FACULTAD DE DERECHO

RÉGIMEN JURÍDICO APLICABLE A LA INTELIGENCIA ARTIFICIAL: ANÁLISIS DEL REGLAMENTO (UE) 2024/1689 DE INTELIGENCIA ARTIFICIAL Y SUS CONSECUENCIAS PARA EL DERECHO CIVIL ESPAÑOL

Autor: Tomás Moré Sebastián

4ºE-5 FIPE

DERECHO CIVIL

Tutora: Guillermina Yanguas Montero

MADRID

[ABRIL 2025]

RESUMEN

El TFG tiene por objeto analizar el Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de Inteligencia Artificial y por el que se modifican los Reglamentos (CE) n.° 300/2008, (UE) n.° 167/2013, (UE) n.° 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Ley de Inteligencia Artificial), y los desafíos derivados de la utilización de tecnologías de Inteligencia Artificial. Se analizará cómo afecta esta norma al Derecho español, con especial mención a las obligaciones que impone y a la atribución de responsabilidad en caso de daños causados por sistemas de Inteligencia Artificial de alto riesgo.

Palabras clave:

ABSTRACT

This paper aims to analyse Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence and amending Regulations (EC) No. 300/2008, (EU) No. 167/2013, (EU) No. 168/2013, (EU) No. 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU (EU) 2019/2144. No. 167/2013, (EU) No. 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), and the challenges arising from the use of artificial intelligence technologies. It will also explore how this regulation affects Spanish law, with special mention to the obligations it imposes and the attribution of liability in case of damage caused by high-risk artificial intelligence systems.

Keywords:

TABLA DE CONTENIDOS

I. INTRODUCCIÓN

1. Objeto de estudio

El objeto de estudio de este Trabajo de Fin de Grado (TFG) es el **análisis del régimen jurídico aplicable a la Inteligencia Artificial (IA)**, en especial del en el **Reglamento (UE) 2024/1689** del Parlamento Europeo y del Consejo (Reglamento de Inteligencia Artificial). Este Reglamento establece normas armonizadas dentro de la Unión Europea para la regulación de la IA. El presente trabajo examina sus implicaciones específicas para el **Derecho civil español** con el objetivo de evaluar cómo esta nueva normativa encaja en el marco jurídico nacional y su posible impacto para los actores involucrados.

Partiendo de un análisis conceptual sobre los Sistemas de Inteligencia Artificial, sus tipos y sus sectores de aplicación, el trabajo aborda la regulación objeto de estudio desde los **desafíos éticos y jurídicos** que plantea la expansión de la IA, tales como la **privacidad y protección de datos**, la **transparencia**, los **sesgos y la discriminación**, y la **responsabilidad** en la toma de decisiones autónomas.

El análisis del Reglamento se centra en la **atribución de responsabilidad civil** en situaciones donde los sistemas de IA, especialmente aquellos considerados de **alto riesgo** según la clasificación del Reglamento de Inteligencia Artificial, causen daños. En este sentido, el TFG explora las **obligaciones que impone** el Reglamento a los desarrolladores, operadores y empresas usuarias de estas tecnologías, y cómo estas obligaciones afectan a la **protección de los consumidores y usuarios** en el contexto de la evaluación de la conformidad de Sistemas de Inteligencia Artificial.

El TFG también incluye un estudio sobre las **propuestas legislativas complementarias** en el ámbito europeo, como las propuestas de **Directivas** relativas a la responsabilidad por daños causados por productos defectuosos y la adaptación de las normas de responsabilidad civil extracontractual a los avances en IA. Esto permite una comparación de cómo el Reglamento (UE) 2024/1689 se inserta en un marco regulatorio

más amplio, dirigido a hacer frente a los riesgos asociados con el uso de tecnologías de IA, y puede ser útil de cara a establecer predicciones sobre futuros desarrollos en la cuestión.

2. Justificación del tema

La **Inteligencia Artificial se ha desarrollado de forma frenética en los últimos años impulsada por la aparición de modelos de lenguaje generativos como ChatGPT o Copilot.** Puede apreciarse un interés entre las grandes empresas tecnológicas en lograr su adaptación a múltiples sectores clave, desde la sanidad hasta el transporte y las finanzas. En este punto, ya no se trata de una cuestión meramente teórica, es un hecho que la Inteligencia Artificial está transformando profundamente la manera en que las sociedades operan. Sin embargo, su integración plantea **importantes desafíos jurídicos y éticos**, ya que muchos sistemas de IA no solo influyen en la toma de decisiones, sino que también pueden causar **daños directos o indirectos** a las personas o bienes. Esta cuestión abre un campo crucial para el Derecho, que debe adaptarse a la nueva realidad tecnológica y establecer marcos claros para regular la IA, sin por ello ser excesivamente restrictivo y causar un retroceso en la innovación tecnológica.

El **Reglamento (UE) 2024/1689**, conocido como "Reglamento de Inteligencia Artificial," es una de las primeras normativas a nivel mundial que establece un marco legal comprensivo para regular la IA. La importancia de este Reglamento radica en que se trata de una **regulación pionera**, cuyo objetivo es **armonizar las reglas** sobre IA en todos los Estados miembros de la Unión Europea y garantizar un uso seguro, ético y responsable de esta tecnología. Dada su novedad y su amplio alcance, resulta fundamental analizar sus disposiciones para comprender su **impacto** en los ordenamientos nacionales. En este trabajo, el análisis se centra en el **Derecho civil español**.

3. Explicación de la estructura

La estructura seguida para analizar la cuestión descrita consta de cuatro apartados. El primer apartado aborda el concepto de Inteligencia Artificial desde una perspectiva general, e incluye una definición y un análisis de los distintos tipos de IA y sus principales aplicaciones en sectores relevantes tales como la sanidad, o el transporte. Tras esta definición, se presentan los principales **retos jurídicos y éticos** que se asocian al uso de la IA.

Posteriormente, en el segundo apartado, se lleva a cabo un **análisis exhaustivo** del Reglamento (UE) 2024/1689, también conocido como la **Ley de Inteligencia Artificial**, y su interrelación con otras normativas y propuestas legislativas en el ámbito de la **responsabilidad civil**. El objetivo principal de esta sección es comprender el alcance y las implicaciones de esta nueva regulación para los Estados miembros, así como su papel en la creación de un marco normativo armonizado para la IA dentro de la Unión Europea.

En el tercer apartado se estudia el **impacto directo** que el **Reglamento (UE) 2024/1689** tiene sobre el **Derecho civil español**, especialmente en lo que respecta a la **responsabilidad civil** y la **protección de consumidores y usuarios** en el contexto de la Inteligencia Artificial. El objetivo es analizar cómo la normativa europea influye en el marco jurídico español y qué **adaptaciones** serán necesarias para asegurar una correcta implementación de la regulación.

El trabajo finaliza con una síntesis de las principales **conclusiones** derivadas del análisis del **Reglamento (UE) 2024/1689** y su impacto en el **Derecho civil español**,

4. Objetivos del trabajo

La **finalidad** de este trabajo de fin de grado (TFG) es analizar en profundidad el **Reglamento (UE) 2024/1689 sobre Inteligencia Artificial** y su impacto en el **Derecho civil español**. Se estudian las implicaciones éticas y jurídicas que surgen de la intersección de la tecnología y el derecho en este ámbito. Este Reglamento es la primera legislación exhaustiva sobre una materia tan clave y compleja como la IA

En un momento en que la Inteligencia Artificial se está convirtiendo en una parte integral de nuestra vida cotidiana, resulta esencial examinar no solo cómo se aplican las normas, sino también cómo estas pueden contribuir a aumentar la confianza en la innovación tecnológica sin por ello limitarla. El trabajo pretende identificar y analizar los desafíos que plantea la IA en el ámbito jurídico, así como la atribución de responsabilidad en caso de daños causados por sistemas de alto riesgo, así como los dilemas éticos asociados a la **privacidad**, la **transparencia** y los **sesgos algorítmicos**. Este enfoque permite una reflexión sobre la necesidad de equilibrar la innovación tecnológica con la salvaguarda de los derechos de los ciudadanos.

