REVISTA DE

PRIVACIDAD Y DERECHO DIGITAL

DIRECTOR • D. PABLO GARCÍA MEXÍA

AÑO IX • MAYO-AGOSTO 2024 • NÚMERO 34

ISSN: 2444-5762

SISTEMAS DE IA PROHIBIDOS, DE ALTO RIESGO, DE LIMITADO RIESGO, O DE BAJO O NULO RIESGO[1] (*)

PROHIBITED, HIGH-RISK, LIMITED RISK, OR MINIMAL OR NO RISK AI SYSTEMS

Por MARTÍN MARÍA RAZQUIN LIZARRAGA

Catedrático de Derecho Administrativo.
Universidad Pública de Navarra

(*) Este trabajo se recibió el 28 de mayo de 2024 y fue aceptado en septiembre.

1 Este trabajo se enmarca en el Proyecto "Biometría, Derecho Administrativo y Datos -BIODATA", PID2021-125170NB-I00, financiado por MCIN/AEI/10.13039/501100011033/ y por FEDER Una manera de hacer Europa, del que soy investigador principal.

REVISTA DE
PRIVACIDAD Y DERECHO DIGITAL

RESUMEN

La LIA regula los sistemas de IA desde la perspectiva del riesgo y por tanto los clasifica en sistemas de IA prohibidos, de alto riesgo o de bajo o nulo riesgo. Quedan prohibidas las prácticas de IA que supongan riesgos inadmisibles, salvo algunas excepciones. La categoría principal es la de sistemas de IA de alto riesgo, que vienen enumerados en función de su inclusión en normas sobre productos o en razón de criterios horizontales. Se permite su introducción en el mercado, comercialización y uso de los sistemas de IA de alto riesgo, previo cumplimiento de diversos requisitos y obligaciones relevantes para los operadores. Por el contrario, los sistemas de bajo o nulo riesgo quedan sólo vinculados al cumplimiento voluntario de códigos de conducta.

PALABRAS CLAVE: *Ley de Inteligencia Artificial, inteligencia artificial, sistemas de IA prohibidos, sistemas de IA de alto riesgo.*

ABSTRACT

AIA regulates AI systems from a risk perspective and therefore classifies them into prohibited, high risk, or low or no risk AI systems. AI practices that suppose unacceptable risks are prohibited, with some exceptions. The main catogery is high-risk AI systems, which are listed based on their inclusion in products standard or based on horizontal criteria. High-risk systems are permitted to be introduced into the market, commercialized and used, subject to compliance with various relevant requirements and obligations for operators. On the contrary, minimal or no risk systems are only linked to voluntary compliance with codes of conduct.

KEY WORDS: *Artificial Intelligence Act, artificial intelligence, prohibited AI systems, high-risk AI systems.*

SUMARIO

I.- INTRODUCCIÓN: ENFOQUE Y OBJETO DE LA LIA

Una explicación de los sistemas de IA desde la perspectiva del riesgo, no puede entenderse sin atender al enfoque que la LIA efectúa de su regulación sobre la inteligencia artificial. Sólo desde esta premisa inicial e integral podrá comprenderse por qué se ha adoptado la perspectiva del riesgo y los condicionamientos que se irán estableciendo respecto de cada una de las categorías de sistemas de IA. Y dicha premisa es el valor superior de la persona sobre la inteligencia artificial[2].

2 La posición de la Comisión Europea, ahora plasmada en la LIA, proviene de los diversos documentos que ha ido aprobando hasta llegar a su propuesta de LIA. Por citar los

El Considerando 1 de al LIA muestra este enfoque "de una inteligencia artificial (IA) centrada en el ser humano y fiable, garantizando al mismo tiempo un elevado nivel de protección de la salud, la seguridad y los derechos fundamentales consagrados en la Carta de los Derechos Fundamentales de la Unión Europea (en lo sucesivo, «Carta»), incluidos la democracia, el Estado de Derecho y la protección del medio ambiente, proteger frente a los efectos perjudiciales de los sistemas de IA en la Unión, así como brindar apoyo a la innovación". Asimismo, su Considerando 2 insiste en esta premisa: "El presente Reglamento debe aplicarse de conformidad con los valores de la Unión consagrados en la Carta, lo que facilitará la protección de las personas físicas, las empresas, la democracia, el Estado de Derecho y la protección del medio ambiente y, al mismo tiempo, impulsará la innovación y el empleo y convertirá a la Unión en líder en la adopción de una IA fiable". Y en su Considerando 6 afirma que "como requisito previo, la IA debe ser una tecnología centrada en el ser humano. Además, debe ser una herramienta para las personas y tener por objetivo último aumentar el bienestar humano".

más cercanos y relevantes: la Comunicación "Generar confianza en la inteligencia artificial centrada en el ser humano" (COM 2019-640, de 8.4.2019), el "Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza" (de 19.2.2020) y "Una Estrategia Europea de Datos" (de 19.2.2020). Así la Comunicación citada afirma lo siguiente: "La Estrategia europea de IA y el plan coordinado dejan claro que la confianza es un requisito previo para garantizar un enfoque de la IA centrado en el ser humano: la IA no es un fin en sí mismo, sino un medio que debe servir a las personas con el objetivo último de aumentar su bienestar. Para ello, la fiabilidad de la IA debe estar garantizada. Los valores en los que se basan nuestras sociedades han de estar plenamente integrados en la evolución de la IA. La Unión se fundamenta en los valores de respeto de la dignidad humana, la libertad, la democracia, la igualdad, el Estado de Derecho y el respeto de los derechos humanos, incluidos los derechos de las personas pertenecientes a minorías. Estos valores son comunes a las sociedades de todos los Estados miembros, en las que prevalecen el pluralismo, la no discriminación, la tolerancia, la justicia, la solidaridad y la igualdad. Además, la Carta de los Derechos Fundamentales de la UE reúne, en un único texto, los derechos individuales, civiles, políticos, económicos y sociales de que gozan los ciudadanos de la UE" (pág. 2).

Sólo teniendo en cuenta este enfoque ético[3] se podrán entender las prácticas prohibidas de IA, por ser consideradas como inadmisibles para el ser humano, así como los requisitos y obligaciones impuestas para los sistemas de IA de alto riesgo, principalmente, la evaluación de impacto.

Partiendo de dicha premisa, procede advertir desde el principio que la LIA no es una ley general de los sistemas de IA, puesto que no regula todos ellos y menos todos sus aspectos. El enfoque adoptado del riesgo provoca que la regulación de la LIA sea limitada en cuanto a su objeto, puesto que disciplina sólo aquellos sistemas de IA que tengan riesgo inadmisible o elevado, dejando fuera de su regulación el resto de los sistemas de IA. Dentro de este riesgo se incluyen también los sistemas y modelos de IA de uso general, conocidos como IA generativa, en la medida que son sistemas que dada su amplitud, profundidad y alcance pueden comportar un elevado riesgo. No obstante, en la presente exposición no se van a analizar estos sistemas y modelos de IA de uso general, por ser objeto de otro Estudio.

Sin embargo, en contrapartida, la LIA afecta a todos los sectores que utilizan o pueden utilizar un sistema de IA[4].

El enfoque del riesgo viene ya precisado en el artículo 1.2 LIA que, aunque afirma como primer objetivo, el de establecer normas armonizadas para la introducción, la puesta en servicio y la utilización de sistemas de IA en la Unión, sin embargo, seguidamente precisa que va a establecer prohibiciones de determinadas

3 SALAZAR GARCÍA, I., explica que el primer reto ético de la inteligencia artificial es el del ser humano como centro, siendo ésta "la premisa de las principales organizaciones e instituciones, a nivel internacional, que estudian la regulación ética y normativa de la IA" ("Retos actuales de la ética en la Inteligencia Artificial", en *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Aranzadi, Cizur Menor, 2022, pág. 59).

4 SIMÓN CASTELLANO resalta la concepción y visión amplia de la LIA por su enfoque horizontal, de modo que no hay un *numerus clausus* de sectores a los que afecte ("Allende una teoría general de las garantías jurídicas para una inteligencia artificial confiable", en *Derecho Público de la Inteligencia Artificial*, Fundación Manuel Giménez Abad, Zaragoza, 2023, pág. 119).

prácticas de IA y los requisitos específicos para los sistemas de IA de alto riesgo y obligaciones para los operadores de dichos sistemas.

Una segunda advertencia es necesaria. La LIA parte de la relación a veces inescindible entre sistemas de IA y productos. Como es bien sabido, la LIA es la primera regulación sobre IA de Europa, e incluso del mundo (a salvo de China y de EEUU). Sin embargo, la UE ha venido aprobando, desde hace tiempo, un importante número de disposiciones normativas relativas a los productos, como puede verse en la larga relación contenida en el Anexo I de la LIA. Pues bien, el artículo 2.2 limita, de forma casi total[5], la aplicación de la LIA para aquellos sistemas de IA de alto riesgo que estén asociados a productos que estén sometidos al sistema de armonización total de la UE, en concreto, la referenciada en la sección B del Anexo I. El art. 2.3 contiene, también, algunas exclusiones, como la relativa a los sistemas de IA en materia de seguridad nacional, fines militares o de defensa.

Así pues, la LIA no regula la "fabricación" o producción de sistemas de IA, sino solamente su comercialización (introducción al mercado o posterior) y su uso, pero sólo cuando genere riesgo inadmisible o alto[6]. La producción o fabricación que esté integrada en productos queda sometida a las regulaciones mencionadas en el Anexo I. Incluso cabe añadir que más propiamente la LIA regula las posibles consecuencias que se pueden derivar o pueden ser provocadas por el uso de un sistema de IA, es decir, sus efectos, que es lo que explica el enfoque del riesgo que adopta. Como afirma el Considerando 27 de la LIA, ésta adopta

5 Sólo se les aplican los arts. 6.1, 102 a 109 (disposiciones finales sobre modificación de diversos Reglamentos y Directivas europeos), y 112 (evaluación y revisión), y el art. 57 (espacios controlados de pruebas para la IA) en la medida en que los requisitos de las LIA estén integrados en la legislación de armonización.

6 Como indica SIMÓN CASTELLANO, "la prohibición por riesgo inadmisible no se refiere al diseño tecnológico, sino a la introducción en el mercado, la puesta en servicio y el uso de los sistemas de inteligencia artificial en cuestión en el conjunto de la Unión" ("Allende una teoría general", op. cit, pág. 123).

el enfoque basado en el riesgo que "es la base de un conjunto proporcionado y eficaz de normas vinculantes".

Además, aunque el objetivo de la LIA sea evitar riesgos inadmisibles o condicionar los riesgos elevados para las personas, falta una regulación de los derechos y garantías de los ciudadanos respecto de los sistemas de IA[7], y ello a pesar de que el art. 2.1 g) extiende el ámbito de aplicación a las personas afectadas que estén ubicadas en la UE. Sin embargo, en la fase legislativa se han introducido tres preceptos en orden a paliar esta deficiencia. Por un lado, desde una perspectiva más general, toda persona física o jurídica tiene derecho a presentar una reclamación ante una autoridad de vigilancia del mercado si considera que se ha infringido la LIA (art. 85). Por otro, las personas afectadas tienen derecho a exigir una explicación de las decisiones tomadas individualmente por los responsables del despliegue, cuando entiendan que tienen un efecto perjudicial para su salud, su seguridad o sus derechos fundamentales (art. 86)[8]. Y, por último, se aplica la Directiva (UE) 2019/1937 sobre protección del denunciante a las denuncias de infracciones de la LIA (art. 87).

Por último, la regulación de la LIA se acerca, a veces, más a una disposición de principios, que a una regulación precisa y concreta, por más que la propuesta de la Comisión Europea haya recibido aportaciones en la fase legislativa en orden a su mayor grado de precisión y aplicabilidad directa[9]. Ello provoca que la

7 Así lo advierten COTINO et al. ("Un análisis crítico constructivo de la propuesta de Reglamento de la Unión Europea por la que se establecen normas armonizadas sobre la Inteligencia Artificial (Artificial Intelligence Act)", *Diario La Ley, sección Ciberderecho*, 2 de julio de 2021) y FERNÁNDEZ HERNÁNDEZ ("La futura regulación europea de la inteligencia artificial: objetivos, principios y pautas", en *Claves de inteligencia artificial y derecho*, La Ley, Madrid, 2022, págs. 129-131).

8 Bien es verdad que este derecho es subsidiario, puesto que se aplica únicamente cuando no esté previsto de otro modo en el Derecho de la UE (apartado 3 del art. 86).

9 HUERGO LORA, A., enfatiza en esta cuestión, poniendo el ejemplo que la LIA exige un sistema de gestión de riesgos que admite determinados riesgos residuales y también que la evaluación de conformidad es como regla general un autocontrol ("Gobernar con algoritmos,

LIA precisará de una concreción posterior a través de una diversidad de actuaciones de la Comisión europea o del Comité Europeo de Inteligencia Artificial[10].

II.- LA NOCIÓN DE RIESGO: CLASIFICACIÓN DE LOS SISTEMAS DE IA

La LIA adopta el enfoque del riesgo como noción central y vertebral de su regulación[11]. El riesgo viene definido en el punto 2 del artículo 3 LIA como "la combinación de la probabilidad de que se produzca un perjuicio y la gravedad de dicho perjuicio"[12].

El Diccionario de la Lengua Española define el riesgo como contingencia o proximidad de un daño. Y considera como sinónimos los de "peligro, amenaza, ventura, risco" y como antónimo el de seguridad. Así puede verse cómo el art. 9.5 LIA se refiere a riesgos asociados a cada peligro.

gobernar los algoritmos", *El Cronista del Estado Social y Democrático de Derecho*, núm. 100, 2022, pág. 89).

10 HERNÁNDEZ PEÑA, J. C., indica al respecto que "el reglamento recoge normas de intensidad regulatoria reducida e incluso escasa. Por tanto, es esperable que corresponda al comité -al menos parcialmente- contribuir a completar el programa normativo sectorial. Por tanto, explícitamente se le atribuye la aprobación de normas de Soft Law, con el fin de contribuir a uniformar prácticas administrativas de los Estados miembros" ("Organización y gobernanza de la inteligencia artificial: marco general", en *Inteligencia artificial y sector público. Retos, límites y medios*, Tirant lo blanc, Valencia, 2023, pág. 614).

11 COTINO HUESO, L., afirma que este enfoque del riesgo que adopta la propuesta de LIA supone "la mayor imposición de obligaciones y garantías cuanto mayor riesgo implique el tratamiento de datos o el sistema de IA" ("Nuevo paradigma en la garantía de los derechos fundamentales y una nueva protección de datos frente al impacto social y colectivos de la inteligencia artificial", en *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Aranzadi, Cizur Mayor, 2022, pág. 93).

