Unión Europea

EL REGLAMENTO DE IA: EL PRIMER PASO DEL CAMINO HACIA UNA REGULACIÓN COMPLETA DE LA INTELIGENCIA ARTIFICIAL

Salvador Espinosa de los Monteros Pérez-Brotóns
y Gonzalo Sanz Setién

Abogados del Área de Mercantil de Uría Menéndez (Madrid)

El Reglamento de IA: el primer paso del camino hacia una regulación completa de la inteligencia artificial

En medio de la vertiginosa digitalización que caracteriza a la Cuarta Revolución Industrial, las autoridades europeas han sido pioneras en la regulación de la inteligencia artificial con la aprobación de un nuevo Reglamento por el que se establecen normas armonizadas en este ámbito para la Unión Europea.

El presente artículo analiza los aspectos esenciales del Reglamento que tendrán un impacto significativo en la regulación de la inteligencia artificial en otras jurisdicciones, como en Estados Unidos o en China. El Reglamento de inteligencia artificial busca alcanzar un delicado equilibrio entre la protección de los derechos fundamentales y el apoyo a la innovación. Esta nueva normativa define los conceptos clave de la materia, establece un sistema de control basado en el riesgo e impone una serie de requisitos y obligaciones sobre los sistemas de inteligencia artificial y sus operadores, además de fijar un concreto régimen sancionador.

PALABRAS CLAVE:
Inteligencia artificial, Enfoque basado en el riesgo, Seguridad, Innovación, Derechos fundamentales.

EA AI Act: a first step towards full regulation

In the midst of the meteoric rise of digitalisation that epitomises the Fourth Industrial Revolution, European authorities are leading the way when it comes to regulating AI by adopting a new piece of legislation that provides the harmonised rules in this area for the European Union.

This article analyses the main aspects of the EU AI Act, which will affect how AI is regulated in other jurisdictions, such as the United States and China. The Act seeks to strike a delicate balance between protecting fundamental rights and fostering innovation. This new piece of legislation defines the key concepts in the field, establishes a risk-based control system and imposes a series of requirements and obligations on AI systems and their operators; it also includes specific penalties.

KEYWORDS:
Artificial Intelligence (AI), Risk-Based Approach, Security, Innovation, Fundamental Rights.

FECHA DE RECEPCIÓN: 3-6-2024

FECHA DE ACEPTACIÓN: 23-6-2024

Espinosa de los Monteros Pérez-Brotóns, Salvador; Sanz Setién, Gonzalo (2024). El Reglamento de IA: el primer paso del camino hacia una regulación completa de la inteligencia artificial. *Actualidad Jurídica Uría Menéndez*, 65, pp. 180-196 (ISSN: 1578-956X).

1. Introducción

Después de un largo proceso de negociaciones, el Consejo de la Unión Europea (el "Consejo de la UE") y el Parlamento Europeo (el "Parlamento") alcanzaron el 9 de diciembre de 2023 un histórico acuerdo sobre una propuesta normativa en materia de inteligencia artificial ("IA"). Meses más tarde, el 13 de marzo de 2024, el Parlamento aprobó, en primera lectura, la propuesta de "Reglamento por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión" y, el 21 de mayo de 2024, el Consejo dio definitivamente luz verde a la primera norma global sobre IA mediante la aprobación del Reglamento 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.º 300/2008, (UE) n.º 167/2013, (UE) n.º 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 (publicado en el DOUE de 12 de julio de 2024 —el "Reglamento"—).

El Reglamento se enmarca en el contexto de la "Estrategia Europea de IA", mediante la que las instituciones europeas aspiran a convertir a la Unión Europea ("UE" o "Unión") en un centro de excelencia mundial en materia de IA, siendo la primera norma jurídica de calado sobre la materia, tal y como ya logró la UE en otros ámbitos, como la protección de datos.

Como abordaremos más adelante, el Reglamento busca compaginar el fomento de la inversión y la protección de los derechos fundamentales, la democracia y el Estado de derecho, para lo que ha adoptado un enfoque basado en el riesgo, que pretende limitar las obligaciones aplicables a aquellos sistemas de IA que presenten un riesgo bajo.

1.1. Antecedentes en la normativa europea

La rápida evolución de la IA en los últimos años ha puesto de manifiesto la necesidad de dotar a esta tecnología de un marco jurídico propio. A tal efecto, la UE ha venido otorgando a la IA un papel político primordial, a través de un enfoque que garantice una IA segura y fiable, centrada en el ser humano y que respete los valores europeos, establecidos en la Carta de Derechos Fundamentales de la UE.

El debate a nivel europeo sobre la materia comenzó ya en febrero de 2017, cuando el Parlamento instó a la Comisión Europea (la "Comisión") a evaluar el impacto de la IA y preparar una estrategia europea para regular esta tecnología. Asimismo, como prueba de la gran importancia que otorga a la IA, el Parlamento continuó con su actividad en este campo, y adoptó durante los años 2020 y 2021 una serie de resoluciones en las que instaba de nuevo a la Comisión a establecer un marco jurídico de principios éticos para el desarrollo, despliegue y uso de la IA, la robótica y las nuevas tecnologías.

En la misma línea, el Consejo Europeo ha insistido también en la necesidad de adoptar normas comunes sobre IA. Baste resaltar a estos efectos sus Conclusiones de 19 de octubre 2017 y 11 de febrero de 2019, así como particularmente las del 9 de junio de 2020, en las que pedía a la Comisión que presentase propuestas concretas para garantizar que la IA se usara de un modo proporcionado y que minimizara sus riesgos.

