La intervención jurídico-pública en la IA

Autora:
Raquel Del Río Torralbo
5º Doble grado de Derecho y Business Analytics
Área Derecho Administrativo
Directora:
Prof. Maria Burzaco Samper

Dedicado a mi madre

Las palabras no pueden expresar lo mucho que te debo, gracias por cuidarme y apoyarme en todas mis decisiones (sobre todo en aquellas que sabias que no eran las correctas), no hubiese llegado aquí sin ti.

Resumen

La IA se ha erigido como uno de los principales pilares sobre los que se va a desarrollar la tecnología en el corto y medio plazo. Las grandes opciones que abre la introducción de esta tecnología también llevan aparejada una serie de riesgos, que han de ser tratados de forma clara y directa. El máximo exponente de este tratamiento del riesgo y de los efectos de la IA se puede observar en el Futuro Reglamento de IA (FRIA) y en las guías e informes que lo han desarrollado. A través de este trabajo, se analizará la futura legislación y se expondrá el enfoque de clasificación basado en el riesgo y aquellas obligaciones asociadas a éste. Asimismo, dada la importancia de la protección de datos en el desarrollo de esta tecnología, el trabajo contendrá numerosas referencias a la legislación de esta materia.

En este trabajo se ha seguido una metodología para el análisis jurídico de la IA, que consta de los siguientes pasos: una presentación de la delimitación de la IA, una exposición de la clasificación de los sistemas IA y los problemas asociados a la misma con un enfoque basado en la protección de datos y, finalmente, un análisis completo de las obligaciones que deben cumplimentar las partes vinculadas a los sistemas IA.

Abstract

Artificial Intelligence has emerged as one of the main pillars on which technology will develop in the short and medium term. The great options opened by the introduction of this technology also entail a series of risks, which must be dealt clearly and directly. The best example of this treatment of risk and the effects of AI can be seen in the Future Regulation of Artificial Intelligence (FRIA) and in the guides and reports that have developed it. Through this paper, the future legislation will be analyzed, and the risk-based classification approach and those obligations associated with it will be exposed. Also, given the importance of data protection in the development of this technology, the paper will contain numerous references to the legislation in this area.

This paper has followed a methodology for the legal analysis of AI, which consists of the following steps: a presentation of the delimitation of AI, a presentation of the classification of AI systems and the problems associated with it with a data protection-based approach and, finally, a complete analysis of the obligations to be fulfilled by the parties linked to AI systems.

Palabras Claves

IA, IA de alto riesgo, obligaciones, clasificación, FRIA, Libro Blanco, transparencia, cumplimiento, garantía, conformidad, identificación biométrica, riesgo inaceptable, riesgo limitado, riesgo mínimo, Chat GPT

Key Words

Artificial Intelligence, High Risk AI, requirements, classification, FRIA, White Paper, transparency, warranty, compliance, fulfillment, biometric identification, Unacceptable Risk, Low Risk, Minimal Risk, Chat GPT

Índice

Índice de siglas

AEPD Agencia Española de Protección de Datos

AIPD Agencia Italiana de Protección de Datos

CE Constitución Española

CEDH Convenio Europeo de Derechos Humanos

CESE Comité Económico y Social Europeo

EM Estado Miembro

FRIA Futuro Reglamento de la IA

IA Inteligencia Artificial

LGT Ley 58/2003, de 17 de diciembre, General Tributaria

LOPD Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales

LPI Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia

RGPD Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos)

TC Tribunal Constitucional

UE Unión Europea

CAPÍTULO I. INTRODUCCIÓN

1. OBJETO DE LA INVESTIGACIÓN

La eclosión de la Inteligencia Artificial (en adelante IA) ha cambiado la forma en la que la sociedad interactúa con la tecnología. Las aplicaciones que surgen de la IA son incontables y a día de hoy no se puede poner un límite claro a las posibilidades que pueden desarrollarse. No obstante, un avance tecnológico de tal envergadura también lleva aparejado una serie de cuestiones legales que deben ser abordados.

