sistema—. En suma, todos los sistemas que están diseñados para funcionar con cierto grado de independencia de las acciones humanas cumplen la condición de autonomía en la definición de un sistema de IA.

Las consecuencias jurídicas de esta conceptualización de los sistemas de IA son obvias: los sistemas que carezcan de inferencia y autonomía no serán IA, al menos a los efectos del RIA, y, por tanto, no estarán sujetos a sus exigencias, lo que supone que quedan al margen de las mismas importantes herramientas predictivas que vienen siendo utilizadas con frecuencia por los poderes públicos españoles y que afectan a derechos fundamentales de las personas; por volver a mencionar una bien conocida, *VioGén*, que se puso en marcha ya en julio de 2007 por el Ministerio del Interior y consiste en formularios informatizados para practicar y administrar las evaluaciones de riesgo de la mujer que denuncie un caso de violencia de género, así como las

optimizar las recomendaciones de contenido personalizadas en los motores de búsqueda y el rendimiento de los vehículos autónomos.

³¹ En lugar de aprender de los datos, aprenden del conocimiento, incluidas las reglas, los hechos y las relaciones, codificados por expertos humanos; sobre esa base pueden «razonar» a través de motores deductivos o inductivos o utilizando operaciones como la clasificación, la búsqueda, la coincidencia y el encadenamiento. Al utilizar la inferencia lógica para sacar conclusiones, estos sistemas aplican la lógica formal, reglas predefinidas u ontologías a nuevas situaciones. Por ejemplo, los sistemas expertos de primera generación destinados al diagnóstico médico, que se desarrollan codificando el conocimiento de una serie de expertos médicos y que están destinados a extraer conclusiones de un conjunto de síntomas de un paciente determinado.

funcionalidades precisas para llevar a cabo el seguimiento de dichos casos y la implementación de las medidas de seguridad y protección policial acordes con los niveles de riesgo resultantes³².

VioGén es, en suma, un ejemplo de sistema predictivo que no encaja en el concepto de sistema de IA previsto en el RIA y «explicado», quizá no con la suficiente claridad, en el Considerando 12 del RIA y en las Directrices, que excluyen los sistemas que se basan en las reglas definidas únicamente por las personas físicas para ejecutar operaciones y no generan ningún tipo de aprendizaje a partir de ese proceso; así, *VioGén*, por más que haya tenido algunos ajustes, ha venido funcionando esencialmente igual durante 18 años, a partir de las reglas introducidas por las personas que lo han diseñado, sin que los más de 700 000 casos en los que se ha aplicado le hayan permitido aprender de ese enorme banco de datos de circunstancias que deberían tenerse en cuenta a la hora de valorar el riesgo de reincidencia o formular nuevas categorizaciones de riesgo que demandaran medidas distintas a las hasta ahora previstas³³.

V. UNA REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL MÁS ORIENTADA AL MERCADO QUE A LA GARANTÍA DE LOS DERECHOS FUNDAMENTALES

Comenzábamos estas páginas aludiendo a los tres modelos de capitalismo digital basados en diferentes teorías sobre la relación entre los mercados, el Estado y los derechos individuales y colectivos y ahí se partía de que la Unión Europea habría elegido un modelo que, en teoría, tiene como motor los derechos fundamentales y no, como en Estados Unidos, el mercado. También se matizaba que los modelos no son puros y que no todo lo decide el mercado en Estados Unidos ni todo en Europa está orientado a la garantía de los derechos.

En las líneas siguientes analizaremos la atención a los derechos y al mercado en las previsiones contenidas en el RIA y lo primero que cabe destacar es que, si bien encontramos 96 referencias a los derechos fundamentales, ya

³² Sobre *VioGén*, Presno Linera (2023), San Martín Segura (2023b) y, especialmente, Martínez Garay (2024: 25-72).

³³ Quiero agradecer al profesor David San Martín Segura el análisis crítico de la definición de sistemas de IA incluida en el RIA y desarrollada en las Directrices, así como las aclaraciones, a partir del sistema *Bosco* (usado en España para resolver las solicitudes del bono eléctrico), sobre la exclusión del concepto de IA de los sistemas basados en las normas definidas únicamente por personas físicas para ejecutar automáticamente operaciones.

desde el primer párrafo del preámbulo se deja claro que «el objetivo del presente Reglamento es mejorar el funcionamiento del mercado interior mediante el establecimiento de un marco jurídico uniforme, en particular para el desarrollo, la introducción en el mercado, la puesta en servicio y la utilización de sistemas de inteligencia artificial en la Unión», es decir, el Reglamento está orientado al *mercado*, palabra que aparece en 356 ocasiones, y a la seguridad de los productos de IA, lo que ocurre es que dicha orientación se pretende que sea compatible con «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, incluidos la democracia, el Estado de Derecho y la protección del medio ambiente».

Además, y como es sabido, el RIA no es una suerte de declaración iusfundamental en materia de IA y el reconocimiento de los derechos a presentar una reclamación y a demandar una explicación (arts. 85 y 86) es una excepción introducida por el Parlamento Europeo que no aparecía en la propuesta inicial de la Comisión Europea ni en la Posición Común del Consejo y que, como se dirá más adelante, tiene un alcance más retórico que práctico (López-Tarruella Martínez, 2024: 893-918).

1. LA EXCLUSIÓN DE DETERMINADOS SISTEMAS DE INTELIGENCIA ARTIFICIAL DE LA SUJECIÓN AL REGLAMENTO

Acabamos de ver que todo lo que no responda al concepto de sistema de IA queda fuera del ámbito de aplicación del RIA, pero, además, también quedan excluidos sistemas que sí son IA, pero que el legislador europeo ha querido eximir de las obligaciones propias de estos. Así, y conforme al art. 2.3, no estarán sujetos al RIA los sistemas en la medida en que se introduzcan en el mercado, se pongan en servicio o se utilicen, con o sin modificaciones, exclusivamente con fines militares, de defensa o de seguridad nacional, independientemente del tipo de entidad que lleve a cabo estas actividades.

En el considerando 24 se trata de justificar estas exclusiones argumentando que, por lo que respecta a los fines militares y de defensa, dicha exclusión está justificada tanto por el art. 4.2 del Tratado de Funcionamiento de la Unión Europea (TUE) como por las especificidades de la política de defensa de los Estados miembros y de la política común de defensa de la Unión, que están sujetas al derecho internacional público, que, por lo tanto, es el marco jurídico más adecuado para la regulación de los sistemas de IA en el contexto del uso de la fuerza letal y de otros sistemas de IA en el contexto de las actividades militares y de defensa.