En vista de estas solicitudes, la Comisión adoptó en abril de 2019 el primer instrumento de *soft law* sobre IA con la publicación de las *Directrices éticas para una IA fiable*, en las que ya ponía el foco en la necesidad del respeto a los derechos fundamentales. Posteriormente, puso en marcha el proceso legislativo mediante la publicación, en febrero de 2020, del *Libro Blanco sobre la inteligencia artificial*, y el comienzo, en esa misma fecha, de una consulta pública sobre la materia. Además, en abril de 2021, la Comisión publicó una evaluación de impacto, en la que identificó algunos de los principales problemas a abordar en relación con los sistemas de IA, tales como su opacidad, complejidad, comportamiento autónomo, etc.

Como puede verse, estas primeras aproximaciones reflejan ya las preocupaciones clave de las instituciones europeas, que se centran en garantizar la fiabilidad de los sistemas de IA y minimizar el riesgo de vulneración de derechos fundamentales. De hecho, los resultados de la consulta pública antes mencionada fueron precisamente la base del primer borrador del Reglamento.

1.2. La IA en otras jurisdicciones

Sin perjuicio del liderazgo normativo que la UE ha tratado de ejercer en materia de IA, otras jurisdicciones no se han mostrado ajenas a esta tecnología y han venido ofreciendo respuestas de distinto calado e intensidad normativa. Por citar algunos ejemplos:

i. Estados Unidos partió de un enfoque liberal, más permisivo con el desarrollo de estos sistemas, para después aumentar su intervención. Así, en octubre de 2023, la Casa Blanca publicó un compendio de principios para orientar el diseño, uso e implementación de

sistemas automatizados, con el objetivo de proteger los derechos de sus ciudadanos en la era de la IA. En la misma línea, el presidente Joe Biden promulgó en esas mismas fechas una orden ejecutiva que establecía, por primera vez, obligaciones sobre los sistemas de IA, buscando proteger la intimidad de sus ciudadanos y fomentar la competencia en este mercado.

ii. China ha sido, desde 2017, pionera en la creación de una estrategia nacional de IA, que se ha centrado en fomentar el desarrollo del país como un *hub* tecnológico en materia de inteligencia artificial. En particular, el Gobierno del país asiático se ha mostrado muy receptivo a las preocupaciones de la industria, buscando así tomar la delantera en la carrera por el desarrollo de *large language models*. A nivel normativo, este enfoque se ha plasmado en una ley general reguladora de la IA y en un reglamento sobre la gestión de los servicios de IA generativa que, pese a ser instrumentos novedosos, adoptan un enfoque poco restrictivo.

iii. En cuanto a Reino Unido, el Gobierno ha optado por un enfoque basado en principios normativos más que en nuevas obligaciones de *compliance*. Así, su estrategia se ha centrado en promover la innovación, regulando la IA a través de leyes existentes, y en promover un marco general que pueda ser interpretado y aplicado por los reguladores sectoriales.

iv. A nivel internacional cabe resaltar, entre otras, las iniciativas adoptadas por la OCDE (a través de la adopción de sus *Principios no vinculantes sobre la IA* en 2019 y que se ha actualizado en mayo de 2024), la UNESCO (que aprobó en noviembre de 2021 la *Recomendación sobre la ética de la IA*, adoptada por los 193 Estados miembros), el G7 (que publicó los *Principios rectores internacionales del proceso de Hiroshima para un sistema avanzado de IA*) y el Consejo de Europa (que adoptó el primer tratado internacional sobre IA, denominado *Convenio Marco del Consejo de Europa sobre Inteligencia Artificial y derechos humanos, democracia y Estado de derecho*).

2. Objetivos y ámbito de aplicación

2.1. Objetivos

Como hemos adelantado en la sección anterior, el enfoque europeo en materia de IA se ha centrado en un objetivo principal: garantizar un elevado nivel de protección de la salud, la seguridad, los derechos fundamentales, la democracia, el Estado de derecho y el medio ambiente.

Naturalmente, esta aproximación ha regido los pasos de las instituciones europeas con respecto al Reglamento, que, conforme a las palabras de la propia Comisión, tiene como objetivos específicos: (i) garantizar que la IA se use en el mercado europeo de forma fiable y segura; (ii) mejorar la gobernanza y la aplicación efectiva de la normativa vigente en materia de derechos fundamentales; (iii) facilitar el desarrollo de un mercado único de IA; y (iv) garantizar la seguridad jurídica, para impulsar la inversión e innovación en IA.

2.1.1. SEGURIDAD Y DERECHOS FUNDAMENTALES

Como resulta evidente, y resalta el Reglamento ya en sus considerandos, los sistemas de IA presentan importantes implicaciones éticas y de seguridad, lo que hace especialmente relevante su solidez técnica y fiabilidad. A modo de ejemplo, el propio Reglamento cita algunos casos de usos "preocupantes" de los sistemas de IA, como pueden ser los sistemas de *social scoring*, sistemas de identificación biométrica remota en tiempo real en espacios públicos, bases de datos de reconocimiento facial, sistemas para la supervisión y detección de comportamientos prohibidos durante exámenes, etc.

