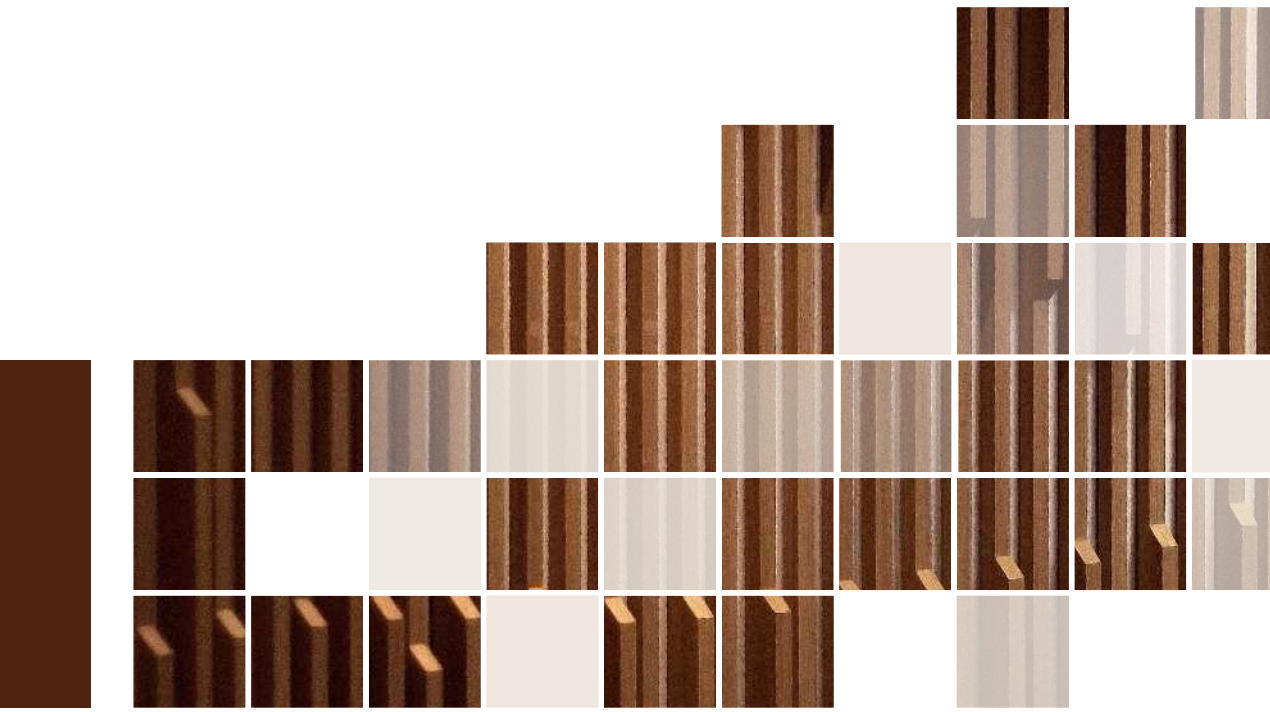
TEMAS

Guía práctica del Reglamento Europeo de Inteligencia Artificial

Manual de referencia para conocer, entender y aplicar la AI Act

Carlos Fernández Hernández



LA LEY

ARANZADI LA LEY, S.A.U.
C/ Collado Mediano, 9
28231 Las Rozas (Madrid)
e-mail: clienteslaley@aranzadilaley.es
Acceso a Soporte: https://areacliente.aranzadilaley.es/solicitud_alta_area_cliente
https://www.aranzadilaley.es

Primera edición: Enero 2025

Depósito Legal: M-684-2025
ISBN versión impresa: 978-84-10292-48-2
ISBN versión electrónica: 978-84-10292-49-9

Diseño, Preimpresión e Impresión: ARANZADI LA LEY, S.A.U.
Printed in Spain

ÍNDICE SISTEMÁTICO

PREFACIO

Las Naciones Unidas, a través de su Organismo de alto nivel para el asesoramiento en materia de inteligencia artificial, afirmaban en septiembre de 2024 que la inteligencia artificial (IA) constituye «el siguiente gran salto en el avance científico» de la humanidad.

No faltan quienes consideran que la innovación es una ideología (entre otros, y en España, Suárez o Aibar), o quienes tildan de «solucionismo tecnológico» (Morozov) la apelación a la tecnología digital para afrontar los problemas de la sociedad actual.

No me sumo desde luego a ellos, a la hora de propugnar que resulta inimaginable prescindir de la tecnología digital, y en especial, de la IA, para hacer frente a retos como el cambio climático, el cuidado de la salud o la pobreza (que todavía en amplias regiones del mundo aqueja a tantos millones de personas). Las ganancias en términos de eficiencia y de productividad que dimanan de la IA están fuera de toda duda, y aplicadas a la resolución de esos problemas, justifican más que sobradamente su utilización.

Europa arrastra un notable atraso respecto de los líderes mundiales, especialmente los Estados Unidos, en lo que se refiere al avance en inteligencia artificial, lo que resulta preocupante, no únicamente por lo recién afirmado —siendo eso ya de por sí bastante—, sino porque según datos de la Comisión Europea, solo la IA generativa implicará un valor en los años inmediatos de entre 2,4 y 4 billones (no anglosajones) de euros.

La propia Comisión Europea indica que la inversión en capital-riesgo en IA ha pasado de 2.300 millones de euros en 2022 a 24.000 millones de euros en 2023, si bien el 90% de dicha inversión se ha llevado a cabo en los Estados Unidos. Otro dato más, en esta misma línea, nos indica que de entre las cien mayores empresas tecnológicas mundiales, solamente existen 11 europeas, mientras que los Estados Unidos cuentan con 59. El muy conocido Informe Draghi, aparecido también en septiembre de 2024, ha hecho notar que alrededor del 70% de los modelos de IA generativa producidos desde 2017 se han desarrollado en los Estados Unidos; así como que Europa se encuentra muy atrás en términos de productividad respecto de los Estados Unidos, hasta el punto de que la renta disponible por habitante ha crecido

el doble en ese país que en la Unión Europea en los últimos 25 años, estando a la vez constatado que es precisamente la tecnología digital la que fundamentalmente ha propiciado este desfase.