5. Método

Para estudiar la cuestión descrita, se han consultado una pluralidad de fuentes y se ha empleado el análisis jurídico, la investigación bibliográfica y la reflexión crítica. Para el primer capítulo, que trata cuestiones relativas a la definición de Inteligencia Artificial y los principales desafíos éticos y jurídicos que plantea, se ha recurrido fundamentalmente a la doctrina especializada que ha analizado esta cuestión. En los capítulos posteriores, se analizan las disposiciones del Reglamento de Inteligencia Artificial y otra legislación europea asociada.

Se estudian los criterios de clasificación de sistemas de IA según su nivel de riesgo que establece, además de las obligaciones impuestas a los operadores y desarrolladores de IA, así como los mecanismos de responsabilidad establecidos. Este análisis permite identificar las implicaciones directas del Reglamento en el marco jurídico español, y analizando cómo estas normas afectan a la regulación de la responsabilidad civil y a la protección de los derechos de los consumidores y usuarios.

Finalmente, se exponen conclusiones sobre el posible impacto del Reglamento en el marco normativo español una vez entre plenamente en vigor. Esta reflexión busca mantener en todo momento un enfoque proactivo, éticamente responsable y orientado hacia el bienestar de los ciudadanos a la hora de analizar legislaciones de este tipo.

II. LA INTELIGENCIA ARTIFICIAL: CONCEPTOS CLAVE

1. CONCEPTO DE INTELIGENCIA ARTIFICIAL, TIPOS Y APLICACIONES

1.1. Concepto de Inteligencia Artificial

La Inteligencia Artificial (IA) se ha convertido en uno de los campos más transformadores de la tecnología en las últimas décadas, generando un impacto significativo en múltiples sectores de la sociedad. En términos generales, podemos decir que la IA hace referencia a la capacidad de máquinas o sistemas informáticos para realizar tareas que tradicionalmente requieren "inteligencia humana". Desde un punto de vista más técnico, el HLEG (High-Level Expert Group on Artificial Intelligence) la define como "sistemas de software (y posiblemente también de hardware) diseñados por humanos que, ante un objetivo complejo, actúan en la dimensión física o digital percibiendo su entorno mediante la adquisición de datos, interpretando los datos estructurados o no estructurados recogidos, razonando, o procesando la información, derivada de estos datos y decidiendo la(s) mejor(es) acción(es) a tomar para alcanzar el objetivo dado. Los sistemas de IA pueden utilizar reglas simbólicas o aprender un modelo numérico, y también pueden adaptar su comportamiento analizando cómo se ve afectado el entorno por sus acciones anteriores."[1].

La definición del HLEG es especialmente relevante, además, porque ha sido utilizada como base por el Joint Research Centre (JRC), que es el principal asesor en cuestiones científicas para la Comisión Europea (Sofia Samoili & Montserrat Lopez Cobo & Emilia Gomez & Giuditta De Prato & Fernando Martinez-Plumed & Blagoj Delipetrev, 2020. "AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence), lo que puede ayudar a entender la perspectiva del legislador europeo en esta materia.

Por otro lado, Russell y Norwig[2] clasifican las principales definiciones doctrinales sobre el concepto de IA siguiendo una división entre el pensamiento/acción humana, basado en hipótesis y confirmaciones mediante experimentos; y el pensamiento/acción racional, combinación de matemáticas e ingeniería, pero cabe resaltar que según va progresando la tecnología, esta división va dejando paso a sistemas de IA cada vez más capaces de ofrecer tanto un enfoque humano como uno racional.

1.2. Tipos de Inteligencia Artificial

[1] HLEG. (2019). *Una Definición de la Inteligencia Artificial: Principales Capacidades y disciplinas Científicas*. Comisión Europea. Pág. 8

[2] Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: a modern approach*. Tercera Edición. Pearson. Pág.

En términos generales, los sistemas de Inteligencia Artificial se dividen en dos grandes categorías diferenciadas por su funcionamiento: IA débil e IA fuerte

La principal diferencia entre ambos tipos reside en su funcionamiento. La IA débil se limita a simular la inteligencia humana en un rango limitado y definido previamente para lograr un fin concreto.[3] Un ejemplo de IA débil es Siri, el asistente virtual de Apple que simula precisamente a un asistente humano y ayuda al usuario en sus tareas diarias, pero no es capaz de tomar decisiones autónomas y su eficacia disminuye cuando se le pregunta algo fuera de sus funciones predeterminadas.

La IA fuerte, en cambio, es capaz de replicar la toma de decisiones a través del "machine learning" o aprendizaje automático para aprender de experiencias pasadas y tomar decisiones autónomas.[4] Un ejemplo de IA fuerte son los prototipos de vehículos autónomos como los de Waymo, cuya IA ha demostrado ser capaz de predecir trayectorias de vehículos en escenarios complejos y las interacciones entre ellos en la carretera.[5]

La distinción es relevante precisamente porque esa autonomía que caracteriza a la IA fuerte puede justificar que se le aplique un marco legal más restrictivo que a la IA débil.[6] De hecho, así parece ser, ya que como veremos posteriormente en mayor profundidad, el Reglamento de Inteligencia Artificial establece una clasificación de riesgos asociados a usos específicos de la IA de la que parece intuirse una mayor preocupación del legislador europeo por la IA fuerte. Veremos como un gran número de usos asociados a los sistemas de IA fuerte son considerados como de alto riesgo (conducción autónoma), o riesgo inaceptable (actuación policial predictiva), mientras que otros usos más propios de la IA débil son considerados como de riesgo mínimo o nulo (videojuegos)

1.3. Principales Sectores de Aplicación

El legislador también tiene en cuenta los sectores de aplicación de los Sistemas de Inteligencia Artificial para determinar su riesgo. Si bien, es muy probable que la

[3]Martinez, R. (2018). *Artificial intelligence: Distinguishing between types & definitions*. Nevada Law Journal, 19(3), 1016-1037.
[4]Martinez, R. (2018). *Artificial intelligence: Distinguishing between types & definitions*. Nevada Law Journal, 19(3), 1016-1037.
[5]Waymo. VectorNet: *Predicting behaviour to help the Waymo Driver make better decisions*. (s.f.). Recuperado el 17 de Octubre de 2024 de https://waymo.com/blog/2020/05/vectornet/
[6]Martinez, R. (2018). *Artificial intelligence: Distinguishing between types & definitions*. Nevada Law Journal, 19(3), 1016-1037.

Inteligencia Artificial se acabe extendiendo a prácticamente todas las áreas de trabajo, los sectores más sensibles actualmente son: medicina, transporte, finanzas, y defensa.

En la medicina, se prevé el uso de Inteligencia Artificial para, entre otras funciones, realizar diagnósticos, sugerir tratamientos, predecir enfermedades o monitorizar pacientes. Los beneficios de su implementación incluyen el avance de los tratamientos actuales, la disminución de sus costes, y el aumento de la seguridad de los hospitales.[7] Por otro lado, la sensibilidad de este sector radica tanto en el gran impacto que tienen las decisiones médicas sobre la vida de los pacientes, y a que los sistemas de IA dedicados a este sector, necesariamente operan con grandes bases de datos médicos personales; lo que necesariamente genera dudas desde el punto de vista de salvaguardar la privacidad del paciente. Por ello, desde los Comités de ética asistencial y clínica (CEA) y los Comités de investigación (CEI) se aboga por el uso de sistemas que anonimicen los datos a los que tengan acceso y que tengan salvaguardias para evitar su reidentificación posterior. También se recomienda mantener en todo momento la supervisión humana de las decisiones y diagnósticos realizados por los sistemas de Inteligencia Artificial, incluso si su nivel de fiabilidad es alta.[8]

En el campo del transporte, ya hemos mencionado como distintas plataformas como Waymo están logrando grandes avances a la hora de diseñar vehículos capaces de conducción autónoma. La prioridad en este campo siempre ha de ser la seguridad del conductor del vehículo autónomo, la del resto de vehículos en la vía y la de los peatones. Por ello, la aplicación de la IA en este sector plantea importantes dilemas con respecto a su seguridad y la atribución de responsabilidad por los daños derivados de decisiones tomadas por sistemas de Inteligencia Artificial; dilemas que las instituciones de la Unión Europea tienen especialmente presentes.