Por ello, el objetivo a este respecto del Reglamento es garantizar que los sistemas de IA sean resilientes en caso de dificultades y frente a intentos de alterar su funcionamiento, así como que tengan capacidad para impedir usos ilícitos por terceros y minimizar daños potenciales.

En definitiva, el Reglamento muestra la plena conciencia que tienen las instituciones europeas acerca de la posibilidad de que un mal uso de estos sistemas tenga consecuencias muy perjudiciales a nivel social, buscando ofrecer soluciones y mecanismos de gobernanza y de protección *ex ante* para intentar que estas posibles externalidades negativas queden todo lo minimizadas que sea posible.

2.1.2. INVERSIÓN

Pese a su preocupación por los riesgos, el Reglamento demuestra también que la UE es consciente del alto potencial de los sistemas de IA, que ofrecen grandes oportunidades para el desarrollo económico y social. De este modo, la norma busca alcanzar un complicado equilibrio entre seguridad e inversión, que permita, por un lado, garantizar el respeto a los derechos fundamentales y la minimización de riesgos y, por otro, fomentar la creación de riqueza y el desarrollo tecnológico.

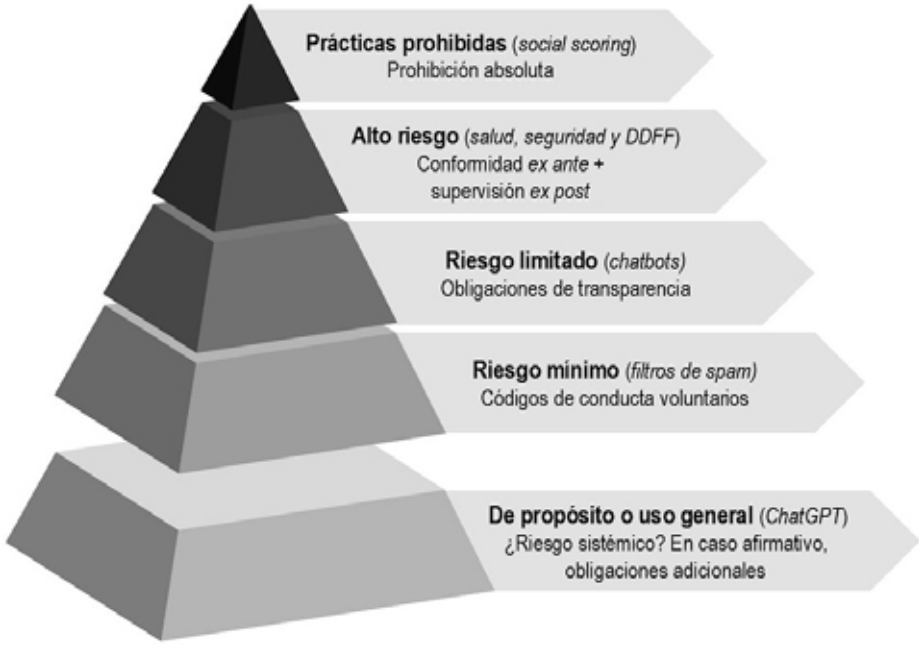
Así, el Reglamento prohíbe, con algún matiz, la imposición de restricciones al desarrollo, la comercialización y la utilización de sistemas de IA, creando un mercado único y homogéneo para este tipo de tecnologías. Gracias a ello, los sistemas de IA podrán circular por toda la UE sin restricciones y desplegarse con facilidad en sectores muy diversos.

Con esta misma idea, algunos de los Estados miembros de la UE han ido adelantándose a la aprobación definitiva del Reglamento. Por poner un ejemplo cercano, a finales de 2023 España aprobó el Real Decreto 817/2023, que establecía un entorno controlado de pruebas (o *sandbox*) para el ensayo del cumplimiento de la propuesta de Reglamento. De este modo, nuestro país creaba un marco normativo que permitía que los potenciales afectados por el Reglamento fuesen adaptándose de un modo más sencillo y progresivo a la nueva normativa, sin que la regulación impactase de un modo tan intenso en su actividad de producción.

2.1.3. ENFOQUE BASADO EN RIESGOS

Como es bien sabido, conciliar seguridad e inversión no siempre es fácil, pues a menudo la imposición de obligaciones excesivamente onerosas puede dificultar o ralentizar la apuesta por el desarrollo de este tipo de tecnologías en aquellos territorios que optan por enfoques normativos más estrictos.

Para intentar alcanzar este objetivo, la apuesta de la UE ha sido optar por un enfoque basado en riesgos: a mayor riesgo, mayor número de obligaciones. Así, el Reglamento adapta el tipo y contenido de sus normas a la intensidad y alcance de los riesgos que presentan los sistemas de IA sobre los que se aplica. Para ello, establece los siguientes niveles de riesgo:



Se trata, en definitiva, de buscar un marco jurídico sólido, pero flexible (sin imponer obligaciones cuando no se crea que es estrictamente necesario) y, sobre todo, que pueda resistir el paso del tiempo. Aunque, lógicamente, aún es pronto para saber si la UE ha acertado en este último punto, que es clave en una tecnología con potencial de desarrollo exponencial, no cabe duda de que resultará particularmente relevante al evaluar el éxito del Reglamento en un futuro.