A pesar de todo ello, difícilmente se puede negar que la IA presenta serios riesgos. Riesgos medioambientales, en términos de alto consumo de energía motivados por la inmensa voracidad de datos de estos sistemas y modelos; riesgos reputacionales, derivados de malos usos, que pueden redundar en serios perjuicios para determinadas organizaciones; o riesgos jurídicos, que pueden implicar responsabilidades legales, bien por quiebras de seguridad, por violaciones de privacidad, por discriminaciones o por lesiones de normativa de propiedad intelectual o industrial, por solo citar los que quizá resulten más patentes. Ya advertía de estos problemas el gran experto en IA y profesor de Berkeley Stuart Russell, cuando en 2019 avisaba a las empresas tecnológicas de que justo por ello, y tarde o temprano, terminarían siendo reguladas.

Y en este ámbito, el de la regulación digital, qué duda cabe de que Europa constituye la referencia global por excelencia. El propio Informe Draghi menciona que la Unión Europea cuenta con alrededor de 100 normas reguladoras de la tecnología digital, de entre las que obviamente destacan el Reglamento general de protección de Datos, y desde el uno de agosto de 2024, el Reglamento europeo de inteligencia artificial (RIA), al que esta obra se dedica.

Como el libro que aquí se prologa explica con todo detalle, el RIA se ancla fundamentalmente sobre dos bases. La primera es la regulación de la IA como un producto: baste un simple ejemplo, como pudiera ser un televisor. Ello obedece a que uno de los grandes objetivos de esta norma es precisamente el de la seguridad de las personas, en realidad de los consumidores europeos, frente a cualesquiera daños a su integridad y a su salud que su uso pudiera generar. Siendo a su vez una de las consecuencias de lo anterior el que los destinatarios del RIA sean personas, ya físicas o ya jurídicas, quienes, desde su entrada en vigor, y en la medida en que desarrollen o utilicen modelos o sistemas de IA, quedarán por él vinculados. Esta última es de hecho una diferencia crítica entre el RIA y el Convenio Marco del Consejo de Europa sobre inteligencia artificial de mayo de 2024; este Convenio no tiene como destinatarios a personas, sino a Estados, que podrán firmarla y ratificarla, pero que son libres para dar cauce ulterior a sus contenidos en regulaciones que puedan desarrollarla.

La segunda gran base de anclaje del RIA es su regulación de vocación horizontal, a su vez comprensible en un doble sentido. Horizontal por cuanto el Reglamento regula los riesgos que deriven de esta tecnología con independencia del sector o de la actividad en que se aplique. Como ya a estas alturas suele ser bien sabido, el RIA establece riesgos inaceptables, que llevan a prohibir el uso de esta tecnología cuando se generan; riesgos altos, que implican el corazón de su reglamentación y las mayores exigencias; riesgos medios, que no evitan sin embargo que proveedores

y usuarios de modelos o sistemas que los generen queden también sujetos a algunas reglas; y riesgos mínimos, que quedan excluidos del alcance del RIA. Y el segundo sentido en el que la horizontalidad de este Reglamento se despliega es el de resultar aplicable tanto a las tecnologías de IA digamos clásica o analítica; y asimismo a los sistemas y modelos de IA generativa, que como se sabe constituye la última generación de estas tecnologías, y la que ha propiciado su extensión prácticamente indiscriminada en la sociedad. En algún otro lugar la he calificado de «Ford T» de la inteligencia artificial.

Estos dos grandes rasgos del Reglamento europeo son los que permiten trazar una clara frontera de distinción entre el mismo y otro modelo regulatorio determinante en el mundo, como es el norteamericano. Este último, básicamente dimanante del Decreto Presidencial de 30 de octubre de 2023, se centra esencialmente en los riesgos que la IA pueda plantear en materia de seguridad nacional, y no como simple producto de gran consumo. Por otro lado, este foco en la seguridad del Estado evita una regulación de la tecnología en sí, ya se trate de IA analítica o de IA generativa, que prescinda del concreto fin que motive el uso concreto de una u otra.

Es claro que los riesgos de la IA no pueden afrontarse únicamente desde la regulación. En una interesante obra dedicada a los accidentes de aviación, el profesor de la Universidad de Bristol John Downer nos recuerda que el muy reducido riesgo de la aviación no deriva en modo alguno primordialmente de las normas que la regulan, sino del exhaustivo y concienzudo análisis que las compañías aéreas que sufren desastres efectúan de las causas de los mismos, causas que a su vez comparten con otras líneas aéreas. El propio Stuart Russell nos recuerda también cómo, en su experiencia, tratar de aplicar remedios para que una IA existente resulte segura (*make AI safe*), «ni ha funcionado nunca, ni funcionará jamás»; siendo en cambio lo correcto tratar de, *desde el diseño* —podríamos decir en la terminología creada en Canadá a finales de los años 90 y adoptada por la legislación europea—, y en continuo diálogo entre todos los afectados por esta tecnología, ya se trate de diseñadores, ya de gestores empresariales, ya de gobiernos, ya de los propios consumidores, construir una IA que sea segura (*make safe AI*).

Sea como fuere, y aun cuando insuficiente, la regulación no deja de constituir un paso en el sentido correcto. Y es también innegable que el RIA reforzará lo que constituye uno de sus objetivos esenciales, como es el de reforzar los derechos y libertades de los europeos frente a los usos de sistemas y modelos de IA. En esto, por cierto, la norma europea contrasta fuertemente con el otro gran modelo regulatorio mundial, el de la República Popular China, cuyo baremo principal a la hora de estimar aceptable un sistema o modelo no es otro que su compatibilidad con los principios del partido que sustenta a ese Estado.