12 Cabe recordar en este sentido la normativa legal y reglamentaria sobre prevención de riesgos laborales. Así el artículo 4 de la Ley de prevención de riesgos laborales define el riesgo laboral en los siguientes términos: "Se entenderá como «riesgo laboral» la posibilidad de que un trabajador sufra un determinado daño derivado del trabajo. Para calificar un riesgo desde el punto de vista de su gravedad, se valorarán conjuntamente la probabilidad de que se produzca el daño y la severidad del mismo".

BECK, U., acuñó el concepto de la sociedad del riesgo en la época de la postmodernidad y la postindustrialización, puesto que desde hace tiempo la sociedad vive en una situación continuada y permanente de riesgo, que plantea incluso el dilema de la política tecnológica[13]. También ESTEVE PARDO, J., ha profundizado sobre la sociedad del riesgo, señalando que los riesgos son riesgos tecnológicos, que se deben a la intervención humana, y advierte cómo la legislación europea, especialmente la relativa a la seguridad y calidad industrial, pretende dominar, encauzar y unificar, en especial, a partir del denominado nuevo enfoque en el que ha optado por la armonización de normas técnicas[14].

La UE conoce desde hace tiempo el concepto de riesgo, así como el principio de precaución, como, por ejemplo, muestra el Derecho ambiental. Por ejemplo, La Directiva 2004/35/CE, sobre responsabilidad medioambiental en relación con la prevención y reparación de daños medioambientales está repleta de referencias al riesgo, e incluso con menciones al riesgo significativo o al bajo riesgo. Así también la Directiva 2011/92/UE, sobre evaluación de impacto ambiental ha ido incorporando referencias a los riesgos ambientales que sirven como criterio para la realización de una EIA.

También es un concepto esencial en los actos legislativos de armonización a que se refiere el Anexo I de la LIA. A título ejemplo, cabe reparar en la Directiva 2006/42/CE sobre máquinas, (derogada y sustituida por el Reglamento (UE) 2023/1230, a partir de

13 BECK, U. afirma que "los riesgos y peligros de hoy se diferencian esencialmente de los de la Edad Media (que a menudo se les parecen exteriormente) por la globalidad de su amenaza (seres humanos, animales, plantas) y por sus causas modernas. Son riesgos de la modernización. Son un producto global de la maquinaria del progreso industrial y son agudizados sistemáticamente con su desarrollo ulterior" (p. 33). Más adelante señala que los riesgos tienen que ver con la previsión, con lo que todavía no ha llegado (p. 48). Y frente al autocontrol, defiende una generalización, con ciertas garantías jurídicas, de ciertas capacidades de influencia de la subpolítica (*La sociedad del riesgo. Hacia una nueva modernidad*, Paidós, Barcelona, 2010, p. 371).

14 ESTEVE PARDO, J., se refiere a los riesgos tecnológicos derivados de la acción humana y no de causas naturales (p. 29) y al avance del Derecho europeo en aprobar normas de armonización para superar las barreras técnicas (*Técnica, riesgo y Derecho. Tratamiento del riesgo tecnológico en el Derecho ambiental,* Ariel, Barcelona, 1999, págs. 170-171).

14 de enero de 2027). Este Reglamento recoge un concepto de riesgo muy similar al de la LIA: "«riesgo»: una combinación de la probabilidad y la gravedad de una lesión o de un daño a la salud que pueda surgir en una situación peligrosa" (Anexo III, Parte A). Además, cabe referirse a la similitud respecto del control sobre aquellos productos introducidos en el mercado que entrañen riesgos para la salud o la seguridad de las personas (art. 45 del Reglamento). Por su parte, el Reglamento (UE) 2019/1020, de productos contiene una definición similar de riesgo y diferencia entre producto que presenta un riesgo y producto que presenta un riesgo grave (art. 3, apartados 18, 19 y 20).

También en el ámbito de la protección de datos personales, el RGPD utiliza el concepto de riesgo, e incluso de alto riesgo. Por un lado, en todo momento los responsables del tratamiento deben tener en cuenta los riesgos, en especial, en cuanto a su seguridad (art. 32), pero, además, en el caso de que implique un alto riesgo, deberán realizar una evaluación de impacto relativa a la protección de datos (art. 35 RGPD)[15].

El Reglamento (UE) 2022/2065 (conocido como DSA) también se refiere en numerosas ocasiones al riesgo, e incluso establece cuatro categorías de sistémico (Considerando 80).

Ya el Libro Blanco sobre la Inteligencia Artificial de la Comisión Europea de 2020 apuntaba como elemento central de una futura regulación el riesgo, que se configura como punto de equilibrio entre el fomento de la IA y la protección de las personas[16].

15 El art. 35 citado es muy ilustrativo respecto de la noción de alto riesgo: "Cuando sea probable que un tipo de tratamiento, en particular si utiliza nuevas tecnologías, por su naturaleza, alcance, contexto o fines, entrañe un alto riesgo para los derechos y libertades de las personas físicas, el responsable del tratamiento realizará, antes del tratamiento, una evaluación del impacto de las operaciones de tratamiento en la protección de datos personales. Una única evaluación podrá abordar una serie de operaciones de tratamiento similares que entrañen altos riesgos similares" (apartado 1).

16 MORAL SORIANO, L., señala que "el riesgo es la piedra angular del modelo de gobernanza ética, tal y como está recogida en el Libro Blanco de la Comisión, reflejada también en la propuesta que hace el Parlamento Europeo sobre el Marco ético de la IA, y por supuesto en el borrador de Reglamento sobre la IA de la Comisión. Un enfoque basado en los

Por tanto, como indica el Considerando 26 de la LIA, es la intensidad y el alcance de los riesgos que pueden generar los sistemas de IA, lo que determina que éstos sean clasificados en tres tipos: a) sistemas o prácticas prohibidas; y b) sistemas de alto riesgo; y c) sistemas de bajo o nulo riesgo[17]. Ello provoca que los sistemas de bajo o nulo riesgo queden, en la práctica, fuera de la regulación de la LIA[18]. La doctrina, y también la Comisión europea en su página web, consideran como una categoría diferenciada la de los sistemas de IA de riesgo limitado, referida a aquellos sistemas para los que se imponen obligaciones de información y transparencia (art. 50 LIA)[19]. Respecto de los sistemas o modelos de IA de uso general se contempla otra categoría, la del peligro sistémico, definido en el art. 3.65 LIA, que no va a ser objeto de análisis aquí, por corresponder su examen a otro Estudio.

riesgos asegura, como sostiene la Comisión, una intervención proporcionada: la equidistancia necesaria entre el principio de precaución y el principio de innovación, si es que la UE quiere ser un *hub* de IA" ("Modelos de gobernanza global de la inteligencia artificial", en *Inteligencia artificial y Derecho. El jurista ante los retos de la era digital,* Aranzadi, Cizur Menor, 2021, pág. 248). HERNÁNDEZ PEÑA, en idéntica línea, apunta la vinculación de los riesgos con el principio de precaución, y afirma que esta posición está alineada con lo dispuesto en el RGPD (*El marco jurídico de la inteligencia artificial. Principios, procedimientos y estructuras de gobernanza*, Aranzadi, Cizur Menor, 2022, pág. 100).

17 COTINO HUESO, L., lo ejemplifica muy bien, indicando que "teóricamente el AIA supone un sistema de semáforo. Rojo: prohíbe algunos usos de IA (art. 5). Amarillo: fija algunos usos de "alto riesgo" (art. 6 y Anexos II y III). Verde: no es obligatorio cumplir la regulación del AIA" ("Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos. Mejor regular bien que prohibir mal", en *El Cronista del Estado Social y Democrático de Derecho*, núm. 100, 2022, pág. 72).

18 FERNÁNDEZ HERNÁNDEZ afirma que para la Comisión Europea la gran mayoría de los sistemas de IA implican un riesgo bajo o mínimo, con base en el documento de la Comisión titulado Nuevas normas sobre la inteligencia artificial: preguntas y respuestas, aunque a la vista del Anexo III se pone de relieve que el campo de aplicabilidad de la LIA es amplísimo ("La futura regulación europea de la inteligencia artificial: objetivos, principios y pautas", *Claves de inteligencia artificial y derecho*, op. cit., p. 144).

19 Es por ello que SIMÓN CASTELLANO se refiera a cuatro clases de sistemas de IA, añadiendo la categoría de sistema de IA de riesgo limitado para referirse a estos supuestos del art. 50 ("Allende una teoría general", op. cit, págs. 124-125).

Así pues, puede hablarse de una construcción piramidal[20] de cinco escalones, en los que se encuentran cada una de las categorías de sistemas de IA:

1) Sistemas de IA prohibidos, ya que el riesgo es inadmisible.
2) Sistemas de IA de alto riesgo, permitidos pero sometidos al cumplimiento de requisitos y obligaciones.
3) Sistemas de IA de riesgo sistémico, referidos a los sistemas de IA de uso general, que son permitidos bajo cumplimiento de los requisitos establecidos en la LIA.
4) Sistemas de IA de riesgo limitado, permitidos y únicamente sujetos a obligaciones de transparencia e información.
5) Sistemas de IA de bajo o nulo riesgo, excluidos en la práctica de la regulación de la LIA; y respecto de los cuales los operadores pueden cumplir voluntariamente los códigos de conducta.

Conviene tener en cuenta, asimismo, el sistema de gestión de riesgos al que se refiere el art. 9, y que será abordado más adelante.

La noción central de riesgo requiere que el sistema de IA se ajuste a lo previsto en la LIA. Así el art. 79 regula el procedimiento nacional aplicable a los sistemas de IA que presenten riesgo, que exige una evaluación a fin de comprobar si el sistema se ajusta a las exigencias de la LIA. Dicho precepto debe completarse con lo dispuesto en el art. 82 que se dirige a aquellos sistemas de alto riesgo que, a pesar de ser conformes con la LIA, sin embargo, presentan riesgos, en orden a que los proveedores adopten las medidas correctoras pertinentes.

Por último, la clasificación tiene una relevante consecuencia en materia de sanciones, en orden a la determinación de su cuantía (art. 99. 3 y 4).

20 Para una explicación de la pirámide de los puestos que cada categoría ocupa en ella véase CHRISTAKIS, T. y KARRATHANASIS, T., "Tools for Navigating the EU AI Act (2) Visualisation Pyramid", AI Regulation Papers 24-03-5, AI-Regulation.com, March 8th, 2024.

III.- SISTEMAS DE IA PROHIBIDOS

Dentro de la escala de riesgo antes expuesta la LIA examina, en primer lugar, aquellos sistemas que suponen riesgos inaceptables y, por tanto, deben estar prohibidos. Sin embargo, incluso dentro de los sistemas prohibidos se efectúan algunas excepciones que permiten su utilización. El art. 1.2 de la LIA señala que establece "prohibiciones de determinadas prácticas de IA" (letra b)). Así pues, estas prácticas de sistemas de IA no se pueden comercializar ni usar, salvo que concurra alguna de las excepciones que permitan su realización.

Téngase en cuenta que lo dispuesto en el Capítulo II, compuesto únicamente por el art. 5, ambos titulados "Prácticas de IA prohibidas", será de aplicación a los 6 meses de la fecha de entrada en vigor de la LIA, que se producirá a los 20 días de su publicación en el DOUE (art. 113 LIA).

Como indica el Considerando 45 de la LIA, ésta no afecta a las prácticas prohibidas por otras normas del Derecho de la Unión Europea, singularmente en materia de protección de datos, de no discriminación, de protección de consumidores y sobre competencia. Y así el art. 5.8 LIA afirma que "El presente artículo no afectará a las prohibiciones aplicables cuando una práctica de IA infrinja otro acto legislativo de la Unión".

Resulta necesario apuntar la idea, repetida en diversos lugares por la LIA, de su complementariedad con el RGDP (y con la Directiva 2016/680) y con las demás disposiciones de la UE que dispongan normas de protección de las personas o de la competencia.

El art. 5.1 LIA contempla ocho supuestos de prácticas de IA prohibidas. Debe advertirse que se trata, por un lado, de una regulación que constituye un *numerus clausus*, dado que no se contiene una disposición que determine su posible modificación, como ocurre en el art. 7 LIA para los sistemas de alto riesgo.

Tampoco el art. 97 LIA al referirse a los actos delegados de la Comisión contiene mención alguna al art. 5.

Por otra parte, la redacción de los diferentes supuestos del art. 5.1, a pesar de la introducción durante la fase legislativa de ciertas precisiones, adolece de numerosos conceptos jurídicos indeterminados, que dejan un campo abierto, tal vez demasiado extenso, en orden a su interpretación. Y debe advertirse que la interpretación tendrá una incidencia muy relevante porque supondrá, en su caso, la consideración de un sistema de IA como prohibido o como de algo riesgo. En este sentido el Considerando 28 señala que la IA tiene usos beneficiosos y también perversos, cuando lleva a cabo prácticas perjudiciales e incorrectas de manipulación, explotación y control social, que son las que deben resultar prohibidas.

En tercer lugar, en la fase legislativa ha cobrado importancia la determinación de los sistemas biométricos, en orden a su introducción y mayor precisión, pero también en las últimas lecturas se ha realizado una interpretación más restrictiva de los supuestos de prohibición.

Con estas precisiones cabe examinar cada uno de los ocho supuestos legales de prohibición.

III.1.- USOS QUE SUPONGAN MANIPULACIÓN O ENGAÑO O ALTERACIÓN DEL COMPORTAMIENTO

La letra a) del art. 5.1 recoge como primer supuesto de usos prohibidos el siguiente: "la introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que se sirva de técnicas subliminales que trasciendan la conciencia de una persona o de técnicas deliberadamente manipuladoras o engañosas con el objetivo o el efecto de alterar de manera sustancial el comportamiento de una persona o un colectivo de personas, mermando de manera apreciable su capacidad para tomar una decisión informada y haciendo que tomen una

decisión que de otro modo no habrían tomado, de un modo que provoque, o sea razonablemente probable que provoque, perjuicios considerables a esa persona, a otra persona o a un colectivo de personas".

Su explicación se contiene en el Considerando 29, donde se examinan las posibilidades de la IA para manipular a las personas e inducirlas en su comportamiento, en suma, que provocan la pérdida de autonomía de las personas. Incluso se describen algunas técnicas de manipulación y de componentes subliminales, o las interfaces cerebro-máquina.

Esta letra a) debe ser interpretada con arreglo al Considerando 29, que permite efectuar varias observaciones. La primera consistente en que, dado que el enfoque es el de riesgo, en este caso de un riesgo inadmisible, no tiene importancia la conducta del proveedor o del responsable del despliegue, sino que "el perjuicio se derive de las prácticas de manipulación o explotación que posibilita la IA".

La segunda que, dado que se trata de evitar la anulación de la autonomía personal[21], no están prohibidas aquellas prácticas que cuenten con el consentimiento explícito de las personas o de sus representantes legales.

En tercer lugar, debe darse una relación de causalidad entre el sistema de IA y el objetivo o efecto que provoca la prohibición[22].