En lo que hace referencia a los fines de seguridad nacional, la exclusión estaría justificada tanto por el hecho de que sigue siendo responsabilidad exclusiva de los Estados miembros de conformidad con el art. 4.2 del TUE como por la naturaleza específica y las necesidades operativas de las actividades de seguridad nacional y por las normas nacionales específicas aplicables a dichas actividades. Como excepción, si estos sistemas excluidos se utilizaran temporal o permanentemente fuera de estos ámbitos con otros fines (por ejemplo, con fines civiles o humanitarios, de garantía del cumplimiento del derecho o de seguridad pública), entrarían en el ámbito de aplicación del RIA.

Siguiendo con las exclusiones, de acuerdo con los apdos. 6 y 8 del mismo art. 2, no se aplicará a los sistemas o modelos de IA, incluidos sus resultados de salida, desarrollados y puestos en servicio específicamente con la investigación y el desarrollo científicos como única finalidad; tampoco se aplicará a ninguna actividad de investigación, prueba o desarrollo relativa a sistemas de IA o modelos de IA antes de su introducción en el mercado o puesta en servicio. Las pruebas en condiciones reales no estarán cubiertas por esa exclusión.

El considerando 25 fundamenta esta exclusión en la pretensión de «apoyar la innovación, respetar la libertad de ciencia y no socavar la actividad de investigación y desarrollo», objetivos sin duda loables, pero que liberan de las exigencias del RIA a actividades en las que no es descartable imaginar un riesgo («la combinación de la probabilidad de que se produzca un perjuicio y la gravedad de dicho perjuicio» en palabras del art. 3.2 del RIA) cierto de lesión de derechos fundamentales cuya eventual tutela no debería quedar en las manos de los comités de ética de las universidades y de los organismos de investigación.

2. LA PROHIBICIÓN, UN TANTO PACATA, DE CIERTAS PRÁCTICAS DE INTELIGENCIA ARTIFICIAL

El art. 5 del RIA prohíbe determinadas prácticas de IA debido, según el considerando 28, a que «son sumamente perjudiciales e incorrectas y deben estar prohibidas, pues van en contra de los valores de la Unión de respeto de la dignidad humana, la libertad, la igualdad, la democracia y el Estado de Derecho y de los derechos fundamentales consagrados en la Carta, como el derecho a la no discriminación, a la protección de datos y a la intimidad y los derechos del niño». Estas prohibiciones, en pocas palabras, afectarán a los sistemas que aprovechen los puntos vulnerables de las personas para la manipulación y el uso de técnicas subliminales; a los sistemas de puntuación social con fines públicos y privados; a la actuación policial predictiva individual basada únicamente en la elaboración de perfiles de personas; a la extracción

no selectiva de internet o de un circuito cerrado de televisión de imágenes faciales con el fin de crear o ampliar bases de datos; a los sistemas de reconocimiento de emociones en el lugar de trabajo y en las instituciones educativas, a menos que sea por razones médicas o de seguridad; a la categorización biométrica de las personas físicas para deducir o inferir su raza, opiniones políticas, afiliación sindical, creencias religiosas o filosóficas, u orientación sexual, y a la identificación biométrica remota, en tiempo real, en lugares de acceso público por parte de la policía, con excepciones limitadas.

A la vista de estas prohibiciones, el resultado final nos parece un tanto decepcionante³⁴ porque, para expresarlo en pocas palabras, la tutela de los derechos se somete a la voluntad de no poner demasiados obstáculos a las posibilidades de avance técnico de la Unión Europea, un avance que no era precisamente rápido antes del Reglamento; además, varias de las prohibiciones —así, las de los sistemas que permitan alterar de manera sustancial el comportamiento de una persona o un colectivo de personas, mermando de manera apreciable su capacidad para tomar una decisión informada; las de los sistemas que exploten alguna de las vulnerabilidades de una persona física o un determinado colectivo de personas derivadas de su edad o discapacidad, o de una situación social o económica específica, y las de los sistemas que provoquen un trato perjudicial o desfavorable hacia determinadas personas físicas o colectivos de personas— ya se podrían amparar en normas vigentes.

Otras prohibiciones parecen haberse quedado cortas; por ejemplo, limitando la de los sistemas que infieren emociones al ámbito laboral y educativo y, además, con matices, pues estaría justificada la introducción de un sistema de estas características si está destinado a ser instalado o introducido en el mercado por motivos de seguridad.

Lo mismo cabría decir de la prohibición que afecta a los sistemas de categorización biométrica que clasifiquen individualmente a las personas físicas sobre la base de sus datos biométricos para deducir o inferir su raza, opiniones políticas, afiliación sindical, convicciones religiosas o filosóficas, vida sexual u orientación sexual, pues «esta prohibición no incluye el etiquetado o filtrado de conjuntos de datos biométricos adquiridos lícitamente, como imágenes, basado en datos biométricos, ni la categorización de datos biométricos en el ámbito de la garantía del cumplimiento del Derecho»³⁵. Esta excepción es más laxa que la prevista en el Reglamento General de Protección de Datos, cuyo art. 9 prohíbe, con carácter general y a expensas de unas excepciones que se

³⁴ Un balance más positivo o «realista» en Mantelero (2024: 65) y Mir Puigpelat (2024: 75); más negativo, en Jakubowska *et al.* (2024).

³⁵ Al respecto, Escajedo San Epifanio (2024: 183-235).

recogen a continuación, tratar datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o la orientación sexual de una persona física.

3. UN ENFOQUE DE LA GOBERNANZA DE LA INTELIGENCIA ARTIFICIAL BASADO EN LA ACEPTACIÓN DEL RIESGO INDUSTRIAL

Si buscamos en el texto del RIA la palabra riesgo aparece 586 veces y en su forma plural otras 181, con lo que en total se puede encontrar en 767 ocasiones, lo que parece colocarlo en el centro de una regulación que, como se acaba de decir, lo concibe como la combinación de la probabilidad de que se produzca un perjuicio y la gravedad de dicho perjuicio. Y es que el RIA asume expresamente que debemos convivir con ciertos niveles de riesgo como precio que pagar por las ventajas que nos ofrecerán los sistemas de IA, lo que ya marca una diferencia con el Reglamento General de Protección de Datos, que, como recuerda Mantelero (2024: 57), parte más de la aversión al riesgo que de su aceptación o mitigación.