Carlos Fernández Hernández ha sido un jurista a lo largo y ancho de su carrera. Más recientemente ha comenzado a ser también un estudioso del Derecho y de la regulación digitales. Escribe profusa y competentemente sobre estos temas, que

enseña en algunas de las mejores universidades del país. Me consta además de primerísima mano, que conoce como muy pocas personas en Europa los pormenores de la gestación, y de la estructura y contenidos del RIA. Estamos por todo ello, ante una persona de competencia ideal para escribir esta obra.

Este libro dedicado al Reglamento europeo de inteligencia artificial aportará pues a los estudiosos y prácticos del Derecho, y a cualesquiera otras personas interesadas en él, los conocimientos necesarios para conocer en detalle esta norma señera. Su estilo claro facilitará su manejo. Mientras que la concisión propia de su carácter de manual no impedirá una comprensión adecuada y profunda del tema. Agradezco al autor el honor de prologarlo. Aunque sobre todo le felicito por tan excelente trabajo.

Pablo García Mexía, PhD

Codirector del Posgrado sobre privacidad e IA (UAM)

Director de Derecho digital en Herbert Smith Freehills

Letrado de las Cortes Generales

PRÓLOGO

Llegó la Revolución Industrial y el mundo se transformó completamente a partir de una innovación tecnológica: la máquina de vapor. Cambió la forma de transportarnos, producir y distribuir lo fabricado, así como las relaciones sociales. Muchos oficios desaparecieron. Otros, totalmente desconocidos, surgieron. Y todo en apenas 50 años. Igual ocurrió con los ordenadores en los 80s, pero la transformación ya fue mucho más rápida...

Ahora mismo vivimos inmersos en otra revolución, la de la Inteligencia Artificial. En los primeros años del nuevo milenio se hablaba de un horizonte temporal de mitad de siglo para ver el impacto de estas tecnologías; en 2010 los cambios se anunciaban en décadas, en 2020 se aceleró a lustros y, actualmente, finalizando el 2024, hablamos de las novedades que aparecerán en los próximos meses e incluso semanas. El tsunami (concepto muy manoseado, pero no se me ocurre una analogía mejor) de la IA generativa se llevará por delante muchas profesiones (prefiero no especificar para no herir sensibilidades) y transformará profundamente otras (por ejemplo, la mía, la educativa). Podemos plantear todas las pegas que queramos, pero la tecnología nunca se para, siempre avanza. Existen importantes retos y dilemas éticos, legales y sociales, que requieren un marco regulador robusto y adaptado a las circunstancias actuales y, sobre todo, futuras. En este contexto se ha desarrollado el *Reglamento Europeo de Inteligencia Artificial (RIA)*, una normativa pionera que busca establecer el equilibrio entre la innovación tecnológica y la protección de los derechos fundamentales de los ciudadanos.

Este manual tiene como objetivo ofrecer una guía práctica, clara y comprensible sobre el RIA, desglosando sus aspectos más relevantes y proporcionando una herramienta de consulta indispensable para quienes deseen entender y aplicar sus disposiciones. A lo largo de sus capítulos, el lector encontrará un análisis detallado de los principios, obligaciones y responsabilidades que introduce el Reglamento, así como las implicaciones en su adaptación práctica a diferentes sectores.

Hay que tener en cuenta que este reglamento no solo aborda las obligaciones de los proveedores y usuarios de sistemas de IA de alto riesgo, sino que también dedica especial atención a aspectos cruciales como la protección de los derechos humanos, la prevención de prácticas abusivas y el impulso de una IA centrada en

las personas. Con este manual se busca facilitar la comprensión de una normativa compleja pero necesaria, para que empresas, profesionales del derecho, desarrolladores y todos los interesados en la IA puedan navegar el nuevo ecosistema regulatorio con confianza y conocimiento.

En el manual se exploran las consideraciones previas que hacen de este reglamento una norma única en su clase, como su estructura abierta a modificaciones futuras y su alineación con los estándares de seguridad y transparencia más elevados. Se examinan las previsiones de entrada en vigor del Reglamento y los pasos que deberán tomar, tanto actores públicos como privados, para garantizar su correcta implementación. También aborda las exclusiones previstas en el Reglamento, como los sistemas utilizados con fines militares o exclusivamente personales, y se estudian las categorías de IA prohibidas por su potencial peligro para la sociedad.

Este manual presta una especial atención a microempresas, PYMES y startups, ofreciendo una guía detallada sobre sus obligaciones con el fin de fomentar la innovación sin imponer cargas regulatorias desproporcionadas. Sin dejar de lado los requerimientos del Reglamento que impactan a los grandes operadores de sistemas de IA y las medidas que deben adoptar para cumplir con ellos, incluyendo el despliegue de sistemas de gestión de riesgos, la transparencia y la trazabilidad de los resultados de la IA.

Por todo lo comentado, este libro ya es extraordinario por su contenido. Si nos fijamos en el autor y conocemos su trayectoria ¡el interés y su utilidad crece mucho más! Porque hay una enorme cantidad de «expertos en IA» que basan su conocimiento en la lectura de dos recortes de periódico (siempre ocurre así cuando una disciplina crece muy rápido), pero Carlos se ha convertido en un referente en el campo del Derecho y la IA aplicada, como lo demuestran sus innumerables actividades, publicaciones y continuas ganas de estar al día. Su experiencia en este campo le ha consolidado en un puesto destacado, siendo una figura habitual en foros, conferencias y grupos de expertos. Es autor de contrastadas publicaciones sobre derecho digital, IA y regulación, contribuyendo positivamente al debate actual sobre cómo conciliar el avance de la tecnología y el desarrollo sostenible.