En las finanzas, la implementación de sistemas de IA se está llevando a cabo especialmente en áreas como la concesión de préstamos; la prevención del fraude; predicción de variables económicas y financieras; la gestión de carteras y valoración de activos, y la gestión y trato con los clientes bancarios. El principal beneficio de la aplicación de la IA en este sector es su precisión a la hora de identificar sucesos con base

[7] Haleem, Abid. Javaid, Mohd (2019). Current status and applications of Artificial Intelligence (AI) in the medical field: An overview. *Current Medicine Research and Practice Journal*, 9 (6), 231-237. https://www.sciencedirect.com/science/article/abs/pii/S235208171930193X

[8] Ruiz, F. J. B. (2022). Aplicaciones de la inteligencia artificial al ámbito biosanitario: Protección de datos y privacidad. Implicaciones éticas y legales. En *Anales de la Cátedra Francisco Suárez*, 56, 245-268. Universidad de Granada

en las directrices dadas con el apoyo de grandes bases de datos. Por otro lado, consideraciones similares a las expuestas respecto de la privacidad en el uso de la IA en el campo de la medicina pueden hacerse en este caso, y además, la dificultad a la hora de interpretar los resultados en ciertos supuestos puede reducir su efectividad.[9] Estos problemas de interpretación se derivan de que muchas veces estos sistemas operan como "algoritmos de caja negra", es decir, aquellos sistemas cuya toma de decisiones es tan compleja que ni sus programadores son capaces de predecir la respuesta concreta que va a dar el sistema ante determinadas preguntas, incluso si son modelos altamente fiables.[10]

Por último, el uso de sistemas de Inteligencia Artificial en armas de defensa plantea especiales consideraciones éticas en el caso de Sistemas de Armas Autónomos (SAA) capaces de planificar ataques y llevarlos a cabo sin necesidad de un operador humano. Por un lado, estos sistemas disminuyen la sobreexposición de soldados a riesgos en operaciones militares y disminuyen el tiempo de respuesta ante un posible ataque, pero también puede llevar a la desensibilización de la comunidad internacional ante conflictos armados al sustituir a los combatientes humanos por robots, y relegar la toma de decisiones estratégicas a sistemas de IA que, como hemos visto, funcionan como cajas negras y no son siempre predecibles incluso para sus propios creadores.[11]

2. DESAFÍOS JURÍDICOS Y ÉTICOS

A medida que la tecnología avanza y la Inteligencia Artificial se integra en sectores clave, como la salud o la seguridad, surgen preguntas fundamentales sobre cómo regular su uso, quién es responsable de sus decisiones y cómo se puede evitar que su implementación infrinja derechos fundamentales. Este apartado del trabajo aborda las principales preocupaciones legales y éticas asociadas a la IA, centrándose en cuestiones como la privacidad de los datos, la transparencia de los algoritmos, los sesgos y la discriminación en las decisiones automatizadas.

2.1. Privacidad y Protección de Datos

[9] Banco de España. (2022). *Inteligencia artificial y finanzas: una alianza estratégica*. https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosOcasionales/22/Fich/do2222.pdf

[10] Gràcia, X. G. & Sancho-Gil, J. M. (2021). Artificial intelligence in education: Big data, black boxes, and technological solutionism. *International journal of media, technology and lifelong learning*, 17 (2). ttps://doi.org/10.7577/seminar.4281

[11] Moliner Juan. (2021). Desafíos éticos en la aplicación de la inteligencia artificial a los sistemas de defensa. *Revista DIECISIETE*, 4. DOI: 10.36852/2695-4427_2021_04.06

La cantidad de datos personales recopilados y procesados por los sistemas de Inteligencia Artificial ha incrementado especialmente desde la introducción de modelos de lenguaje generativos como Chat GPT. Estos modelos requieren de acceso continuo a grandes cantidades de datos tanto externos como derivados de sus interacciones con usuarios para su constante entrenamiento. Esto genera preocupaciones sobre cómo se gestionan, almacenan y utilizan dichos datos, y si se están respetando los derechos de los individuos en materia de privacidad.

Estas cuestiones han llevado a los principales investigadores a abogar por la integración de pausas éticas que ayuden a mitigar estos riesgos y del aprendizaje federado en el desarrollo de sistemas de Inteligencia Artificial, lo que permitiría descentralizar su proceso de aprendizaje y evitar riesgos a la vez que se aumenta la colaboración en el proceso de desarrollo de estos sistemas.[12]

Además, el uso de privacidad diferencial ha probado ser eficaz a la hora de limitar la información revelada de sujetos individuales durante el análisis de datos grupales recopilados[13]. Este método se basa en la adición de "ruido estadístico" en las computaciones para mantener la privacidad de individuos cuyos datos personales forman parte del conjunto de datos sin afectar a la utilidad del conjunto[14]

Sin embargo, y si bien en el espacio europeo la normativa de protección de datos recogida en el RGPD incluye disposiciones relevantes para proteger los datos personales de los ciudadanos de la UE en supuestos similares al presente, todas estas sugerencias propuestas por la doctrina se mantienen como algo optativo para las empresas desarrolladoras de sistemas de Inteligencia Artificial, por lo que estos mismos investigadores han urgido a la comunidad internacional a colaborar y armonizar los mecanismos regulatorios empleados para salvaguardar la privacidad en un contexto de

[12] Kumar Anil. (2023). Advances in Data Protection and Artificial Intelligence: Trends and Challenges. *International Journal of Advanced Engineering Technologies and Innovations*, 1(01), 294-319. https://ijaeti.com/index.php/Journal/article/view/392

[13] Dwork, C. (2008). Differential Privacy: A Survey of Results. En Agrawal, M., Du, D., Duan, Z., Li, A. (eds) *Theory and Applications of Models of Computation*.. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-79228-4_1

[14] Dwork, C. (2008). Differential Privacy: A Survey of Results. En Agrawal, M., Du, D., Duan, Z., Li, A. (eds) *Theory and Applications of Models of Computation*.. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-79228-4_1

implantación de la Inteligencia Artificial, y fomentar el desarrollo de las tecnologías mencionadas.[15]

2.2. Transparencia

En relación con el punto anterior, la falta de transparencia en la forma de operar de sistemas de Inteligencia Artificial, especialmente en cuanto al uso y toma de decisiones automatizadas que impactan directamente en los individuos y en la sociedad, es otra de las principales barreras que pueden generar rechazo a su adopción a gran escala. No en vano, la transparencia es una herramienta que actúa como garantía de múltiples derechos fundamentales y principios democráticos y además nos permite evaluar el buen funcionamiento de los sistemas de Inteligencia Artificial.

Por tanto, la transparencia algorítmica ha de ser mayor en proporción al riesgo, opacidad y tipo asociado a la IA en cuestión.[16] Esta transparencia algorítmica puede interpretarse como un principio de notificación al usuario de distintos factores como pueden ser el hecho de que está interactuando con una IA; de la adopción de decisiones por un sistema de IA y la opción de no usarlas o aportarles información de quienes son los proveedores de las IA y su finalidad.[17]

Recuérdese que el artículo 13 RGPD obliga cuando se obtengan de un interesado datos personales relativos a él, a facilitar información sobre "a) la identidad y los datos de contacto del responsable y, en su caso, de su representante; b) los datos de contacto del delegado de protección de datos, en su caso;" Y, también, "e) los destinatarios o las categorías de destinatarios de los datos personales, en su caso;"

Especialmente sensible es la ausencia en los marcos regulatorios actuales de una obligación de transparencia sobre el sector público a fin de que exista un registro de que sistemas de IA existen y su impacto en las Administraciones Públicas.