Encontramos diferentes pruebas de esta distinta actitud ante el riesgo como precio que pagar por las ventajas de la IA en el articulado del Reglamento; así, y por mencionar dos, en el art. 7, se prevé que, entre los criterios que tener en cuenta para suprimir la condición de sistema de alto riesgo, estará «la probabilidad de que el despliegue del sistema de IA resulte beneficioso para las personas, los colectivos o la sociedad en general, y la magnitud de este beneficio, incluidas posibles mejoras en la seguridad de los productos» (art. 7.2.b); por su parte, conforme al art. 9.5, «las medidas de gestión de riesgos mencionadas en el apartado 2, letra d), considerarán aceptables los riesgos residuales pertinentes asociados a cada peligro, así como el riesgo residual general de los sistemas de IA de alto riesgo».

A este respecto, cabe recordar que se parte de la premisa de que la clasificación de un sistema de IA como «de alto riesgo», que son a los que se dedica más atención, debe limitarse a los sistemas que tengan un efecto perjudicial considerable en la salud, la seguridad y los derechos fundamentales de las personas de la Unión, y dicha limitación debe reducir al mínimo cualquier posible restricción del comercio internacional (considerando 46), es decir, la calificación que implica una mayor limitación —al margen de los sistemas directamente prohibidos— exige que, para someterlo a sus mandatos, el perjuicio sea «considerable» y que la limitación que se imponga afecte lo mínimo posible al mercado. Ese mismo considerando empieza diciendo que la introducción en el mercado de la Unión, la puesta en servicio o la utilización de sistemas

de IA de alto riesgo deben supeditarse al cumplimiento por su parte de determinados requisitos obligatorios, los cuales deben garantizar que los sistemas de IA de alto riesgo disponibles en la Unión o cuyos resultados de salida se utilicen en la Unión no planteen riesgos inaceptables para intereses públicos importantes de la Unión, reconocidos y protegidos por el derecho de la Unión.

Parece, pues, que el Reglamento se alinea con una visión de los sistemas de IA como productos seguros tal y como los entendió la derogada Directiva 2001/95/CE del Parlamento Europeo y del Consejo, de 3 de diciembre de 2001, relativa a la seguridad general de los productos, en cuyo art. 2.d consideraba producto seguro cualquier producto que, en condiciones de utilización normales o razonablemente previsibles, incluidas las condiciones de duración y, si procede, de puesta en servicio, instalación y de mantenimiento, no presente riesgo alguno o únicamente riesgos mínimos, compatibles con el uso del producto y considerados admisibles.

Lo dicho, y lo que se añadirá a continuación, pone de relieve que el objetivo del Reglamento no es primordialmente proteger los derechos fundamentales, sino fomentar el desarrollo de la IA en un contexto de debilidad industrial del sector en Europa (Mantelero, 2024: 57).

4. LA CONFIANZA EN NORMAS ARMONIZADAS Y ESPECIFICACIONES COMUNES TODAVÍA POCO DESARROLLADAS

Como explica Álvarez García (2024: 397), la regulación de la IA en la Unión Europea ha seguido la técnica del nuevo enfoque armonizador, de tal manera que las instituciones europeas han renunciado a la ordenación de este campo por ellas solas, para apelar a una colaboración con los sujetos privados³⁶, y las dos grandes categorías que prevé al respecto el RIA son las normas armonizadas (art. 40) y las especificaciones comunes (art. 41).

³⁶ Abunda este autor recordando que un análisis de la técnica armonizadora del nuevo enfoque revela que cuenta con tres grandes pilares: en primer lugar, la existencia de unos requisitos esenciales obligatorios que debe respetar un producto, y que se fijan directamente por el acto legislativo aprobado por las instituciones europeas (en el caso de la inteligencia artificial, por un reglamento); en segundo lugar, la regulación de las especificaciones técnicas que facilitan la justificación de la conformidad del producto con dichos requisitos imperativos, y, en tercer lugar, la existencia de controles que permiten acreditar en última instancia que el producto cumple con dichas exigencias esenciales obligatorias (399); previamente, sobre la técnica armonizadora del nuevo enfoque y su aplicación a la regulación de la IA, Álvarez García y Tahiri Moreno (2023).

La primera consecuencia es que los sistemas de IA de alto riesgo o los modelos de IA de uso general que sean conformes con normas armonizadas, o partes de ellas, cuyas referencias estén publicadas en el *Diario Oficial de la Unión Europea*, se presumirá que son conformes con los requisitos establecidos para los sistemas de alto riesgo o, en su caso, con las obligaciones establecidas para los modelos de uso general.

La cuestión es que hay motivos para pensar que esa confianza en el desarrollo vigente de las normas armonizadas es excesiva porque la normalización técnica de la IA todavía «está en sus primeros momentos de vida» (Álvarez García, 2024: 415). A este respecto, el estudio *Harmonising Artificial Intelligence: The role of standards in the UE AI Regulation*³⁷ sostiene que todavía existe una discordancia significativa entre las normas armonizadas que se consideran necesarias en el ámbito de la IA y los desarrollos de especificaciones técnicas que se han producido de manera efectiva hasta ahora en Europa; en segundo lugar, haría falta, y eso está por ver, que en la elaboración de estas normas participen de manera activa especialistas en el ámbito de la salud, la seguridad y los derechos fundamentales.

En segundo lugar, la Comisión podrá adoptar actos de ejecución por los que se establezcan especificaciones comunes para los requisitos establecidos para los sistemas de alto riesgo y para los modelos de uso general, y lo cierto es que, a día de hoy, poco se sabe sobre cuáles son los contornos y cómo están llamadas a operar las especificaciones comunes. Tampoco está clara cuál es la relación jurídica existente entre estas especificaciones y las normas armonizadas ni cómo se resolverían eventuales conflictos entre ambas (Álvarez García, 2024: 418 y 424).

5. LA EVALUACIÓN DE LA CONFORMIDAD DE LOS SISTEMAS DE ALTO RIESGO CONFIADA, ESENCIALMENTE, A LOS PROPIOS PROVEEDORES

En quinto lugar, el RIA confía un extraordinario poder de control del riesgo a los propios proveedores de los sistemas de IA, es decir, y según los define el propio Reglamento, a las personas físicas o jurídicas, autoridades públicas, órganos u organismos que desarrollen un sistema de IA o un modelo de IA de uso general o para quienes se desarrolle un sistema de IA o un modelo de IA de uso general y lo introduzcan en el mercado o lo pongan en servicio con su propio nombre o marca, previo pago o gratuitamente (art. 3.3); así, conforme al art. 16 del RIA, son los proveedores los que velarán por que sus sistemas de

³⁷ <https://is.gd/w6xyZ1> (consulta: 21-4-2025).