Incluso para el caso de prácticas comerciales y legítimas habrá de demostrarse que incurren en esta manipulación, puesto que por sí mismas, siempre que cumplan el Derecho aplicable, no son prácticas de manipulación perjudiciales a los efectos del art. 5 LIA.

21 Así las Directrices para una IA fiable de 2019 recogían como primer principio ético el respeto de la autonomía humana.

22 HERNÁNDEZ PEÑA señala que "un aspecto objeto de controversia es la causalidad entre la exposición a un sistema de IA y el impacto sobre el comportamiento, de forma tal que se llegue a distorsionar o perturbar el comportamiento" (*El marco jurídico…*, op. cit., pág. 124).

III.2.- USOS DE MANIPULACIÓN DE PERSONAS VULNERABLES

La letra b) del art. 5.1 contiene un segundo supuesto de este tenor literal: "la introducción en el mercado, la puesta en servicio o la utilización de un sistema de IA que explote alguna de las vulnerabilidades de una persona física o un determinado colectivo de personas derivadas de su edad o discapacidad, o de una situación social o económica específica, con la finalidad o el efecto de alterar de manera sustancial el comportamiento de dicha persona o de una persona que pertenezca a dicho colectivo de un modo que provoque, o sea razonablemente probable que provoque, perjuicios considerables a esa persona o a otra".

Es fácil de percibir que este segundo supuesto constituye una variante agravada del primero[23]. Aquí se acentúa la prohibición por tratarse de personas vulnerables, en razón de su edad, discapacidad, situación social o económica, o pertenencia a una minoría étnica o religiosa.

III.3.- USOS PARA LA EVALUACIÓN O CLASIFICACIÓN DE PERSONAS FÍSICAS

El tercer supuesto de uso prohibido de IA se explica en la letra c) del art. 5.1. Dicho supuesto puede ser dividido en dos partes: la premisa y sus circunstancias, toda vez que no todos los usos que incurran en lo que se determina en la premisa quedan prohibidos, sino solamente aquellos que incurran en alguna de las situaciones perjudiciales o desfavorables para las personas. Así también se trata de proteger no sólo a las personas físicas individualmente consideradas sino también a los colectivos de personas, que pueden padecer estas evaluaciones o clasificaciones negativas o despectivas como tales colectivos. Se trata de evitar sistemas de "crédito social" (*social scoring*), como los existentes en China[24].

23 De ahí que también sea examinada conjuntamente con el primer supuesto en el Considerando 29 LIA.

24 Así lo indica HERNÁNDEZ PEÑA (*El marco jurídico…*, págs. 125-127).

La premisa es la siguiente: "la introducción en el mercado, la puesta en servicio o la utilización de sistemas de IA para evaluar o clasificar a personas físicas o a colectivos de personas durante un período determinado de tiempo atendiendo a su comportamiento social o a características personales o de su personalidad conocidas, inferidas o predichas…".

Como indica el Considerando 31 sólo deben prohibirse "los sistemas de IA que impliquen esas prácticas inaceptables de puntuación y den lugar a esos resultados perjudiciales o desfavorables. Esa prohibición no debe afectar a prácticas lícitas de evaluación de las personas físicas que se efectúen para un fin específico de conformidad con el Derecho de la Unión y nacional".

Tal como se ha advertido más arriba, de nuevo se trata de un supuesto limitado, puesto que debe concurrir alguna de las dos condiciones de perjuicio a las personas o colectivos que se relatan en el citado art. 5.1 c) y que dicen así:

"i) un trato perjudicial o desfavorable hacia determinadas personas físicas o colectivos de personas en contextos sociales que no guarden relación con los contextos donde se generaron o recabaron los datos originalmente,

ii) un trato perjudicial o desfavorable hacia determinadas personas físicas o colectivos de personas que sea injustificado o desproporcionado con respecto a su comportamiento social o la gravedad de este".

Así pues, cualquier otra práctica de IA de evaluación o clasificación de personas o colectivos de personas que no comporte estos efectos perjudiciales o desfavorables no se encuentra prohibida.

III.4.- USOS PARA EVALUACIONES DE RIESGOS DE COMISIÓN DE DELITOS

La LIA acoge como cuarto supuesto de prácticas prohibidas de IA la relativa a la realización de evaluaciones de riesgos de personas físicas en orden a la comisión de delitos. En concreto la

letra d) del art. 5.1 dice así: "la introducción en el mercado, la puesta en servicio para este fin específico o el uso de un sistema de IA para realizar evaluaciones de riesgos de personas físicas con el fin de valorar o predecir el riesgo de que una persona física cometa un delito basándose únicamente en la elaboración del perfil de una persona física o en la evaluación de los rasgos y características de su personalidad".

La explicación de esta prohibición se contiene en el Considerando 42 LIA: "En consonancia con la presunción de inocencia, las personas físicas de la Unión siempre deben ser juzgadas basándose en su comportamiento real. Las personas físicas nunca deben ser juzgadas a partir de comportamientos predichos por una IA basados únicamente en la elaboración de sus perfiles, en los rasgos o características de su personalidad, como la nacionalidad, el lugar de nacimiento, el lugar de residencia, el número de hijos, el nivel de endeudamiento o el tipo de vehículo, sin una valoración humana y sin que exista una sospecha razonable, basada en hechos objetivos comprobables, de que dicha persona está implicada en una actividad delictiva".

Este precepto plantea diversas cuestiones. En primer lugar, la prohibición alcanza tanto a los sistemas de IA que se hayan creado con este fin específico, como también a aquellos que permitan en su utilización conseguir este resultado. En segundo término, la referencia a delitos excluye las infracciones administrativas. Y, en tercer término, se acota a la elaboración de perfiles o rasgos o características de la personalidad, cuando este sea el fundamento único del sistema de IA, lo que puede significar que en otro caso (es decir, cuando no sea "únicamente") dicho sistema no está prohibido.

De nuevo este supuesto contiene una excepción consistente en que no están prohibidos los usos de IA en el caso de personas sospechosas, cuando la sospecha se fundamente en hechos objetivos y verificables. Como señala el propio precepto: "esta prohibición no se aplicará a los sistemas de IA utilizados para apoyar la

valoración humana de la implicación de una persona en una actividad delictiva que ya se base en hechos objetivos y verificables directamente relacionados con una actividad delictiva".

III.5.- USOS DE RECONOCIMIENTO FACIAL

La letra e) del art. 5.1 LIA comienza con los supuestos de usos prohibidos de sistemas biométricos, que ha sido una de las grandes preocupaciones de la fase legislativa.

El art. 3.34 LIA define qué se entiende por datos biométricos: "los datos personales obtenidos a partir de un tratamiento técnico específico, relativos a las características físicas, fisiológicas o conductuales de una persona física, como imágenes faciales o datos dactiloscópicos". Estos datos se encuadran dentro de las categorías especiales de datos del art. 9 del RGPD[25] y del art. 10 de la Directiva 2016/680 (a la que se remite el apartado 37 del art. 3 LIA), cuyo tratamiento está prohibido salvo que concurra alguna de las bases jurídicas expresadas en dichos preceptos.

Como señala el Considerando 14, este concepto de datos biométricos debe interpretarse conforme a la normativa europea de protección de datos antes citada. Este Considerando diferencia los diferentes usos de los datos biométricos: 1) Autenticación; 2) Identificación; 3) Categorización; y 4) Reconocimiento de emociones.

Así el primer supuesto está dedicado al reconocimiento facial: "la introducción en el mercado, la puesta en servicio para este fin específico o el uso de sistemas de IA que creen o amplíen bases de datos de reconocimiento facial mediante la extracción

25 El art. 4 apartado 14 del RGPD contienen la siguiente definición de datos biométricos: "datos personales obtenidos a partir de un tratamiento técnico específico, relativos a las características físicas, fisiológicas o conductuales de una persona física que permitan o confirmen la identificación única de dicha persona, como imágenes faciales o datos dactiloscópicos".

no selectiva de imágenes faciales de internet o de circuitos cerrados de televisión".

La LIA no prohíbe cualquier reconocimiento facial, sino solamente aquel que se de en alguna de las condiciones descritas en esta letra e). Como señala el Considerando 43 se protege a las personas frente a la vigilancia masiva, por lo que se hace referencia a una extracción no selectiva.

Esta consideración de vigilancia masiva explica las condiciones limitativas de este supuesto para la extracción de imágenes: internet o circuitos cerrados de televisión.

Pero, además, lo que se prohíbe es la creación o ampliación de bases de datos con estas imágenes no selectivas, no el uso inmediato de las imágenes que no comporte incorporación de las mismas a una base de datos.

III.6.- SISTEMAS DE INFERENCIA DE EMOCIONES

Este sexto supuesto de uso prohibido recoge un nuevo caso de uso de los datos biométricos, ahora con la finalidad de detectar o deducir emociones. Como señala el Considerando 44 de la LIA, su prohibición se fundamenta en que pueden tener resultados discriminatorios e invadir los derechos y libertades de las personas afectadas, en especial, en determinados contextos[26].

El art. 3 apartado 39 lo define así: "«sistema de reconocimiento de emociones»: un sistema de IA destinado a distinguir o inferir las emociones o las intenciones de las personas físicas a partir de sus datos biométricos"[27].

26 El Parlamento Europeo impulsó la redacción de un informe sobre esta cuestión, que le fue entregado en septiembre de 2021 y lleva como título "Biometric Recognition and Behaviour Detection", siendo sus autores externos WENDEHORST, C., y DULLER, Y.

27 El Considerando 18 LIA enumera las emociones que pueden reconocerse: "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

La redacción de la letra f) del art. 5.1 LIA es la siguiente: "la introducción en el mercado, la puesta en servicio para este fin específico o el uso de sistemas de IA para inferir las emociones de una persona física en los lugares de trabajo y en los centros educativos".

La inferencia de emociones se limita en este supuesto a los lugares de trabajo o centros educativos, no a otros ámbitos, y ello porque el carácter intrusivo de estos sistemas de inferencia de emociones tiene un resultado inadmisible en estos ámbitos.

Nuevamente, esta disposición recoge una excepción a este supuesto de prohibición: "excepto cuando el sistema de IA esté destinado a ser instalado o introducido en el mercado por motivos médicos o de seguridad".

Ambos motivos de excepción deben tener como punto común la protección de las personas, de ahí que se trata de aspectos de salud (motivos médicos) o de seguridad (prevención de riesgos laborales, por ejemplo).

III.7.- USOS DE CATEGORIZACIÓN BIOMÉTRICA

El apartado 40 del art. 3 LIA define qué se entiende por categorización biométrica: "un 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".

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".

Como señala el Considerando 16 de la LIA, estas categorías 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".

No tienen esta consideración determinados usos por su nota de accesoriedad a otro uso o servicio principal, tal como observa el considerando 16, con algún ejemplo.

Pues bien, los sistemas de IA que tengan como finalidad específica o permitan el uso de sistemas de categorización biométrica están prohibidos, tal como establece la letra g) del art. 5.1 LIA: "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"[28].

Como puede verse, la protección alcanza a las finalidades o usos que afectan a datos de categoría especial del RGPD, que son los datos especiales y más relevantes de las personas[29].

También aquí este supuesto contiene una excepción: "esta prohibición no incluye el etiquetado o filtrado de conjuntos de datos biométricos adquiridos lícitamente, como imágenes, basado en datos biométricos ni la categorización de datos biométricos en el ámbito de la garantía del cumplimiento del Derecho".

El Considerando 30 señala como supuestos excluidos de la prohibición "la clasificación de imágenes en función del color del

28 Así se explica en el Considerando 30 LIA.

29 Procede recordar que ya el RGPD prohíbe el tratamiento de datos personales con esta finalidad en su art. 9.1: "Quedan prohibidos el tratamiento de datos personales que revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o las orientación sexuales de una persona física".

pelo o del color de ojos, que pueden utilizarse, por ejemplo, en el ámbito de la garantía del cumplimiento del Derecho".

III.8.- SISTEMAS DE IDENTIFICACIÓN BIOMÉTRICA REMOTA EN TIEMPO REAL EN ESPACIOS DE ACCESO PÚBLICO

Este supuesto se revela como uno de los más debatidos dentro del uso de sistemas biométricos. Venía incorporado en la propuesta de la Comisión europea, pero ha recibido varias aportaciones en la fase legislativa, que por un lado precisan este supuesto y por otro lo acotan.

Por otra parte, la LIA se ha hecho eco de la crítica social a los sistemas masivos de control de las personas, uno de cuyos ejemplos es el de China, que se trata de evitar a toda costa en la Unión Europea.

El supuesto requiere, en primer término, una explicación del mismo, que viene dada en las definiciones que recoge la LIA. El apartado 35 del art. 3 LIA explica, por un lado, lo que es la identificación biométrica: "el reconocimiento automatizado de características humanas de tipo físico, fisiológico, conductual o psicológico para determinar la identidad de una persona física comparando sus datos biométricos con los datos biométricos de personas almacenados en una base de datos".

El Considerando 15 de la LIA se ocupa de definir con mayor precisión el concepto de identificación biométrica y su diferenciación con la verificación o autenticación biométricas[30]. Es esencial

30 Resulta, por su importancia, obligado transcribir el Considerando 15 citado: "El concepto de «identificación biométrica» a que hace referencia el presente Reglamento debe definirse como el reconocimiento automatizado de características humanas de tipo físico, fisiológico o conductual, como la cara, el movimiento ocular, la forma del cuerpo, la voz, la entonación, el modo de andar, la postura, la frecuencia cardíaca, la presión arterial, el olor o las características de las pulsaciones de tecla, a fin de determinar la identidad de una persona comparando sus datos biométricos con los datos biométricos de personas almacenados en una base de datos de referencia, independientemente de que la persona haya dado o no su consentimiento. Quedan excluidos los sistemas de IA destinados a la verificación biométrica, que comprende la autenticación, cuyo único propósito es confirmar que una persona física

entender estas diferencias para comprender que la LIA sólo se refiere en el art. 5 a los sistemas de identificación y no a los de verificación y autenticación.

Y luego los apartados 41 a 44 del art. 3 LIA explican con detalle qué se entiende por cada uno de los conceptos adicionales:

1) Sistema de identificación biométrica remota (se explica en el Considerando 17).
2) Sistema de identificación biométrica remota en tiempo real (se explica en el Considerando 17).
3) Sistema de identificación biométrica en diferido (se explica en el Considerando 17).
4) Espacio de acceso público (se explica en el Considerando 19).