Por último, en todo supuesto en el que intervengan decisiones automatizadas, la doctrina distingue distintos principios éticos y mejores prácticas relacionadas con la

---

[15] Kumar Anil. (2023). Advances in Data Protection and Artificial Intelligence: Trends and Challenges. *International Journal of Advanced Engineering Technologies and Innovations*, 1(01), 294-319. https://ijaeti.com/index.php/Journal/article/view/392

[16] Cotino, L. (2022). Transparencia y explicabilidad de la inteligencia artificial y "compañía" (comunicación, interpretabilidad, inteligibilidad, auditabilidad, testabilidad, comprobabilidad, simulabilidad…). Para qué, para quién y cuánta. *Transparencia y explicabilidad de la inteligencia artificial*, 29-70. https://dialnet.unirioja.es/servlet/articulo?codigo=8709893

[17] Cotino, L. (2023). Qué concreta transparencia e información de algoritmos e inteligencia artificial es la debida. *Revista española de la transparencia*, (16), 17-63.

notificación sobre los datos personales utilizados y los procesos lógico-técnicos que utiliza el sistema para tomar dichas decisiones.[18] Dichos principios han sido recogidos en normas como el art. 22 del RGPD, según el cual "Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar.". Otro ejemplo podría ser el art. 64.4.d) del Estatuto de los Trabajadores, según el cual los Comités de Empresa tienen derecho a "Ser informado por la empresa de los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de Inteligencia Artificial que afectan a la toma de decisiones que pueden incidir en las condiciones de trabajo, el acceso y mantenimiento del empleo, incluida la elaboración de perfiles."

Ambos ejemplos muestran que tanto el legislador europeo como el nacional se muestran receptivos a la codificación de estos principios éticos y mejores prácticas, especialmente en supuestos particularmente sensibles como los relativos al empleo.

2.2. Sesgos y Discriminación

La última preocupación ética que se analizará en este apartado reside en la preocupación de que los sistemas de IA, aunque desarrollados con el objetivo de mejorar la eficiencia y la precisión en la toma de decisiones, puedan perpetuar o incluso amplificar los sesgos existentes en los datos que utilizan. Esto puede dar lugar a decisiones discriminatorias que afecten de manera injusta a determinados grupos o individuos, especialmente en los sectores claves discutidos como el acceso a servicios financieros, o la atención médica.

Estos sesgos pueden ser producto del mero análisis de datos por el sistema sin establecer relaciones causa-efecto entre ellas, o identificar circunstancias subyacentes a problemas sociales complejos, pudiendo incluso llegar a considerar a las personas de raza negra como simplemente más propensas a la delincuencia[19], o disminuir las probabilidades de ascenso laboral de las mujeres si en la empresa en la que se implementa

[18] Cotino, L. (2023). Qué concreta transparencia e información de algoritmos e inteligencia artificial es la debida. *Revista española de la transparencia*, (16), 17-63.

[19] Belloso Martín, N. (2022). La problemática de los sesgos algorítimicos (con especial referencia a los de género). ¿Hacia un derecho a la protección a los sesgos? En F. H. (Coordinador), *Inteligencia Artificial y filosfía del Derecho,*45-69. Murcia: Ediciones Laborum S.L.

el sistema había un patrón en el que los hombres eran ascendidos desproporcionadamente por encima de sus compañeras de trabajo.[20]

Eliminar estos sesgos puede conseguirse si evitamos que la Inteligencia Artificial pueda tomar decisiones en base a características protegidas (sexo, raza…), pero esto puede llevar a una pérdida de eficacia del sistema. Por ello, también se propone introducir barreras en su programación y entrenamiento que, establezca consideraciones técnicas, sociales, legales y éticas que permitan al sistema reconocer la discriminación y ser consciente de ella.[21]

Podemos conectar esta cuestión con las preocupaciones expresadas respecto a la transparencia, puesto que si somos capaces de entender como una Inteligencia Artificial ha tomado una decisión discriminatoria, es más probable que podamos corregir esos problemas en su código o entrenamiento para así convertir algoritmos discriminatorios en no discriminatorios, aunque también tenemos que tener en cuenta que a veces los elementos discriminatorios pueden ser introducidos por el propio usuario.

III. ANÁLISIS DE LA REGULACIÓN DE LA UNIÓN EUROPEA SOBRE INTELIGENCIA ARTIFICIAL

1. EL REGLAMENTO (UE) 2024/1689 (LEY DE INTELIGENCIA ARTIFICIAL)

1.1. Antecedentes y Proceso de Elaboración del Reglamento de Inteligencia Artificial

El Reglamento de Inteligencia Artificial es sobre todo, una regulación sobre riesgos. Por ello, se enmarca en un proceso de intervención pública progresiva en sectores altamente tecnificados como es la digitalización en este caso. En este sentido, regulaciones como el RGPD son un claro antecedente para el Reglamento de Inteligencia Artificial al anticipar la necesidad "de minimizar los riesgos que produce una actividad con riesgo inherente a la que no se puede o quiere renunciar"[22]

[20] Aránguez Sánchez, T. (2022). Sesgos sexistas de los algoritmos e Inteligencia Artificial. En T. A. Sánchez, & O. Olariu, *Algoritmo, teletrabajo y otros grandes temas del feminismo digital*, 71-86. Madrid: Dykinson S.L.

[21] Ferrer, X., Van Nuenen, T., Such, J. M., Coté, M., & Criado, N. (2021). Bias and discrimination in AI: a cross-disciplinary perspective. *IEEE Technology and Society Magazine*, 40(2), 72-80. https://arxiv.org/abs/2008.07309

[22] Miranzo-Diaz, J. (2024). El Reglamento de Inteligencia Artificial de la Unión Europea: regulación de riesgos y sistemas deestandarización. A&C - Revista de Direito Administrativo & Constitucional, 24(96), 43-78 https://doi.org/10.21056/aec.v24i96.1932

Su proceso de elaboración ha sido particularmente extenso, con múltiples años de debate sobre su texto antes de su aprobación. Los principales hitos en este proceso son los siguientes:

Como antecedentes, en 2018 la Comisión Europea publicó dos Comunicaciones relacionadas con la IA ("Inteligencia Artificial Para Europa" [23]y "Plan Coordinado Sobre Inteligencia Artificial"[24]) y en 2020 el "Libro Blanco sobre la Inteligencia Artificial"[25]. Todos estos documentos nos revelan los principios que guían la estrategia europea en materia de Inteligencia Artificial. Estos son:

1. Crear un ecosistema de confianza y excelencia.
2. Fomentar un enfoque ético, transparente y centrado en el ser humano.
3. Implementar un marco regulador para garantizar la seguridad y los derechos fundamentales.
4. Promover la colaboración público-privada y aumentar inversiones en IA.
5. Promover alianzas estratégicas entre los Estados miembros y el sector privado.

En Abril de 2021 la Comisión Europea presentó una propuesta de un marco regulador sobre Inteligencia Artificial que recogiese los principios fundamentales para abordar las implicaciones humanas y éticas de esta tecnología. En 2022 la Comisión Europea introdujo a su vez dos Directivas sobre responsabilidad por daños causados por productos defectuosos, y en materia de IA, ambas aplicables a la Inteligencia Artificial si se aprueban. Dentro del proceso de enmiendas y propuestas destacan las planteadas por el Consejo y el Parlamento en junio de 2023, tras las cuales se logró un acuerdo en el Consejo en diciembre de 2023. En febrero de 2024 se crea la Oficina Europea de Inteligencia Artificial. Por último, el Parlamento aprobó el Reglamento de Inteligencia Artificial en 2024, entrando en vigor el 1 de agosto de este año.[26]

Este breve análisis refleja la complejidad de regular los riesgos de una tecnología en constante evolución, con implicaciones profundas para la sociedad, la economía y los derechos fundamentales. La experiencia de elaborar este reglamento no solo evidencia el

[23] Comisión Europea. (2018) *Inteligencia artificial para Europa.* https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52018DC0237

[24] Comisión Europea. (2018) *Plan coordinado sobre la inteligencia artificial.* https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52018DC0795

[25] Comisión Europea. (2020) *Libro Blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza.* https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52020DC0065

[26] Historic timeline. EU Artificial Intelligence Act. (2024). Recuperado el 16 de Diciembre de 2024 de https://artificialintelligenceact.eu/developments/

compromiso de la Unión Europea por liderar la regulación ética y responsable de la inteligencia artificial, sino que también sienta un precedente global al abordar por primera vez de manera proactiva los riesgos de la IA mientras se impulsa su desarrollo sostenible, posicionando a Europa como referente en la gobernanza de tecnologías emergentes.