IA de alto riesgo cumplan los requisitos previstos en la sección 2; los que contarán con un sistema de gestión de la calidad que cumpla lo dispuesto en el art. 17; los que conservarán la documentación a que se refiere el art. 18 y, por no abundar, los que se asegurarán de que los sistemas de IA de alto riesgo sean sometidos al procedimiento pertinente de evaluación de la conformidad a que se refiere el art. 43 antes de su introducción en el mercado o puesta en servicio.

Ese art. 43 prevé, en su apdo. 2, que, «en el caso de los sistemas de IA de alto riesgo mencionados en el anexo III, puntos 2 a 8, los proveedores se atenderán al procedimiento de evaluación de la conformidad fundamentado en un control interno a que se refiere el anexo VI³⁸, que no contempla la participación de un organismo notificado», que podría ser una entidad pública o privada, y no es, desde luego, irrelevante que se deje en manos de los propios proveedores la verificación de que se ajustan a las exigencias del RIA sistemas tan importantes y que inciden de manera directa en derechos fundamentales y libertades públicas como:

[...] los destinados a ser utilizados como componentes de seguridad en la gestión y el funcionamiento de infraestructuras críticas; las herramientas de IA utilizadas en el sistema educativo y en la formación profesional, que pueden determinar el acceso al mismo y la vida profesional de una persona;

los sistemas destinados a ser utilizados para la contratación o la selección de personas físicas o para tomar decisiones que afecten a las condiciones de las relaciones de índole laboral o a la promoción o rescisión de relaciones contractuales de índole laboral;

los sistemas destinados a evaluar la admisibilidad de las personas físicas para beneficiarse de servicios y prestaciones esenciales de asistencia pública, incluidos los servicios de asistencia sanitaria, así como para conceder, reducir o retirar dichos servicios y prestaciones o reclamar su devolución, así como los destinados a evaluar la solvencia de personas físicas o establecer su calificación crediticia, la evaluación de riesgos y la fijación de precios en relación con las personas físicas

³⁸ 1. El procedimiento de evaluación de la conformidad fundamentado en un control interno es el procedimiento de evaluación de la conformidad basado en los puntos 2, 3 y 4. 2. El proveedor comprueba que el sistema de gestión de la calidad establecido cumple los requisitos establecidos en el art. 17. 3. El proveedor examina la información de la documentación técnica para evaluar la conformidad del sistema de IA con los requisitos esenciales pertinentes establecidos en el capítulo III, sección 2. 4. Asimismo, el proveedor comprueba que el proceso de diseño y desarrollo del sistema de IA y la vigilancia poscomercialización de este a que se refiere el art. 72 son coherentes con la documentación técnica.

en el caso de los seguros de vida y de salud y los destinados a la evaluación y la clasificación de las llamadas de emergencia realizadas por personas físicas o para el envío o el establecimiento de prioridades en el envío de servicios de primera intervención en situaciones de emergencia;

los sistemas destinados a ser utilizados para evaluar el riesgo de que una persona física sea víctima de delitos, los polígrafos y herramientas similares, los que evalúan la fiabilidad de las pruebas durante la investigación o el enjuiciamiento de delitos, los destinados a evaluar el riesgo de que una persona física cometa un delito o reincida y los que sirven para elaborar perfiles de personas físicas durante la detección, la investigación o el enjuiciamiento de delitos;

los sistemas para la gestión de la migración, el asilo y el control fronterizo;

los sistemas destinados a ser utilizados por una autoridad judicial, o en su nombre, para ayudar en la investigación e interpretación de hechos y de la ley o para ser utilizados en una resolución alternativa de litigios y

los destinados a ser utilizados para influir en el resultado de una elección o referéndum o en el comportamiento electoral de personas físicas que ejerzan su derecho de voto en elecciones o referendos.

El propio RIA trata de justificar esta sorprendente y peligrosa dejación de controles externos en la escasa capacidad técnica existente en la actualidad, por lo que «procede limitar, al menos en la fase inicial de aplicación del presente Reglamento, el alcance de las evaluaciones externas de la conformidad a los sistemas de IA de alto riesgo que no están asociados a productos. En consecuencia, el proveedor es quien, por norma general, debe llevar a cabo la evaluación de la conformidad de dichos sistemas bajo su propia responsabilidad, con la única excepción de los sistemas de IA que están destinados a utilizarse para la biometría» (considerando 125).

6. LAS INSUFICIENCIAS DE LA EVALUACIÓN DEL IMPACTO DE LOS SISTEMAS DE ALTO RIESGO EN LOS DERECHOS FUNDAMENTALES

El art. 27 del RIA prevé que, antes de desplegar uno de los sistemas de IA de alto riesgo a que se refiere el art. 6, apdo. 2, con excepción de los sistemas de IA de alto riesgo destinados a ser utilizados en el ámbito enumerado en el anexo III, punto 2³⁹, los responsables del despliegue que sean

³⁹ Sistemas de IA destinados a ser utilizados como componentes de seguridad en la gestión y el funcionamiento de las infraestructuras digitales críticas, del tráfico rodado o del suministro de agua, gas, calefacción o electricidad.

organismos de derecho público, o entidades privadas que prestan servicios públicos, y los responsables del despliegue de sistemas de IA de alto riesgo a que se refiere el anexo III, punto 5, letras b) y c)⁴⁰, llevarán a cabo una evaluación del impacto que la utilización de dichos sistemas puede tener en los derechos fundamentales⁴¹.

Esta previsión parece un instrumento adecuado para, como dice el RIA (considerando 96), «garantizar eficazmente la protección de los derechos fundamentales» pero cabe oponer algunos reparos a su articulación; así, en primer lugar, y sin que se ofrezcan más explicaciones, quedan fuera de la evaluación los sistemas destinados a ser utilizados como componentes de seguridad en la gestión y el funcionamiento de las infraestructuras digitales críticas, del tráfico rodado o del suministro de agua, gas, calefacción o electricidad; en segundo lugar, están obligados a llevar a cabo la evaluación los responsables del despliegue que sean organismos de derecho público, o entidades privadas que prestan servicios públicos, pero no las entidades privadas que no llevan a cabo esa prestación de servicio público, con lo que las últimas quedan excluidas de la obligación sin que tampoco sepamos a qué obedece esta exención; en tercer lugar, el propio art. 27 contempla en su apdo. 5 que la Oficina de IA elaborará un modelo de cuestionario, también mediante una herramienta automatizada, «a fin de facilitar que los responsables del despliegue cumplan sus

⁴⁰ b) Sistemas de IA destinados a ser utilizados para evaluar la solvencia de personas físicas o establecer su calificación crediticia, salvo los sistemas de IA utilizados al objeto de detectar fraudes financieros. c) Sistemas de IA destinados a ser utilizados para la evaluación de riesgos y la fijación de precios en relación con las personas físicas en el caso de los seguros de vida y de salud.