La letra h) del art. 5.1 LIA impone la prohibición del "uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de garantía del cumplimiento del Derecho". El fundamento de esta prohibición se recoge en el Considerando 32: "El uso de sistemas de IA para la identificación biométrica remota «en tiempo real» de personas físicas en espacios de acceso público con fines de garantía del cumplimiento del Derecho invade de forma especialmente grave los derechos y las libertades de las personas afectadas, en la medida en que puede afectar a la vida privada de una gran parte de la población, provocar la sensación de estar bajo una vigilancia constante y disuadir indirectamente a los ciudadanos de ejercer su libertad de reunión y otros derechos fundamentales. Las imprecisiones técnicas de los sistemas de IA destinados a la identificación biométrica remota de las personas físicas pueden dar lugar a resultados sesgados y tener efectos discriminatorios. Tales posibles resultados sesgados y efectos discriminatorios son especialmente pertinentes por lo que respecta a la edad, la

concreta es la persona que dice ser, así como la identidad de una persona física con la finalidad exclusiva de que tenga acceso a un servicio, desbloquee un dispositivo o tenga acceso de seguridad a un local".

etnia, la raza, el sexo o la discapacidad. Además, la inmediatez de las consecuencias y las escasas oportunidades para realizar comprobaciones o correcciones adicionales en relación con el uso de sistemas que operan «en tiempo real» acrecientan el riesgo que estos conllevan para los derechos y las libertades de las personas afectadas en el contexto de actividades de garantía del cumplimiento del Derecho, o afectadas por estas".

Puede advertirse de la "excepcionalidad" de la propia salvedad, que provoca que la LIA impida incluso la obtención de determinadas consecuencias de su uso cuando éste sea posible de forma excepcional, tal como indica el inciso final del apartado 3: "Dicha autoridad no podrá adoptar ninguna decisión que produzca efectos jurídicos adversos para una persona exclusivamente sobre la base de los resultados de salida del sistema de identificación biométrica remota «en tiempo real»".

Como se desprende de la literalidad del precepto, debe tratarse de identificación biométrica (no verificación o autenticación) y, además, deben concurrir tres aspectos esenciales:

1) Que la identificación sea remota. De ahí que no alcance a la identificación en cercanía, es decir, la que se realice con consciencia de la persona física como cuando se acerca voluntariamente a un sistema de identificación. Ello hace referencia, por ejemplo, a las grabaciones indiscriminadas de imágenes de las personas.
2) Que la identificación sea en tiempo real. Por tanto, su diferenciación con la identificación en diferido, tal como explica el Considerando 17.
3) Que la identificación sea en espacios de acceso público, lo que no alcanza a otro tipo de espacios de acceso no público, sean de titularidad pública o privada[31]. Aquí se hace

31 Así lo indica la definición de espacio de acceso público contenida en el art. 3.44 LIA, tras la precisión añadida en el procedimiento legislativo. En cambio, HERNÁNDEZ PEÑA consideraba que "No parece que la definición incluya instalaciones privadas de uso público,

> referencia especial al acceso a las calles o lugares de tránsito público, como también aeropuertos o estaciones de transporte terrestre o marítimo.

Así pues, en este supuesto caracterizado con las notas antedichas, el uso de la identificación biométrica recibe una prohibición absoluta.

Sin embargo, la LIA contempla un supuesto excepcional que permite utilizar la identificación biométrica remota en espacios de acceso público con fines de garantía del cumplimiento del Derecho. Este supuesto excepcional constituye una "lex specialis", aplicable frente a lo dispuesto en la normativa de protección de datos personales, singularmente, el art. 10 de la Directiva (UE) 2016/680. Por tanto, lo dispuesto en el artículo 10 citado, así como en el art. 9.1 del RGPD sólo se aplica a los demás tratamientos de datos personales diferentes de la citada excepción (apartado final de la letra h) y Considerando 39).

No obstante, esta excepción se somete a unos requisitos muy estrictos, con el objetivo de limitar al máximo estos usos con la finalidad expresada, en orden a evitar un Estado policial de continuada vigilancia sobre los ciudadanos.

La LIA establece importantes límites y condiciones para permitir el uso de este tipo de prácticas.

El más importante y previo es su aceptación por los Estados. Para poder hacer uso de la posibilidad contenida en la excepción antedicha, los Estados deben introducir en su normativa interna la posibilidad de utilizar, total o parcialmente, esta excepción (apartado 5, y también apartado 2, segundo párrafo)[32].

como podría ser el caso de universidades o clínicas privadas, así como comercios abiertos al público, entre otros" (*El marco jurídico…*, op. cit., pág. 130).

32 Como señala el Considerando 37 este uso excepcional es posible "cuando el Estado miembro de que se trate haya decidido contemplar expresamente la posibilidad de autorizarlo en las normas detalladas de su Derecho nacional, y en la medida en que lo haya contemplado". Y el Estado miembro puede acogerse a la totalidad de los objetivos de la excepción o sólo a algunos de ellos.

Y si así lo hicieran, comunicarán a la Comisión las normas adoptadas al respecto en el plazo de 30 días siguientes.

Por otra parte, la regulación que efectúa la LIA sobre esta excepción de uso de sistemas de identificación biométrica es una regulación de mínimos y por tanto obligada para los Estados. No obstante, éstos podrán adoptar leyes más restrictivas sobre el uso de sistemas de identificación biométricas (apartado 5 *in fine*).

Veamos, pues, los límites para el uso de esta excepción.

En primer lugar, no basta cualquier fin de garantía del cumplimiento del Derecho, sino que debe perseguir alguno de los objetivos específicos establecidos en el apartado 1), párrafo primero, de la letra h), siempre que se cumplan los límites y condiciones que se fijan en los apartados siguientes. Como expresa el Considerando 33, la salvedad persigue lograr un interés público esencial cuya importancia compense los riesgos.

Y es preciso constatar que el uso de estos sistemas de IA constituye tratamiento de datos biométricos (Considerando 38, párrafo 1º)[33].

Aquí comienza el acotamiento de los tres supuestos excepcionales previstos. En primer lugar, se contiene un límite general, aplicable a todos ellos, "en la medida en que dicho uso sea estrictamente necesario". Y seguidamente se describe cada uno de los objetivos que permiten salvar la prohibición.

Los límites y condiciones cabe enumerarlos del siguiente modo:

1) Concurrencia de uno de los tres objetivos señalados en la letra h) del apartado 1.
2) Necesidad de una finalidad específica: confirmar la identidad de la persona que constituya el objetivo específico, y con las condiciones de las letras a) y b) del apartado 2.

33 De ahí que la regulación de la LIA sea calificada como *lex specialis* respecto del art. 10 de la Directiva (UE) 2016/680.

3) Evaluación de impacto de protección de datos y registro en la base de datos de la UE[34].

4) Autorización previa[35] de autoridad judicial o de autoridad administrativa independiente[36]. A tal fin se fijan condiciones para la concesión de dicha autorización previa.

5) Notificación del uso a la autoridad de vigilancia del mercado y a la autoridad nacional de protección de datos. Estas autoridades deben efectuar informes anuales a la Comisión sobre este uso. La Comisión realizará informes anuales con dicha información agregada.

Los objetivos que habilitan para el uso de esta excepción son los tres siguientes:

"i) la búsqueda selectiva de víctimas concretas de secuestro, trata de seres humanos o explotación sexual de seres humanos, así como la búsqueda de personas desaparecidas,

ii) la prevención de una amenaza específica, importante e inminente para la vida o la seguridad física de las personas

34 Excepcionalmente el registro en la base de datos puede efectuarse *a posteriori*.

35 El Considerando 33 excluye de esta autorización previa determinados controles de identidad: "Además, el presente Reglamento debe preservar la capacidad de las autoridades garantes del cumplimiento del Derecho, de control fronterizo, de la inmigración o del asilo para llevar a cabo controles de identidad en presencia de la persona afectada, de conformidad con las condiciones establecidas en el Derecho de la Unión y en el Derecho nacional para estos controles. En particular, las autoridades garantes del cumplimiento del Derecho, del control fronterizo, de la inmigración o del asilo deben poder utilizar sistemas de información, de conformidad con el Derecho de la Unión o el Derecho nacional, para identificar a las personas que, durante un control de identidad, se nieguen a ser identificadas o no puedan declarar o demostrar su identidad, sin que el presente Reglamento exija que se obtenga una autorización previa. Puede tratarse, por ejemplo, de una persona implicada en un delito que no quiera revelar su identidad a las autoridades garantes del cumplimiento del Derecho, o que no pueda hacerlo debido a un accidente o a una afección médica".

36 Excepcionalmente puede utilizarse el sistema con carácter de urgencia sin autorización previa, siempre que la misma se solicite inmediatamente dentro del plazo de 24 horas. No obstante, también se contempla que en caso de denegación de la autorización debe interrumpirse de inmediato el uso y además desecharse y suprimirse los datos obtenidos (Considerando 35).

físicas[37] o de una amenaza real y actual o real y previsible de un atentado terrorista,

iii) la localización o identificación de una persona sospechosa de haber cometido una infracción penal a fin de llevar a cabo una investigación o un enjuiciamiento penales o de ejecutar una sanción penal por alguno de los delitos mencionados en el anexo II[38] que en el Estado miembro de que se trate se castigue con una pena o una medida de seguridad privativas de libertad cuya duración máxima sea de al menos cuatro años[39]".

IV.- SISTEMAS DE IA DE ALTO RIESGO

La LIA presta especial y amplia atención a la regulación de los sistemas de IA de alto riesgo. Frente al art. 5 que es la única disposición dedicada a los sistemas de IA prohibidos, el Capítulo III relativo a los sistemas de IA de alto riesgo está compuesto de 44 artículos, e incluso los Capítulos siguientes se refieren en gran medida a este tipo de sistemas. Ello permite concluir que la regulación de los sistemas de IA de alto riesgo constituye el eje central de la LIA.

37 Según el Considerando 33 también se engloba en esta excepción la perturbación grave de las infraestructuras críticas, definidas en el art. 2.4 de la Directiva (UE) 2022/2557.

38 El Anexo II enumera los siguientes delitos: "terrorismo, trata de seres humanos, explotación sexual de menores y pornografía infantil, tráfico ilícito de estupefacientes o sustancias psicotrópicas, tráfico ilícito de armas, municiones y explosivos, homicidio voluntario, agresión con lesiones graves, tráfico ilícito de órganos o tejidos humanos, tráfico ilícito de materiales nucleares o radiactivos, secuestro, detención ilegal o toma de rehenes, delitos que son competencia de la Corte Penal Internacional, secuestro de aeronaves o buques, violación, delitos contra el medio ambiente, robo organizado o a mano armada, sabotaje, y participación en una organización delictiva implicada en uno o varios de los delitos enumerados en esta lista".

39 El número de 4 años es un mínimo, que puede ser elevado por los Estados miembros, tal como señala expresamente el párrafo final del Considerando 33.

Dentro del Capítulo III la clave se encuentra en su primera Sección que determina qué sistemas de IA se clasifican como de alto riesgo, que viene complementada por los Anexos I y III. A partir de ello, se fijan, luego, los requisitos (Sección 2) y las obligaciones (Secciones 3, 4 y 5).

Los sistemas de IA de alto riesgo se pueden introducir en el mercado y utilizar, siempre que reúnan los requisitos exigidos. Sólo en aquellos supuestos en que el riesgo no supere la evaluación de conformidad o, en su caso, la evaluación de impacto (art. 27), los sistemas de IA de alto riesgo no podrán ser utilizados.

IV. 1.- CLASIFICACIÓN DE SISTEMA DE ALTO RIESGO

IV.1.A.- La noción de alto riesgo: criterios de determinación

Como se ha advertido más arriba, la LIA parte de la noción de riesgo para su regulación, de modo que hay cuatro clases de sistemas de IA, además de la de riesgo sistémico para los sistemas de uso general: a) sistemas prohibidos por ser inadmisibles dado su riesgo; b) sistemas que se pueden autorizar con los requisitos y obligaciones que se determinan por tener un riesgo alto; c) sistemas de riesgo limitado, sujetos a obligaciones de transparencia; y d) sistemas de IA que quedan prácticamente fuera de la regulación de la LIA, dado que su nivel de riesgo es bajo o nulo. Procede recordar que la noción de riesgo es un cálculo de probabilidad, tal como define el art. 3 LIA. Y es dicho cálculo en razón de la intensidad y alcance de los riesgos el que provoca su clasificación, en este caso, su encuadramiento como sistemas de alto riesgo (Considerando 26).

La primera idea que se desprende de la LIA es que la clasificación de un sistema como de alto riesgo es restrictiva, es decir, sólo cabe su encuadre en esta categoría cuando el sistema se encuentre incluido dentro de las determinaciones expresadas en su art. 6. Lo enuncia de forma clara el Considerando 46 al

afirmar que "La clasificación de un sistema de IA como «de alto riesgo» debe limitarse a aquellos sistemas de IA que tengan un efecto perjudicial considerable en la salud, la seguridad y los derechos fundamentales de las personas de la Unión, y dicha limitación debe reducir al mínimo cualquier posible restricción del comercio internacional".

El concepto de "alto riesgo" se encuentra vinculado a la "magnitud" o "gravedad" de sus consecuencias o perjuicios adversos (Considerandos 47 y 48), en relación con los derechos fundamentales, y con especial atención a los menores y también al medio ambiente. Se trata de que los productos que se introduzcan en el mercado o se comercialicen sean seguros y conformes.

La dificultad de definir la noción de alto riesgo justifica que el art. 6.5 ordene a la Comisión la aprobación de Directrices, junto con ejemplos prácticos, de cuáles sean sistemas de alto riesgo y cuáles no lo sean.

La LIA diferencia entre dos tipos de sistemas de IA para su encuadramiento como sistemas de alto riesgo. Por un lado, los que están vinculados o forman parte de un producto, y, por otro, los que son independientes de un producto. Como señala el art. 6.2, hay dos supuestos de sistemas de alto riesgo: a) los del apartado 1, vinculados a productos; y b) los del apartado 2, independientes de productos y relacionados en el Anexo III.

IV.1.B.- Sistemas de IA de alto riesgo vinculados a un producto: el Anexo I

El primer supuesto de sistemas de IA de alto riesgo es el de aquellos sistemas que son componentes de un producto o incluso constituyen por sí mismos un producto. Los Considerandos de la LIA ponen diversos ejemplos: componentes de robots de uso en fábricas o en sanidad (Considerando 47) o ciertos productos como "máquinas, juguetes, ascensores, equipo y sistemas de protección para uso en atmósferas potencialmente

explosivas, equipos radioeléctricos, equipos a presión, equipos de embarcaciones de recreo, instalaciones de transporte por cable, aparatos que queman combustibles gaseosos, productos sanitarios, productos sanitarios para diagnóstico *in vitro, automoción y aviación"* (Considerando 50).

El art. 6.1 LIA comienza con una advertencia que indica la amplitud de este supuesto: "Con independencia de si se ha introducido en el mercado o se ha puesto en servicio sin estar integrado en los productos que se mencionan en las letras a) y b)". De ahí que se utilice el concepto de vinculación al producto, puesto que no precisa estar integrado en el mismo, sino solamente que pueda ser utilizado en relación con un producto.