Tienen delante un gran libro, que responde a una necesidad actual y escrito por un experto en el tema, por ello sólo me queda agradecer a Carlos la oportunidad de hacer el prólogo para tan excelente obra.

¡Que ustedes lo disfruten!

Emilio Soria Olivas

Catedrático de Universidad, UV

Fundador de IDAL, http://idal.uv.es

12.5.2. Sistemas de IA, incluidos los sistemas de uso general, que generen contenidos sintéticos de audio, imagen, vídeo o texto

Preludia el Cdo. 133 que existe una diversidad de sistemas de IA que puede generar grandes cantidades de contenidos sintéticos que para las personas cada vez es más difícil distinguir del contenido auténtico generado por seres humanos.

La amplia disponibilidad y las crecientes capacidades de dichos sistemas tienen importantes repercusiones en la integridad del ecosistema de la información y en la confianza en este, haciendo surgir nuevos riesgos de desinformación y manipulación a escala, fraude, suplantación de identidad y engaño a los consumidores.

Por tanto, explica, en vista de estos efectos, del rápido desarrollo tecnológico y la necesidad de nuevos métodos y técnicas para asegurar la trazabilidad del origen de la información, procede exigir a los proveedores de tales sistemas que **integren soluciones técnicas que permitan marcar, en un formato legible por máquina, y detectar que el resultado de salida ha sido generado o manipulado por un sistema de IA** y no por un ser humano.

Dichas técnicas y métodos deben ser lo **suficientemente fiables, interoperables, eficaces y sólidos**, en la medida en que sea técnicamente viable, teniendo en cuenta las técnicas disponibles o una combinación de dichas técnicas, como marcas de agua, identificación de metadatos, métodos criptográficos para demostrar la procedencia y la autenticidad del contenido, métodos de registro, impresiones dactilares u otras técnicas, según proceda.

A la hora de aplicar esta obligación, añade este Considerando, los proveedores también **deben tener en cuenta las especificidades y las limitaciones de los diferentes tipos de contenidos** y los avances tecnológicos y del mercado pertinentes en ese ámbito, tal como se refleja en el estado de la técnica generalmente reconocido.

Dichas técnicas y métodos pueden implantarse a nivel de sistema de IA o a nivel de modelo de IA, incluidos modelos de IA de uso general que generan contenidos, facilitando así el cumplimiento de esta obligación por parte del proveedor posterior del sistema de IA.

Para garantizar la proporcionalidad, conviene prever que esta obligación de marcado **no se aplique a los sistemas de IA que desempeñen una función de apoyo a la edición estándar o no alteren sustancialmente los datos de entrada** facilitados por el responsable del despliegue o su semántica.

Añade el Cdo. 135 que, sin perjuicio del carácter obligatorio y de la plena aplicabilidad de las obligaciones de transparencia, la Comisión podrá también **fomentar y facilitar la elaboración de códigos de buenas prácticas** a escala de la Unión, a fin de facilitar la aplicación eficaz de las obligaciones en materia de detección y etiquetado de contenidos generados o manipulados de manera artificial, también para apoyar disposiciones prácticas para que, según proceda, los mecanismos de detección sean accesibles y facilitar la cooperación con otros agentes de la cadena de valor, difundiendo los contenidos o comprobando su autenticidad y procedencia, a

fin de que el público pueda distinguir efectivamente los contenidos generados por IA.

Establece a este respecto el número 2 del art. 50 que los proveedores de sistemas de IA, entre los que se incluyen los sistemas de IA de uso general, que generen contenido sintético de audio, imagen, vídeo o texto, velarán por que los resultados de salida del sistema de IA **estén marcados en un formato legible por máquina** y que sea posible detectar que han sido generados o manipulados de manera artificial.

Los proveedores velarán por que sus **soluciones técnicas sean eficaces, interoperables, sólidas y fiables** en la medida en que sea técnicamente viable, teniendo en cuenta las particularidades y limitaciones de los diversos tipos de contenido, los costes de aplicación y el estado actual de la técnica generalmente reconocido, según se refleje en las normas técnicas pertinentes.

Pero **esta obligación no se aplicará** en la medida en que los sistemas de IA desempeñen una función de apoyo a la edición estándar o no alteren sustancialmente los datos de entrada facilitados por el responsable del despliegue o su semántica, o cuando estén autorizados por ley para detectar, prevenir, investigar o enjuiciar delitos.

Finalmente, el Cdo. 136 recuerda que las obligaciones impuestas a los proveedores y a los responsables del despliegue de determinados sistemas de IA destinadas a permitir que se detecte y divulgue que los resultados de salida de dichos sistemas han sido generados o manipulados de manera artificial, resultan especialmente pertinentes para facilitar la aplicación efectiva del Reglamento (UE) 2022/2065 (Reglamento de Servicios Digitales).

Esto se aplica en particular en lo referente a las obligaciones de los prestadores de plataformas en línea de muy gran tamaño o de motores de búsqueda en línea de muy gran tamaño **para detectar y mitigar los riesgos sistémicos que pueden surgir de la divulgación de contenidos que hayan sido generados o manipulados de manera artificial**, en particular el riesgo de los efectos negativos reales o previsibles sobre los procesos democráticos, el discurso cívico y los procesos electorales, como a través de la desinformación.

Por ello, concluye este Considerando, la exigencia de etiquetar los contenidos generados por sistemas de IA con arreglo al el Reglamento se entiende **sin perjuicio de la obligación prevista en el artículo 16, apartado 6, del Reglamento de servicios digitales** para los prestadores de servicios de alojamiento de datos de tratar las notificaciones que reciban sobre contenidos ilícitos en virtud del artículo 16, apartado 1, de dicho Reglamento[116], y no debe influir en la evaluación y la decisión sobre el carácter ilícito del contenido de que se trate. Dicha evaluación debe realizarse únicamente con referencia a las normas que rigen la legalidad del contenido.