⁴¹ A tal fin, los responsables del despliegue llevarán a cabo una evaluación que consistirá en: a) una descripción de los procesos del responsable del despliegue en los que se utilizará el sistema de IA de alto riesgo en consonancia con su finalidad prevista; b) una descripción del período de tiempo durante el cual se prevé utilizar cada sistema de IA de alto riesgo y la frecuencia con la que está previsto utilizarlo; c) las categorías de personas físicas y colectivos que puedan verse afectados por su utilización en el contexto específico; d) los riesgos de perjuicio específicos que puedan afectar a las categorías de personas físicas y colectivos determinadas con arreglo a la letra c) del presente apartado, teniendo en cuenta la información facilitada por el proveedor con arreglo al art. 13; e) una descripción de la aplicación de medidas de supervisión humana, de acuerdo con las instrucciones de uso, y f) las medidas que deben adoptarse en caso de que dichos riesgos se materialicen, incluidos los acuerdos de gobernanza interna y los mecanismos de reclamación. Sobre la evaluación de impacto algorítmico en los derechos fundamentales, Simón Castellanos (2023); en particular, sobre el art. 27 RIA, Chaveli Donet (2024: 496-533).

obligaciones en virtud del presente artículo de manera simplificada», lo que pone en evidencia, valga la expresión, el riesgo de que el riesgo sea autoevaluado; en cuarto lugar, encontramos otra muestra del poder que el RIA atribuye a los propios proveedores de los sistemas de alto riesgo, pues el responsable del despliegue podrá basarse en evaluaciones de impacto relativas a los derechos fundamentales realizadas previamente o a evaluaciones de impacto existentes realizadas por los proveedores; en quinto lugar, una vez realizada la evaluación, el responsable del despliegue notificará sus resultados a la autoridad de vigilancia del mercado, pero no se prevé la publicación de estos.

7. LA POSIBILIDAD DE EXPORTAR SISTEMAS DE INTELIGENCIA ARTIFICIAL PROHIBIDOS EN LA UNIÓN EUROPEA POR SER LESIVOS DE DERECHOS FUNDAMENTALES

Quizá uno de los datos clave para entender la importancia que el RIA otorga al mercado y no tanto a la garantía de los derechos, aunque no sea el mercado interior sino el exterior y no sean los derechos concernidos los de quienes residen en la EU sino los de las personas que viven fuera de sus fronteras, es la posibilidad de exportar a países extracomunitarios sistemas de IA cuya introducción en el mercado⁴², comercialización⁴³ o puesta en servicio⁴⁴ estará prohibida en la UE conforme a las previsiones del art. 5 RIA: en síntesis, sistemas que provoquen manipulación y engaño, la explotación de vulnerabilidades, la puntuación social que origine desigualdades, el reconocimiento de emociones en lugares de trabajo e instituciones educativas, la categorización biométrica para deducir determinadas características protegidas y, con límites, los sistemas de reconocimiento biométrico en tiempo real.

Llegamos a esta conclusión a partir de la lectura del precepto que delimita el ámbito de aplicación (art. 2) y de las definiciones incluidas en el art. 3, que tienen en cuenta siempre la incidencia de los sistemas de IA o de los resultados de salida generados por dichos sistemas en el espacio de la Unión Europea,

⁴² Introducción en el mercado: la primera comercialización en el mercado de la Unión de un sistema de IA o de un modelo de IA de uso general (art. 3.9).

⁴³ Comercialización: el suministro de un sistema de IA o de un modelo de IA de uso general para su distribución o utilización en el mercado de la Unión en el transcurso de una actividad comercial, previo pago o gratuitamente (art. 3.10).

⁴⁴ Puesta en servicio: el suministro de un sistema de IA para su primer uso directamente al responsable del despliegue o para uso propio en la Unión para su finalidad prevista (art. 3.11).

y ello al margen de si los proveedores están establecidos o ubicados en la Unión o en un tercer país; en el caso de los responsables del despliegue de sistemas, de los representantes autorizados, de los importadores y de los distribuidores, se les aplica el RIA, y, por tanto, sus prohibiciones y limitaciones, si operan en el mercado interior, pero quedan exentos de estas si sus actividades consisten en introducir en los mercados de países no comunitarios, comercializar en ellos o poner en servicio en dichos Estados sistemas de IA o resultados prohibidos o limitados por el RIA por ser directamente lesivos de derechos fundamentales o por suponer un alto riesgo para ellos.

8. LOS DERECHOS, MÁS BIEN RETÓRICOS, PARA PRESENTAR UNA RECLAMACIÓN Y PARA RECIBIR UNA EXPLICACIÓN

La sección 4 («Vías de recurso») del capítulo IX («Vigilancia poscomercialización, intercambio de información, vigilancia del mercado») del RIA se introdujo al paso del proyecto de Reglamento por el Parlamento Europeo, y en los arts. 85 y 86 contempla el derecho a presentar una reclamación ante una autoridad de vigilancia del mercado por parte de toda persona física o jurídica que tenga motivos para considerar que se ha infringido lo dispuesto en el Reglamento y el derecho de toda persona que se vea afectada por una decisión que el responsable del despliegue adopte basándose en los resultados de salida de un sistema de IA de alto riesgo que figure en el anexo III, con excepción de los sistemas enumerados en su punto 2, y que produzca efectos jurídicos o le afecte considerablemente del mismo modo, de manera que considere que tiene un efecto perjudicial para su salud, su seguridad o sus derechos fundamentales, para obtener del responsable del despliegue explicaciones claras y significativas acerca del papel que el sistema de IA ha tenido en el proceso de toma de decisiones y los principales elementos de la decisión adoptada⁴⁵.