1.2. Análisis del Contenido del Reglamento de Inteligencia Artificial

Dada su concepción como una regulación sobre riesgos, es lógico que uno de los pilares centrales del Reglamento de Inteligencia Artificial (RIA) es su enfoque basado en la clasificación de riesgos. Este modelo jerarquiza los riesgos en diferentes categorías y permite distinguir entre Sistemas de IA, estableciendo obligaciones específicas según la naturaleza y el impacto de cada categoría.

1.2.1. Clasificación de Riesgos en el Reglamento de Inteligencia Artificial

En primer lugar, encontramos los **Sistemas de IA prohibidos** por constituir un riesgo no permitido. Estos Sistemas son definidos en el Capítulo II del Reglamento de IA e incluyen aquellos que explotan vulnerabilidades de una persona física o un determinado colectivo de personas, y emplean manipulación subliminal o engañosa para alterar sustancialmente el comportamiento de las personas. (Art. 5 RIA a y b RIA) También incluyen Sistemas que tengan como finalidad el "social scoring" (Art. 5c RIA), la predicción de delitos (Art. 5d RIA), la creación o ampliación de bases de datos de reconocimiento facial (Art. 5e RIA), y la inferencia de emociones en lugares de trabajo o educativos, salvo en casos médicos o de seguridad. (Art. 5f RIA). Por último, se prohíben también aquellos Sistemas dedicados a la categorización biométrica con la finalidad de inferir o deducir categorías especiales de datos, y aquellos Sistemas dedicados a la identificación biométrica remota en tiempo real y en espacios de acceso público (Art. 5g y h RIA) salvo y en la medida en que dicho uso sea estrictamente necesario para alcanzar objetivos como la prevención de amenazas graves o la búsqueda de víctimas de delitos específicos, siempre bajo estrictas condiciones legales y de proporcionalidad.

En segundo lugar, encontramos los **Sistemas de IA de alto riesgo** definidos en el Capítulo III del Reglamento de IA. Un sistema de IA se considerará de alto riesgo si está destinado a ser utilizado como componente de seguridad de un producto regulado por los actos legislativos de armonización de la Unión (Anexo I RIA) o es uno de dichos productos, y dicho producto o sistema debe someterse a una evaluación de conformidad por terceros antes de ser introducido en el mercado. Además de lo anterior, los sistemas

enumerados en el Anexo III se clasifican automáticamente como de alto riesgo, salvo excepciones. (Art. 6.1 RIA)

Los sistemas clasificados como de alto riesgo en el Anexo III están agrupados según sus ámbitos de uso (Art. 6.2 RIA):

En biometría, aquellos dedicados a la identificación biométrica remota (excepto para simplemente verificar la identidad de una persona), la categorización biométrica según atributos sensibles, y el reconocimiento de emociones. (Anexo III.1 RIA)

En infraestructuras críticas, aquellos Sistemas utilizados como componentes de seguridad en la gestión y el funcionamiento de las infraestructuras digitales críticas, del tráfico rodado, o cadenas de suministros. (Anexo III.2 RIA)

En educación y formación profesional, aquellos Sistemas utilizados para la determinación de acceso o admisión a instituciones educativas, la evaluación de resultados de aprendizaje para orientar procesos educativos, y la supervisión de comportamientos prohibidos en exámenes. (Anexo III.3 RIA)

En cuestiones de empleo, aquellos Sistemas utilizados para la contratación, promoción, rescisión o selección de personas y asignación de tareas. (Anexo III.4 RIA)

En cuestiones relativas al acceso a servicios esenciales, aquellos Sistemas dedicados a la evaluación de elegibilidad para asistencia pública o sanitaria, la calificación crediticia (salvo para la detección de fraudes), la evaluación de riesgos en seguros de vida y salud. (Anexo III. 5 RIA)

En cuestiones relativas al cumplimiento del Derecho, aquellos Sistemas dedicados a la evaluación del riesgo de que una persona sea víctima de un crimen por las autoridades competentes o la elaboración de perfiles durante la detección, la investigación o el enjuiciamiento de delitos. (Anexo III. 6 RIA)

En cuestiones relacionadas con la migración, aquellos Sistemas diseñados para la evaluación de riesgos (seguridad, salud, migración irregular), el análisis de solicitudes de asilo, visado o residencia y la detección e identificación de personas. (Anexo III. 7 RIA)

En cuestiones relacionadas con la administración de justicia, aquellos Sistemas dedicados 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. (Anexo III. 8 RIA)

Sin embargo, existen excepciones a esta clasificación automática para aquellos Sistemas que no plantean riesgos significativos para la salud, seguridad o derechos fundamentales, y cumplen cualquiera de las siguientes condiciones; realizan tareas procedimentales limitadas; mejoran una actividad humana sin reemplazar el juicio humano; detectan patrones o desviaciones sin sustituir decisiones humanas previas ni influir sustancialmente en ellas; o realizan tareas preparatorias para decisiones humanas en ámbitos del anexo III. No obstante, siempre se considerarán de alto riesgo cuando impliquen elaboración de perfiles de personas físicas. (Art. 6.3 RIA)

En tercer lugar, encontramos los **Sistemas de riesgo limitado**, entre los que se incluyen "chatbots" o Inteligencias Artificiales generadoras de contenido como "Chat GPT". Estos Sistemas deben cumplir con ciertas obligaciones de transparencia ante sus usuarios y el público general:

Los Sistemas de IA deben informar a las personas físicas de que están interactuando con una IA, salvo que esto sea evidente para un usuario razonablemente informado y considerando el contexto. Existe una excepción para sistemas de IA autorizados por ley para detectar, prevenir, investigar o enjuiciar delitos, con sujeción a las garantías adecuadas para los derechos y libertades de terceros, salvo que estos sistemas estén a disposición del público para denunciar un delito penal. (Art. 50.1 RIA)

Los Sistemas de IA que generen contenido (audio, video, texto, imágenes) deben etiquetar los resultados como generados o manipulados artificialmente, en un formato legible por máquina. Existen también excepciones para sistemas que actúen como apoyo a la edición estándar sin alterar significativamente los datos de entrada, o sistemas autorizados para fines legales como la prevención o investigación de delitos en los mismos términos que en el caso anterior. (Art. 50.2 RIA).

Los Sistemas de IA dedicados al reconocimiento de emociones o a la categorización biométrica deben informar a las personas expuestas al sistema sobre su funcionamiento. y tratarán sus datos personales de conformidad con la legislación aplicable. También está prevista una excepción para sistemas autorizados con fines legales en los mismos términos que los anteriores supuestos. (Art. 50.3 RIA)

Los Sistemas de IA que generen o manipulen imágenes o contenidos de audio o vídeo que constituyan una ultrasuplantación ("deepfakes" en inglés) deben informar públicamente que el contenido ha sido generado o manipulado por IA. Existen

excepciones para casos de Sistemas autorizados con fines legales y aquellos Sistemas dedicados a la creación de contenidos creativos (arte, sátira, ficción), donde basta con una notificación adecuada que no interfiera en el disfrute de la obra. Si un texto generado por IA se publica para informar sobre asuntos de interés público, se debe divulgar su origen artificial, salvo que haya sido revisado o controlado editorialmente y exista una responsabilidad editorial asumida por una persona física o jurídica. (Art. 50.4 RIA)

Por último, encontramos los **Sistemas de IA de riesgo mínimo o nulo**, que son todos aquellos Sistemas que no puedan encuadrarse en ninguna de las categorías anteriores (filtros de correo o videojuegos, por ejemplo). Estos sistemas no están regulados por el reglamento y por lo tanto pueden usarse sin restricciones.