La Directiva de máquinas establecía la definición de componente de seguridad[40]. En el futuro dicha definición será sustituida por la contenida en el Reglamento (UE) 2023/1230: "un componente físico o digital, incluido el software, de un producto incluido en el ámbito de aplicación del presente Reglamento que esté diseñado o destinado a desempeñar una función de seguridad y que se introduzca en el mercado por separado, cuyo fallo o funcionamiento defectuoso ponga en peligro la seguridad de las personas, pero que no sea necesario para que dicho producto funcione o cuyos componentes normales puedan ser sustituidos para que dicho producto funcione" (art. 3, apartado 3).

Ahora el art. 2.14 LIA define los componentes de seguridad en los siguientes términos: "un componente de un producto o un sistema que cumple una función de seguridad para dicho producto o

40 Su art. 2 c) contiene la siguiente definición de componente de seguridad: "componente:
- que sirva para desempeñar una función de seguridad,
- que se comercialice por separado;
- cuyo fallo y/o funcionamiento defectuoso ponga en peligro la seguridad de las personas, y
- que no sea necesario para el funcionamiento de la máquina o que, para el funcionamiento de la máquina, pueda ser reemplazado por componentes normales.
En el anexo V figura una lista indicativa de componentes de seguridad que podrá actualizarse con arreglo al artículo 8, apartado 1, letra a)".

sistema, o cuyo fallo o defecto de funcionamiento pone en peligro la salud y la seguridad de las personas o los bienes".

Los productos son los regulados en la normativa de la UE, que viene explicitada en el Anexo I de la LIA. Así pues, este Anexo, dividido en dos Secciones A y B, recoge la doble lista de actos legislativos de armonización de la UE, que obedece a la categorización de productos sometidos al denominado *old approach* o sistema de armonización total o a los componentes de seguridad de productos sometidos al nuevo enfoque o *New Framework Legislation*[41]. Procede recordar que a los sistemas de IA referidos a la Sección B sólo se les aplica la LIA de forma muy limitada, únicamente en lo relativo a sus arts. 6.1, 102 a 109 (disposiciones finales sobre modificación de diversos Reglamentos y Directivas europeos), y 112 (evaluación y revisión), y el art. 57 (espacios controlados de pruebas) (art. 2.2).

Por tanto, para que un sistema de IA se encuadre en el art. 6 como de alto riesgo, se exige, en primer lugar (letra a) que su regulación armonizada esté comprendida en el Anexo I, en definitiva, constituye un "numerus clausus" de actos legislativos.

Como segundo requisito, el art. 6.1 requiere que el sistema de IA "deba someterse a una evaluación de la conformidad de terceros para su introducción en el mercado o puesta en servicio con arreglo a los actos legislativos de armonización de la Unión enumerados en el anexo I" (letra b)).

La exigencia de ambos requisitos (deben reunirse las dos condiciones) provoca que sólo aquellos productos para los que la regulación armonizada de la UE requiera evaluación de conformidad serán considerados como sistemas de alto riesgo. Ello supone que, si dicha regulación armonizada no exige dicha evaluación de conformidad, por considerar que el producto no entraña un riesgo elevado a los efectos de dicha normativa, no puede ser considerado como sistema de IA de alto riesgo.

41 Véase al respecto HERNÁNDEZ PEÑA, *El marco jurídico*..., op. cit., págs. 136-138.

Asimismo, es preciso tener en cuenta lo dispuesto en el art. 74.3, que permite en relación con los actos legislativos de la Sección A del Anexo I, que los Estados opten por designar otra autoridad pertinente como autoridad de vigilancia del mercado.

Finalmente, conviene advertir que este apartado 1 y las obligaciones unidas al mismo se aplicarán a los 36 meses desde la entrada en vigor de la LIA.

IV.1.C.- Sistemas de alto riesgo independientes de un producto: el Anexo III

IV.1.C.1.- Sistemas de IA de alto riesgo del Anexo III

El segundo supuesto de sistemas de IA de alto riesgo es el de aquellos sistemas incluidos en el Anexo III de la LIA. Así lo afirma el apartado 2 del art. 6: "...se considerarán de alto riesgo los sistemas de IA contemplados en el anexo III". Como puede verse, la LIA no se refiere a sectores sino a sistemas de IA[42].

El Anexo III enumera los sistemas de IA de alto riesgo a que se refiere el artículo 6, apartado 2, es decir, los que no están vinculados a productos conforme a las normas normalizadas del Anexo I. Así lo señala el Considerando 52: "En cuanto a los sistemas de IA independientes, a saber, aquellos sistemas de IA de alto riesgo que no son componentes de seguridad de productos, o que son productos en sí mismos, deben clasificarse como de alto riesgo si, a la luz de su finalidad prevista, presentan un alto riesgo de ser perjudiciales para la salud y la seguridad o los derechos fundamentales de las personas, teniendo en cuenta tanto la gravedad del posible perjuicio como la probabilidad de que se produzca, y se utilizan en varios ámbitos predefinidos especificados en el presente Reglamento".

42 HERNÁNDEZ PEÑA considera acertado que la propuesta de la Comisión Europea se apartara "de un enfoque basado exclusivamente en sectores (banca, defensa, salud, etc.), que podría suponer costes elevados y desproporcionados en aplicaciones y sistemas con un riesgo bajo para los derechos fundamentales, la salud o la seguridad de las personas" (*El marco jurídico...*, op. cit., pág. 139).

1. Biometría

El Considerando 54 afirma que los datos biométricos constituyen una categoría de datos personales sensibles, y, en definitiva, remite su aceptación a las disposiciones del RGPD y de la Directiva (UE) 2016/680. De ahí el inciso de que "su uso esté permitido por el Derecho de la Unión o nacional aplicable". A tal efecto la LIA parece entender que, en principio, los sistemas biométricos están autorizados por la normativa de protección de datos (con el cumplimiento de los requisitos en ella establecidos) y solamente califica como de alto riesgo algunos sistemas biométricos específicos por entender que pueden dar lugar a resultados sesgados y tener efectos discriminatorios[43].

Por tanto, en el caso de los sistemas biométricos de IA se imponen tres niveles: 1) Prohibición absoluta de los sistemas de identificación remota en tiempo real en espacios públicos (con la excepción ya explicada para casos de garantía de cumplimiento del Derecho); 2) sistemas de alto riesgo especificados en el punto 1 del Anexo III; y 3) demás sistemas biométricos, que son considerados como de bajo o nulo riesgo.

Los tres sistemas biométricos de alto riesgo son los siguientes:

a) Sistemas de identificación biométrica remota. Aquí se incluyen la identificación biométrica remota que no se desarrolle en espacios de acceso público y la identificación biométrica remota en diferido; y se excluyen tanto la identificación biométrica que no sea remota, por contar con la participación de las personas físicas, como la verificación biométrica cuya única finalidad sea confirmar que una persona física concreta es la persona que afirma ser. El apartado 36 del art. 3 define la verificación biométrica

43 Esta regulación da solución, cuando menos parcial, a la denuncia de COTINO HUESO en el sentido de que la regulación de la LIA y del RGPD se ignoran ("Sistemas de inteligencia artificial...", op. cit., pág. 75). Además, debe tenerse en cuenta las cautelas introducidas en la fase legislativa en la LIA, especialmente, la evaluación de impacto introducida por el art. 27 de la LIA, que habrá de compaginarse con la prevista en el RGPD.

en los siguientes términos: "la verificación automatizada y uno-a-uno, incluida la autenticación, de la identidad de las personas físicas mediante la comparación de sus datos biométricos con los datos biométricos facilitados previamente". Y la verificación biométrica es completamente diferente de la identificación biométrica, que aparece definida en el art. 3.35, más arriba transcrito. Ello provoca que en la biometría haya tres niveles: prohibición, permisión en caso de alto riesgo sometida a requisitos y obligaciones, y permisión por ser de bajo o nulo riesgo, sujeta únicamente a códigos de conducta voluntarios.

b) Sistemas de IA destinados a ser utilizados para la categorización biométrica en función de atributos o características sensibles o protegidos basada en la inferencia de dichos atributos o características. Su límite se encuentra en la prohibición de estos sistemas de categorización cuando se refieran a aquellos atributos personales referidos en la letra g) del art. 5.1 LIA.

c) Sistemas de IA destinados a ser utilizados para el reconocimiento de emociones. Aquí se incluyen los sistemas que se desarrollen en ámbitos diferentes de los lugares de trabajo o centros educativos, en cuyo caso estarían prohibidos.

2. Infraestructuras críticas

El Considerando 55 ofrece una explicación más completa de este punto. En primer término, qué debe entenderse por infraestructuras críticas[44], para lo que es necesario acudir al Anexo de la Directiva (UE) 2022/2557, modificado por el Reglamento Delegado (UE) 2023/2450. La LIA considera de alto riesgo los sistemas vinculados a las infraestructuras digitales críticas, al tráfico rodado y al suministro de agua, gas, calefacción o electricidad.

44 Conforme al art. 2.4 de la Directiva se considera infraestructura crítica: "un elemento, instalación, equipo, red o sistema, o parte de un elemento, instalación, equipo, red o sistema, que es necesario para la prestación de un servicio esencial".

Se trata de componentes de seguridad, no necesarios para el funcionamiento del sistema, pero que pueden dar lugar en caso de fallo o defecto de funcionamiento a riesgos para la salud y seguridad de las personas.

3. Educación y formación profesional

Dentro del ámbito educativo se consideran sistemas de IA de alto riesgo sólo algunos de los que pueden ser utilizados, dado que la gran mayoría fomentarán la formación digital y la adquisición de capacidades digitales.

La característica común a los cuatro supuestos catalogados como de alto riesgo en el ámbito educativo viene explicitada en el Considerando 56: evitar la discriminación de las personas, en especial de las pertenecientes a colectivos vulnerables (mujeres, mayores, discapacitadas, de determinado origen racial o étnico, o personas con una determinada orientación sexual).

Con base en este objetivo de no discriminación se consideren de alto riesgo, en resumen, cuatro casos:

a) Determinación del acceso o admisión a centros educativos[45].

b) Evaluación de resultados de aprendizaje.

c) Evaluación del nivel de educación adecuado a recibir o acceder.

d) Control de exámenes.

4. Empleo, gestión de los trabajadores y acceso al autoempleo

También en este punto, el objetivo de la clasificación de determinados sistemas como de alto riesgo obedece a una finalidad

45 Parece que se ha tenido presente el caso del algoritmo utilizado en el Reino Unido de valoración de estudiantes para acceso a las Universidades y que tuvo que ser retirado a causa de los sesgos en contra de los estudiantes de las escuelas públicas (HERNÁNDEZ PEÑA, *El marco jurídico…*, op. cit., pág. 141).

de evitar la discriminación, sobre todo de determinados colectivos vulnerables (Considerando 57).

Y así se consideran sistemas de IA de alto riesgo los dos siguientes:

a) Acceso al empleo: anuncios, solicitudes de empleo y evaluación de candidatos.

b) Condiciones de trabajo, promoción y despido: cuando se trate de tomar como base comportamientos individuales o características personales para supervisión o evaluación[46].

5. Acceso a servicios privados esenciales y a servicios y prestaciones públicos esenciales y disfrute de estos servicios y prestaciones

Se trata de garantizar el acceso y el disfrute de servicios, tanto privados como públicos, que se consideran esenciales para las personas.

Entre los servicios públicos se encuentran todos aquellos que pueden englobarse bajo el amplio epígrafe de sanidad y asistencia social (asistencia sanitaria, prestaciones de seguridad social contributiva y no contributiva, dependencia, maternidad, enfermedad, dependencia, asistencia social y ayudas a la vivienda). Así también se incluye la solicitud y respuesta a llamadas de emergencia realizadas por personas físicas en servicios como policía, bomberos y servicios de asistencia médica, y en sistemas de triaje de pacientes en el contexto de la asistencia sanitaria de urgencia.

En los servicios privados esenciales, se incluyen los relativos a evaluaciones de solvencia crediticia y para seguros de vida y de salud.

46 Téngase en cuenta la letra d) el art. 64.4 del Texto Refundido del Estatuto de los Trabajadores, introducida por el Real Decreto Ley 9/2021 y confirmada por la Ley 12/2021, de 28 de septiembre, que reconoce como un derecho del comité de empresa: "Ser informado por la empresa de los parámetros, reglas e instrucciones en los que se basan los algoritmos o sistemas de inteligencia artificial que afectan a la toma de decisiones que pueden incidir en las condiciones de trabajo, el acceso y mantenimiento del empleo, incluida la elaboración de perfiles".

Procede recordar que el Considerando 58 incluye una importante cautela: "No obstante, el presente Reglamento no debe obstaculizar el desarrollo y el uso de enfoques innovadores en la Administración, que podrían beneficiarse de una mayor utilización de sistemas de IA conformes y seguros, siempre y cuando dichos sistemas no conlleven un alto riesgo para las personas jurídicas y físicas".

6. Garantía del cumplimiento del Derecho

También en ese supuesto se precisa con carácter previo que el uso de sistemas de IA con esta finalidad esté autorizado por el Derecho de la UE o el de los Estados miembros. Además, los sistemas de IA sólo pueden ser utilizados por las autoridades encargadas del cumplimiento del Derecho o en su nombre, bien sean estatales o de la UE.

Cinco son los sistemas de IA considerados como de alto riesgo[47]:

a) Sistemas utilizados para la evaluación del riesgo de que una persona física sea víctima de delitos.

b) Polígrafos o herramientas similares.

c) Sistemas de IA para evaluar la fiabilidad de las pruebas durante la investigación o el enjuiciamiento de delitos.

47 El Considerando 59 explica los motivos de su inclusión como sistemas de alto riesgo: "Dado su papel y su responsabilidad, las actuaciones de las autoridades garantes del cumplimiento del Derecho que implican determinados usos de los sistemas de IA se caracterizan por un importante desequilibrio de poder y pueden dar lugar a la vigilancia, la detención o la privación de libertad de una persona física, así como tener otros efectos negativos sobre los derechos fundamentales consagrados en la Carta. En particular, si el sistema de IA no está entrenado con datos de buena calidad, no cumple los requisitos adecuados en términos de rendimiento, de precisión o de solidez, o no se diseña y prueba debidamente antes de introducirlo en el mercado o ponerlo en servicio, es posible que señale a personas de manera discriminatoria, incorrecta o injusta. Además, podría impedir el ejercicio de importantes derechos procesales fundamentales, como el derecho a la tutela judicial efectiva y a un juez imparcial, así como el derecho a la defensa y a la presunción de inocencia, sobre todo cuando dichos sistemas de IA no sean lo suficientemente transparentes y explicables ni estén suficientemente bien documentados".

d) Sistemas de IA para evaluar la probabilidad de que una persona física cometa un delito o reincida en la comisión de un delito atendiendo no solo a la elaboración de perfiles de personas físicas mencionada en el artículo 3, punto 4, de la Directiva (UE) 2016/680 o para evaluar rasgos y características de la personalidad o comportamientos delictivos pasados de personas físicas o colectivos[48].

e) Sistemas de IA para elaborar perfiles de personas físicas durante la detección, la investigación o el enjuiciamiento de delitos.

