

EL DERECHO DE LA UNIÓN EUROPEA ANTE EL DESARROLLO TECNOLÓGICO: EL REGLAMENTO DE INTELIGENCIA ARTIFICIAL

EUROPEAN UNION LAW AND TECHNOLOGICAL DEVELOPMENT: THE ARTIFICIAL INTELLIGENCE REGULATION*

Jesús R. Mercader Uguina**

Universidad Carlos III de Madrid

SUMARIO: I. ¿Inteligencia o racionalidad artificial?, 1. Definiendo jurídicamente la IA, 2. La IA como sistema asociado a una máquina, 3. Autonomía, adaptación, consecución de objetivos, inferencia de resultados, lógica predictiva e influencia en el entorno. Los otros rasgos definidores de la IA. –II. Sistemas de IA prohibidos en el ecosistema laboral, 1. Delimitación de sistema de IA prohibido, 2. Prohibición de sistemas de reconocimiento de emociones en el lugar de trabajo, 3. Prohibición de sistemas de categorización biométrica que clasifiquen individualmente a las personas físicas. –III. Sistemas de alto riesgo en lo laboral, 1. Ámbito de aplicación de los sistemas de alto riesgo en lo laboral. – IV. La responsabilidad por el uso indebido de los sistemas de IA. –Bibliografía citada.

RESUMEN

El artículo analiza el nuevo Reglamento de Inteligencia Artificial de la Unión Europea (RIA), centrándose en su definición jurídica, en los sistemas de IA prohibidos en el ámbito laboral y en los sistemas de alto riesgo relacionados con el empleo. Asimismo, se examina el régimen de responsabilidad por el uso indebido de la IA, destacando la necesidad de protección de los derechos fundamentales, la transparencia algorítmica y las obligaciones de empleadores y proveedores. El trabajo subraya los desafíos regulatorios y técnicos derivados de la autonomía y opacidad de los sistemas de IA, así como el impacto en las relaciones laborales.

ABSTRACT

This article examines the new European Union Artificial Intelligence Regulation (RIA), focusing on the legal definition of AI, the prohibition of certain AI systems in the workplace, and the regulation of high-risk AI systems related to employment. It also analyzes the liability regime for the misuse of AI, emphasizing the protection of fundamental rights, algorithmic transparency, and the obligations of employers and providers. The work highlights the regulatory and technical challenges posed by the autonomy and opacity of AI systems, as well as their impact on labor relations.

* Recibido el 1 de julio de 2025. Aprobado el 14 de julio de 2025.

** Catedrático de Derecho del Trabajo y de la Seguridad Social.

Palabras clave: Inteligencia Artificial. Derecho Laboral. Reglamento de la Unión Europea. Riesgos Tecnológicos.

Key words: Artificial Intelligence. Labor Law. European Union Regulation. Technological Risks.

I. ¿INTELIGENCIA O RACIONALIDAD ARTIFICIAL?

1. DEFINIENDO JURÍDICAMENTE LA IA.

1. La Inteligencia Artificial (“IA”) es la capacidad que tienen las máquinas de imitar la manera en la que los seres humanos aprenden e incluso razonan. Es una nueva racionalidad artificial¹.

Que los pilares maestros de la IA son los datos y la capacidad de computación es una evidencia fuera de toda duda. Una noción, la de “dato”, que aparece precisamente definida Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724 (Reglamento de Gobernanza de Datos), entendiendo por tal: “*toda representación digital de actos, hechos o información, así como su recopilación, incluso como grabación sonora, visual o audiovisual*”. Los datos son el aire de los nuevos ecosistemas productivos, su nutriente fundamental y se han convertido en un recurso esencial para el crecimiento económico, la creación de empleo y el progreso social.

Por su parte, el formidable desarrollo de la capacidad de computación que marca fronteras inimaginables con los ordenadores cuánticos convierte en irresistible esta lógica transformadora que tiene en el trabajo uno de sus centros de gravedad². Gracias a la evolución de procesadores, la llegada de unidades de procesamiento especializado y a la computación cuántica resulta posible resolver en 200 segundos una tarea que a un superordenador convencional le tomaría 10 mil años realizar. A todo ello se unen los nuevos algoritmos de aprendizaje profundo: Métodos

¹ De estas y otras muchas cuestiones vinculadas con este conjunto de transformaciones, me he ocupado en mi monografía, Jesús. R. MERCADER UGUINA, *Algoritmos e inteligencia artificial en el derecho digital del trabajo*, Valencia, Tirant lo Blanch, 2022. Posteriormente, en “En busca del empleador invisible: algoritmos e inteligencia artificial en el derecho digital del trabajo”, *El Cronista del Estado Social y Democrático de Derecho*, 2022, n.º 100 (Ejemplar dedicado a: Inteligencia artificial y derecho), pp.136-145. De los efectos en materia laboral durante la tramitación del RIA me ocupé en los trabajos anteriores y en “*El Reglamento de Inteligencia Artificial entra en la recta final, una primera lectura en clave laboral*”, *Revista General de Derecho del Trabajo y de la Seguridad Social*, 2024, n.º 67. Y, más recientemente tras su aprobación en “*El Reglamento de Inteligencia Artificial: frecuentemos el futuro*”, *Brief AEDTSS* 42/2024 y en “*Los usos de alto riesgo en el ámbito laboral de la IA y la autocertificación*”, *El Foro de Labos*, 9 de mayo de 2024. “El principio de proporcionalidad como límite al control laboral basado en la inteligencia artificial”, *Trabajo y Derecho: nueva revista de actualidad y relaciones laborales*, 2024, n.º 19 (monográfico Impacto Laboral de la Inteligencia Artificial) e “*Inteligencia Artificial y relaciones laborales*”, *Cuadernos de derecho transnacional*, 2024, vol. 16, n.º 2, pp. 1114-1128. Algunas reflexiones y contenidos de este trabajo son tributarios de esos antecedentes.

² Como señala el Preámbulo del Real Decreto 317/2025, de 15 de abril, por el que se regula la concesión directa de subvenciones a entidades de referencia en el ámbito de las comunicaciones cuánticas, con el fin de fomentar su colaboración en el marco de la adenda del componente 16, Estrategia Nacional de Inteligencia Artificial, del Plan de Recuperación, Transformación y Resiliencia, financiado por la Unión Europea - Next Generation EU: “*Las tecnologías cuánticas suponen una importante revolución tecnológica al aprovechar las leyes de la mecánica cuántica, que describen el comportamiento de partículas a nivel subatómico, para desarrollar nuevos paradigmas de computación, algoritmos y seguridad de las comunicaciones. Estas tecnologías no solo representan avances científicos, sino que tienen el potencial de transformar sectores clave de la economía, la sociedad y la investigación, marcando el comienzo de una nueva revolución tecnológica*”.

como las redes neuronales convolucionales (CNN) y las redes neuronales recurrentes (RNN) han permitido mejoras espectaculares en tareas de reconocimiento de imágenes, voz y patrones complejos.