Pues bien, en lo que se refiere al derecho a presentar la reclamación, el art. 85.2 dispone que, «de conformidad con el Reglamento (UE) 2019/1020, tales reclamaciones se tendrán en cuenta a la hora de llevar a cabo actividades de vigilancia del mercado y se tramitarán de conformidad con los procedimientos específicos establecidos con este fin por las autoridades de vigilancia del mercado», es decir, remite al Reglamento relativo a la vigilancia del mercado y la conformidad de los productos, que no contempla ninguna regulación de las reclamaciones. Como recuerda López-Tarruella (2024: 899), no existe en ese Reglamento una obligación de la autoridad de vigilancia del mercado de

⁴⁵ Sobre su tramitación, López-Tarruella (2024: 895-897).

informar al reclamante sobre el curso y el resultado de la reclamación. Y tampoco existe una obligación como la establecida por el Tribunal de Justicia de la Unión Europea en relación con el Reglamento General de Protección de Datos de otorgar una respuesta de fondo a dicha reclamación⁴⁶. Lo único que establece el art. 85.2 RIA, y reitera el art. 11.3 del Reglamento 2019/1020, es que dichas reclamaciones «se tendrán en cuenta», pero no impone una explicación de los motivos por los cuales la reclamación no se toma en consideración y ni siquiera existe una obligación de la autoridad de vigilancia del mercado de adoptar una decisión respecto a dicha reclamación.

Y por lo que atañe al derecho a recibir una explicación clara y significativa acerca del papel que el sistema de alto riesgo ha tenido en el proceso de toma de decisiones y los principales elementos de la decisión adoptada, el propio precepto ya excluye de dicho derecho los sistemas destinados a ser utilizados como componentes de seguridad en la gestión y el funcionamiento de las infraestructuras digitales críticas, del tráfico rodado o del suministro de agua, gas, calefacción o electricidad.

En segundo lugar, y como también explica López-Tarruella (2024: 915), si bien el art. 86 no lo establece expresamente, del análisis de lo previsto en el art. 78 en relación con los arts. 25.5, 52.6 y 53.1.b, el responsable del despliegue no tiene obligación de proporcionar información que pueda considerarse un secreto comercial o que esté protegida por derechos de propiedad intelectual e industrial, a la hora de ofrecer explicaciones sobre la decisión adoptada a partir de los resultados del sistema de IA, y, en el caso de que el responsable del despliegue considere necesario proporcionarla, se deberán respetar la confidencialidad de la información y los datos obtenidos.

9. LA AUSENCIA DE UNOS PRINCIPIOS OBLIGATORIOS COMUNES A TODOS LOS SISTEMAS DE INTELIGENCIA ARTIFICIAL

A su paso por el Parlamento Europeo en 2023 se incorporaron al proyecto de RIA unos «Principios generales aplicables a todos los sistemas de IA» (art. 4 bis, nuevo), de alto riesgo o no, incluidos los modelos fundacionales, en la línea de las previsiones del Reglamento General de Protección de Datos; dichos principios eran los siguientes: intervención y vigilancia humanas, solidez y seguridad

⁴⁶ En la sentencia de 7 de diciembre de 2023, C-26/22, *SCHUFA*, resolvió que el procedimiento de reclamación no se asemeja al de una petición, sino que se concibe como un mecanismo capaz de proteger de manera eficaz los derechos y los intereses de los interesados (apdo. 58).

técnicas, privacidad y gobernanza de datos, transparencia; diversidad, no discriminación y equidad, y bienestar social y medioambiental. Sin embargo, esta enmienda no prosperó y, si tenemos en cuenta lo que prevé el considerando 1, se «impide que los Estados miembros impongan restricciones al desarrollo, la comercialización y la utilización de sistemas de IA, a menos que el presente Reglamento lo autorice expresamente»⁴⁷.

En suma, a los sistemas de IA que no sean de alto riesgo únicamente se les pueden aplicar «los códigos de conducta para la aplicación voluntaria de requisitos específicos» del art. 95, «destinados a fomentar la aplicación voluntaria de alguno o de todos los requisitos establecidos en el capítulo III, sección 2, a los sistemas de IA que no sean de alto riesgo, teniendo en cuenta las soluciones técnicas disponibles y las mejores prácticas del sector que permitan la aplicación de dichos requisitos». Así pues, y en palabras de Mir Puigpelat (2024: 67-68), en relación con este ingente abanico de sistemas de IA que existen y, sobre todo, que se desarrollarán en el futuro, importa más lo que no dice el RIA que lo que dice: al someterlos únicamente a códigos voluntarios de conducta, el RIA otorga total libertad a su desarrollo.

Por contraste, el Convenio Marco del Consejo de Europa sobre inteligencia artificial, derechos humanos, democracia y Estado de derecho sí prevé, en su capítulo III, unos «Principios relacionados con las actividades dentro del ciclo de vida de los sistemas de inteligencia artificial: dignidad humana y autonomía individual, transparencia y supervisión, rendición de cuentas y responsabilidad, igualdad y no discriminación, privacidad y protección de los datos personales, fiabilidad e innovación segura (arts. 7 a 13).

10. LA EXCESIVA DILACIÓN DE LA ENTRADA EN VIGOR DE ALGUNAS EXIGENCIAS RELATIVAS A SISTEMAS YA INTRODUCIDOS EN EL MERCADO O PUESTOS EN SERVICIO

El art. 111 RIA contiene unas previsiones específicas sobre los sistemas de IA ya introducidos en el mercado o puestos en servicio y los modelos de IA de uso general también introducidos en el mercado conforme a las cuales, y sin perjuicio de que se aplique el art. 5, relativo a las prohibiciones, exigible desde el 2 de febrero de 2025, los sistemas de IA que sean componentes de los sistemas informáticos de gran magnitud establecidos en virtud de los actos legislativos enumerados en el anexo X y aplicables en el espacio de

⁴⁷ Sobre el devenir de esa enmienda, Cotino Hueso y Simón Castellanos (2024: 781-783).

libertad, seguridad y justicia, que se hayan introducido en el mercado o se hayan puesto en servicio antes del 2 de agosto de 2027, deberán estar en conformidad con el presente Reglamento a más tardar el 31 de diciembre de 2030.

Llama la atención que, por una parte, se consideren como ya introducidos en el mercado sistemas que han sido introducidos después de la aprobación del RIA —antes del 2 de agosto de 2027—, y, sobre todo, que dichos sistemas no precisen ajustarse a las prescripciones del propio Reglamento hasta más de seis años después de dicha aprobación.

Por si no fuera suficiente, el apdo. 2 del art. 111 dispone que, «en cualquier caso, los proveedores y los responsables del despliegue de los sistemas de IA de alto riesgo destinados a ser utilizados por las autoridades públicas adoptarán las medidas necesarias para cumplir los requisitos y obligaciones del presente Reglamento a más tardar el 2 de agosto de 2030», dejando un extraordinario espacio temporal para que sistemas de tanta relevancia como los previstos para ser empleados por los poderes públicos sigan al margen de las prescripciones reglamentarias hasta bien entrado el año 2030.