El objeto de este trabajo pivotará sobre un análisis detallado sobre los desafíos jurídicos que se han abordado en el Futuro Reglamento de IA (FRIA), ponderando las soluciones que se han ofrecido y sus limitaciones frente a los nuevos avances tecnológicos. Se considerará, dada su notable relación, la regulación de protección de datos y sus efectos en la materia.

2. ANTECEDENTES

Al tratarse de una materia tan novedosa y con poca regulación, no existe aún abundante cantidad doctrinal. No obstante, la importancia cada vez mayor de la IA ha propiciado la proliferación de un debate en torno a las cuestiones jurídicas que afectan, o podrían afectar, a esta materia.

Desde el plano comunitario el Reglamento General de Protección de Datos (RGPD) se erige como la principal referencia normativa, en concreto, la regulación de las decisiones automatizadas y el tratamiento de datos en general afectan directamente al ejercicio de las actividades desarrollados por la IA. Adicionalmente, el FRIA, aunque todavía no en vigor, se pretende convertir en la principal fuente legislativa en torno a la IA.

Por su parte, desde el plano nacional la Ley 15/2022 es la primera en incluir una referencia directa sobre la IA. Desde dicha norma, se demanda a las Administraciones Públicas garantizar que los algoritmos que vayan a ser empleados para tomar decisiones

cuenten con unos criterios de minimización de sesgos, transparencia y rendición de cuentas adecuados.

En definitiva, las regulaciones y estudios relativos a la IA están en auge habiéndose producido ya importantes antecedentes y desarrollos legislativos claves para el desarrollo de la IA en Europa.

3. OBJETIVOS PERSEGUIDOS

A través del desarrollo de este trabajo se persiguen una serie de objetivos. Principalmente, se pretende mostrar una visión completa de la situación legislativa de la IA, localizando las principales normas españolas y europeas. Asimismo, se tratará de identificar los principales problemas que se derivan de la clasificación por riesgos de la IA. Finalmente, se llevará a cabo un análisis completo de las obligaciones previstas en el FRIA.

4. METODOLOGÍA

Este trabajo ha sido estructurado sobre tres pilares fundamentales: el marco jurídico de la IA, la clasificación de la IA y los problemas asociados a la misma y las obligaciones impuestas a los sistemas IA de alto riesgo.

En primer lugar, se ha expuesto la evolución de la delimitación de la IA a lo largo de estos últimos años. Una vez establecida una definición concreta, se ha acudido a la regulación europea donde se ha llevado a cabo un estudio sobre los estudios y leyes que han perfilado la legislación comunitaria. Tras ello, el trabajo se ha remitido a la legislación nacional para ponderar las guías y normas que han definido en España el impulso europeo de la materia.

En segundo lugar, se ha presentado la clasificación de riesgos de la IA. Dentro de este capítulo se han expuesto las categorías y se han valorado los principales conflictos que afectan esta delimitación, De esta manera, se ha considerado necesario analizar la situación de la identificación biométrica, la cual actualmente cuenta tanto con supuestos encuadrados en dos categorías como con actividades no reguladas en ningún sector en concreto. Por otro lado, también se ha examinado la sobredimensión de la categoría de

alto riesgo, ésta ha aglutinado el mayor número de supuestos provocando un desequilibrio entre las materias reguladas. Finalmente, se ha estimado la situación de los procesadores de lenguaje (como puede ser Chat GPT) en el FRIA que puede ser insuficiente para proteger los derechos de los usuarios.

Por último, se ha realizado un análisis de las obligaciones de la categoría de IA de alto riesgo, al ser la sección con el mayor número de requisitos, con la que se ha pretendido identificar la pertinencia de las mismas y la posibilidad de materializarlas. Con dicho objetivo en mente se han localizado las obligaciones generales, las obligaciones previstas para los proveedores y las obligaciones previstas a los distribuidores.

CAPÍTULO II. MARCO JURÍDICO DE LA IA

1. DELIMITACIÓN JURÍDICA DE LOS SISTEMAS IA

La constante evolución y desarrollo de los sistemas de IA ha provocado la aparición de un sinfín de herramientas y soluciones.