2. Pero, jurídicamente, ¿qué debemos entender por IA? Según el artículo 3.1 del 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 ("RIA"), entendemos por tal, *"un sistema basado en una máquina diseñado para funcionar con distintos niveles de autonomía, que puede mostrar capacidad de adaptación tras el despliegue y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar información de salida, como predicciones, contenidos, recomendaciones o decisiones, que puede influir en entornos físicos o virtuales"*. La versión que se ha incorporado al RIA coincide con la establecida por la OCDE en la versión actualizada en 2024 de sus principios sobre la IA³, y que ha sido acogida por el Convenio Marco sobre Inteligencia Artificial y Derechos Humanos, Democracia y Estado de Derecho, de 17 de mayo de 2024⁴.

Para la adecuada comprensión de esta definición es imprescindible acudir al Considerando 12 del RIA, del que podemos extraer unas primeras pautas. A ello se añade el borrador de directrices *"Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689"* ("Directrices para la definición de la IA"), publicado por la Comisión Europea, sobre la definición de sistema de IA en las que se explica la definición ofrecida por el regulador en el RIA⁵. Y es que el art. 96.1 f) del RIA exige que la Comisión elabore unas directrices sobre la aplicación práctica del Reglamento y, en particular, sobre la aplicación de la definición de sistema de IA que figura en el artículo 3.1, que tendrán en cuenta el estado de la técnica generalmente reconocido en materia de IA.

Establecen las referidas Directrices que el RIA "no se aplica a todos los sistemas, sino sólo a aquellos que se ajustan a la definición de "sistema de IA" en el sentido del artículo 3, apartado 1, de la Ley de la IA. La definición de sistema de IA es, por tanto, clave para comprender el ámbito de aplicación del RIA". De igual forma, las mismas aclaran que "la definición de un sistema de IA adopta una perspectiva basada en el ciclo de vida que abarca dos fases principales: la fase previa al despliegue o de "construcción" del sistema y la fase posterior al despliegue o de "uso" del sistema". No obstante, no se exige que estos todos elementos establecidos en la definición del artículo 3.1 del RIA estén presentes de forma continua a lo largo de las fases del ciclo de vida de los sistemas de IA. Por el contrario, la definición reconoce que determinados elementos pueden aparecer en una fase, pero no persistir a lo largo de ambas. Este enfoque para definir un sistema de IA refleja la complejidad y diversidad de los sistemas de IA, garantizando que la definición se ajusta a los objetivos de la Ley de IA al dar cabida a una amplia gama de sistemas de IA.

³ "An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment". <https://oecd.ai/en/ai-principles>:

⁴ "For the purposes of this Convention, "artificial intelligence system" means a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations or decisions that may influence physical or virtual environments. Different artificial intelligence systems vary in their levels of autonomy and adaptiveness after deployment". <https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treaty-num=225>:

⁵ Anexo a la *Communication to the Commission Approval of the content of the draft Communication from the Commission - Commission Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act)*, C(2025) 924 final. <https://digital-strategy.ec.europa.eu/es/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application>

2. LA IA COMO SISTEMA ASOCIADO A UNA MÁQUINA

3. Según el art. 3.1 del RIA, un sistema de IA es un sistema basado en una máquina. Todos los sistemas de IA requieren de máquinas para permitir su funcionamiento. Las máquinas se convierten, por tanto, en el punto de referencia esencial para construir estos sistemas. Así lo ponen de manifiesto el Considerando 12 del Reglamento (UE) 2023/1230 del Parlamento Europeo y del Consejo, de 14 de junio de 2023, relativo a las máquinas, y por el que se derogan la Directiva 2006/42/CE del Parlamento Europeo y del Consejo y la Directiva 73/361/CEE del Consejo.

“(12) Recientemente se han introducido en el mercado máquinas más avanzadas, que no dependen tanto de los operadores humanos. Estas máquinas trabajan en tareas definidas y en entornos estructurados, pero pueden aprender a realizar nuevas acciones en este contexto y hacerse más autónomas. Otras mejoras que ya se han incorporado o cabe esperar que se incorporen a las máquinas tienen que ver con el tratamiento de la información en tiempo real, la resolución de problemas, la movilidad, los sistemas de sensores, el aprendizaje, la adaptabilidad, y la capacidad de operar en entornos no estructurados (por ejemplo, obras de construcción). El informe de la Comisión sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica, de 19 de febrero de 2020, establece que la aparición de nuevas tecnologías digitales, como la inteligencia artificial, el internet de las cosas y la robótica, entraña nuevos retos para la seguridad de los productos. El informe concluye que la legislación vigente en materia de seguridad de los productos, incluida la Directiva 2006/42/CE, presenta una serie de resquicios jurídicos a este respecto que deben corregirse. Por tanto, el presente Reglamento debe cubrir los riesgos para la seguridad derivados de las nuevas tecnologías digitales”.

Las Directrices para la definición de la IA vienen a precisar que el término “máquina” puede entenderse referido tanto a los componentes de hardware como de software que permiten el funcionamiento del sistema de IA. Los componentes de hardware se refieren a los elementos físicos de la máquina, como unidades de procesamiento, memoria, dispositivos de almacenamiento, unidades de red e interfaces de entrada/salida, que proporcionan la infraestructura para el cálculo. Los componentes de software, por su parte, abarcan el código informático, las instrucciones, los programas, los sistemas operativos y las aplicaciones que gestionan la forma en que el hardware procesa los datos y realiza las tareas.

Precisan las citadas Directrices que “todos los sistemas de IA están basados en máquinas, ya que necesitan máquinas para funcionar, como la formación de modelos, el procesamiento de datos, los modelos predictivos y la toma de decisiones automatizada a gran escala. Todo el ciclo de vida de los sistemas avanzados de IA se basa en máquinas que pueden incluir muchos componentes de hardware o software. El elemento “basado en máquinas” en la definición de sistema de IA subraya el hecho de que los sistemas de IA deben ser impulsados computacionalmente y basarse en operaciones de máquinas”. El término “basado en máquinas” abarca una amplia variedad de sistemas informáticos. Por ejemplo, los sistemas de computación cuántica emergentes más avanzados en la actualidad, que representan una desviación significativa de los sistemas de computación tradicionales, constituyen sistemas basados en máquinas, a pesar de sus principios operativos únicos y del uso de la mecánica cuántica.

En este contexto, debe destacarse que los sistemas de IA pueden funcionar como un producto independiente o como un componente de otro producto, con independencia de que se encuentre integrado físicamente en el producto (integrado) o contribuya a su funcionalidad sin estar presente dicha integración (no integrado).