VI. A MODO DE BREVE CONCLUSIÓN

Es de sobra conocido que la Unión Europea va muy por detrás de Estados Unidos y de China en materia de investigación, desarrollo e innovación de la inteligencia artificial y no solo aspira a mejorar su posición en este ámbito, sino también a hacerlo de una manera que resulte compatible con un elevado nivel de protección de la salud, la seguridad y los derechos fundamentales consagrados en la Carta, incluidos la democracia, el Estado de derecho y la protección del medio ambiente, frente a los efectos perjudiciales de los sistemas de IA (art. 1 RIA).

Este ambicioso objetivo se enfrenta a importantes dificultades y una no menor es la articulación de un sistema jurídico de gobernanza de la IA que, por la propia naturaleza del objeto para regular, va a necesitar constantes adaptaciones y frecuentes revisiones, para lo que ya se atribuye un relevante papel a la Comisión Europea a través de la adopción de «derecho blando», algo que, probablemente, sea inevitable, pero que puede amparar una eventual, e ilegítima, relajación de las obligaciones que impone el Reglamento con el argumento de no frenar el desarrollo tecnológico europeo y el desarrollo del propio mercado interno.

Y, sin necesidad de esperar a las directrices de la Comisión, ya se advierte en el contenido del Reglamento una clara opción en favor del mercado interior

y también del exterior, a costa de una menor protección de la salud, la seguridad y los derechos fundamentales: varios sistemas que, en principio, tendrían que estar sujetos a las prohibiciones y limitaciones del RIA quedan fuera de su ámbito de aplicación; se articula una prohibición pacata de ciertas prácticas de IA; el enfoque de la gobernanza de la inteligencia artificial basado en la aceptación del riesgo industrial no parece el más adecuado para la garantías de los derechos fundamentales; se confía en normas armonizadas y especificaciones comunes que todavía están poco desarrolladas; se deja, en buena medida, la evaluación de la conformidad de los sistemas de alto riesgo confiada a los propios proveedores; la evaluación del impacto de los sistemas de alto riesgo sobre los derechos fundamentales parece insuficiente; se permite, en aras de la expansión de la industria europea de la IA en el mercado exterior, exportar sistemas de inteligencia artificial prohibidos en la Unión Europea por ser lesivos de derechos fundamentales; el reconocimiento de los derechos a presentar una reclamación y a recibir una explicación es más retórico que efectivo; se advierte la ausencia de unos principios obligatorios comunes a todos los sistemas de inteligencia artificial, y, finalmente, se ha permitido una excesiva dilación de la entrada en vigor de algunas exigencias relativas a sistemas ya introducidos en el mercado o puestos en servicio.

Una visión menos crítica y, tal vez, más realista tendrá en cuenta, primero, que, con todas sus carencias, este derecho europeo de la IA es ya nuestro derecho, pero no agota el derecho de la IA en nuestro país, especialmente, en materia de derechos y libertades; así, el propio RIA destaca, en su considerando 63, que el hecho de que un sistema de IA sea clasificado como de alto riesgo no debe interpretarse como indicador de que su uso sea lícito con arreglo a otros actos del derecho de la Unión o del derecho nacional. Y, por citar otro ejemplo, el art. 5.2 RIA prevé que el uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público deberá cumplir las garantías y condiciones necesarias y proporcionadas de conformidad con el derecho nacional, en particular en lo que respecta a las limitaciones temporales, geográficas y personales. Así pues, corresponde singularmente al legislador nacional llevar a cabo las reformas necesarias en las normas que desarrollan derechos fundamentales para ajustarlas a las nuevas exigencias derivadas de la generalización de sistemas de IA que entran de lleno en el objeto y contenido de las libertades garantizadas.

En segundo lugar, un texto con el contenido moderado del RIA es posible que consiga una mayor «efecto Bruselas», es decir, un mayor efecto de emulación jurídica en las próximas legislaciones sobre el uso de la IA en otros países (Bradford, 2020: 25-65). Veremos.

Bibliografía

- Álvarez García, V. (2024). La aplicación de las normas armonizadas y de las especificaciones comunes en el ámbito de la inteligencia artificial (artículos 40 y 41 Reglamento). En L. Cotino Hueso y P. Simón Castellanos (dirs.). *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea* (pp. 397-426). Madrid: Aranzadi.
- Álvarez García, V. y Tahiri Moreno, J. (2023). La regulación de la inteligencia artificial en Europa a través de la técnica armonizadora del nuevo enfoque. *Revista General de Derecho Administrativo*, 63, 1-56.
- Balaguer Callejón, F. (2022). *La constitución del algoritmo*. Zaragoza: Fundación Manuel Giménez Abad.
- Barrio Andrés, M. (2024). *Comentarios al Reglamento Europeo de Inteligencia Artificial*. Madrid: La Ley. Disponible en: <https://doi.org/10.69592/3045-6681-N1-OCTUBRE-2024-ART-4>.
- Bradford, A. (2020). *The Brussels Effect: How the European Union Rules the World*. Oxford: Oxford University Press. Disponible en: <https://doi.org/10.1093/oso/9780190088583.001.0001>.
- Bradford, A. (2023). *Digital Empires. The Global Battle to Regulate Technology*. Oxford: Oxford University Press. Disponible en: <https://doi.org/10.1093/oso/9780197649268.001.0001>.
- Bustos Gisbert, R. (2024). El constitucionalista europeo ante la inteligencia artificial: reflexiones metodológicas de un recién llegado. *Revista Española de Derecho Constitucional*, 131, 146-178. Disponible en: <https://doi.org/10.18042/cepc/redc.131.05>.
- Chaveli Donet, E. (2024). La evaluación de impacto de derechos fundamentales por quienes despliegan sistemas de inteligencia artificial en el Reglamento. En L. Cotino Hueso y P. Simón Castellanos (dirs.). *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea* (pp. 495-533). Madrid: Aranzadi.
- Cotino Hueso, L. y Simón Castellanos, P. (2024). *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea*. Madrid: Aranzadi.
- Escajedo San Epifanio, L. (2024). El reconocimiento biométrico en el reglamento de inteligencia artificial: exenciones, prohibiciones y especialidades de alto riesgo. En L. Cotino Hueso y P. Simón Castellanos (dirs.). *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea* (pp. 183-235). Madrid: Aranzadi.
- Fajardo del Castillo, T. (2024). *El soft law en el Derecho Internacional y Europeo: su capacidad para dar respuesta a los desafíos normativos actuales*. Valencia: Tirant lo Blanch.
- Gómez-Pérez, A. (2023). *Inteligencia artificial y lengua española*. Madrid: Real Academia Española.
- González Álvarez, J. L., López Ossorio, J. J., Rodríguez Díaz, M. y Urruela Cortés C. (2018). Sistema de Seguimiento Integral de los Casos de Violencia de Género. *Sistema VioGén. Behavior & Law Journal*, 4 (1), 29-40.
- Green, B. (2022). The flaws of policies requiring human oversight of government algorithms. *Computer law & security review*, 45, 105681. Disponible en: <https://doi.org/10.1016/j.clsr.2022.105681>.
- Hildebrandt, M. (2015). *Smart Technologies and the End(s) of Law. Novel Entanglements of Law and Technology*. Cheltenham: Edward Elgar Publishing.
- Huergo Lora, A. y Díaz González, G. (2025). *The EU Regulation on artificial intelligence: a commentary*. Milán: Wolters Kluwer.