1.2.2. Modelos de IA de Uso General

Aparte de la clasificación de Sistemas de IA en función de su riesgo, el Reglamento de Inteligencia Artificial contiene también una regulación especifica para los **modelos de IA de uso general y modelos de IA de uso general y riesgo sistémico**. Estos modelos son aquellos que no tienen una finalidad concreta, sino que se le pueden darle múltiples usos sin estar limitados a una función específica.

El Considerando 110 del Reglamento determina como de riesgo sistémico aquellos modelos que puedan producir efectos negativos reales o razonablemente previsibles en relación con accidentes graves, perturbaciones de sectores críticos, salud y seguridad públicas; efectos negativos reales o razonablemente previsibles sobre los procesos democráticos; y la difusión de contenidos ilegales, falsos o discriminatorios. Un modelo es de riesgo sistémico cuando tenga capacidades de alto impacto, lo que se presume siempre que la potencia de cálculo empleada en su preparación supere los 10^25 FLOPS (cálculos matemáticos que puede hacer por segundo una CPU y GPU).

Los proveedores de estos modelos deben mantener actualizada la documentación técnica a fin de facilitarla a la Oficina de IA y a las autoridades pertinentes. (Art. 53.1a RIA). Esta documentación deberá contener como mínimo una descripción general del modelo, las tareas que vaya a realizar, políticas de uso razonable, fecha de lanzamiento y distribución, arquitectura y número de parámetros, modalidad, y licencia. También deberán incluir una descripción de los elementos técnicos del modelo que incluya como mínimo las herramientas e infraestructura necesarias para su integración, las especificaciones del diseño y proceso de entrenamiento, los datos de entrenamiento,

prueba y validación, el consumo de energía y los recursos computacionales utilizados. Por último, los modelos con riesgo sistémico deben incluir también una descripción detallada de las estrategias de evaluación; sus métodos, criterios y resultados, incluyendo las limitaciones de estos. También deben incluir una descripción de pruebas adversas internas o externas y ajustes realizados al modelo, además de una explicación de como los componentes del software interactúan y se integran en el procesamiento general . (Anexo XI RIA).

Asimismo, deben compartir información clave sobre el sistema con los proveedores de sistemas de IA que tengan la intención de integrar el modelo de IA de uso general en sus sistemas (art. 53.1b RIA). Esta información incluirá como mínimo los elementos del Anexo XI más la manera en que el modelo interactúa o puede utilizarse para interactuar con el hardware o el software que no formen parte del propio modelo, y sin tener que incluir el consumo de energía del modelo. Estas obligaciones deben compaginarse con el respeto a la propiedad intelectual y secretos comerciales, y no aplicarán a los proveedores de modelos de IA que se divulguen con arreglo a una licencia libre y de código abierto.

Además de estas obligaciones, los modelos de IA de uso general con riesgo sistémico deben cumplir una serie de medidas de seguridad adicionales relacionadas con la evaluación y mitigación de riesgos; la monitorización y notificación de incidentes graves y medidas correctivas a la Oficina de IA y la autoridad pertinente; y el mantenimiento de un nivel adecuado de protección en ciberseguridad. Los proveedores pueden usar códigos de buenas prácticas o normas armonizadas para demostrar conformidad. Si no las utilizan, deberán demostrar el cumplimiento por otros medios de prueba alternativa. La confidencialidad de la información y documentación, incluidos secretos comerciales, se protegerán conforme al artículo 78. (Arts. 55 y 92 RIA).

1.2.3. Vigilancia y Poscomercialización

Los proveedores deben establecer un sistema de **vigilancia poscomercialización** para Sistemas de IA de alto riesgo. Dicho sistema deberá compilar los datos que puedan recopilarse sobre el funcionamiento del Sistema durante su vida útil y su interacción con otros sistemas a fin de evaluar el cumplimiento de los requisitos establecidos en el Reglamento (Art. 72 RIA). También deberán notificar a las autoridades de vigilancia del mercado de los Estados miembros cualquier incidente grave vinculado al Sistema de IA en un plazo máximo de 15 días, 10 días si el incidente está vinculado al fallecimiento de

una persona, o 48h si se trata de un incidente generalizado. Posteriormente, el proveedor deberá cooperar con las autoridades y realizar una evaluación de riesgos del incidente e integrar medidas correctoras (Art. 73 RIA).

Si una autoridad de vigilancia del mercado recibe información sobre un incidente grave durante una prueba en condiciones reales podrá, en su territorio, suspenderla o requerir al proveedor o responsable del despliegue que realicen modificaciones a dichas pruebas. (Art. 76.3 RIA). Asimismo, si considera que un Sistema de IA presenta riesgos, la autoridad de vigilancia del mercado realizará una evaluación para verificar su cumplimiento del Reglamento, prestando especial atención a riesgos para colectivos vulnerables y derechos fundamentales. En caso de incumplimiento, requerirá al operador tomar medidas correctoras, retirar o recuperar el sistema en un plazo máximo de 15 días hábiles; y si el operador no cumple este requerimiento, podrá tomar todas las medidas provisionales adecuadas para prohibir o restringir la comercialización o puesta en servicio del Sistema de IA.

Estas medidas provisionales se considerarán justificadas si ninguna otra autoridad de vigilancia del mercado ha planteado objeciones tras 3 meses. En caso contrario, la Comisión consultará a la autoridad de vigilancia del mercado del Estado miembro y a los operadores involucrados para evaluar la medida nacional. En un plazo de seis meses, o de 60 días si se trata de prácticas de IA prohibidas según el artículo 5, decidirá si la medida está justificada y notificará su decisión tanto a la autoridad correspondiente como a las demás autoridades de vigilancia del mercado. (Art. 81 RIA)

La autoridad informará y cooperará con organismos nacionales pertinentes, y los operadores deberán colaborar con las autoridades. El operador se asegurará de que se adopten todas las medidas correctoras adecuadas en relación con todos los sistemas de IA afectados que haya comercializado en la Unión. (Art. 79.2 y 79.4-5 RIA)

Por último, el Reglamento contempla medidas específicas para abordar incumplimientos en la forma de **sanciones** que deben ser efectivas, proporcionadas y disuasorias.

Toda persona física o jurídica que tenga motivos para considerar que se ha infringido lo dispuesto en el Reglamento de Inteligencia Artificial podrá presentar reclamaciones ante la autoridad de vigilancia del mercado pertinente (Art. 85 RIA).

En el caso de incumplimientos del Art. 5 (Prácticas Prohibidas), las sanciones previstas serán multas administrativas de hasta 35.000.000 EUR o, si el infractor es una empresa, de hasta el 7 % de su volumen de negocios mundial total correspondiente al ejercicio financiero anterior, si esta cuantía fuese superior (o inferior si es una PYME) (Art. 99.3 RIA).

El incumplimiento de las obligaciones establecidas en los Arts. 16, 22-24, 26, 31, 33.1, 33.3, 33.4, 34 y 50 estará sancionado con multas administrativas de hasta 15.000.000 EUR o, si el infractor es una empresa, de hasta el 3 % de su volumen de negocios mundial total correspondiente al ejercicio financiero anterior, si esta cuantía fuese superior (o inferior si la empresa es una PYME) (Art. 99.4 RIA).