3. AUTONOMÍA, ADAPTACIÓN, CONSECUCIÓN DE OBJETIVOS, INFERENCIA DE RESULTADOS, LÓGICA PREDICTIVA E INFLUENCIA EN EL ENTORNO. LOS OTROS RASGOS DEFINIDORES DE LA IA

4. Junto al elemento anterior (un sistema basado en una máquina), la definición de IA que utiliza el artículo 3.1 del RIA incorpora seis rasgos definidores: (1) que está diseñado para funcionar con distintos niveles de autonomía; (2) que puede mostrar capacidad de adaptación tras su despliegue; (3) y que, para objetivos explícitos o implícitos; (4) infiere, a partir de la entrada que recibe, cómo generar salidas (5) como predicciones, contenidos, recomendaciones o decisiones (6) que pueden influir en entornos físicos o virtuales.

5. Siguiendo con la definición del artículo 3.1 del RIA, los sistemas de IA están diseñados para funcionar con cierto grado de independencia con respecto a la actuación humana y tienen ciertas capacidades para funcionar sin intervención humana (Considerando 12 del RIA). Es decir, no son sistemas que repiten reglas, sino que pueden valorar situaciones y tomar decisiones de forma autónoma. Esa referencia a *“cierto grado de independencia de acción”* excluye del concepto a los sistemas que están diseñados para funcionar únicamente con participación e intervención del ser humano, ya sea directa, por ejemplo, mediante controles manuales, o indirecta, por ejemplo, mediante controles automatizados basados en sistemas que permiten a los seres humanos delegar o supervisar las operaciones del sistema.

Según las Directrices para la definición de la IA, las nociones de autonomía e inferencia de los sistemas de IA van de la mano: la capacidad de inferencia de un sistema de IA (es decir, su capacidad para generar resultados como predicciones, contenidos, recomendaciones o decisiones que puedan influir en entornos físicos o virtuales) es clave para lograr su autonomía. En todo caso, la capacidad de operar con cierto nivel de autonomía es una condición necesaria para calificar a un sistema como de IA. Por otra parte, esta autonomía de la máquina respecto de las instrucciones humanas dota a estos sistemas de un componente de imprevisibilidad en cuanto al resultado ofrecido.

A este respecto, las Directrices para la definición de la IA consideran que los siguientes sistemas presentan un cierto grado de independencia: Un sistema que requiere información de entrada proporcionada manualmente para generar un resultado de salida por sí mismo, sin que este resultado de salida sea controlado manualmente, o especificado explícitamente por un humano. Un sistema experto que sigue una delegación de automatización de procesos por parte de humanos y que es capaz, basado en la información de entrada proporcionada por un humano, de producir un resultado de salida por sí mismo, como una recomendación.

En suma, concluyen las Directrices, “el nivel de autonomía es una condición necesaria para determinar si un sistema puede considerarse un sistema de IA. Todos los sistemas diseñados para funcionar con un grado razonable de independencia de las acciones cumplen la condición de autonomía en la definición de sistema de IA”.

6. De igual modo, los sistemas de IA pueden mostrar capacidad de adaptación tras su despliegue o puesta en funcionamiento. Se refiere a una capacidad de autoaprendizaje que permite al sistema cambiar mientras está en uso y producir nuevos resultados ante los mismos datos de entrada (Considerando 12 del RIA).

Según las Directrices para la definición de la IA, los conceptos de autonomía y de capacidad de adaptación son distintos, pero están muy relacionados y representan dos dimensiones diferentes de la funcionalidad de un sistema de IA. Además, esa característica les hace poder hacer cosas para las que no han sido explícitamente programados y entrenados, e incluso desviarse del comportamiento esperado. Sin embargo, las Directrices explican también, que el uso de la expresión *“puede”* en relación con este elemento de la definición indica que un sistema puede, pero no necesariamente tiene, que poseer esa capacidad de adaptación o autoaprendizaje. Por tanto, no tiene por qué ser una característica decisiva para considerar si un sistema es o no de

IA. La capacidad de aprendizaje de los sistemas de IA después de su introducción en el mercado o puesta en servicio es objeto de atención en dos preceptos del Reglamento, concretamente, en el artículo 15.4 del RIA, al regular la exigencia de precisión, solidez y ciberseguridad de los sistemas de IA de alto riesgo y en el artículo 43.4 del RIA, que regula la evaluación de conformidad a la que deben ser sometidos los sistemas de IA de alto riesgo⁶.

7. Los sistemas de IA se dirigen a alcanzar objetivos implícitos o explícitos. La referencia a objetivos explícitos o implícitos se refiere a que los sistemas de IA no funcionan conforme a reglas, sino a objetivos que pueden ser definidos y explícitos o implícitos, consecuencia de su programación y su proceso de aprendizaje. Los objetivos explícitos, indican Directrices para la definición de la IA, se refieren a fines claramente definidos que se encuentran directamente codificados por el desarrollador del sistema. Los objetivos implícitos, por su parte, son aquellos que no están explícitamente establecidos pero que se deducen de su comportamiento o de sus implicaciones subyacentes y que pueden derivarse de sus datos de entrenamiento o de la interacción del sistema con su entorno.

El objetivo de un sistema de IA está muy vinculado a su finalidad, en los términos en los que se encuentra definida por el artículo 3.12 del RIA, entendida como *"el uso para el que un proveedor concibe un sistema de IA, incluidos el contexto y las condiciones de uso concretos, según la información facilitada por el proveedor en las instrucciones de uso, los materiales y las declaraciones de promoción y venta, y la documentación técnica"*. Esta finalidad debe aparecer reflejada en la documentación técnica que debe acompañar a los sistemas de IA de alto riesgo, con el contenido que se describe en el Anexo IV del Reglamento.

8.. El cuarto elemento de un sistema de IA es que debe ser capaz de inferir, a partir de la entrada que recibe, cómo generar salidas. El Considerando 12 del RIA aclara que "[una] característica clave de los sistemas de IA es su capacidad de inferir". Como se explica con más detalle en dicho considerando, los sistemas de IA deben distinguirse de "los sistemas de software o enfoques de programación tradicionales más sencillos y no deben abarcar los sistemas que se basan en las reglas definidas únicamente por personas físicas para ejecutar automáticamente operaciones". Esta capacidad de inferir es, por tanto, una condición clave e indispensable que distingue a los sistemas de IA de otros tipos de sistemas.

El Considerando 12 del RIA también explica que "esta capacidad de inferencia se refiere al proceso de obtención de resultados, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos y virtuales, y a una capacidad de los sistemas de IA para derivar modelos o algoritmos, o ambos, a partir de entradas o datos". Esta interpretación del concepto de "inferencia" no contradice la norma ISO/IEC 22989, que define la inferencia "como el razonamiento mediante el cual se derivan conclusiones a partir de premisas conocidas", y esta norma incluye una nota específica de la IA que afirma: "[e]n la IA, una premisa es un hecho, una regla, un modelo, una característica o datos brutos".