El Considerando 59 excluye de estos supuestos los sistemas de IA que se utilicen en la lucha contra el blanqueo de capitales.

7. Migración, asilo y gestión del control fronterizo

De nuevo aquí el punto de partida es el mismo que en los supuestos 1 y 6, que el uso de los sistemas de IA esté permitido. Y al igual que en el cumplimiento del Derecho (punto 6), estos sistemas de IA de alto riesgo sólo pueden ser utilizados por las autoridades públicas competentes o en su nombre, bien sean estatales o de la UE.

Los casos de sistemas de IA de alto riesgo son los siguientes[49]:

a) Polígrafos y herramientas similares.

48 Así se han utilizado sistemas de valoración con este fin, como el Compass en Estados Unidos o el RISCANVI en Cataluña (HERNÁNDEZ PEÑA, *El marco jurídico…*, op. cit., pág. 146).

49 El objetivo de esta inclusión es el siguiente: "Los sistemas de IA empleados en la migración, el asilo y la gestión del control fronterizo afectan a personas que con frecuencia se encuentran en una situación especialmente vulnerable y que dependen del resultado de las actuaciones de las autoridades públicas competentes. Por este motivo, es sumamente importante que los sistemas de IA que se utilicen en estos contextos sean precisos, no discriminatorios y transparentes, a fin de garantizar que se respeten los derechos fundamentales de las personas afectadas y, en particular, su derecho a la libre circulación, a la no discriminación, a la intimidad personal y la protección de los datos personales, a la protección internacional y a una buena administración" (Considerando 60).

b) Sistemas de IA para evaluar un riesgo, por ejemplo, un riesgo para la seguridad, la salud o de migración irregular, que plantee una persona física que tenga la intención de entrar en el territorio de un Estado miembro o haya entrado en él.

c) Sistemas de IA para ayudar a las autoridades públicas competentes a examinar las solicitudes de asilo, visado o permiso de residencia y las reclamaciones conexas con el fin de determinar si las personas físicas solicitantes reúnen los requisitos necesarios para que se conceda su solicitud, con inclusión de la evaluación conexa de la fiabilidad de las pruebas.

d) Sistemas de IA para detectar, reconocer o identificar a personas físicas, con excepción de la verificación de documentos de viaje.

8. Administración de justicia y procesos democráticos

Por un lado, se introducen como sistemas de IA de alto riesgo aquellos relacionados con la administración de justicia y, por tanto, utilizados por una autoridad judicial directa o indirectamente. Son los sistemas que vayan a ser utilizados para ayudar a una autoridad judicial en la investigación e interpretación de hechos y de la ley, así como en la garantía del cumplimiento del Derecho a un conjunto concreto de hechos, o a ser utilizados de forma similar en una resolución alternativa de litigios. Como se indica, son sistemas que ayudan al juez, pero que no le sustituyen, dado que la reserva de humanidad debe ser aquí prevalente. Pero, incluso, en estos supuestos de colaboración, que no de sustitución, también deben cumplirse los requisitos y obligaciones de los sistemas de IA de alto riesgo.

El segundo supuesto hace referencia a los sistemas democráticos, en orden a evitar la influencia en los votantes en los referendos o

elecciones[50]. Por eso se excluyen de esta consideración aquellos sistemas que no afectan directamente a las personas físicas.

IV.1.C.2.- Excepciones a la aplicación del Anexo III: la evaluación del proveedor

Como acaba de exponerse, los sistemas de IA relacionados en el Anexo III son de alto riesgo. Sin embargo, el art. 6.3 permite que determinados sistemas incluidos en dicho Anexo III no sean considerados de alto riesgo.

La razón de esta exclusión se encuentra en el nivel de riesgo, que se define como no "importante"[51]: "No obstante lo dispuesto en el apartado 2, un sistema de IA a que se refiere el anexo III no se considerará de alto riesgo cuando no plantee un riesgo importante de causar un perjuicio a la salud, la seguridad o los derechos fundamentales de las personas físicas, también al no influir sustancialmente en el resultado de la toma de decisiones" (art. 6.3).

Así pues, la falta de magnitud o gravedad del riesgo y, sobre todo, el respeto a la autonomía de las personas puede provocar que el sistema de IA, a pesar de estar incluido en la relación del Anexo III, pueda no ser considerado de alto riesgo.

Sin embargo, algunos sistemas siempre van a ser de alto riesgo, como por ejemplo aquellos que lleven a cabo la elaboración de perfiles de personas físicas. Se entiende por elaboración de perfiles: "toda forma de tratamiento automatizado de datos personales consistente en utilizar datos personales para evaluar determinados aspectos personales de una persona física, en particular para analizar o predecir aspectos relativos al rendimiento

50 A tal respecto cabe hacer referencia a las experiencias de influencia en procesos electorales de la conocida empresa *Cambridge Analytica*.

51 La noción de riesgo importante aparece también mencionada en el apartado 6, con relación a la posible adición o modificación de las condiciones de excepción a la consideración de un sistema del Anexo III como no de alto riesgo.

profesional, situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, ubicación o movimientos de dicha persona física" (art. 4.4 RGPD).

Para la aplicación de la excepción antedicha de no consideración como de alto riesgo, se requieren dos requisitos:

1) Que se dé una o varias de las cuatro condiciones[52] señaladas en el citado apartado 3. Por tanto, bastaría que se diera sólo una de ellas.
2) Que el proveedor documente una evaluación previa de la que se concluya que, a pesar de estar incluido el sistema en el Anexo III, no es de alto riesgo por concurrir alguna de las condiciones señaladas.

En definitiva, es el proveedor quien, previa evaluación, declara que su sistema de IA no es de alto riesgo, por darse alguna de las condiciones enumeradas en este artículo.

La problemática derivada del debate de esta disposición se puede ver plasmada en el apartado 6, donde se permite a la Comisión la revisión de estas condiciones en función de la evolución tecnológica y del mercado, aunque sin que pueda reducir el nivel global de protección de las personas. La Comisión puede adoptar tres conductas:

1) Modificar las condiciones.
2) Añadir nuevas condiciones.
3) Suprimir las condiciones existentes. Pero sólo si la supresión es necesaria para mantener el nivel de protección de las personas.

52 El Considerando 53 explica la relación de cada una de esas cuatro condiciones con el concepto general de no influir sustancialmente en la toma de decisiones o no perjudicar dichos intereses sustancialmente.

La excepcionalidad de esta medida de exclusión de sistemas de IA del Anexo III va acompañada de una cautela regulada en el art. 80, referida al supuesto en el que el proveedor califique erróneamente un sistema de IA como excluido. En estos supuestos si la autoridad de vigilancia del mercado constata que el sistema de IA es de alto riesgo exigirá el cumplimiento de los requisitos y obligaciones de la LIA.

IV.1.D.- Modificación del Anexo III

El apartado 1 del art. 7 LIA habilita a la Comisión para modificar los supuestos contemplados en el Anexo III o para añadir otros nuevos. La redacción de este apartado se ajusta más bien a la previsión de la propuesta de la Comisión consistente en la adición, puesto que la posibilidad de modificación se introdujo en el procedimiento legislativo.

Desde dicha perspectiva, la posible adición de nuevos casos de uso de sistemas de alto riesgo del Anexo III requiere el cumplimiento de las dos condiciones expresadas en el citado apartado 1. Merece la pena detenerse en la segunda condición que es la equivalencia de riesgo, que implica la adición de nuevos casos por el hecho de que su riesgo es equivalente o mayor al de los sistemas ya contemplados en el Anexo III. A tal efecto, el apartado 2 del art. 7 refiere nada menos que once criterios, que permitan a la Comisión evaluar esta segunda condición a los efectos de la inclusión de un sistema de IA en el Anexo III, tal como reza el título de este artículo 7.

Por último, el apartado 3 del art. 7 permite a la Comisión suprimir sistemas de IA enumerados en el Anexo III cuando se den las dos condiciones expresadas en dicho precepto y que hacen ver que se ha reducido el nivel de riesgo (el riesgo ya no es considerable) y que no se ha bajado el nivel general de protección de las personas.

IV.2.- REQUISITOS DE LOS SISTEMAS DE IA DE ALTO RIESGO

Una vez definidos los sistemas de alto riesgo, la LIA pasa a establecer cuáles son los requisitos obligatorios para su introducción en el mercado y comercialización.

Los sistemas de alto riesgo del apartado 1 en relación con el Anexo I y del apartado 2 del art. 6 en relación con el Anexo III, no convierten a dichos sistemas en sistemas que pueden ser puestos en uso directamente. Lo expresa con rotundidad el Considerando 63: "El hecho de que un sistema de IA sea clasificado como un sistema de IA de alto riesgo en virtud del presente Reglamento no debe interpretarse como indicador de que su uso sea lícito con arreglo a otros actos del Derecho de la Unión o del Derecho nacional compatible con el Derecho de la Unión, por ejemplo, en materia de protección de los datos personales o la utilización de polígrafos y herramientas similares u otros sistemas para detectar el estado emocional de las personas físicas. Todo uso de ese tipo debe seguir realizándose exclusivamente en consonancia con los requisitos oportunos derivados de la Carta y de los actos aplicables del Derecho derivado de la Unión y del Derecho nacional. No debe entenderse que el presente Reglamento constituye un fundamento jurídico para el tratamiento de datos personales, incluidas las categorías especiales de datos personales, en su caso, salvo que el presente Reglamento disponga específicamente otra cosa".

Es, por ello, que la Sección 2 del Capítulo III establece los requisitos que debe cumplir un sistema de alto riesgo para poder ser introducido en el mercado y utilizado. Así pues, el operador que debe cumplir con estos requisitos es el proveedor, tal como impone el art. 16. a). Para el proveedor ésta es su primera obligación: velar porque sus sistemas de IA de alto riesgo cumplan los requisitos de la Sección 2. Además, tanto el importador como el distribuidor están obligados a verificar que el proveedor ha cumplido con estos requisitos (arts. 23 y 24) e incluso el proveedor inicial puede ser sustituido por nuevos proveedores (art. 25).

La LIA establece varios requisitos que debe cumplir un sistema de alto riesgo para ser introducido en el mercado y su posterior comercialización.

Estos requisitos persiguen mitigar los riesgos y garantizar la fiabilidad de los sistemas de IA. A tal efecto cabe recordar que el Considerando 27 efectúa un resumen de las Directrices éticas para una IA fiable, aprobadas en 2019 por el Grupo Independiente de Expertos de Alto Nivel sobre IA. Los siete requisitos para una IA fiable[53] pasan ahora a convertirse en requisitos obligatorios para los sistemas de IA de alto riesgo.

El art. 8.2 permite un cumplimiento de requisitos efectuado de forma integrada en cuanto a procedimientos y documentos para los sistemas de IA de productos de la Sección A del Anexo I. Como señala el Considerando 64, la LIA completa lo dispuesto en los actos legislativos referidos en dicha Sección A, lo que comportará una aplicación simultánea y complementaria de diversos actos legislativos, cuando menos, el propio del producto y la LIA. En estos casos el proveedor de un producto que incorpore un sistema de IA de alto riesgo puede efectuar la evaluación del producto y simultáneamente la del sistema de IA incorporado al mismo. De forma específica se recoge la integración del sistema de riesgos en el art. 9.10.

Los requisitos de los sistemas de IA de alto riesgo son los siguientes:

1) Implantación de un sistema de gestión de riesgos.
2) Realización de prácticas de gobernanza de datos para el caso de entrenamiento de modelos de IA.
3) Elaboración de la documentación técnica del sistema de IA.

53 Las Directrices para una IA fiable de 2019 señalaban los siguientes siete requisitos clave: 1) acción y supervisión humana; 2) solidez técnica y seguridad; 3) gestión de la privacidad y de los datos; 4) transparencia; 5) diversidad, no discriminación y equidad; 6) bienestar social y ambiental; y 7) rendición de cuentas.

4) Trazabilidad asegurada mediante un registro automático de eventos.
5) Transparencia y comunicación de información a los responsables del despliegue.
6) Vigilancia humana.
7) Nivel adecuado de precisión, solidez y ciberseguridad.

IV.2. A.- Implantación de un sistema de gestión de riesgos

El art. 9.1 LIA recoge este requisito: "Se establecerá, implantará, documentará y mantendrá un sistema de gestión de riesgos en relación con los sistemas de IA de alto riesgo". Este precepto[54] utiliza diversos verbos que suponen cada uno de ellos exigencias para el proveedor, y que van ligados a otros requisitos y obligaciones que la LIA les impone, así como al ciclo de vida.

Además, el art. 17.1 LIA obliga al proveedor a incluir dentro del sistema de gestión de la calidad el sistema de gestión de riesgos (letra g)).

La idea principal es que este requisito constituye no sólo una exigencia previa, sino también una obligación continuada, toda vez que el sistema de gestión de riesgos no termina con su establecimiento e implantación, sino que debe ser mantenido a lo largo de la vida del sistema de IA[55].

En todo caso, el apartado 3 del art. 9 acota los riesgos a que se refiere este precepto: "Los riesgos a que se refiere el presente artículo son únicamente aquellos que pueden mitigarse o eliminarse razonablemente mediante el desarrollo o el diseño del

54 Para un examen completo y amplio del precepto me remito a SCHUETT, J., "Risk management in the Artificial Intelligence Act", en European Journal of Risk Regulation, 2023, págs. 1-19.

55 HERNÁNDEZ PEÑA lo califica de sistema de *compliance* o cumplimiento normativo (*El marco jurídico…*, op. cit., págs. 149-150).

sistema de IA de alto riesgo o el suministro de información técnica adecuada".

Por ello, es relevante la vinculación entre sistema de gestión de riesgos y ciclo de vida del sistema de IA[56]. Mientras un sistema de IA esté en uso debe mantenerse el sistema de gestión de riesgos.

Este requisito está plenamente justificado dado que, como se ha dicho reiteradamente, la regulación de la LIA está basada en el riesgo y por tanto se trata de controlar y minimizar los riesgos desde el inicio hasta el final de la vida del sistema de IA. Es, por ello, que el apartado 2 de este artículo 9 lo califique como "proceso iterativo continuo", sometido a revisiones y actualizaciones, e incluso relate las etapas de gestión de riesgos.

Estas etapas están divididas en función del tipo de riesgo: riesgo evaluable y previsible, y riesgo que podría surgir bien por una estimación de estas eventualidades no conocidas pero que podrían surgir por un uso adecuado o inadecuado del sistema (uso indebido razonablemente previsible) o bien derivados de la ejecución del sistema de vigilancia poscomercialización.