(116) Artículo 16. Mecanismos de notificación y acción.
1. Los prestadores de servicios de alojamiento de datos establecerán mecanismos que permitan que cualquier persona física o entidad les notifique la presencia en su servicio de

12.5.3. Sistemas de reconocimiento de emociones y sistemas de clasificación biométrica

Según explica el Cdo. 132, es preciso notificar a las personas físicas cuando estén expuestas a sistemas de IA que, mediante el tratamiento de sus datos biométricos, puedan determinar o inferir sus emociones o intenciones o incluirlas en una serie de categorías específicas que pueden referirse a aspectos como el sexo, la edad, el color del pelo, el color de ojos, los tatuajes, los rasgos personales, el origen étnico o las preferencias e intereses personales.

En consecuencia, establece el número 3 del art. 50 que los responsables del despliegue de un sistema de reconocimiento de emociones o de un sistema de categorización biométrica deberán informar del funcionamiento del sistema a las personas físicas expuestas a él y tratarán sus datos personales de conformidad con los Reglamentos (UE) 2016/679 (RGPD) y (UE) 2018/1725, sobre tratamiento de datos

---

elementos de información concretos que esa persona física o entidad considere contenidos ilícitos. Dichos mecanismos serán de fácil acceso y manejo, y permitirán el envío de notificaciones exclusivamente por vía electrónica.
2. Los mecanismos mencionados en el apartado 1 serán de tal naturaleza que faciliten el envío de notificaciones suficientemente precisas y adecuadamente fundamentadas. Con ese fin, los prestadores de servicios de alojamiento de datos adoptarán las medidas necesarias para habilitar y facilitar el envío de notificaciones que contengan todos los elementos siguientes:
a) Una explicación suficientemente motivada de los motivos por los que una persona física o entidad considera que la información en cuestión es contenido ilícito;
b) una indicación clara de la localización electrónica exacta de esa información, como por ejemplo el o los URL exactos y, en su caso, información adicional que permita identificar el contenido ilícito adaptado al tipo de contenido y al tipo concreto de servicio de alojamiento de datos;
c) el nombre y una dirección de correo electrónico de la persona física o entidad que envíe la notificación, excepto en el caso de información que se considere que implica uno de los delitos a que se refieren los artículos 3 a 7 de la Directiva 2011/93/UE;
d) una declaración que confirme que la persona física o entidad que envíe la notificación está convencida de buena fe de que la información y las alegaciones que dicha notificación contiene son precisas y completas.
3. Se considerará que las notificaciones a que se refiere el presente artículo proporcionan un conocimiento efectivo o permiten ser consciente, a los efectos del artículo 6, del elemento de información concreto de que se trate, cuando permitan a un prestador diligente de servicios de alojamiento de datos determinar, sin un examen jurídico detallado, que la información o la actividad pertinentes son ilícitas.
4. Cuando la notificación contenga información de contacto electrónica de la persona física o entidad que la envíe, el prestador de servicios de alojamiento de datos enviará, sin dilación indebida, un acuse de recibo de la notificación a dicha persona física o entidad.
5. El prestador también notificará a esa persona física o entidad, sin dilación indebida, su decisión respecto de la información a que se refiera la notificación e incluirá información sobre las vías de recurso respecto de esa decisión.
6. Los prestadores de servicios de alojamiento de datos tratarán las notificaciones que reciban a través de los mecanismos a que se refiere el apartado 1 y adoptarán sus decisiones respecto de la información a que se refieran tales notificaciones, en tiempo oportuno y de manera diligente, no arbitraria y objetiva. Cuando utilicen medios automatizados para dicho tratamiento o toma de decisión, incluirán información sobre dicho uso en la notificación a que se refiere el apartado 5.

personales por las instituciones, órganos y organismos de la Unión, y con la Directiva (UE) 2016/680, según corresponda.

Recordemos que, según el apartado 39 del art. 3, se consideran «sistema de reconocimiento de emociones» a los sistemas de IA destinados a distinguir o inferir las emociones o las intenciones de las personas físicas a partir de sus datos biométricos, y que, según el apartado 40 del mismo artículo, se considera «sistema de categorización biométrica» a los sistemas de IA destinados 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.

Esta obligación, sin embargo, **no se aplicará** a los sistemas de IA utilizados para la categorización biométrica y el reconocimiento de emociones que **hayan sido autorizados por ley** para detectar, prevenir e investigar delitos, con sujeción a las garantías adecuadas para los derechos y libertades de terceros y de conformidad con el Derecho de la Unión.

En estos casos, sus datos personales se tratarán de conformidad con los Reglamentos (UE) 2016/679 y (UE) 2018/1725 y con la Directiva (UE) 2016/680, según corresponda.

12.5.4. Sistemas que generan o manipulan imágenes o audio de personas, haciéndolas pasar por auténticas (ultrafalsificación o deep fakes)

Detalla el Cdo. 134 que los responsables del despliegue que utilicen un sistema de IA para generar o manipular un contenido de imagen, audio o vídeo generado o manipulado por una IA que se asemeje notablemente a personas, objetos, lugares, entidades o sucesos reales y que puede inducir a una persona a pensar erróneamente que son auténticos o verídicos (ultrasuplantaciones o Deep fakes) **deben hacer público, de manera clara y distinguible,** que este contenido ha sido creado o manipulado de manera artificial **etiquetando** los resultados de salida generados por la IA en consecuencia e indicando su origen artificial.

Recordemos que según el apartado 60 del art. 3 del Reglamento, se considera «ultrasuplantación» a todo contenido de imagen, audio o vídeo generado o manipulado por una IA que se asemeja a personas, objetos, lugares, entidades o sucesos reales y que puede inducir a una persona a pensar erróneamente que son auténticos o verídicos.