9. Los sistemas de IA pueden generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones. La capacidad de obtener resultados de salida, como predicciones, contenidos, recomendaciones o decisiones basados en los datos de entrada, utilizando técnicas de aprendizaje automático (*machine learning*) o basadas en la lógica o el conocimiento, es fundamental para identificar un sistema de IA y diferenciarlo de otras formas de software.

Las Directrices para la definición de la IA señalan que la forma más habitual de resultado obtenido por un sistema de IA son las predicciones. A estos efectos se considera predicción la estimación de un valor desconocido (*output*) a partir de un valor conocido que se proporciona

⁶ Un riguroso estudio de esta materia y, en particular, de la definición de IA, puede encontrarse en FERNÁNDEZ HERNÁNDEZ, C., *Guía Práctica del Reglamento de Inteligencia Artificial*, Madrid, Wolters Kluwer, 2025.

al sistema (input). Aunque los sistemas de software se vienen utilizando desde hace décadas para realizar predicciones, los sistemas de IA basados en técnicas de aprendizaje automático son capaces de generar predicciones que revelan patrones ocultos en los datos y de realizar predicciones precisas en entornos muy dinámicos y complejos.

El Considerando 12 del RIA también explica que la definición de sistema de IA debe distinguir los sistemas de IA de "los sistemas de software o enfoques de programación tradicionales más sencillos y no debe abarcar los sistemas que se basan en las reglas definidas únicamente por personas físicas para ejecutar operaciones automáticamente". Precisas las Directrices que algunos sistemas tienen la capacidad de inferir de forma limitada pero, sin embargo, pueden quedar fuera del ámbito de aplicación de la definición de sistema de IA debido a su capacidad limitada para analizar patrones y ajustar de forma autónoma su producción.

10. El quinto elemento de la definición de sistema de IA del artículo 3.1 del RIA es que el sistema infiera "*cómo generar resultados tales como predicciones, contenidos, recomendaciones o decisiones que puedan influir en entornos físicos o virtuales*". La capacidad de un sistema de generar resultados, como predicciones, contenidos y recomendaciones, basándose en los datos de entrada que recibe y utilizando el aprendizaje automático y enfoques lógicos y basados en el conocimiento, es fundamental para lo que hacen los sistemas de IA y lo que distingue a esos sistemas de otras formas de software. La capacidad de generar resultados y el tipo de resultados que el sistema puede generar es fundamental para entender la funcionalidad y el impacto de un sistema de IA.

Los resultados de los sistemas de IA pertenecen a cuatro grandes categorías enumeradas en el artículo 3,1 del RIA: predicciones, contenidos, recomendaciones y decisiones.

Como se ocupan de precisar las Directrices, cada categoría difiere en su nivel de implicación humana.

Las *predicciones* son una de las salidas más comunes que producen los sistemas de IA y que menos implicación humana requieren. Una predicción es una estimación de un valor desconocido (la salida) a partir de valores conocidos suministrados al sistema (la entrada). Los sistemas de software se han utilizado durante décadas para generar predicciones. Los sistemas de IA que utilizan el aprendizaje automático son capaces de generar predicciones que descubren patrones complejos en los datos y realizan predicciones precisas en entornos muy dinámicos y complejos.

Por *contenido* se entiende la generación de nuevo material por parte de un sistema de IA. Puede incluir texto, imágenes, vídeos, música y otras formas de salida. Cada vez hay más sistemas de IA que utilizan modelos de aprendizaje automático (por ejemplo, basados en tecnologías de Transformación Generativa Preentrenada (GPT)) para generar contenidos.

Las *recomendaciones* son sugerencias de acciones, productos o servicios específicos a los usuarios basadas en sus preferencias, comportamientos u otros datos.

Las *decisiones* son conclusiones o elecciones realizadas por un sistema. Un sistema de IA que produce una decisión automatiza procesos que tradicionalmente se gestionan mediante el juicio humano. Un sistema de este tipo implica un proceso totalmente automatizado por el que se produce un determinado resultado en el entorno que rodea al sistema sin ninguna intervención humana.

11. El sexto elemento de la definición de un sistema de IA es que sus resultados "*pueden influir en entornos físicos o virtuales*". Este elemento debe entenderse en el sentido de que los sistemas de IA no son pasivos, sino que influyen activamente en el entorno en los que se despliegan. La referencia a "entornos físicos o virtuales" indica que la influencia de un sistema de IA puede afectar tanto a objetos físicos tangibles (por ejemplo, un brazo robótico) como a entornos virtuales, incluidos espacios digitales, flujos de datos y ecosistemas de software.

II. SISTEMAS DE IA PROHIBIDOS EN EL ECOSISTEMA LABORAL

1. DELIMITACIÓN DE SISTEMA DE IA PROHIBIDO

12. El desarrollo de la IA está asociado, de manera inescindible, a los múltiples riesgos que conlleva su incorporación a la dinámica económica y social en su más amplio sentido. Así lo pone de manifiesto la RIA que articula su regulación, precisamente, sobre una aproximación desde la idea de “riesgo”, definida como *“la combinación de la probabilidad de que se produzca un daño y la gravedad de dicho daño”* (de forma que, a mayor probabilidad de que se produzca el daño y mayor la gravedad de este, mayor riesgo, y viceversa). El RIA establece una jerarquía de riesgos en función del uso de la IA y sobre las categorías detectadas, establece una serie de obligaciones cuyas proyecciones sobre lo laboral resultan más que evidentes.

También señala la norma que, a la vez, la IA también puede menoscabar los intereses públicos y los derechos fundamentales que protege el Derecho de la Unión, de una forma que puede ser tangible o intangible e incluir perjuicios físicos, psíquicos, sociales o económicos (Considerando 5 del RIA). En concreto, el Considerando 28 del RIA recuerda que: *“Al margen de los múltiples usos beneficiosos de la IA, esta también puede utilizarse indebidamente y proporcionar nuevas y poderosas herramientas para llevar a cabo prácticas de manipulación, explotación y control social. Dichas prácticas son sumamente perjudiciales 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”*.

Sobre esta base, el artículo 5 del RIA, prohíbe una serie de sistemas de IA que se consideran una clara amenaza para la seguridad, los medios de vida y los derechos de las personas. Aunque el Reglamento prevé que su entrada en vigor se producirá a los veinte días de su publicación en el Diario Oficial de la Unión Europea y que será aplicable veinticuatro meses a partir de la fecha de su entrada en vigor (artículo 113, primer párrafo), el Considerando 179 advierte que *“Teniendo en cuenta el riesgo inaceptable asociado a determinadas formas de uso de la IA”*, las prohibiciones, así como las disposiciones generales del Reglamento, deben aplicarse ya a los seis meses a partir de la fecha de entrada en vigor del Reglamento, es decir, el 2 de febrero de 2025. Una entrada en vigor que deberá tenerse en cuenta a efectos de las sanciones previstas en el artículo 99 del RIA.