Procede detenerse en el primer caso, referido a los riesgos conocidos y previsibles derivados de una utilización del sistema de IA conforme a su finalidad. Se exige, primero, que se determinen y analicen estos riesgos, con especial atención a si afectan a menores o personas vulnerables (apartado 9). Además, se impone que se diseñen y se adopten medidas adecuadas para la gestión de estos riesgos. Los apartados 4 y 5 se refieren a estas medidas en orden a que minimicen los riesgos o evalúen aquellos que pueden ser considerados admisibles por su carácter residual.

56 PALMA ORTIGOSA, A., diferencia dos fases sustanciales en el ciclo de vida: la fase diseño y la fase de despliegue y toma de decisiones, que analiza posteriormente ("El ciclo de vida de los sistemas de inteligencia artificial. Aproximación técnica de las fases presentes durante el diseño y despliegue de los sistemas algorítmicos", en *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Aranzadi, Cizur Menor, 2022, págs. 37-48).

Para la determinación de estas medidas, los sistemas de IA de alto riesgo deben ser sometidos a pruebas, bien mediante banco de pruebas previo o incluso en condiciones reales, que deben ser realizadas en todo caso de forma previa a su introducción en el mercado y también, si procede, en cualquier momento del proceso de desarrollo (apartados 6-8).

IV.2.B.- Gobernanza de datos

El art. 10 impone el requisito de la gobernanza de datos[57]. Este precepto obliga a diferenciar entre:

1) Sistemas de IA: sólo se les aplica lo dispuesto en los apartados 2 a 5 para los conjuntos de datos de prueba.
2) Modelos de IA: suponen la utilización de conjuntos de datos para entrenamiento del modelo, y son el objeto de regulación de este precepto.
3) Modelos de IA de uso (o propósito) general, regulados de forma singular en el Capítulo V.

Así pues, este artículo se refiere, principalmente, a los modelos de IA, puesto que se aplica en el caso de sistemas de IA solo a los conjuntos de datos de prueba (apartado 6).

El Considerando 67 explica el objetivo de este precepto: "Los datos de alta calidad y el acceso a datos de alta calidad desempeñan un papel esencial a la hora de proporcionar una estructura y garantizar el funcionamiento de muchos sistemas de IA, en especial cuando se emplean técnicas que implican el entrenamiento de modelos, con vistas a garantizar que el sistema de IA de alto riesgo funcione del modo previsto y en condiciones de seguridad y no se convierta en una fuente de algún tipo de discriminación prohibida por el Derecho de la Unión".

57 HERNÁNDEZ PEÑA enfatiza en la importancia de la gobernanza de datos para los sistemas de IA de alto riesgo, dado que "parte importante de los riesgos para los derechos fundamentales de los sistemas de IA derivan de los datos utilizados para entrenar o validar los modelos" (*El marco jurídico…*, op. cit., págs. 150-151).

Se somete a los conjuntos de datos a prácticas de gobernanza y gestión de datos a fin de asegurar la calidad de datos y evitar o eliminar la introducción de sesgos.

En todo caso, la gobernanza de datos debe garantizar la aplicación de los principios de la normativa de protección de datos personales respecto de estos datos: minimización, anonimización o cifrado[58], puesto que, como señala el Considerando 69, "El derecho a la intimidad y a la protección de datos personales debe garantizarse a lo largo de todo el ciclo de vida del sistema de IA".

Téngase también en cuenta que se impone al responsable del despliegue la obligación de asegurarse de que los datos de entrada sean pertinentes y suficientemente representativos (art. 26.2).

Aparecen dos cuestiones relevantes en relación con los datos. La primera es el control para la evitación de la aparición de sesgos en los sistemas de IA (letras f) y g) del art. 10.2), que incluso permite el tratamiento excepcional de datos personales de categoría especial en orden a detectar y corregir sesgos (apartado 5)[59]. La aparición de sesgos se ha mostrado en diversos casos de uso de inteligencia artificial tanto en el sector público como en el sector

58 Así lo afirma el Considerando 69: "El derecho a la intimidad y a la protección de datos personales debe garantizarse a lo largo de todo el ciclo de vida del sistema de IA. A este respecto, los principios de minimización de datos y de protección de datos desde el diseño y por defecto, establecidos en el Derecho de la Unión en materia de protección de datos, son aplicables cuando se tratan datos personales. Las medidas adoptadas por los proveedores para garantizar el cumplimiento de estos principios podrán incluir no solo la anonimización y el cifrado, sino también el uso de una tecnología que permita llevar los algoritmos a los datos y el entrenamiento de los sistemas de IA sin que sea necesaria la transmisión entre las partes ni la copia de los datos brutos o estructurados, sin perjuicio de los requisitos en materia de gobernanza de datos establecidos en el presente Reglamento".

59 El Considerando 67 explica esta preocupación por los sesgos: "Los sesgos, por ejemplo, pueden ser inherentes a los conjuntos de datos subyacentes, especialmente cuando se utilizan datos históricos, o generados cuando los sistemas se despliegan en entornos del mundo real. Los resultados de los sistemas de IA dependen de dichos sesgos inherentes, que tienden a aumentar gradualmente y, por tanto, perpetúan y amplifican la discriminación existente, en particular con respecto a las personas vulnerables pertenecientes a determinados colectivos, en particular colectivos raciales o étnicos".

privado, con resultados inaceptables para determinadas personas o categorías de personas[60]. Se trata de salvaguardar los derechos fundamentales, principalmente, aquí la igualdad frente a la discriminación directa e indirecta por medio de algoritmos[61]. Para ello jugará un papel fundamental la evaluación de impacto relativa a los derechos fundamentales exigida por el art. 27[62].

La segunda cuestión es la de la representatividad de los datos. El art. 10 LIA exige que los datos sean suficientemente representativos (apartado 3) y que carezcan de errores, y que tengan en cuenta el entorno geográfico, contextual, conductual o funcional (apartado 4).

IV.2.C.- Documentación técnica

La documentación técnica del sistema de IA debe ser elaborada antes de su introducción al mercado, y posteriormente actualizada (art. 11). Su finalidad es doble: facilitar a las autoridades

60 BELLOSO MARTÍN indica que "los sesgos no ocurren de manera espontánea. Los algoritmos de IA basados en datos, no producen sesgos pero sí pueden reproducirlos sin la adecuada intervención humana y ello en tres de las fases principales: en la recolección de los datos, porque tales datos recopilados reflejen prejuicios ya existentes; en la preparación de datos de entrenamiento (a la hora de seleccionar y procesar los atributos que le proporcionamos al algoritmo); y en la toma de decisiones (Las propuestas y decisiones que se adoptan a lo largo de todo el ciclo de vida del desarrollo inteligente)" (pág. 48). Y más adelante pone ejemplos de casos de discriminación algorítmica por el uso de sesgos ("La problemática de los sesgos algorítmicos (con especial referencia a los de género). ¿Hacia un derecho a la protección contra los sesgos?", en *Inteligencia artificial y filosofía del Derecho*, Laborum. Murcia, 2022, págs. págs. 49-51), con especial atención a la discriminación por género (págs. 55-61). Por su parte, SORIANO ARNANZ, A., se refiere a la introducción de sesgos en la base de datos, en la clasificación, los aprendidos tras la puesta en marcha, y los derivados de las decisiones humanas que intervienen en el uso y funcionamiento de los sistemas algorítmicos ("La aplicación del marco jurídico europeo en materia de igualdad y no discriminación al uso de aplicación de inteligencia artificial", en *Nuevas normatividades: inteligencia artificial, derecho y género,* Aranzadi, Cizur Menor, 2021, págs. 65-68).

61 SORIANO ARNANZ, A., llama la atención sobre la necesidad de atender a ambos tipos de discriminación, tanto la directa como la indirecta ("La aplicación del marco jurídico...", op. cit., págs. 73-78.

62 Así lo sostiene SIMÓN CASTELLANO, "Las evaluaciones de impacto algorítmico en los derechos fundamentales: hacia una efectiva minimización de sesgos", en *Algoritmos abiertos y que no discriminen en el sector público*, Tirant lo blanch, Valencia, 2023, pág. 45-46.

competentes y organismos notificados el ejercicio de sus competencias, y permitir la comprobación del cumplimiento de los requisitos exigidos para los sistemas de IA de alto riesgo[63].

Debe indicar que se cumplen los requisitos de la Sección 2, y ser clara y completa. Su contenido mínimo viene fijado en el Anexo IV, que puede ser modificado por la Comisión (apartado 3).

Se prevé una documentación simplificada para las pymes. Y, asimismo, cuando el sistema de IA esté vinculado a un producto en el caso del Anexo I, la documentación será única, para el producto y para el sistema de IA.

El art. 77 LIA se refiere a un supuesto especial de solicitud de esta documentación por parte de las Autoridades competentes en materia de protección de datos, en orden a determinar si el sistema de IA cumple con la protección de los derechos fundamentales. Incluso de entender que no es suficiente dicha documentación podrá solicitar de la autoridad de vigilancia del mercado que el sistema de IA de alto riesgo se someta a pruebas.

IV.2.D.- Trazabilidad y registro

La idea fundamental de ciclo de vida o de proceso iterativo continuo comporta la exigencia de trazabilidad de los sistemas de IA (art. 12). En todo momento debe ser posible detectar y constatar los eventos en la ejecución del sistema de IA[64]. Por ello, el art. 12 impone el requisito de que los sistemas de IA de alto riesgo permitan técnicamente el registro de eventos en todo momento.

63 Así lo indica HERNÁNDEZ PEÑA (*El marco jurídico…*, op. cit., pág. 153).

64 SIMÓN CASTELLANO señala que la trazabilidad se refiere a que "cualquier acción que lleve a cabo un usuario del sistema…quedará registrado y dejará un rastro que, en el futuro y en caso de ser necesario, podrá ser examinado", por lo que estas acciones deben quedar registradas ("Allende una teoría general…", op. cit., págs. 134-135).

Para garantizar una adecuada trazabilidad, el registro tiene que tener unas capacidades de registro de eventos que permitan la detección de riesgos, la vigilancia poscomercial, y la vigilancia de su funcionamiento por parte de los responsables del despliegue. Incluso se imponen unos registros de eventos específicos para el caso de los sistemas biométricos (apartado 3).

IV.2.E.- Transparencia e información

El art. 13 LIA impone la exigencia de transparencia e información a los proveedores respecto de los responsables del despliegue[65]. Hay que advertir que no se trata de una transparencia general[66], sino que, por el contrario, se limita a determinar las relaciones entre el proveedor, en cuanto diseñador del sistema de IA, y los responsables del despliegue, en cuanto personas que utilizan el sistema[67].

A tal efecto los sistemas de IA de alto riesgo deben ir acompañados de instrucciones, que contendrán una información concisa, completa, correcta, clara, pertinente, accesible y comprensible. Incluso se impone un contenido mínimo de dichas instrucciones (apartado 3).

65 SIMÓN CASTELLANO eleva el marco de la transparencia dándole un alcance más general, diferenciando cinco subcategorías: simulabilidad, descomponibilidad, legibilidad, auditabilidad y publicidad activa. Y considera como categoría diferente, aunque próxima, a la explicabilidad que se divide a su vez en tres subcategorías: inteligilidad, comprensibilidad e interpretabilidad ("Allende una teoría general...", op. cit., págs. 127-133).

66 COTINO HUESO califica la transparencia de la propuesta de LIA como "transparencia interna", dado que no se refiere a las personas afectadas ("Transparencia y explicabilidad de la inteligencia artificial. Elementos conceptuales, generales y de género", en *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo blanch, Valencia, 2022, pág. 47).

67 El Considerando 72 señala en su primer párrafo lo siguiente: "A fin de abordar las preocupaciones relacionadas con la opacidad y complejidad de determinados sistemas de IA y ayudar a los responsables del despliegue a cumplir sus obligaciones en virtud del presente Reglamento, debe exigirse transparencia respecto de los sistemas de IA de alto riesgo antes de su introducción en el mercado o su puesta en servicio".

El Considerando 72 detalla el objeto de estas instrucciones: ayudar a utilizar el sistema y tomar decisiones con conocimiento de causa, y también elegir correctamente el sistema que se va a utilizar. Para ello se refiere a que las instrucciones contengan ejemplos prácticos y se redacten en la lengua que decida el Estado miembro. Por su parte, el art. 15.3 LIA exige que indiquen los niveles de precisión y los parámetros de evaluación de los sistemas de IA.

Por tanto, parece que la LIA engloba dentro de este requisito tanto la transparencia como la explicabilidad de los sistemas de IA[68], aunque limitada a las relaciones entre proveedor y responsable del despliegue. Para obligaciones más generales es preciso acudir a lo dispuesto en el art. 50 que es aplicable, también, a los sistemas de alto riesgo.

IV.2.F.- Supervisión humana

El art. 14 LIA impone la vigilancia humana[69] que supervise el funcionamiento del sistema de IA de alto riesgo, plasmando en el plano normativo las previsiones de la Declaración Europea sobre los Derechos y los Principios Digitales para la Década Digital[70]. Dicha supervisión debe estar prevista en el diseño del

68 Diversos autores, entre los que cabe destacar a COTINO HUESO, han insistido en las diferencias conceptuales y prácticas entre transparencia y explicabilidad (COTINO HUESO, L. y CASTELLANOS CLARAMUNT, J., *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo blanch, Valencia, 2022).

69 PONCE SOLÉ, J., considera la supervisión humana como una técnica menos drástica que la reserva de humanidad, puesto que permite el uso de un sistema de IA, aunque sometido a supervisión o vigilancia humana ("Reserva de humanidad y supervisión humana de la Inteligencia artificial", en *El Cronista del Estado Social y Democrático de Derecho*, núm. 100, 2022, pág. 64).

70 Esta Declaración en su Capítulo III, apartado 6, sobre Libertad de elección, se refiere a las interacciones con algoritmos y sistemas de inteligencia artificial, y dado que, la inteligencia artificial debe ser un instrumento al servicio de las personas, establece el compromiso de "velar por que los sistemas algorítmicos se basen en conjuntos de datos adecuados para evitar la discriminación y permitir la supervisión humana de todos los resultados que afecten a la seguridad y los derechos fundamentales de las personas" (letra c)). También en el Capítulo

sistema de IA y por tanto éste debe contar con un interfaz humano-máquina[71].

La supervisión se encomienda a una persona física. Como señala el art. 26.2 el responsable del despliegue debe designar para esta tarea personas físicas que tengan la competencia, formación y autoridad necesarias.

Es aquí donde cobran relevancia las instrucciones que el proveedor debe facilitar al responsable del despliegue, y que éste deberá transmitir a la persona supervisora. La vigilancia exige que la persona supervisora entienda el sistema de IA y su funcionamiento, hasta el extremo de que pueda decidir intervenir en dicho funcionamiento e, incluso, interrumpirlo (letra e) del apartado 4).

Las medidas de vigilancia las define el proveedor antes de la introducción del sistema de IA en el mercado o de su puesta en funcionamiento, y pueden venir integradas en el propio sistema y actuar desde el momento inicial o a lo largo de su funcionamiento (apartado 3).