La obligación establecida en el apartado 4 del art. 50 no se aplicará cuando la ley autorice el uso de los contenidos así generados para para detectar, prevenir, investigar o enjuiciar delitos.

En su virtud, establece el número 4 del art. 50 que los responsables del despliegue de un sistema de IA que genere o manipule imágenes o contenidos de audio o vídeo que constituyan una ultrasuplantación deberán:

17.2. OBJETO DE LOS ESPACIOS CONTROLADOS DE PRUEBA PARA LA IA

Conforme al número 9 del art. 57, el establecimiento de espacios controlados de pruebas para la IA tiene por objeto contribuir a los siguientes objetivos:

a. **Mejorar la seguridad jurídica** para lograr el cumplimiento del o, en su caso, de otras disposiciones de Derecho de la Unión y nacional aplicable;

b. **Apoyar el intercambio de mejores prácticas** mediante la cooperación con las autoridades que participan en el espacio controlado de pruebas para la IA;

c. **Fomentar la innovación y la competitividad** y facilitar el desarrollo de un ecosistema de la IA;

d. **Contribuir a un aprendizaje normativo** basado en datos contrastados;

e. **Facilitar y acelerar el acceso al mercado de la Unión de los sistemas de IA**, en particular cuando los proporcionen pymes, incluidas las empresas emergentes.

La participación en el espacio controlado de pruebas para la IA debe centrarse en **cuestiones que generen inseguridad jurídica** y que, por lo tanto, dificulten que los proveedores y los proveedores potenciales innoven y experimenten con la IA en la Unión y contribuir a un aprendizaje normativo basado en datos contrastados.

Por consiguiente, la supervisión de los sistemas de IA en el espacio controlado de pruebas para la IA **debe comprender su desarrollo, entrenamiento, prueba y validación antes** de su introducción en el mercado o puesta en servicio, así como el concepto de «modificación sustancial» y su materialización, que puede hacer necesario un nuevo procedimiento de evaluación de la conformidad.

Cualquier riesgo significativo detectado durante el proceso de desarrollo y prueba de estos sistemas de IA debe dar lugar a la adopción de medidas de reducción adecuadas y, en su defecto, a la suspensión del proceso de desarrollo y prueba.

A estos efectos, según el número 6 del art. 57, las autoridades competentes proporcionarán, en su caso, orientación, supervisión y apoyo dentro del espacio controlado de pruebas para la IA con vistas a determinar los riesgos, en particular para los derechos fundamentales, la salud y la seguridad, a las pruebas y a las medidas de reducción y su eficacia en relación con las obligaciones y los requisitos del Reglamento y, cuando proceda, de otras disposiciones de Derecho de la Unión y nacional cuya observancia se supervise en el espacio controlado de pruebas.

Además, las autoridades competentes proporcionarán a los proveedores y proveedores potenciales que participen orientaciones sobre las expectativas en materia de regulación y la manera de cumplir los requisitos y obligaciones establecidos en el Reglamento.

Los espacios controlados de pruebas para la IA establecidos en virtud del Reglamento deben entenderse sin perjuicio de otros actos legislativos que permitan el establecimiento de otros espacios controlados de pruebas encaminados a garantizar el cumplimiento de actos legislativos distintos del Reglamento. Cuando proceda, las

autoridades competentes pertinentes encargadas de esos otros espacios controlados de pruebas deben ponderar las ventajas de utilizarlos también con el fin de garantizar el cumplimiento del Reglamento por parte de los sistemas de IA.

Previo acuerdo entre las autoridades nacionales competentes y los participantes en el espacio controlado de pruebas para la IA, las pruebas en condiciones reales también podrán gestionarse y supervisarse en el marco del espacio controlado de pruebas para la IA.

Finalmente, el Cdo. 147 anima a la Comisión a que facilite, en la medida de lo posible, el acceso a las instalaciones de ensayo y experimentación a organismos, grupos o laboratorios establecidos o acreditados con arreglo a la legislación de armonización de la Unión pertinente y que realicen tareas en el marco de la evaluación de la conformidad de productos o dispositivos regulados por dicha legislación. Tal es el caso, en particular, en lo que respecta a los paneles de expertos, los laboratorios especializados y los laboratorios de referencia en el ámbito de los productos sanitarios, de conformidad con los Reglamentos (UE) 2017/745 y (UE) 2017/746.

17.3. TRATAMIENTO DE DATOS PERSONALES EN LOS ENTORNOS CONTROLADOS DE PRUEBAS

El Reglamento presta particular atención a esta cuestión, desde el Cdo. 140, en el que se explica que **se debe proporcionar la base jurídica** para que los proveedores y los proveedores potenciales en el espacio controlado de pruebas para la IA utilicen datos personales recabados para otros fines para desarrollar determinados sistemas de IA en favor del interés público en el espacio controlado de pruebas para la IA, únicamente en determinadas condiciones, de conformidad con el artículo 6, apartado 4, y el artículo 9, apartado 2, letra g), del Reglamento (UE) 2016/679 y los artículos 5, 6 y 10 del Reglamento (UE) 2018/1725, y sin perjuicio de lo dispuesto en el artículo 4, apartado 2, y el artículo 10 de la Directiva (UE) 2016/680.

Por ello, en primer lugar dispone que **siguen siendo aplicables** en estos entornos las demás obligaciones de los responsables del tratamiento y los derechos de los interesados en virtud del Reglamento (UE) 2016/679, el Reglamento (UE) 2018/1725 y la Directiva (UE) 2016/680.

En particular, el Reglamento no debe ofrecer una base jurídica en el sentido del artículo 22, apartado 2, letra b), del Reglamento (UE) 2016/679 y del artículo 24, apartado 2, letra b), del Reglamento (UE) 2018/1725 (en ambos casos, tratamiento automatizado de datos, incluida la elaboración de perfiles, que produzca efectos jurídicos o equivalentes con base en el Derecho de la Unión).