2.2. Ámbito de aplicación

Una vez explicado el origen de la norma, así como los principales objetivos que persigue, el siguiente paso para entender el Reglamento es aclarar su ámbito de aplicación, tanto objetivo como subjetivo. A tal efecto, en esta sección abordaremos primero la definición de sistemas de IA que plantea el Reglamento, para después analizar quiénes son los sujetos obligados por la norma.

2.2.1. SISTEMAS DE INTELIGENCIA ARTIFICIAL: ¿QUÉ ES LA IA?

Aunque pueda parecer una pregunta sencilla, el intenso debate al respecto protagonizado por las instituciones europeas demuestra la gran cantidad de matices y dificultades que entraña la definición de IA. Finalmente, el legislador europeo ha optado por una definición amplia y tecnológicamente neutra, que sigue de forma prácticamente idéntica la usada por la OCDE, precisamente con el fin de "*facilitar la convergencia a escala internacional y una amplia aceptación*". Concretamente, se entiende por "*sistema de IA*", a efectos del Reglamento:

"*Un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que puede influir en entornos físicos o virtuales*".

Como se puede observar, la definición pretende distinguir estos sistemas de los sistemas de *software* o programación tradicionales, excluyendo así aquellos sistemas que ejecutan automáticamente operaciones a partir de instrucciones específicas. La clave para el legislador europeo se halla aquí en la capacidad de inferencia, que se construye a través del aprendizaje automático y que trasciende de un mero tratamiento básico de datos. Además, naturalmente, incluye la referencia a que los sistemas de IA son sistemas "*basados en una máquina*", en el sentido de que se ejecutan en y por una máquina.

2.2.2. SUJETOS OBLIGADOS: REGLAMENTO DE PRODUCTO

La gran importancia que desde los medios de comunicación se le da habitualmente a tecnologías como la IA hace que a menudo se generen expectativas sobre su regulación. No obstante, debe aclararse que el Reglamento es una norma "de producto", que pretende establecer una serie de requisitos y obligaciones para aquellas empresas que introduzcan en el mercado europeo, pongan en servicio, importen, distribuyan, fabriquen y desplieguen sistemas de IA. Por consiguiente, el impacto del Reglamento en el día a día de las personas físicas será bastante limitado, sin perjuicio de su capacidad de hacer valer los derechos que el Reglamento les otorga y de poder beneficiarse de mayores garantías de fiabilidad y seguridad al usar sistemas de IA.

Por otra parte, y debido al fenómeno mundial de la deslocalización empresarial, el Reglamento opta por un enfoque territorial muy amplio, que pretende abarcar todos los sistemas de IA que afecten (de forma directa o indirecta) a personas que se hallen en el territorio de la UE. De este modo, y sin ánimo de entrar en un detalle exhaustivo que excede del ámbito de este artículo, se puede afirmar que la norma impone obligaciones sobre:

i. Proveedores que introduzcan en la UE o pongan en servicio sistemas de IA o modelos de IA de uso general, con independencia de si están establecidos o ubicados en la UE. A tal efecto, el Reglamento considera *proveedores* a las personas, autoridades u otros órganos (a) que desarrollen sistemas de IA o modelos de IA de uso general; (b) para los que un

tercero desarrolle un sistema de IA o un modelo de IA de uso general y lo introduzca en el mercado; o (c) que pongan en servicio un sistema de IA con su propio nombre o marca.

ii. Responsables del despliegue, si están establecidos o ubicados en la UE o si los *outputs* del sistema se utilizan en la UE, y entendiendo por tal a personas, autoridades u otros órganos que usen un sistema de IA bajo su propia autoridad (excepto si el uso se produce en el marco de una actividad no profesional).

iii. Importadores y distribuidores en la UE de sistemas de IA.

iv. Fabricantes de otros productos que introduzcan o pongan en servicio en la UE un sistema de IA junto con su propio producto y con su propio nombre o marca.

Como excepción, el Reglamento excluye de su ámbito de aplicación, entre otros, los sistemas y modelos de IA (i) en el ámbito militar, de defensa o seguridad nacional (con independencia de la entidad que lleve a cabo dichas actividades); (ii) desarrollados específicamente y puestos en servicio únicamente con fines de investigación y desarrollo científicos; y (iii) divulgados con arreglo a licencias libres y de código abierto (excepto si se trata de sistemas de alto riesgo, prohibidos o que estén destinados a interactuar directamente con personas físicas).

3. Prácticas de IA prohibidas

Tal y como hemos mencionado en la sección 2.1.3, el legislador europeo considera que algunos sistemas de IA presentan riesgos inaceptables para los derechos fundamentales y los valores de la UE, y en consecuencia establece una "lista negra" de sistemas de IA "prohibidos", que se ha ido ampliando respecto a la propuesta inicial de la Comisión.

Así, no se permite la introducción en el mercado o el uso de sistemas de IA que:

i. Se sirvan de técnicas subliminales o deliberadamente manipuladoras o engañosas para alterar el comportamiento de las personas.

ii. Exploten vulnerabilidades derivadas de la edad, la discapacidad o la situación social o económica.

iii. Clasifiquen a individuos o grupos basándose en su comportamiento social o características personales (*i. e.*, *social scoring*).

iv. Evalúen el riesgo de que una persona cometa delitos penales basándose únicamente en la elaboración de perfiles o en rasgos y características de su personalidad.

v. Creen o amplíen bases de datos de reconocimiento facial a través de la búsqueda no selectiva en internet o en grabaciones de circuitos cerrados de televisión.

vi. Reconozcan las emociones de una persona en lugares de trabajo o centros educativos, salvo por motivos médicos o de seguridad.

vii. Infieran, con sistemas de categorización biométrica, la raza, las opiniones políticas, la afiliación sindical, las creencias religiosas o filosóficas, la vida sexual o la orientación sexual de las personas (excepto el etiquetado o filtrado de datos biométricos con fines policiales).

viii. Utilicen la identificación biométrica remota en tiempo real en espacios públicos (con excepciones, como la búsqueda de víctimas o sospechosos de secuestro o explotación sexual).