Un importante elemento para asegurar la aplicación homogénea del Reglamento, en relación con las prácticas prohibidas, se recoge en el número 1 del artículo 96, según el cual *“La Comisión elaborará directrices sobre la aplicación práctica del Reglamento y, en particular, sobre: ... b) las prácticas prohibidas a que se refiere el artículo 5”*. Estas directrices se hicieron públicas, de forma aun no oficial, por medio del anexo a la Comunicación de la Comisión por la que se establecen directrices sobre prácticas prohibidas de 4 de febrero de 2025 (*“Directrices sobre sistemas de IA prohibidos”*)⁷. En cuanto a la aplicación de las prohibiciones establecidas por el artículo 5 del RIA, las directrices señalan que la misma es competencia de las autoridades de vigilancia del mercado que deberán estar designadas el 2 de agosto de 2025⁸.

⁷ *Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)*, C (2025) 884 final, <https://digital-strategy.ec.europa.eu/es/library/commission-publishes-guidelines-prohibited-artificial-intelligence-ai-practices-defined-ai-act>.

⁸ En España, esa designación aparece contemplada en el Anteproyecto de ley para el buen uso y gobernanza de la IA aprobado el 11 de marzo de 2025. El objetivo del Capítulo IV (Infracciones y sanciones) desarrolla este marco punitivo, incluyendo la determinación de multas y su eventual aplicación a autoridades y organismos públicos.

2. PROHIBICIÓN DE SISTEMAS DE RECONOCIMIENTO DE EMOCIONES EN EL LUGAR DE TRABAJO

13. En este ámbito quedan directamente proscritos los sistemas de reconocimiento de emociones⁹. El RIA expresamente prohíbe *“la introducción en el mercado, la puesta en servicio o la utilización de sistemas de IA para inferir las emociones de una persona física en los ámbitos de la aplicación de la ley (...) en lugares de trabajo (...)”* (art. 5.1 f) del RIA).

Con explica el Considerando 18 del RIA *“el concepto de «sistema de reconocimiento de emociones» a que hace referencia el presente Reglamento debe definirse como un sistema de IA destinado a distinguir o deducir las emociones o las intenciones de las personas físicas a partir de sus datos biométricos. El concepto se refiere a emociones o intenciones como la felicidad, la tristeza, la indignación, la sorpresa, el asco, el apuro, el entusiasmo, la vergüenza, el desprecio, la satisfacción y la diversión. No incluye los estados físicos, como el dolor o el cansancio, como, por ejemplo, los sistemas utilizados para detectar el cansancio de los pilotos o conductores profesionales con el fin de evitar accidentes. Tampoco incluye la mera detección de expresiones, gestos o movimientos que resulten obvios, salvo que se utilicen para distinguir o deducir emociones. Esas expresiones pueden ser expresiones faciales básicas, como un ceño fruncido o una sonrisa; gestos como el movimiento de las manos, los brazos o la cabeza, o características de la voz de una persona, como una voz alzada o un susurro”*.

Las Directrices sobre sistemas de IA prohibidos proponen los siguientes ejemplos. Así, un sistema de IA que infiere que un empleado está descontento, triste o enojado con los clientes (por ejemplo, a partir de gestos corporales, un ceño fruncido o la ausencia de una sonrisa) es un sistema de reconocimiento de emociones. Sin embargo, utilizar sistemas de reconocimiento de IA para inferir la fatiga de un piloto o conductor profesional, alertarlo y sugerirle cuándo frenar para evitar accidentes no es reconocimiento de emociones, ya que el reconocimiento de emociones no incluye estados físicos como el dolor o la fatiga.

14. Explican las Directrices sobre sistemas de IA prohibidos que la noción de “lugar de trabajo” debe interpretarse en sentido amplio. En principio se refiere a cualquier espacio físico o virtual específico donde las personas físicas desempeñan tareas y responsabilidades asignadas por su empleador o por la organización a la que están asignadas, como en el caso del trabajo por cuenta propia. Esto incluye cualquier entorno donde se realiza el trabajo y puede abarcar desde oficinas interiores, fábricas y almacenes hasta espacios de acceso público como tiendas, estadios o museos, pasando por lugares al aire libre o automóviles, así como lugares de trabajo temporales o móviles.

Esta noción se considera independiente de la condición de empleado, contratista, aprendiz, voluntario, etc. Las Directrices precisan que el concepto de *“lugar de trabajo”* del artículo 51 f) del RIA también debe entenderse aplicable a los candidatos durante el proceso de selección y contratación, en consonancia con otras disposiciones Reglamento que abordan la comercialización, la puesta en servicio o el uso de sistemas de IA en el ámbito del empleo, la gestión de los trabajadores y el acceso al trabajo por cuenta propia, dado que existe un desequilibrio de poderes y la naturaleza intrusiva del reconocimiento de emociones puede aplicarse ya en la fase de contratación.

Las Directrices sobre sistemas de IA prohibidos señalan, por ejemplo, que está prohibido el uso de cámaras web y sistemas de reconocimiento de voz por parte de un *call center* para rastrear las emociones de sus empleados, como el enfado. Pero si estos sistemas solo se implementan para fines de formación del personal, y si sus resultados no se comparten con los

⁹ Tema del que se ha preocupado MUÑOZ RUIZ, A. B., *Biometría y sistemas automatizados de reconocimiento de emociones: Implicaciones Jurídicos-Laborales*, Valencia, Tirant lo Blanch, 2023.

responsables de recursos humanos, de forma que no puedan afectar a la evaluación o la promoción de la persona afectada, pueden considerarse permitidos siempre que no se eluda la prohibición y el uso del sistema de reconocimiento de emociones no afecte la relación laboral. De igual modo, no se considera incluido en la prohibición, el uso de sistemas de reconocimiento de voz por parte de un centro de llamadas para rastrear las emociones de sus clientes, como el enfado o la impaciencia cuando se utilizan para ayudar a los empleados a lidiar con clientes enfadados.

Sin embargo, las Directrices sobre sistemas de IA prohibidos, consideran que se encuentran prohibidos los sistemas de IA que monitorizan el tono emocional cuando se utilizan equipos de trabajo híbridos (de imagen y voz) mediante la identificación e inferencia de emociones a partir de la voz y las imágenes de las videollamadas. Igualmente, se prohíbe el uso de sistemas de IA de reconocimiento de emociones durante el proceso de contratación o durante el período de prueba. Está también prohibido el uso de cámaras en supermercados para rastrear las emociones de sus empleados, como su felicidad. Sin embargo, no se considera prohibido el uso de cámaras en supermercados o bancos para detectar clientes sospechosos, por ejemplo, para inferir que alguien está a punto de cometer un robo, siempre que se garantice que no se esté rastreando a ningún empleado y existan suficientes garantías.