Los proveedores y los proveedores potenciales en el espacio controlado de pruebas para la IA deben proporcionar las garantías adecuadas y cooperar con las autoridades competentes, siguiendo sus indicaciones y actuando con rapidez y de buena fe para mitigar adecuadamente cualquier riesgo considerable para la seguridad, la

salud y los derechos fundamentales que se detecte y pueda surgir durante el desarrollo, las pruebas y la experimentación en dicho espacio.

En el número 10 de este artículo se establece, además, que las autoridades nacionales competentes velarán por que, en la medida en que los sistemas innovadores de IA impliquen el tratamiento de datos personales o estén comprendidos dentro del ámbito de supervisión de otras autoridades nacionales o autoridades competentes que proporcionen o respalden el acceso a los datos, las autoridades nacionales de protección de datos y las demás autoridades nacionales o competentes **estén vinculadas** al funcionamiento del espacio controlado de pruebas para la IA e involucradas en la supervisión de dichos aspectos en la medida en que lo permitan sus respectivas funciones y competencias.

17.3.1. Tratamiento ulterior de datos personales en el espacio controlado de pruebas

Pero, además, el Reglamento presta particular atención al tratamiento ulterior de datos personales para el desarrollo de determinados sistemas de IA en favor del interés público en el espacio controlado de pruebas para la IA, al que dedica todo el art. 59.

Según este precepto, los datos personales recabados lícitamente con otros fines únicamente podrán tratarse en el espacio controlado de pruebas, con el objetivo de desarrollar, entrenar y probar determinados sistemas de IA, **cuando se cumplan todas** las condiciones siguientes:

a. Que los sistemas de IA se desarrollen para que una autoridad pública u otra persona física o jurídica **proteja un interés público esencial** en uno o varios de los siguientes ámbitos:

i. La seguridad y la salud públicas, incluidos la detección, el diagnóstico, la prevención, el control y el tratamiento de enfermedades y la mejora de los sistemas sanitarios,

ii. La protección y mejora de la calidad del medio ambiente, la protección de la biodiversidad, la protección contra la contaminación, las medidas de transición ecológica, la mitigación del cambio climático y las medidas de adaptación a este,

iii. La sostenibilidad energética,

iv. La seguridad y la resiliencia de los sistemas de transporte y la movilidad, las infraestructuras críticas y las redes,

v. La eficiencia y la calidad de la administración pública y de los servicios públicos;

b. Que **los datos tratados resulten necesarios** para cumplir uno o varios de los requisitos mencionados en el capítulo III, sección 2, cuando dichos requisitos

no puedan cumplirse efectivamente mediante el tratamiento de datos anonimizados o sintéticos o de otro tipo de datos no personales;

c. Que existan **mecanismos de supervisión eficaces** para detectar si pueden producirse durante la experimentación en el espacio controlado de pruebas riesgos elevados para los derechos y libertades de los interesados, mencionados en el artículo 35 del Reglamento (UE) 2016/679 y en el artículo 39 del Reglamento (UE) 2018/1725, así como mecanismos de respuesta para mitigar sin demora dichos riesgos y, en su caso, detener el tratamiento;

d. Que los datos personales que se traten en el contexto del espacio controlado de pruebas se encuentren en un **entorno de tratamiento de datos funcionalmente separado, aislado y protegido**, bajo el control del proveedor potencial, y que únicamente las personas autorizadas tengan acceso a dichos datos;

e. Que los proveedores solo puedan compartir los datos recabados originalmente de conformidad con el Derecho de la Unión en materia de protección de datos. Dicho más sencillamente: los datos personales creados en el espacio controlado de pruebas **no pueden salir del espacio controlado de pruebas;**

f. Que el tratamiento de datos personales en el contexto del espacio controlado de pruebas **no dé lugar a medidas o decisiones que afecten a los interesados** ni afecte a la aplicación de sus derechos establecidos en el Derecho de la Unión en materia de protección de datos personales;

g. Que los datos personales tratados en el contexto del espacio controlado de pruebas **se protejan mediante medidas técnicas y organizativas adecuadas** y se eliminen una vez concluida la participación en dicho espacio o cuando los datos personales lleguen al final de su período de conservación;

h. Que los archivos de registro del tratamiento de datos personales en el contexto del espacio controlado de pruebas **se conserven mientras dure la participación** en el espacio controlado de pruebas, salvo que se disponga otra cosa en el Derecho de la Unión o el Derecho nacional;

i. Que se conserve **una descripción completa y detallada del proceso y la lógica subyacentes al entrenamiento, la prueba y la validación** del sistema de IA junto con los resultados del proceso de prueba como parte de la documentación técnica a que se refiere el anexo IV;

j. Que **se publique una breve síntesis del proyecto** de IA desarrollado en el espacio controlado de pruebas, junto con sus objetivos y resultados previstos, en el sitio web de las autoridades competentes; esta obligación no comprenderá los datos operativos sensibles relativos a las actividades de las autoridades garantes del cumplimiento del Derecho, del control fronterizo, de la inmigración o del asilo.

Las condiciones anteriores se entienden **sin perjuicio del Derecho de la Unión o nacional que proscriba el tratamiento de datos personales con fines distintos** de los expresamente mencionados en dichos actos, así como sin perjuicio del Derecho de la Unión o nacional que establezca las bases para el tratamiento de datos personales necesario para desarrollar, probar o entrenar sistemas innovadores de IA o de

cualquier otra base jurídica, de conformidad con el Derecho de la Unión en materia de protección de datos personales.

Por otra parte, cuando ese tratamiento de datos personales en los espacios controlados de pruebas para la IA se lleve a cabo, bajo el control y la responsabilidad de las autoridades garantes del cumplimiento del Derecho, con fines de prevención, investigación, detección o enjuiciamiento de delitos o de ejecución de sanciones penales, incluidas la protección frente a amenazas para la seguridad pública y la prevención de dichas amenazas, **se basará en un Derecho específico, de la Unión o nacional** y cumplirá las condiciones acumulativas anteriores.