En general, el Reglamento considera que estas prácticas son perjudiciales y generan en la ciudadanía sentimientos de vigilancia masiva, además de que pueden resultar discriminatorias. No obstante, el listado no es un *numerus clausus*, ya que la Comisión deberá llevar a cabo una evaluación sobre la necesidad de modificar la lista de prácticas prohibidas una vez al año. De este modo, se intenta garantizar que la norma no quede desactualizada debido a los rápidos avances tecnológicos.

4. Sistemas de IA de alto riesgo

El siguiente escalón en la pirámide de riesgos establecida por el Reglamento es el referido a los sistemas de IA de alto riesgo. A continuación abordamos su clasificación y obligaciones.

4.1. Clasificación

El Reglamento considera como de alto riesgo a aquellos sistemas de IA que presentan riesgos significativos para la salud, la seguridad o los derechos fundamentales. Se trata, por tanto, de sistemas que ofrecen riesgos relevantes, pero no tan grandes como para meritar su prohibición.

A este respecto, cabe mencionar que el legislador europeo ha querido diferenciar los sistemas de IA de alto riesgo en sentido estricto de aquellos productos que encajan en la definición, pero —por el tipo de producto o el sector— se hallan ya regulados por normativa europea armonizada (*e. g.*, productos sanitarios, equipos radioeléctricos). Para estos últimos, el Reglamento establece obligaciones adicionales más ligeras, entendiendo que la salud y la seguridad de los usuarios quedan suficientemente protegidas por la propia normativa específica.

En relación con el resto de sistemas de alto riesgo, el Reglamento identifica una serie de familias de sistemas de IA que pueden causar un perjuicio relevante para la salud, la seguridad o los derechos fundamentales, que listamos a continuación:

i. Identificación biométrica y reconocimiento de emociones, en la medida en que su uso no esté prohibido.

ii. Gestión de infraestructuras críticas (*e. g.*, gestión del tráfico, gas, agua o electricidad).

iii. Educación y formación profesional (*e. g.*, acceso o admisión a centros educativos, detección de comportamientos prohibidos durante los exámenes).

iv. Selección de personal y relaciones laborales (*e. g.*, contratación de personal, evaluación del rendimiento de los trabajadores).

v. Gestión del acceso a servicios esenciales, públicos y privados (*e. g.*, clasificación de llamadas de emergencia, contratación de seguros de vida y salud).

vi. Actividades de las fuerzas y cuerpos de seguridad (*e. g.*, evaluación del riesgo de que una persona sea víctima o cometa un delito, valoración de pruebas y sospechosos).

vii. Migración, asilo y gestión del control fronterizo (*e. g.*, polígrafos, examen de visados, solicitudes de asilo o permisos de residencia).

viii. Administración de justicia y procesos democráticos (*e. g.*, procesos electorales, referéndums o asistencia a los jueces en la investigación e interpretación de la ley y de los hechos).

Sin perjuicio de lo anterior, el hecho de que un sistema se encuadre dentro de estas familias no implicará necesariamente su consideración como de alto riesgo, pues el artículo 6.3 del Reglamento establece un filtro adicional que permite excluir de las obligaciones correspondientes a aquellos sistemas incluidos en el listado que no planteen un riesgo significativo para la salud, la seguridad o los derechos fundamentales. Así, se limitan las obligaciones de aquellos sistemas utilizados principalmente para tareas procedimentales o preparatorias. En todo caso, un sistema de IA siempre se considerará de alto riesgo si elabora perfiles de personas físicas.

4.2. Requisitos y obligaciones

De conformidad con los artículos 8 y siguientes del Reglamento, el hecho de que un sistema de IA sea considerado como de alto riesgo implica, fundamentalmente, que al menos deberá cumplir con los siguientes requisitos:

i. Que el sistema de IA conserve la documentación técnica que demuestre el cumplimiento de los requisitos pertinentes y proporcione a las autoridades competentes la información necesaria para evaluar su conformidad.

ii. Que el sistema de IA tenga implantado un sistema de gestión de riesgos, que deberá mantenerse actualizado y medirá los riesgos previsibles del sistema y las medidas diseñadas para hacerles frente.

iii. Que el sistema de IA permita la llevanza de un registro automático de eventos a lo largo del ciclo de vida del producto que garantice una adecuada trazabilidad.

iv. Que el sistema de IA se diseñe y se desarrolle para que cuente con un nivel de transparencia suficiente para que los responsables del despliegue interpreten y usen correctamente su información de salida.

v. Que el sistema de IA se diseñe y se desarrolle de tal forma que pueda ser vigilado de manera efectiva por personas físicas.

vi. Que el sistema de IA se diseñe y se desarrolle de modo que alcance un nivel adecuado de precisión, solidez, ciberseguridad y funcionamiento uniforme durante todo su ciclo de vida.

vii. Que, en caso de que el sistema de IA utilice técnicas de entrenamiento con datos, se diseñe sobre la base de unas prácticas de gobernanza y de gestión de los datos de entrenamiento, validación y prueba que evite que el producto incurra en sesgos.