El art. 14.5 recoge un supuesto de supervisión humana reforzada para determinados sistemas de identificación biométrica (Considerando 73), en los que la supervisión debe efectuarse por dos personas que lo verifiquen y confirmen por separado. No obstante, este requisito reforzado no se aplica en determinados casos (garantía del cumplimiento del Derecho, migración, control fronterizo o asilo) si el Derecho nacional o de la UE lo consideran desproporcionado (segundo párrafo del apartado 5).

II sobre Solidaridad e inclusión en el epígrafe 6 relativo a Condiciones de trabajo justas y equitativas, fija el compromiso de: "garantizar, en particular, que las decisiones importantes que afecten a los trabajadores cuenten con supervisión humana y que, en general, se los informe de que están interactuando con sistemas de inteligencia artificial" (letra e)).

71 La interfaz humano-máquina (HMI) debe estar incorporada al sistema y hay diversos modelos que permiten su control y supervisión por las personas físicas.

IV.2.G.- Precisión, solidez y ciberseguridad

El diseño y desarrollo de los sistemas de IA de alto riesgo debe garantizar su precisión, solidez y ciberseguridad a lo largo de todo su ciclo de vida (art. 15.1). Y en las instrucciones de uso deben indicarse los niveles de precisión y los parámetros para su evaluación (apartado 3).

La LIA diferencia entre solidez y resistencia. La solidez es *ad intra* del sistema, de modo que sea resistente a errores, fallos e incoherencias. El Considerando 75 pone como ejemplos de la solidez: "Estas soluciones técnicas pueden incluir, por ejemplo, mecanismos que permitan al sistema interrumpir de forma segura su funcionamiento (planes de prevención contra fallos) en presencia de determinadas anomalías o cuando el funcionamiento tenga lugar fuera de determinados límites predeterminados".

La resistencia es *ad extra* del sistema, de forma que se impida el acceso de terceros no autorizados, o terceros maliciosos (Considerando 76), por lo que se deben tener medidas de ciberseguridad. El Considerando 77 se remite en este punto de forma adicional a lo que disponga el futuro Reglamento de ciberseguridad, que deberá ser tenido en cuenta en el procedimiento de evaluación de conformidad previsto en la LIA (Considerando 78). En este momento se halla vigente la Directiva (UE) 2022/2555, relativa a las medidas destinadas a garantizar un elevado nivel común de ciberseguridad en la UE.

La UE ha creado la Agencia de la Unión Europea para la Ciberseguridad (ENISA), que se encuentra regulada en el Reglamento (UE) 2019/881, con el fin de alcanzar un nivel elevado de ciberseguridad, ciberresiliencia y confianza dentro de la Unión, en orden a garantizar el correcto funcionamiento del mercado interior.

V.- SISTEMAS DE IA DE RIESGO LIMITADO

El art. 50 impone obligaciones de transparencia para los sistemas de IA de alto riesgo y, también, para determinados sistemas de IA, aunque no sean de alto riesgo. De ello, se ha derivado la configuración de una nueva categoría de sistemas de IA: la de riesgo limitado.

El Considerando 132 lo explica así: "Determinados sistemas de IA destinados a interactuar con personas físicas o a generar contenidos pueden conllevar riesgos específicos de suplantación o engaño, con independencia de si cumplen las condiciones para ser considerados como de alto riesgo o no".

Esto significa que aquellos sistemas de bajo riesgo que incidan en alguna de las situaciones descritas en el art. 50 están sometidos a las obligaciones de información y transparencia, pero únicamente a éstas. Y, por otra parte, el art. 50 constituye una obligación adicional para los proveedores y responsables del despliegue de sistemas de IA de alto riesgo, añadida a las enumeradas en los arts. 16 y 24 LIA, como advierte el apartado 6 del art. 50: "Los apartados 1 a 4 no afectarán a los requisitos y obligaciones establecidos en el capítulo III...".

El art. 50 prevé cuatro supuestos:

1) Sistemas de IA destinados a interactuar directamente con personas físicas.
2) 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,
3) Sistemas de reconocimiento de emociones o de un sistema de categorización biométrica. Procede advertir que estos sistemas son de alto riesgo (Anexo III. punto 1).
4) Sistemas de IA que generen o manipulen imágenes o contenidos de audio o vídeo que constituyan una ultrafalsificación o que generen o manipulen texto que se publique

con el fin de informar al público sobre asuntos de interés público.

En todos estos casos se impone la obligación, a los proveedores en los dos primeros supuestos y a los responsables del despliegue en los supuestos tercero y cuarto, de informar a las personas sobre dichas circunstancias. También se contemplan excepciones a las obligaciones de transparencia, especialmente, cuando se trata de detección, prevención, investigación y enjuiciamiento de delitos.

Además de estas obligaciones de información, que debe facilitarse de forma clara y distinguible, se prevé la elaboración de códigos de buenas prácticas e, incluso, la adopción por la Comisión de actos de ejecución para especificar normas comunes para el cumplimiento de estas obligaciones (apartado 7 del art. 50).

VI.- SISTEMAS DE IA DE BAJO O NULO RIESGO

La LIA no regula los sistemas de IA de bajo o nulo riesgo, pero se refiere de forma directa a ellos en sus Considerandos 165 y 166. En el Considerando 165 reitera el sometimiento de los sistemas que no sean de alto riesgo a profundizar en la adopción de una inteligencia artificial ética y fiable en la Unión. Para ello, se opta por un sistema de autoregulación mediante códigos de conducta que vayan incorporando los requisitos y obligaciones contempladas en la LIA también para estos sistemas que no son de alto riesgo. Se trata de que se autoimpongan requisitos adicionales, de los que se ofrecen algunos ejemplos: "los elementos de las Directrices éticas de la Unión para una IA fiable[72], la sostenibilidad medioambiental, medidas de alfabetización en

72 El considerando 27 de la LIA señala que las Directrices éticas para una IA fiable son aplicables a todos los sistemas de IA.

materia de inteligencia artificial, la inclusividad y la diversidad en el diseño y el desarrollo de los sistemas de IA, lo que incluye tener en cuenta a las personas vulnerables y la accesibilidad de las personas con discapacidad, la participación de las partes interesadas".

Por su parte el Considerando 166 se refiere a los sistemas de IA asociados a productos, que quedan fuera de la LIA por no ser sistemas de alto riesgo, para los que se exige que sean seguros, remitiéndose a la aplicación, como red de seguridad, del Reglamento (UE) 2023/988[73].

La implicación más relevante de la LIA respecto de los sistemas de IA que no son de alto riesgo se encuentra establecida en su art. 95, referido a los códigos de conducta. Se anima a los proveedores, y también a los responsables del despliegue, a contar con códigos de conducta, que les conduzcan, de forma voluntaria, hacia el cumplimiento de los requisitos y obligaciones establecidos en la LIA.

Por último, conviene no confundir los sistemas de bajo o nulo riesgo con la declaración de un proveedor de que un sistema de IA incluido en el Anexo III como sistema de alto riesgo no es de alto riesgo. Estos supuestos encuentran diversas previsiones en la LIA. En primer lugar, se establece la obligación de registrarlos en la base de datos de la UE, prevista en el art. 71 (art. 49.2). En segundo lugar, se dispone que, cuando un sistema de IA que haya sido considerado que no es de alto riesgo se convierta en un sistema de IA de alto riesgo por cualquier motivo (desarrollo, uso, funcionamiento, modificaciones, etc.) deberá procederse a calificarlo como sistema de IA de alto riesgo y realizar la evaluación de conformidad (art. 80).

73 Este Reglamento regula la seguridad de los productos, para velar por la protección de los consumidores. Es una norma de aplicación supletoria, para todos aquellos productos que no tengan requisitos específicos de seguridad impuestos por el Derecho de la UE.

VII.- CONCLUSIONES

Del examen que acaba de efectuarse se desprenden algunas conclusiones generales sobre la regulación de la LIA respecto de los sistemas de IA.

En primer término, la LIA efectúa una clasificación de los sistemas de IA en cuatro categorías en función de los riesgos. No obstante, se trata de una categorización llena de excepciones.

No sólo el propio concepto de riesgo (inadmisible, elevado o mínimo) constituye una cuestión con contornos a menudo poco precisos, sino también la LIA encierra demasiada imprecisión y una gran vaguedad.

La LIA, a pesar de su ánimo de ser una ley unificadora de un régimen general aplicable a los sistemas de IA, es una norma que recibirá una aplicación desde diversas perspectivas, principalmente la innovación y la protección de los derechos fundamentales, y por organismos con competencias y funciones diversas tanto a nivel europeo (Comité Europeo de Inteligencia Artificial, Oficina de la IA, Comité Europeo de Protección de Datos) como a nivel estatal (autoridad de vigilancia del mercado, autoridad notificantes, autoridades de protección de datos personales).

Así pues, la LIA precisa de una mayor concreción, sobre todo, para una aplicación práctica segura y acorde a sus disposiciones, que se efectuará tanto por actos normativos o de ejecución como por guías, orientaciones o instrucciones (*soft law*).

VIII.- BIBLIOGRAFÍA

BECK, U., *La sociedad del riesgo. Hacia una nueva modernidad,* Paidós, Barcelona, 2010.

BELLOSO MARTÍN, N., "La problemática de los sesgos algorítmicos (con especial referencia a los de género). ¿Hacia un derecho a la protección contra los sesgos?", en *Inteligencia artificial y filosofía del Derecho*, Laborum. Murcia, 2022, págs. 45-78.

CHRISTAKIS, T. y KARRATHANASIS, T., "Tools for Navigating the EU AI Act (2) Visualisation Pyramid", AI Regulation Papers 24-03-5, AI-Regulation.com, March 8th, 2024.

COTINO HUESO, L., "Sistemas de inteligencia artificial con reconocimiento facial y datos biométricos. Mejor regular bien que prohibir mal", en *El Cronista del Estado Social y Democrático de Derecho*, núm. 100, 2022, págs. 68-79.

COTINO HUESO, L., "Transparencia y explicabilidad de la inteligencia artificial. Elementos conceptuales, generales y de género", en *Transparencia y explicabilidad de la inteligencia artificial,* Tirant lo blanch, Valencia, 2022, págs. 25-70.

COTINO HUESO, L. y CASTELLANOS CLARAMUNT, J., *Transparencia y explicabilidad de la inteligencia artificial*, Tirant lo blanch, Valencia, 2022.

COTINO HUESO, L. et al., "Un análisis crítico constructivo de la propuesta de Reglamento de la Unión Europea por la que se establecen normas armonizadas sobre la Inteligencia Artificial (Artificial Intelligence Act)", en *Diario La Ley, sección Ciberderecho*, 2 de julio de 2021.

COTINO HUESO, L., "Nuevo paradigma en la garantía de los derechos fundamentales y una nueva protección de datos frente al impacto social y colectivos de la inteligencia artificial", en *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Aranzadi, Cizur Mayor, 2022, págs. 69-105.

COTINO HUESO, L., "Reconocimiento facial automatizado y sistemas de identificación biométrica bajo la regulación superpuesta de inteligencia artificial y protección de datos", en *Derecho Público de la Inteligencia Artificial*, Fundación Manuel Giménez Abad, Zaragoza, 2023, págs. 347-402.

DE LA QUADRA FERNÁNDEZ DEL CASTILLO, T. (2019). "Derechos fundamentales, democracia y mercado en la edad digital". *Derecho digital e Innovación* núm. 1.

DE MIGUEL ASENSIO, P.A., "Propuesta de Reglamento sobre inteligencia artificial", *La Ley Unión Europea*, núm. 92, mayo 2021.

ESTEVE PARDO, J., *Técnica, riesgo y Derecho. Tratamiento del riesgo tecnológico en el Derecho ambiental,* Ariel, Barcelona, 1999.

FERNÁNDEZ HERNÁNDEZ, F., "La futura regulación europea de la inteligencia artificial: objetivos, principios y pautas", en *Claves de inteligencia artificial y derecho*, La Ley, Madrid, 2022, págs. 115-179.

HERNÁNDEZ PEÑA, J.C., *El marco jurídico de la inteligencia artificial. Principios, procedimientos y estructuras de gobernanza*, Aranzadi, Cizur Menor, 2022.

HERNÁNDEZ PEÑA, J.C., "Organización y gobernanza de la inteligencia artificial: marco general", en *Inteligencia artificial y sector público. Retos, límites y medios*, Tirant lo blanc, Valencia, 2023, pp. 599-630.

HUERGO LORA, A., "Gobernar con algoritmos, gobernar los algoritmos", *El Cronista del Estado Social y Democrático de Derecho,* núm. 100, 2022, págs. 80-89.

MORAL SORIANO, L. "Modelos de gobernanza global de la inteligencia artificial", en *Inteligencia artificial y Derecho. El jurista ante los retos de la era digital,* Aranzadi, Cizur Menor, 2021, págs. 235-258.

PALMA ORTIGOSA, A., "El ciclo de vida de los sistemas de inteligencia artificial. Aproximación técnica de las fases presentes durante el diseño y despliegue de los sistemas algorítmicos", en *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Aranzadi, Cizur Menor, 2022, págs. 29-51.

PONCE SOLÉ, J., "Reserva de humanidad y supervisión humana de la Inteligencia artificial". *El Cronista del Estado Social y Democrático de Derecho*, núm. 100, 2022, págs. 58-67.

SALAZAR GARCÍA, I., "Retos actuales de la ética en la Inteligencia Artificial", en *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Aranzadi, Cizur Menor, 2022, págs. 53-66.

SCHUETT, J., "Risk management in the Artificial Intelligence Act", en European Journal of Risk Regulation, 2023, págs. 1-19.

SIMÓN CASTELLANO, P., "Allende una teoría general de las garantías jurídicas para una inteligencia artificial confiable", en *Derecho Público de la Inteligencia Artificial*, Fundación Manuel Giménez Abad, Zaragoza, 2023, págs. 111-148.

SIMÓN CASTELLANO, P., "Las evaluaciones de impacto algorítmico en los derechos fundamentales: hacia una efectiva minimización de sesgos", en *Algoritmos abiertos y que no discriminen en el sector público*, Tirant lo blanch, Valencia, 2023, págs. 27-56.

SORIANO ARNANZ, A., "La aplicación del marco jurídico europeo en materia de igualdad y no discriminación al uso de aplicación de inteligencia artificial", en *Nuevas normatividades: inteligencia artificial, derecho y género,* Aranzadi, Cizur Menor, 2021, págs. 63-87.

VIDA FERNÁNDEZ, J., "La gobernanza de los riesgos digitales: Desafíos y avances en la regulación de la Inteligencia artificial", *Cuadernos de Derecho Transnacional*, núm. 14-1.

Síganos en Linked

Visite nuestra web e infórmese de las novedades y actividades formativas que realizamos

www.rdu.es

2444-5762