15. La prohibición del artículo 5.1 f) del RIA contiene una excepción explícita de los sistemas de reconocimiento de emociones utilizados en el ámbito laboral y en instituciones educativas por razones médicas o de seguridad, como los sistemas de uso terapéutico. Señalan las Directrices sobre sistemas de IA prohibidos que, dado el objetivo del RIA de garantizar un alto nivel de protección de los derechos fundamentales, esta excepción debe interpretarse de forma restrictiva. De esta manera, se indica que, en particular, debe entenderse que los usos terapéuticos se refieren a los usos de productos sanitarios con marcado CE. Además, esta excepción no comprende el uso de sistemas de reconocimiento de emociones para detectar aspectos generales del bienestar. La monitorización general de los niveles de estrés en el lugar de trabajo no está permitida en el ámbito de la salud o la seguridad. Por ejemplo, un sistema de IA destinado a detectar el síndrome de burnout o la depresión en el lugar de trabajo o en instituciones educativas no estaría cubierto por la excepción y seguiría prohibido.

3. PROHIBICIÓN DE SISTEMAS DE CATEGORIZACIÓN BIOMÉTRICA QUE CLASIFIQUEN INDIVIDUALMENTE A LAS PERSONAS FÍSICAS

16. Igualmente se encuentra prohibida *“la introducción en el mercado, la puesta en servicio para este fin específico o el uso de 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 (...)”* (art. 5.1 g) del RIA). El artículo 3.40 del RIA define como “sistema de categorización biométrica” al sistema de IA “destinado a incluir a las personas físicas en categorías específicas en función de sus datos biométricos, a menos que sea accesorio a otro servicio comercial y estrictamente necesario por razones técnicas objetivas”. Explica el Considerando 16 del RIA que ese concepto debe definirse como la inclusión de personas físicas en categorías específicas en función de sus datos biométricos. Estas categorías específicas pueden referirse a aspectos como el sexo, la edad, el color del pelo, el color de los ojos, los tatuajes, los rasgos conductuales o de la personalidad, la lengua, la religión, la pertenencia a una minoría nacional o la orientación sexual o política.

III. SISTEMAS DE ALTO RIESGO EN LO LABORAL

1. ÁMBITO DE APLICACIÓN DE LOS SISTEMAS DE ALTO RIESGO EN LO LABORAL

17. Pieza esencial del sistema que construye el RIA se asienta en el establecimiento de límites al uso de los sistemas que califica de “alto riesgo” (art. 6.1 y 2 por relación con lo establecido en el Anexo III del RIA). Una noción que, como precisa el Considerando 46 del RIA, se

diseña tomando en cuenta sus efectos, por cuanto incluye entre tales sistemas aquellos que *“tengan consecuencias perjudiciales importantes para la salud, la seguridad y los derechos fundamentales de las personas de la Unión, y dicha limitación reduce al mínimo cualquier posible restricción del comercio internacional, si la hubiera”*.

Lo laboral ocupa, también aquí, un papel protagonista. Y es que, como viene a subrayar el Considerando 48 del RIA, *“la magnitud de las consecuencias adversas de un sistema de IA para los derechos fundamentales protegidos por la Carta es especialmente importante a la hora de clasificar un sistema de IA como de alto riesgo. Entre dichos derechos se incluyen el derecho a la dignidad humana, el respeto de la vida privada y familiar, la protección de datos de carácter personal, la libertad de expresión y de información, la libertad de reunión y de asociación, la no discriminación, (y) los derechos de los trabajadores (...)”*.

La lectura del art. 6.2 en relación con el Anexo III de RIA ratifica de modo concluyente la referida proyección a nuestro campo. Dicho Anexo III incorpora entre los sistemas de IA de alto riesgo los que afecten al *“empleo, gestión de los trabajadores y acceso al autoempleo”* y, en concreto, a) *“Sistemas de IA destinados a ser utilizados para la contratación o la selección de personas físicas, en particular para publicar anuncios de empleo específicos, analizar y filtrar las solicitudes de empleo y evaluar a los candidatos”* y b) *sistemas de IA destinados a utilizarse para tomar decisiones o influir sustancialmente en ellas que afecten a la iniciación, promoción y resolución de relaciones contractuales de índole laboral, a la asignación de tareas basada en la conducta individual o en rasgos o características personales, o al seguimiento y evaluación del rendimiento y la conducta de las personas en el marco de dichas relaciones”*.

El Considerando (57) justifica dicha inclusión sobre la base de que los mismos: *“pueden perpetuar patrones históricos de discriminación, por ejemplo, contra las mujeres, ciertos grupos de edad, las personas con discapacidad o las personas de orígenes raciales o étnicos concretos o con una orientación sexual determinada, durante todo el proceso de contratación y en la evaluación, promoción o retención de personas en las relaciones contractuales de índole laboral. Los sistemas de IA empleados para controlar el rendimiento y el comportamiento de estas personas también pueden socavar sus derechos fundamentales a la protección de los datos personales y a la intimidad”*. Veamos algunos ejemplos.

No obstante, lo anterior, el Considerando 53 del RIA advierte de que pueden existir casos específicos en los que un sistema de IA incluido en alguno de los ámbitos específicos anteriores no se considere de alto riesgo ya que no entraña un riesgo considerable de causar un perjuicio a los intereses jurídicos amparados por dichos ámbitos, por no influir sustancialmente en la toma de decisiones o no perjudicar sustancialmente dichos intereses. En consecuencia, el artículo 6.3 del RIA establece que un sistema de IA no se considerará de alto riesgo si no supone un riesgo significativo de daño para la salud, la seguridad o los derechos fundamentales de las personas físicas, al no influir sustancialmente en el resultado de la toma de decisiones. En suma, como precisa, el citado Considerando 53, por sistema de IA que no influye sustancialmente en el resultado de la toma de decisiones debe entenderse un sistema de IA que no afecta al fondo, ni por consiguiente al resultado, de la toma de decisiones, ya sea humana o automatizada.

18. El RIA incorpora, por ello, un importante sistema de garantías que se anudan a los requisitos generales que deberán cumplir los sistemas de IA de alto riesgo *“teniendo en cuenta sus finalidades previstas, así como el estado actual de la técnica generalmente reconocido en materia de IA”* (art. 8). Ello lleva consigo el establecimiento, implantación, documentación y mantenimiento de un sistema de gestión de riesgos entendido como *“un proceso iterativo continuo planificado y ejecutado durante todo el ciclo de vida de un sistema de IA de alto riesgo, que requerirá revisiones y actualizaciones sistemáticas periódicas”* (art. 9). A ello se une la necesaria *“gobernanza de los datos”* (art. 10), las exigencias de una precisa documentación técnica (art. 11) y la necesidad de garantizar un nivel de trazabilidad del funcionamiento del sistema (art. 12). La transparencia suficiente para que los responsables del despliegue interpreten y usen correctamente la información de salida que incorporen los sistemas de IA, constituye un principio

maestro (art. 13) al que se une la necesidad de que su diseño y desarrollo permita cumplir con el principio de humano al mando (art. 14). En fin, unos sistemas que se diseñarán y desarrollarán de modo que alcancen un nivel adecuado de precisión, solidez y ciberseguridad y funcionen de manera uniforme durante todo su ciclo de vida (art. 15).