No obstante, como veremos a continuación, este listado no comprende todas las consecuencias de que un sistema sea considerado como de alto riesgo. Y es que el Reglamento reconoce que en la puesta en marcha, comercialización y funcionamiento de los sistemas de IA de alto riesgo hay numerosos actores implicados y, por ello, establece obligaciones detalladas y exhaustivas para toda la cadena de valor de la IA. Estas obligaciones, que abordan un amplio abanico de medidas de gobernanza e intervenciones técnicas, deben (i) aplicarse durante las fases de diseño y desarrollo (conformidad *ex ante*) y (ii) supervisarse y mantenerse durante todo el ciclo de vida del producto (supervisión *ex post*).

Así, por un lado, los proveedores de sistemas de IA de alto riesgo deberán fundamentalmente:

i. velar por que los sistemas de IA de alto riesgo cumplan con los requisitos mencionados previamente;

ii. contar con un sistema de gestión de calidad documentado y mantenido en el tiempo;

iii. conservar la documentación técnica y archivos de registro generados automáticamente;

iv. llevar a cabo una vigilancia tras la comercialización y adoptar las medidas correctoras necesarias para garantizar la conformidad del producto;

v. colocar el marcado CE en los sistemas de IA de alto riesgo;

vi. colaborar con las autoridades nacionales competentes; y

vii. cerciorarse de que los sistemas de IA de alto riesgo superen las evaluaciones de conformidad y elaborar una declaración UE de conformidad.

Por otro, los importadores y distribuidores de sistemas de IA de alto riesgo:

i. comprobarán que el sistema de IA ha superado la evaluación de conformidad correspondiente;

ii. custodiarán la documentación técnica;

iii. se asegurarán de que el sistema lleve el marcado CE exigido y vaya acompañado de la declaración UE de conformidad y de las instrucciones de uso; y

iv. colaborarán con las autoridades nacionales competentes en lo que resulte menester.

Por último, los responsables del despliegue de sistemas de IA de alto riesgo:

i. adoptarán las medidas técnicas y organizativas adecuadas para garantizar el uso de los sistemas de IA de alto riesgo con arreglo a sus instrucciones de uso;

ii. encomendarán la supervisión humana a personas físicas que tengan la competencia, formación y autoridad necesarias;

iii. conservarán los archivos de registro que los sistemas de IA de alto riesgo generen automáticamente;

iv. informarán a los representantes de los trabajadores y a los trabajadores afectados antes de poner en servicio o utilizar un sistema de IA de alto riesgo en un lugar de trabajo; y

v. colaborarán con las autoridades nacionales competentes.

vi. Adicionalmente, en caso de ser entidades públicas o de utilizar sistemas de IA para realizar calificaciones crediticias o evaluar los riesgos de contratar seguros de vida y salud, los responsables del despliegue llevarán a cabo una evaluación del impacto de dichos sistemas en los derechos fundamentales.

5. Obligaciones de transparencia

Continuando con los escalones inferiores de la pirámide, se encuentra un tipo de sistema de IA que a menudo genera preocupaciones entre el público general: aquel que plantea riesgos específicos de suplantación de identidad o engaño. Se está hablando, en este caso, de sistemas que interactúan con personas físicas (*e. g.*, *chatbots*), o que generan o manipulan contenidos de imagen, audio o vídeo (*e. g.*, sistemas de generación automática de contenido, *deep fakes*, etc.).

En estos casos, y con independencia de que se les impongan obligaciones adicionales si entrasen en la categoría de sistemas de alto riesgo (conforme a lo visto en el apartado 4.1), el Reglamento se limita a imponerles obligaciones de información y transparencia. Dichas obligaciones, que varían ligeramente en función de la finalidad del sistema (*i. e.*, si está destinado a interactuar directamente con personas físicas, si generan contenido de texto, etc.), esencialmente implican que los proveedores y/o responsables del despliegue deberán informar a los usuarios o consumidores de que están interactuando con una máquina o de que la información o contenido que están viendo ha sido generado de forma artificial.

6. Sistemas de IA de riesgo mínimo y códigos de conducta

Finalmente, cerrando la pirámide, destacan todos aquellos sistemas de IA que no encajan en ninguno de los supuestos anteriores y que se consideran "de riesgo mínimo". Algunos ejemplos son los filtros de *spam*, los sistemas automatizados de recomendación, etc.

Para estos casos, que el Reglamento entiende que presentan un riesgo muy limitado, no se imponen obligaciones de ningún tipo, sino que únicamente se incentiva la adhesión a, y suscripción voluntarias de, códigos de conducta y buenas prácticas, con el objetivo de alentar a los proveedores de este tipo de sistemas a cumplir de forma voluntaria con altos estándares de seguridad y fiabilidad. A tal efecto, el Reglamento prevé la elaboración de códigos de conducta, a los que habrá que sumar los que elaboren las autoridades comunitarias y nacionales en materia de alfabetización sobre IA, destinados todos ellos a las personas encargadas del desarrollo, el manejo y el uso de esta tecnología.