- Jakubowska, E., Noori, K., Hakobyan, M., Iwańska, K., Vieth-Ditlmann, K., Aszodi, N., Membrives Llorens, J., Rodelli, C., Leufer, D., Benaïssa, N. y Fevola, I. (2024). *EU's AI Act fails to set gold standard for human rights*, 3-4-2024. Disponible en: <https://edri.org/our-work/eu-ai-act-fails-to-set-gold-standard-for-human-rights/>.
- Jones, L. (2017). The Right to a Human in the Loop: Political Constructions of Computer Automation and Personhood. *Social Studies of Science*, 47 (2), 216-239. Disponible en: <http://dx.doi.org/10.2139/ssrn.2758160>.
- Lazcoz Moratinos, G. (2024). La vigilancia o supervisión humana en el artículo 14 del Reglamento de inteligencia artificial: Un mero requisito obligatorio para los sistemas de alto riesgo? En L. Cotino Hueso y P. Simón Castellanos (dirs.). *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea* (pp. 681-699). Madrid: Aranzadi.
- Llano Alonso, F. (2024). *Homo Ex Machina. Ética de la inteligencia artificial y Derecho digital ante el horizonte de la singularidad tecnológica*. Valencia: Tirant lo Blanch.
- López-Tarruella Martínez, A. (2024). Derecho a presentar una reclamación y derecho a una explicación. Vías de recurso para los particulares en el Reglamento de inteligencia artificial. En L. Cotino Hueso y P. Lorenzo y Simón Castellanos (dirs.). *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea* (pp. 893-918) Madrid: Aranzadi.
- Mantelero, A. (2024). El reglamento de inteligencia artificial: La respuesta del legislador europeo a los retos de la inteligencia artificial. En L. Cotino Hueso y P. Simón Castellanos (dirs.). *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea* (pp. 53-66). Madrid: Aranzadi.
- Martínez Garay, L. (2024). *Three predictive policing approaches in Spain: VIOGÉN, RISCANVI and VERIPOL: Assessment from a human rights perspective*. Valencia: Universidad de Valencia. Disponible en: <https://doi.org/10.7203/PUV-OA-451-9>.
- McCarthy, J., Minsky, M., Rochester, N. y Shannon, C. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Disponible en: <https://is.gd/7tVnRa>.
- Mir Puigpelat, O. (2024). The AI Act from the Perspective of Administrative Law: Much Ado About Nothing? *European Journal of Risk Regulation*, 16, 63-75. Disponible en: <https://doi.org/10.1017/err.2024.54>.
- Mitchell, M. (2024). *Inteligencia artificial. Guía para seres pensantes*. Madrid: Capitán Swing.
- Obregón Fernández, A. y Lazcoz Moratinos, G. (2021). La supervisión humana de los sistemas de inteligencia artificial de alto riesgo. Aportaciones desde el Derecho Internacional Humanitario y el Derecho de la Unión Europea. *Revista electrónica de estudios internacionales*, 42, 1-29. Disponible en: <https://reei.tirant.com/reei/article/view/2483/2409>.
- Pasquale, F. (2024). *Las nuevas leyes de la robótica. Defender la experiencia humana en la era de la IA*. Barcelona: Galaxia Gutenberg.
- Pérez-Ugona, M. (2024). Análisis comparado de los distintos enfoques regulatorios de la inteligencia artificial en la Unión Europea, EE. UU, China e Iberoamérica. *Anuario Iberoamericano de Justicia Constitucional*, 28 (1), 129-156. Disponible en: <https://doi.org/10.18042/cepc/aijc.28.05>.
- Peters, A. (2011). Soft Law as a New Mode of Governance. En U. Diedrichs, W. Reiners y W. Wessels (eds.). *The dynamics of change in EU Governance* (pp. 21-51). Cheltenham: Edward Elgar Publishing. Disponible en: <http://dx.doi.org/10.2139/ssrn.1876508>.

- Presno Linera, M. Á. (2022). *Derechos fundamentales e inteligencia artificial*. Madrid: Marcial Pons. Disponible en: <https://doi.org/10.2307/jj.4908196>.
- Presno Linera, M. Á. (2023). Policía predictiva y prevención de la violencia de género: el sistema *VioGén*. *Revista de Internet, Derecho y Política*, 39, 1-13. Disponible en: <http://dx.doi.org/10.7238/idp.v0i39.416473>.
- Presno Linera, M. Á. y Meuwese, A. (2024). La regulación de la inteligencia artificial en Europa. *Teoría Y Realidad Constitucional*, 54, 131-161. Disponible en: <https://doi.org/10.5944/trc.54.2024.43310>.
- Russel, S. y Norvig, P. (2008). *Inteligencia Artificial: un enfoque moderno*. Madrid: Pearson Education.
- Samoili, S., López Cobo, M., Delipetrev, B., Martínez-Plumed, F., Gómez, E. y De Prato, G. (2020). *AI WATCH. Defining Artificial Intelligence*. Luxemburg: Publications Office of the European Union. Disponible en: <https://doi:10.2760/382730>.
- San Martín Segura, D. (2023a). *La intrusión jurídica del riesgo*. Madrid: Centro de Estudios Políticos y Constitucionales.
- San Martín Segura, D. (2023b). Prevención algorítmica de la violencia de género: La discrecionalidad policial como decisión tecnológica en el contexto de *VioGén*. *Estudios penales y criminológicos*, 44, 1-35. Disponible en: <https://doi.org/10.15304/epc.44.9013>.
- Senden, L. (2004). *Soft Law in European Community Law*. Oxford: Hart Publishing.
- Simón Castellanos, P. (2023). *La evaluación de impacto algorítmico en los derechos fundamentales*. Madrid: Aranzadi.