17.4. PRUEBA DE LAS ACTIVIDADES REALIZADAS E INFORME DE SALIDA

A petición del proveedor o proveedor potencial del sistema de IA, la autoridad competente aportará una **prueba escrita** de las actividades llevadas a cabo con éxito en el espacio controlado de pruebas.

La autoridad competente también proporcionará un **informe de salida** en el que se detallen las actividades llevadas a cabo en el espacio controlado de pruebas y los resultados y resultados del aprendizaje correspondientes. Los proveedores podrán utilizar esta documentación para demostrar su cumplimiento del Reglamento mediante el proceso de evaluación de la conformidad o las actividades de vigilancia del mercado pertinentes.

A este respecto, las autoridades de vigilancia del mercado y los organismos notificados tendrán en cuenta positivamente los informes de salida proporcionados y las pruebas escritas aportadas por la autoridad nacional competente, con vistas a acelerar los procedimientos de evaluación de la conformidad en una medida razonable.

Según el número 8 de este artículo, con sujeción a las disposiciones de confidencialidad del artículo 78 y con el acuerdo del proveedor o proveedor potencial, la Comisión y el Consejo de IA estarán autorizados a acceder a los informes de salida y los tendrán en cuenta, según proceda, en el ejercicio de sus funciones en virtud del Reglamento. Si tanto el proveedor o proveedor potencial como la autoridad nacional competente dan expresamente su acuerdo para ello, el informe de salida podrá hacerse público a través de la plataforma única de información a que se refiere el presente artículo.

17.5. SUPERVISIÓN DE LOS ENTORNOS CONTROLADOS DE PRUEBAS

Los espacios controlados de pruebas para la IA **no afectarán** a las facultades de supervisión o correctoras de las autoridades competentes que supervisan los espacios controlados de pruebas, tampoco a escala regional o local.

Cualquier riesgo considerable para la salud, la seguridad y los derechos fundamentales detectado durante el proceso de desarrollo y prueba de estos sistemas de IA deberá dar lugar a su adecuada reducción.

Las autoridades nacionales competentes estarán facultadas para suspender temporal o permanentemente el proceso de prueba, o la participación en el espacio controlado de pruebas si no es posible una reducción efectiva, e informarán a la Oficina de IA de dicha decisión. Con el objetivo de apoyar la innovación en materia de IA en la Unión, las autoridades nacionales competentes ejercerán sus facultades de supervisión dentro de los límites del Derecho pertinente y harán uso de su potestad discrecional a la hora de aplicar disposiciones jurídicas en relación con un proyecto específico de espacio controlado de pruebas para la IA (núm. 11).

17.6. RESPONSABILIDAD POR LOS DAÑOS CONSECUENCIA DE LA EXPERIMENTACIÓN REALIZADA

Los proveedores y proveedores potenciales que participen en el espacio controlado de pruebas para la IA **responderán,** con arreglo al Derecho de la Unión y nacional en materia de responsabilidad, **de cualquier daño infligido a terceros** como resultado de la experimentación realizada en el espacio controlado de pruebas.

Sin embargo, esta posible causación de daños **no dará lugar a la imposición de multas administrativas** en virtud de este Reglamento, siempre que los proveedores potenciales respeten el plan específico y las condiciones de su participación y sigan de buena fe las orientaciones proporcionadas por la autoridad nacional competente.

Igualmente, en los casos en que otras autoridades competentes responsables de otras disposiciones del Derecho de la Unión y nacional hayan participado activamente en la supervisión del sistema de IA en el espacio controlado de pruebas y hayan proporcionado orientaciones para el cumplimiento, no se impondrán multas administrativas en relación con dichas disposiciones (núm. 12).

17.7. FUNCIONAMIENTO DE LOS ESPACIOS CONTROLADOS DE PRUEBAS

El Reglamento remite la determinación de las pautas por las que deberá regirse el funcionamiento de estos entornos controlados de pruebas, a los actos de ejecución que deberá adoptar la Comisión.

Según su art. 58, estos actos de ejecución deberán especificar disposiciones detalladas para el establecimiento, el desarrollo, la puesta en práctica, el funcionamiento y la supervisión de los espacios controlados de pruebas para la IA, incluyendo unos principios comunes sobre las siguientes cuestiones:

a) Los criterios de admisibilidad y selección para participar en el espacio controlado de pruebas para la IA;

b) Los procedimientos para la solicitud, la participación, la supervisión, la salida y la terminación del espacio controlado de pruebas para la IA, incluidos el plan del espacio controlado de pruebas y el informe de salida;

c) Las condiciones aplicables a los participantes.

El Reglamento europeo de IA va a regular en profundidad el desarrollo, la comercialización y el uso de esta tecnología en la Unión.

Se trata de una norma extensa y compleja, de alto componente tecnológico, cuyo articulado diseña un intrincado conjunto de obligaciones a cargo tanto de las empresas como de las administraciones públicas que utilicen esta tecnología.

Cumplir el RIA va a exigir, a los profesionales del Derecho y a los de la tecnología, no solo de un estudio detallado de la norma.

También va a requerir de un diálogo permanente entre ellos y con expertos en otros ámbitos, como la ciencia de datos, la informática, la estadística, la sociología y la ética.

Ante ese panorama, el objetivo de este manual es ser una herramienta que ayude a los profesionales que deben aplicar de una u otra manera esta normativa, y a los estudiantes que se inician o continúan en su estudio, a conocer los principales contenidos de esta norma, a partir de un análisis sistemático de su contenido, a fin de facilitar su cumplimiento, al menos en los momentos iniciales de su entrada en vigor y aplicación.

ISBN: 978-84-10292-48-2

LA LEY