7. Modelos de IA de uso general

7.1. Clasificación

El Reglamento complementa su enfoque basado en riesgos con normas específicas para los modelos de IA de uso general, que define como aquellos modelos de IA "*que presentan un grado considerable de generalidad, que son capaces de realizar de manera competente una gran variedad de tareas distintas y que tienen la capacidad para integrarse con posterioridad en otros sistemas o aplicaciones*". Como se puede apreciar, el Reglamento opta así por una definición poco específica, pero centrada en la generalidad y versatilidad que caracterizan a los modelos de IA de uso general.

Esta regulación, incluida en respuesta al auge de los modelos fundacionales, como son los *large language models*, pretende abordar las especificidades de un tipo de modelos con capacidades de gran impacto, en particular en aquellos casos en que pueden llegar a presentar un riesgo sistémico. Así, por ejemplo, el Reglamento refleja su preocupación por el riesgo de que estos modelos generen accidentes graves, perturbaciones en sectores críticos, consecuencias graves para la salud y la seguridad, efectos negativos sobre los procesos democráticos y la seguridad pública y económica, o contenidos ilícitos, falsos o discriminatorios que se puedan difundir con facilidad.

No obstante, considerando que la definición de sistemas de IA de uso general es amplia, el Reglamento establece un umbral de operaciones que activa la presunción de que un modelo tiene capacidades de gran impacto y presenta un riesgo sistémico. Dicho umbral, que deberá irse actualizando, está basado en la cantidad de cálculo utilizado para su entrenamiento, e inicialmente se ha fijado en un FLOP superior a 10^25 (a efectos comparativos, el entrenamiento de Chat GPT 3 requirió un total de 3*10^23 FLOP). En caso de contar con capacidades de gran impacto, el proveedor del modelo deberá notificarlo a la Comisión en un plazo de dos semanas desde que

tenga conocimiento de dicha circunstancia. En todo caso, la presunción antes referida podrá ser desvirtuada a instancia del proveedor, que deberá demostrar que el modelo tiene características específicas que impiden que sea clasificado como un modelo con riesgo sistémico.

7.2. Obligaciones

Del mismo modo que para el resto de sistemas de IA, el Reglamento sigue un enfoque basado en el riesgo, que distingue entre obligaciones genéricas —aplicables a todos los proveedores de modelos de IA de uso general— y obligaciones específicas para los proveedores de modelos de IA de uso general que presenten un riesgo sistémico.

Entre las primeras, destacan:

i. elaboración y mantenimiento de documentación técnica;

ii. puesta a disposición de información y documentación para aquellos proveedores que tengan la intención de integrar el modelo de uso general en sus sistemas de IA;

iii. cumplimiento de la legislación de la Unión en materia de derechos de autor durante el entrenamiento de estos sistemas (*e. g.*, aplicación del marco normativo relativo a la minería de datos); y

iv. puesta a disposición del público de resúmenes del contenido utilizado para su entrenamiento.

Por otro lado, los proveedores de modelos de uso general con riesgo sistémico, deberán además, entre otras:

i. evaluar los modelos de IA de conformidad con protocolos y herramientas normalizados que reflejen el estado de la técnica;

ii. identificar y reducir los riesgos sistémicos que puedan derivarse del desarrollo, la introducción en el mercado o el uso de los modelos de IA de uso general con riesgo sistémico;

iii. presentar información adicional sobre sus estrategias de evaluación y sus resultados, descripción de medidas adoptadas para la alineación y puesta a punto de los modelos, y descripción detallada de la arquitectura del sistema utilizado;

iv. vigilar, documentar y comunicar a la oficina de IA y, en su caso, a las autoridades nacionales competentes, la información pertinente sobre incidentes graves y las posibles medidas correctoras para resolverlos; y

v. velar por que se establezca un nivel adecuado de ciberseguridad para el modelo y la infraestructura física del modelo.

8. Otros

8.1. Apoyo a la innovación

El Reglamento, además de fomentar la seguridad y fiabilidad de los sistemas de IA, busca también impulsar y apoyar la innovación en el mercado de la Unión. Para ello, impone a las autoridades nacionales la obligación de establecer espacios controlados de pruebas (o *sandboxes* regulatorios) para el desarrollo y experimentación de sistemas de IA en entornos regulatorios controlados.

Los *sandboxes* buscan ofrecer un entorno seguro para la experimentación, en el cual desarrollar, entrenar, probar y validar sistemas de IA durante un periodo limitado de tiempo antes de su introducción en el mercado. Para facilitar su puesta en marcha, se prevé expresamente que la Comisión pueda proporcionar soporte técnico, asesoramiento y herramientas para la implementación y operación de estos espacios, asegurando que se cumpla la normativa pertinente y ofreciendo seguridad jurídica a los proveedores que participen en ellos.

Asimismo, y en aras de una mayor flexibilidad, se prevé también que los proveedores de sistemas de IA de alto riesgo puedan realizar pruebas en condiciones reales fuera de los *sandboxes*, siempre que cumplan con ciertas condiciones fijadas en el Reglamento y con el "plan de prueba en condiciones reales" que deberá detallar la Comisión mediante un acto de ejecución.