El requisito de “vigilancia humana” es especialmente relevante en las relaciones laborales, puesto que un trabajador o un grupo de trabajadores se encargarán de ella. Como puso de relieve en su día el CESE, estos trabajadores deben recibir formación sobre cómo llevar a cabo esta tarea. El RIA otorga carácter de principio a lo que denomina “alfabetización en materia de inteligencia artificial” que enuncia en su art. 4 y conceptúa en el art. 3.56 del siguiente modo: *“Los proveedores y responsables del despliegue de sistemas de IA (en el caso laboral, los empleadores) adoptarán medidas para garantizar que, en la mayor medida posible, su personal y demás personas que se encarguen en su nombre del funcionamiento y la utilización de sistemas de IA tengan un nivel suficiente de alfabetización en materia de IA, teniendo en cuenta sus conocimientos técnicos, su experiencia, su educación y su formación, así como el contexto previsto de uso de los sistemas de IA y las personas o los grupos de personas en que se utilizarán dichos sistemas”*. Además, dado que se espera que estos trabajadores puedan ignorar el resultado del sistema de IA o incluso no utilizarlo, deben establecerse medidas de protección frente a las represalias que pueda sufrir en aquellos casos en que la persona trabajadora no tome la correspondiente decisión.

Además de las obligaciones aplicables con carácter general a los sistemas de IA de alto riesgo, el Sección 3 del Título III, titulado “Obligaciones de los proveedores y responsables del despliegue de los sistemas de IA de alto riesgo y de otras partes”, establece las obligaciones que, en relación con los sistemas IA deben cumplir determinados actores que intervienen en su diseño, fabricación, comercialización y uso.

19. En concreto, los responsables del despliegue de los sistemas de IA que incorporen estas técnicas se encuentran obligados a cumplir con un ineludible presupuesto [art. 26.7 y Considerando 92 del RIA]: *“Antes de poner en servicio o utilizar un sistema de IA de alto riesgo en el lugar de trabajo, los responsables del despliegue que sean empleadores informarán a los representantes de los trabajadores y a los trabajadores afectados de que estarán expuestos a la utilización del sistema de IA de alto riesgo. Esta información se facilitará, cuando proceda, con arreglo a las normas y procedimientos establecidos en el Derecho nacional y de la Unión y conforme a las prácticas en materia de información a los trabajadores y sus representantes”*. Una exigencia esencial si atendemos a los términos del Considerando 92 del RIA que considera que *“este derecho de información es accesorio y necesario para el objetivo de protección de los derechos fundamentales que subyace al presente Reglamento (...)”* y, además, se impone *“incluso aunque no se cumplan las condiciones de las citadas obligaciones de información o de información y consulta previstas en otros instrumentos jurídico (la referencia se hace a la Directiva 2002/14/CE)”*.

Las anteriores obligaciones quedan subrayadas por el art. 26.11 del RIA cuando señala que *“sin perjuicio de lo dispuesto en el artículo 50 del presente Reglamento, los responsables del despliegue de los sistemas de IA de alto riesgo a que se refiere el anexo III que tomen decisiones o ayuden a tomar decisiones relacionadas con personas físicas informarán a las personas físicas de que están expuestas a la utilización de los sistemas de IA de alto riesgo (...)”*. Por su parte, el art. 50 RIA establece que: *“los proveedores garantizarán que los sistemas de IA destinados a interactuar directamente con personas físicas se diseñen y desarrollen de forma que las personas físicas de que se trate estén informadas de que están interactuando con un sistema de IA, excepto cuando resulte evidente desde el punto de vista de una persona física razonablemente informada, atenta y perspicaz, teniendo en cuenta las circunstancias y el contexto de utilización”*.

20. Una exigencia que se ve reforzada por el artículo 86 del RIA que reconoce el derecho a recibir una explicación individualizada de aquellos usos de la IA que afecten a una persona trabajadora. Esto supone que *“toda persona que se vea afectada por una decisión que el responsable del despliegue adopte basándose en los resultados 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 pro-*

duzca 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, tendrá derecho a 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 adoptada”.

21. La Sección 5 de este Título III del RIA (“Normas, evaluación de conformidad, certificados y registro”) incorpora la exigencia de autocertificación de los sistemas de alto riesgo, proceso mediante el cual los proveedores de dichos sistemas evalúan y declaran que cumplen con los requisitos y estándares establecidos normativamente. El fundamento de dicha exigencia se extrae del Considerando 125 del RIA que viene a precisar que: *“Dada la complejidad de los sistemas de IA de alto riesgo y los riesgos asociados a ellos, es importante desarrollar un sistema adecuado para el procedimiento de evaluación de la conformidad de los sistemas de IA de alto riesgo en el que participen organismos notificados, denominado «evaluación externa de la conformidad». No obstante, habida cuenta de la experiencia actual de los profesionales que realizan la certificación previa a la comercialización en el campo de la seguridad de los productos y de la distinta naturaleza de los riesgos implicados, 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”.*

Establece el art. 43.2 del RIA que: “En el caso de los sistemas de IA de alto riesgo mencionados en los puntos 2 a 8 del anexo III, los proveedores se atenderán al procedimiento de evaluación de la conformidad fundamentado en un control interno a que se refiere el anexo VI, que no contempla la participación de un organismo notificado”. De este modo, los sistemas de IA laborales quedan incluidos dentro del radio de la autocertificación. Una exigencia que lleva consigo para los proveedores de estos sistemas el cumplimiento de requisitos: (i) La comprobación de que el sistema de gestión de la calidad establecido reúne los requisitos establecidos en el artículo 17 del RIA. (ii) Examinar 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. (iii) Asimismo, comprobar que el proceso de diseño y desarrollo del sistema de IA y la vigilancia poscomercialización del mismo a que se refiere el artículo 72 son coherentes con la documentación técnica. Con todo, el modelo de autocertificación plantea numerosas dudas y hace más necesaria que nunca la consolidación de empresas tecnológicamente responsables.

Igualmente, y antes de poner en servicio una IA laboral, todos los empleadores/empresas deberán registrar dicha IA en una base de datos de la UE, indicando que es una IA bajo su responsabilidad (art. 49 del RIA). Esta base de datos se regula en el art. 71 del RIA y “únicamente contendrá datos personales en la medida en que sean necesarios para la recogida y el tratamiento de información de conformidad con el presente Reglamento. Dicha información incluirá los nombres y datos de contacto de las personas físicas responsables del registro de sistema y que cuenten con autoridad legal para representar al proveedor o al responsable del despliegue, según proceda”.