Por otro lado, la presentación de información falsa o incompleta a las autoridades competentes, lo que implica un incumplimiento de las obligaciones de transparencia estudiada, esta sancionada con multas administrativas de hasta 7 500 000 EUR o, si el infractor es una empresa, de hasta el 1 % del volumen de negocios (o inferior si la empresa es una PYME) (99.5 RIA)

Más allá de los topes delineados, la cuantía en cada caso concreto se fijará tras valorar la naturaleza de la infracción, su gravedad, duración, consecuencias, e impacto. También el tamaño y volumen de negocios del infractor, su cuota de mercado, la existencia de sanciones previas. Por último, se consideran elementos agravantes o atenuantes como los beneficios obtenidos por el infractor, su cooperación, las medidas adoptadas por este para mitigar daños, el grado de responsabilidad, su intencionalidad y sus acciones tomadas para minimizar los perjuicios causados (99.7 RIA).

En el caso de infracciones por instituciones, órganos u organismos de la UE; el incumplimiento de las prohibiciones del Art. 5 conlleva multas administrativas de hasta 1.500.000 EUR. Otras infracciones conllevarán multas de hasta 750.000 EUR. Tras permitir a la institución infractora ser oída, la cuantía en concreto se fijará tras valorar la gravedad de la infracción, su duración, consecuencias, y el impacto en las personas afectadas, además del grado de responsabilidad de la institución, sus acciones tomadas para mitigar daños, la cooperación con el Supervisor Europeo de Protección de Datos, antecedentes de infracciones similares, la notificación de la infracción y el presupuesto anual del organismo involucrado. (Art. 100 RIA).

Si el incumplimiento es meramente formal (falta de registro en la base de datos de la UE, documentación técnica…) la autoridad de vigilancia del mercado podrá requerir su subsanación, y si esta no se produce, restringir o prohibir la comercialización del Sistema de IA de alto riesgo (Art. 83 RIA).

La Comisión también podrá imponer multas a proveedores de modelos de IA de uso general si estos han incumplido una medida de la Comisión, han infringido las disposiciones del Reglamento, no han respondido a una solicitud de información o han facilitado información incompleta o falsa, o no dieron acceso a la Comisión al modelo de IA de uso general o al modelo de IA de uso general con riesgo sistémico para que se lleve a cabo una evaluación. Dichas multas no superarán el 3 % de su volumen de negocios mundial total anual correspondiente al ejercicio financiero anterior o de 15.000.000 EUR, si esta cifra es superior. El TJUE podrá anulas, reducir o aumentar la cuantía fijada por la Comisión. Antes de adoptar una decisión, la Comisión comunicará sus conclusiones preliminares al Consejo de IA y al proveedor del modelo de IA de uso general o del modelo de IA y le dará la oportunidad de ser oído. (Art. 101 RIA).

2. LA PROPUESTA DE DIRECTIVA SOBRE RESPONSABILIDAD CIVIL POR DAÑOS DERIVADOS DE LA INTELIGENCIA ARTIFICIAL Y LA DIRECTIVA SOBRE RESPONSABILIDAD POR LOS DAÑOS CAUSADOS POR PRODUCTOS DEFECTUOSOS

Si bien el Reglamento de Inteligencia Artificial es el principal objeto de estudio en este trabajo, el análisis breve de otras propuestas legislativas dentro de la UE puede facilitar la interpretación de su contenido y posibles evoluciones legislativas en materia de Inteligencia Artificial.

Tanto el Reglamento de Inteligencia Artificial como las propuestas de Directivas sobre normas de responsabilidad civil extracontractual en Inteligencia Artificial y sobre responsabilidad por los daños causados por productos defectuosos neos encajan dentro de la Estrategia Europea de IA, cuyo propósito es consolidar a la Unión Europea como líder mundial en el desarrollo y uso ético de la inteligencia artificial.

Al tratarse de propuestas, el contenido del articulado de ambos documentos puede variar si son finalmente aprobadas, por lo que el propósito de esta sección no

es realizar un análisis exhaustivo de estas, sino aportar contexto al análisis del Reglamento de Inteligencia Artificial

2.1. La Propuesta de DIRECTIVA DEL PARLAMENTO EUROPEO Y DEL CONSEJO Relativa a la Adaptación de las Normas de Responsabilidad Civil Extracontractual a la Inteligencia Artificial.

Publicada en 2022, esta propuesta de Directiva pretende fortalecer el sistema de normas sobre responsabilidad civil para mantener su eficacia ante los avances tecnológicos mediante la inclusión de normas sobre exhibición de pruebas relativas a sistemas de IA de alto riesgo, y sobre la carga de la prueba.

Una de las principales claves de estas normas es la creación de una **presunción de causalidad** refutable que vincula la culpa del demandado con los resultados generados por el sistema de IA, o la falta de los mismos. Esta presunción responde a dificultades ocasionadas por la opacidad y complejidad con la que operan los ya mencionados "algoritmos de caja negra". Para el legislador europeo, este fenómeno puede resultar en una dificultad probatoria excesiva para las victimas de daños ocasionados por Sistemas de Inteligencia Artificial, que podría incluso resultar en indefensión. Por ello, el propósito de esta medida es claro, facilitar al demandante la acreditación de los daños sufridos y disminuir los costes en los que incurriría para ello.[27]

Dicha presunción operará solo en el caso de que se cumplan todas las condiciones establecidas en al Artículo 4 de la propuesta de Directiva. Estas condiciones son:

Culpa del demandado o de su responsable: El demandante debe demostrar la culpa del demandado o de una persona por la que sea responsable, por incumplir un deber de diligencia establecido en el Derecho de la Unión o nacional, cuyo propósito sea proteger frente a los daños ocurridos.

[27] Martí, Ricard. (2023). Reflexiones acerca de la Propuesta de Directiva sobre responsabilidad por daños derivados de la inteligencia artificial y su impacto en el Derecho español de daños. *Revista Aranzadi Doctrinal,* 4. https://dialnet.unirioja.es/servlet/articulo?codigo=8884458

Probabilidad razonable de influencia de la culpa: Debe ser razonablemente probable, dadas las circunstancias, que dicha culpa haya influido en los resultados producidos o en la falta de resultados del sistema de inteligencia artificial (IA).

Relación de causalidad con los daños: El demandante debe demostrar que la información de salida producida por el sistema de IA o la ausencia de esta información fue la causa de los daños sufridos.

Se debe matizar que en casos de demandas de daños dirigidas contra proveedores de sistemas de IA de alto riesgo o contra personas sujetas a las obligaciones del proveedor se dará por cumplida la primera condición si no utiliza datos de entrenamiento, validación y prueba que cumplan los criterios establecidos en el Reglamento de IA; no respeta los requisitos de transparencia establecidos en dicha normativa; no está diseñado para permitir una vigilancia efectiva por personas físicas durante su uso; no alcanza un nivel adecuado de precisión, robustez y ciberseguridad; o no se adoptan medidas correctoras inmediatas para garantizar su conformidad o retirarlo del mercado.

Similarmente, en casos de demandas de daños dirigidas contra usuarios de sistemas de IA de alto riesgo la primera condición se entenderá cumplida si no se utilizaron o supervisaron conforme a las instrucciones de uso establecidas, incluyendo la suspensión o interrupción de su uso cuando fuera necesario, o si fue expuesto a datos de entrada impertinentes para su finalidad prevista.

Puesto que son múltiples sujetos los que intervienen en el desarrollo y funcionamiento de Sistemas de IA, otra de las facilidades para el demandante introducidas por esta Directiva es la obligación de exhibir pruebas a proveedores y usuarios de Sistemas de IA de alto riesgo sospechoso de haber causado daños. Esta obligación existirá, según el Art. 3 del texto legal, si el demandante potencial presenta “hechos y pruebas suficientes para sustentar la viabilidad de una demanda de indemnización por daños y perjuicios”.

En el supuesto de que un demandado incumpla esta obligación, se presumirá el incumplimiento de un deber de diligencia, aunque esta presunción es refutable.

En conclusión, el objetivo de este mecanismo es facilitar al actor identificar los sujetos responsables de los daños producidos de entre proveedores, operadores y usuarios de Sistemas de IA.