8.2. Gobernanza

Para mejorar la gobernanza a escala de la Unión en materia de IA, el Reglamento prevé la creación de:

i. Una oficina de IA, que tendrá como objetivo reforzar la experiencia y las capacidades de la UE en el ámbito de la IA.

ii. Un comité europeo de IA (el "Comité"), compuesto por un representante de cada Estado miembro, que contará con la participación del Supervisor Europeo de Protección de Datos como observador. El Comité proporcionará asesoramiento y asistencia a la Comisión y a los Estados miembros para facilitar la aplicación coherente y eficaz del Reglamento.

iii. Un foro consultivo, compuesto por miembros de la industria, *start-ups*, pymes, la sociedad civil y el mundo académico, que buscará facilitar la compartición de conocimientos técnicos y asesorar al Comité y a la Comisión en materia de IA.

iv. Un grupo de expertos científicos independientes que asesorarán a la oficina de IA en distintos ámbitos de carácter técnico, como por ejemplo en la alerta de posibles riesgos sistémicos planteados por modelos de IA de uso general.

Por otro lado, a nivel nacional se insta a los Estados miembros a designar (al menos) una autoridad notificante y una autoridad de vigilancia del mercado. Dichos organismos podrán proporcionar orientaciones, asesoramiento y supervisión sobre la aplicación del Reglamento.

8.3. Vigilancia del mercado

De conformidad con la normativa de la Unión en materia de vigilancia del mercado y conformidad de los productos, el Reglamento (i) incluye directrices para la vigilancia posterior a la comercialización, el intercambio de información sobre incidentes graves y la vigilancia del mercado en la UE de los sistemas de IA; y (ii) otorga a las autoridades nacionales de vigilancia del mercado la potestad para exigir la subsanación de incumplimientos y, eventualmente, prohibir la comercialización o exigir la retirada del mercado de los sistemas de IA afectados.

8.4. Régimen sancionador

Continuando con la línea mantenida en normativas anteriores, como el Reglamento General de Protección de Datos, el Reglamento establece un estricto régimen sancionador de límites máximos e insta a los Estados miembros a fijar el régimen de sanciones a nivel nacional (dentro de los límites establecido por el Reglamento), incluyendo penas pecuniarias que podrán alcanzar hasta:

i. el mayor de entre 35 millones de euros o el 7 % del volumen de negocios anual total mundial de la empresa, en caso de incumplimiento del régimen de conductas prohibidas;

ii. el mayor de entre 15 millones de euros o el 3 % del volumen de negocios anual total mundial de la empresa, en caso de incumplimiento de otras disposiciones relacionadas con operadores u organismos notificados, o en caso de infracciones de la normativa sobre modelos de IA de uso general; o

iii. el mayor de entre 7,5 millones de euros o el 1 % del volumen de negocios anual total mundial de la empresa, en caso de suministro de información inexacta, incompleta o engañosa a los organismos notificados o a las autoridades nacionales competentes.

8.5. Entrada en vigor

Pese a que el Reglamento entrará en vigor a los veinte días de su publicación en el *Diario Oficial de la Unión Europea*, no será aplicable de forma general hasta que transcurran veinticuatro meses desde su entrada en vigor. Asimismo, y teniendo en cuenta su gran calado en el sector, prevé un régimen progresivo para su aplicabilidad, conforme al cual:

i. Los capítulos I (Disposiciones Generales) y II (Prácticas de Inteligencia Artificial Prohibidas) serán aplicables a los seis meses de la entrada en vigor del Reglamento.

ii. La sección 4 del capítulo III (Autoridades Notificantes y Organismos Notificados), el capítulo V (Modelos de Uso General), el capítulo VII (Gobernanza), el capítulo XII (Sanciones) y el artículo 78 (Confidencialidad en la Vigilancia del Mercado) serán aplicables a los doce meses de la entrada en vigor del Reglamento, a excepción del artículo 101 (Multas a Proveedores de Modelos de IA de Uso General).

iii. Por último, el apartado 1 del artículo 6 (Sistemas de IA de Alto Riesgo Sujetos a Normativa Europea Armonizada) y las obligaciones correspondientes serán aplicables a los treinta y seis meses desde la entrada en vigor del Reglamento.

9. Conclusiones

La aprobación del Reglamento por la Unión supone un primer paso de gigante en la construcción de un marco regulatorio armonizado en Europa para el desarrollo, uso y comercialización de los sistemas de IA. Con esta norma, que es completamente pionera en cuanto al detalle y nivel de regulación que presenta, las instituciones europeas han tomado la delantera regulatoria a nivel mundial, como ya hicieron en casos anteriores.

No obstante, su éxito (que todavía es incierto) dependerá de la medida en que logre combinar sus dos objetivos principales: garantizar una IA segura y confiable, al tiempo que consiga fomentar la inversión y el desarrollo tecnológico en la Unión. Todavía no es posible apreciar en qué medida el Reglamento ha alcanzado estos objetivos, si bien sí se puede destacar lo evidente: estamos ante una norma muy detallada y exhaustiva que, pese a su enfoque basado en riesgos, tiende a primar la seguridad sobre la innovación. Es por ello por lo que existen ya quienes cuestionan tal nivel de detalle en el Reglamento, esgrimiendo que quizás los pormenores podrían haberse aclarado en actos de ejecución para evitar un excesivo formalismo y/o rigidez. En todo caso, solo la propia experiencia práctica determinará si el compendio de requisitos y obligaciones impuestos son realmente necesarios y razonables, o si las instituciones deben adoptar un enfoque más tolerante al riesgo.

Sea como sea, no cabe duda de que esta norma marcará un antes y un después en el desarrollo del derecho sobre los sistemas de IA, que esperamos con indudable interés.