IV. LA RESPONSABILIDAD POR EL USO INDEBIDO DE LOS SISTEMAS DE IA

22. El sistema de responsabilidades por el uso de la IA se encuentra integrado por dos subsistemas de diferente alcance y contenido: de un lado, el subsistema punitivo, basado en sanciones administrativas y asentado sobre una pluralidad de agentes que tutelan tal actividad de control; y, de otro, el subsistema reparador, asentado en la lógica que proporcional el derecho de daños¹⁰.

¹⁰ De este tema me he ocupado en MERCADER UGUINA, J. R., “Sistema de responsabilidades por el uso de la inteligencia artificial. Un enfoque integral”, *Labos*, 2024, vol. 5 (extraordinario), pp. 211-227.

Las dificultades a la hora de aplicar y hacer efectivas las técnicas de control y sanción son múltiples. La autonomía y el carácter opaco de los sistemas de IA, es decir, la dificultad de comprender y explicar cómo han tomado sus decisiones, por las propias características de la tecnología que utilizan, como sucede en el caso de algunos métodos de aprendizaje profundo (*deep learning*), dificulta de modo especial, tanto la proyección de régimen sancionador, como “la prueba no solo de la culpa sino también de la relación de causalidad”¹¹. La interconectividad es otra de las características distintivas de los productos que incorporan sistemas de IA que pueden plantear numerosos problemas a la hora de detectar y controlar su alcance.

23. El texto de la RIA aprobado por el Consejo no introduce disposiciones concretas sobre responsabilidad civil. Impone, ciertamente, un gran número de obligaciones a los así llamados operadores de sistemas de IA, expresión que incluye una gran variedad de sujetos: proveedores, distribuidores e importadores de sistemas de IA, responsables del despliegue de sistemas de IA, fabricantes de productos que introduzcan en el mercado o pongan en servicio un sistema de IA junto con su producto y con su propio nombre o marca; y representantes autorizados de los proveedores de sistemas o modelos de IA. La RIA, sin embargo, se limita a regular la supervisión gubernamental de la actividad de dichos operadores y a imponerles régimen sancionatorio administrativo por la transgresión de aquellas obligaciones en los términos que hemos analizado.

Pero el daño ha de diferenciarse de la infracción administrativa en cuanto que la misma puede servir para castigar la creación de riesgos que llegarán a generar daños o a quedarse, simplemente, en esferas de peligro que no se hayan llegado a concretar en daños efectivos. De este modo, las personas afectadas por una infracción del RIA deben tener también derecho a recibir una compensación por los daños efectivamente sufridos.

24. Las sanciones administrativas están diseñadas para castigar a las entidades que incumplen el reglamento y para disuadir futuras infracciones. Se basan en criterios que buscan asegurar el cumplimiento de las normas y prevenir comportamientos indebidos. En contraste, la indemnización tiene un objetivo reparador, orientado a compensar a las personas afectadas por el daño sufrido como resultado de la infracción. Como expresa el Considerando (168) del RIA, “Se debe poder exigir el cumplimiento del presente Reglamento mediante la imposición de sanciones y otras medidas de ejecución. Los Estados miembros deben tomar todas las medidas necesarias para garantizar que se apliquen las disposiciones del presente Reglamento, incluso estableciendo sanciones efectivas, proporcionadas y disuasorias para las infracciones, lo que incluye respetar el principio de *non bis in idem* (...) El Supervisor Europeo de Protección de Datos debe estar facultado para imponer multas a las instituciones, los órganos y los organismos de la Unión incluidos en el ámbito de aplicación del presente Reglamento”.

Sus efectos parecen, a priori, demoledores. A título de ejemplo, el art. 99.3 RIA establece que: “El no respeto de la prohibición de las prácticas de IA a que se refiere el artículo 5 estará sujeto a 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”. El art. 99.4 establece el régimen sancionador para el incumplimiento de las obligaciones que se establecen por la RIA. A tal efecto precisa que: “El incumplimiento por parte de un sistema de IA de cualquiera de las disposiciones que figuran a continuación en relación con los operadores o los organismos notificados, distintas de los mencionados en el artículo 5, estará sujeto a 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”, resultando a nuestros efectos relevante el apartado e) que proyecta el referido régimen sancionador sobre “las obligaciones de los responsables del despliegue con arreglo al artículo 26”.

¹¹ MARTÍN CASALS, M. “Las propuestas de la Unión Europea para regular la responsabilidad civil por los daños causados por sistemas de inteligencia artificial”, *InDret: Revista para el Análisis del Derecho*, 2023, nº 2, p. 75.

25. Si bien con referencia al RGPD, pero con una proyección que alcanza también al esquema de responsabilidades que resulta aplicable a los sistemas de IA, la Sentencia del Tribunal de Justicia de la Unión Europea de 20 de junio de 2024, C-590/22, ha distinguido las distintas funciones y fines que poseen la indemnización por daños y perjuicios y las sanciones administrativas. El Tribunal de Justicia de la Unión Europea aclara que los criterios utilizados para imponer multas administrativas no deben aplicarse a la determinación del importe de la indemnización. La indemnización debe reflejar el daño real sufrido por la persona, mientras que las multas administrativas cumplen una función diferente, que es principalmente disuasoria y correctiva en relación con la conducta de las entidades responsables. Ciertamente, “al mantener esta distinción, el Tribunal de Justicia de la Unión Europea protege el derecho a una compensación adecuada para las víctimas sin que se vean afectadas por las medidas punitivas destinadas a los infractores. Ello también previene la posibilidad de que las entidades sean penalizadas en exceso por las mismas infracciones, ya que las sanciones administrativas y las indemnizaciones tienen propósitos distintos y deben aplicarse de manera separada”¹².

BIBLIOGRAFÍA CITADA

- FERNÁNDEZ HERNÁNDEZ, C., *Guía Práctica del Reglamento de Inteligencia Artificial*, Madrid, Wolters Kluwer, 2025.
- FIERRO RODRÍGUEZ, D., “Análisis de los daños derivados de infracciones administrativas del Reglamento General de Protección de Datos a la luz del asunto C-590/22”, *Práctica de Derecho de Daños*, 2024, n.º 160.
- MARTÍN CASALS, M., “Las propuestas de la Unión Europea para regular la responsabilidad civil por los daños causados por sistemas de inteligencia artificial”, *InDret: Revista para el Análisis del Derecho*, 2023, n.º 2, p. 75.
- MERCADER UGUINA, J. R., “Sistema de responsabilidades por el uso de la inteligencia artificial. Un enfoque integral”, *Labos*, 2024, vol. 5 (extraordinario), pp. 211-227.
- MUÑOZ RUIZ, A. B., *Biometría y sistemas automatizados de reconocimiento de emociones: Implicaciones Jurídicos-Laborales*, Valencia, Tirant lo Blanch, 2023.

¹² FIERRO RODRÍGUEZ, D., “Análisis de los daños derivados de infracciones administrativas del Reglamento General de Protección de Datos a la luz del asunto C-590/22”, *Práctica de Derecho de Daños*, 2024, n.º 160.