Por último, la vinculación entre esta propuesta y el Reglamento de Inteligencia Artificial es más que evidente. Durante este breve análisis hemos visto como su texto se apoya en múltiples ocasiones en la terminología, estándares de calidad, y obligaciones de transparencia establecidas Reglamento de IA. Es especialmente relevante la influencia de la clasificación de riesgos estudiada en el punto anterior a la hora de determinar el alcance de las presunciones y obligaciones de la Directiva que han sido resaltadas en esta sección. Ello parece sugerir que el impacto del Reglamento de Inteligencia Artificial puede ir más allá de su estricto contenido, y que podría jugar un rol como "piedra base" de la regulación europea en materia de IA, con las consecuencias que ello conlleva a su vez para el Derecho nacional de los Estados Miembros.

2.2. La DIRECTIVA DEL PARLAMENTO EUROPEO Y DEL CONSEJO Sobre Responsabilidad por los Daños Causados por Productos Defectuosos

La Directiva del Parlamento Europeo y del Consejo sobre Responsabilidad por los Daños Causados por Productos Defectuosos, se propuso el 18 de septiembre de 2022 y fue aprobada en 2024. En ella, se propuso garantizar que los derechos de los consumidores estén adecuadamente protegidos frente a los riesgos inherentes a productos defectuosos, incluyendo aquellos que integran Inteligencia Artificial. Su texto legal se extendería a Sistemas de Inteligencia Artificial sin importar su nivel de riesgo al incluir programas informáticos en la definición de productos, aunque no tengan una forma material definida.

La definición de producto defectuoso como producto que no ofrece la seguridad que el público general tiene derecho a esperar es continuista con anteriores desarrollos

legislativos en este ámbito. Para la determinación de esta expectativa de seguridad, la Directiva incluye en su Artículo 6 una serie de criterios, de los cuales destacan los siguientes por su relación con la IA.[28]

El Art. 6.1b) contempla tanto el uso razonablemente previsible como el indebido, por lo que los proveedores de Sistemas de IA deberán de suministrar instrucciones de uso para no correr el riesgo de que se les impute responsabilidad por no haber suministrado esta información.

La evaluación de seguridad también deberá tener en cuenta la capacidad de autoaprendizaje del modelo de IA concreto, lo que puede generar dudas sobre si los errores producidos tras la evaluación y su entrenamiento pueden redundar en su calificación como producto defectuoso. Por ello, también debe prestarse atención a en qué momento el Sistema comienza a funcionar de forma autónoma y su productor deja de ejercer control sobre el mismo, lo que puede ser posterior a su introducción en el mercado o no producirse. Ello no implica que, si un Sistema requiere de una actualización posterior a su salida al mercado, este fuese defectuoso antes de recibir esta actualización.

En consonancia con el Reglamento de Inteligencia Artificial, la Directiva también requiere en sus Arts. 7.1f) y 7.1g) que los proveedores y productores de Sistemas de IA cumplan las exigencias de seguridad y transparencia con los organismos reguladores que allí se establecen. Ello se extiende también a cuestiones de ciberseguridad, ya que la vulnerabilidad de un Sistema ante un ataque informático puede resultar en su calificación como producto defectuoso.

Más allá de estos criterios, la responsabilidad prevista en esta norma se ve limitada por la extensión del daño indemnizable cubierto por ella. Por ejemplo, no será indemnizable el daño puramente económico que deriva en daños de una persona distinta del demandante. Tampoco serán indemnizables los daños morales distintos de daños para la salud psicológica, salvo que estén previstos en el derecho nacional del Estado miembro

[28] Garea, C. (2023). La Propuesta de Directiva sobre responsabilidad por daños ocasionados por productos defectuosos y su aplicación a la inteligencia artificial. En Santos, F.C.S (et al.), *Estudos en homenagem ao Profesor Doutor Antonio Pinto Monteiro. Volume 1: Direito Civil*. Boletim da Facultade de Direito, 179-206. Coimbra: Universidad de Coimbra.

en concreto.[29] Especialmente relevante en este caso es que solo es resarcible las pérdidas económicas derivadas del borrado o corrupción de datos informáticos, no el hecho en si.

Con respecto al régimen de responsabilidad que establece esta Directiva, el demandante debe probar el carácter defectuoso del producto, los daños padecidos y el nexo causal ante ambos. Por tanto, el fundamento básico de la imputación de responsabilidad en este caso es el defecto del producto en cuestión, ya que sin él los daños no serán indemnizables. Dicho defecto es independiente de la actuación negligente o no del operador. Parece por tanto que se ha optado por un régimen de imputación de responsabilidad objetiva limitada[30].

En relación con la imputación de responsabilidad, el Art. 10.2 establece una presunción de que el producto es defectuoso cuando el demandado incumpla la obligación de exhibir pruebas establecida en el Art. 9.1. Dicho artículo funciona con idéntico fundamento y en los mismos términos que la presunción de causalidad estudiada en la propuesta de Directiva sobre Responsabilidad Civil Extracontractual en Materia de IA, lo que implica que es una presunción refutable diseñada para facilitar la prueba del defecto para el demandante. Esta presunción operará también en aquellos supuestos en los que el demandante pruebe que el Sistema de IA causó un daño por un mal funcionamiento manifiesto; o no cumple con las obligaciones de seguridad del Derecho de la UE, lo que incluye evidentemente el Reglamento de Inteligencia Artificial.

Por último, y como contraparte para fomentar la innovación tecnología, el Art. 11 de la Directiva incluye supuestos de exención de responsabilidad por los daños causados por un producto defectuoso.

Según esto, los fabricantes e importadores quedarán exentos de responsabilidad si no introdujeron el producto en el mercado ni lo pusieron en servicio; el defecto no era

[29] Garea, C. (2023). La Propuesta de Directiva sobre responsabilidad por daños ocasionados por productos defectuosos y su aplicación a la inteligencia artificial. En Santos, F.C.S (et al.), *Estudos en homenagem ao Profesor Doutor Antonio Pinto Monteiro. Volume 1: Direito Civil*. Boletim da Facultade de Direito, 179-206. Coimbra: Universidad de Coimbra.

[30] Garea, C. (2023). La Propuesta de Directiva sobre responsabilidad por daños ocasionados por productos defectuosos y su aplicación a la inteligencia artificial. En Santos, F.C.S (et al.), *Estudos en homenagem ao Profesor Doutor Antonio Pinto Monteiro. Volume 1: Direito Civil*. Boletim da Facultade de Direito, 179-206. Coimbra: Universidad de Coimbra.

detectable con los conocimientos disponibles en el momento o responde al cumplimiento de otros requisitos legales.

Los distribuidores podrán eximirse de responsabilidad si demuestran no haber participado en la comercialización del producto, o si el defecto no existía en el momento de su comercialización.

Un operador que modifica productos quedará exento de responsabilidad si el defecto no está relacionado con su modificación.

En conclusión, la Directiva del Parlamento Europeo y del Consejo sobre Responsabilidad por los Daños Causados por Productos Defectuosos refuerza la protección de los consumidores frente a los riesgos asociados a productos inseguros, adaptando el marco normativo existente a los desafíos tecnológicos actuales. Al abordar los daños derivados de productos complejos, como los que integran sistemas de inteligencia artificial, la propuesta favorece una mayor claridad y equidad en la asignación de responsabilidades. Esto no solo beneficia a las víctimas al facilitar la compensación por daños, sino que también incentiva a los fabricantes y operadores a mantener altos estándares de calidad y seguridad en el diseño y desarrollo de sus productos.

IV. IMPACTO DEL REGLAMENTO DE INTELIGENCIA ARTIFICIAL SOBRE EL DERECHO CIVIL ESPAÑOL

1. ATRIBUCIÓN DE RESPONSABILIDAD CIVIL EN CASO DE DAÑOS CAUSADOS POR SISTEMAS DE IA DE ALTO RIESGO