Esta proliferación de recursos plantea una dificultad añadida para el legislador a la hora de crear un marco normativo aplicable, pues la mejora y el perfeccionamiento de las técnicas impide el desarrollo de una definición cerrada de la IA.

Entre las múltiples acepciones de la IA, algunos autores han optado por definir este tipo de sistemas en función del enfoque aplicado. En esta línea, se observa la siguiente aproximación a una definición de un sistema de IA: "*sistema que piensa como un humano, sistema que piensa racionalmente, sistema que actúa imitando el comportamiento humano o sistema que actúa racionalmente imitando el comportamiento humano"* (S.J. y Norvig, P., 2008, pp 2-3).

Esta dificultad a la hora de encontrar una definición aceptada y concreta se puede observar con claridad a través de la evolución de los dictámenes provenientes de la Unión. De esta forma, si bien en el 2017 la UE admitía la incapacidad de formular una acepción única (Comité Económico y Social Europeo, 2017, pp 58-65), ante los avances constantes de la tecnología y las implicaciones legales que se comienzan a vislumbrar, es en un dictamen del CESE de 2018 cuando por primera vez se ofrece una definición preliminar de la IA: "*A los efectos del presente dictamen, consideraremos la Inteligencia Artificial como una disciplina tendente a utilizar las tecnologías digitales para crear sistemas capaces de reproducir de forma autónoma las funciones cognitivas humanas, incluida la captación de datos y formas de comprensión y adaptación*" (Comité Económico y Social Europeo, 2018, p 3).

Esta primera interpretación de la IA se vería completada posteriormente en una Comunicación de la Comisión Europa al propio CESE. En dicho escrito se apostó por una delimitación más concreta de la materia: *"El término «Inteligencia Artificial» (IA) se aplica a los sistemas que manifiestan un comportamiento inteligente, pues son capaces de analizar su entorno y pasar a la acción –con cierto grado de autonomía– con el fin de alcanzar objetivos específicos. Los sistemas basados en la IA pueden consistir*

simplemente en un programa informático o estar incorporada en dispositivos de hardware (p. ej. robots avanzados, automóviles autónomos, drones o aplicaciones del internet de las cosas)" (Comisión Europea, 2020a, p1).

Las definiciones anteriormente expuestas son fruto de la estrategia europea de IA emprendida en 2018 y que culmina con la publicación del Libro Blanco en 2020 (Casado, E.G., 2021).

En dicho estudio se persiguió un concepto de IA preciso, capaz de proveer de la seguridad jurídica necesaria, pero sin caer en una restricción excesiva que limitase los avances tecnológicos. Para ello, la definición distingue entre la capacidad de los algoritmos y sistemas de percibir el entorno, de la responsabilidad humana a la hora de programar el funcionamiento de las herramientas (Fernández, 2020).

En base a estos objetivos, la definición que se incluiría finalmente en el Libro Blanco sería la siguiente : "*Los sistemas de Inteligencia Artificial (IA) son programas informáticos (y posiblemente también equipos informáticos) diseñados por seres humanos que, dado un objetivo complejo, actúan en la dimensión física o digital mediante la percepción de su entorno mediante la adquisición de datos, la interpretación de los datos estructurados o no estructurados, el razonamiento sobre el conocimiento o el tratamiento de la información, fruto de estos datos y la decisión de las mejores acciones que se llevarán a cabo para alcanzar el objetivo fijado."* (Comisión Europea, 2020b, p 20).

Finalizando este recorrido en busca de una definición satisfactoria de la IA, se debe aceptar que la tecnología está en constante crecimiento, y que los múltiples enfoques desde los que se puede explicar la IA impiden la formulación de un concepto concreto y aceptado transversalmente.

En consecuencia, desde la Unión Europea se ha apostado por la delimitación jurídica de los sistemas IA en Europa, abandonando la creación de una definición concreta de la misma. De esta forma, el artículo 3 del futuro Reglamento Europea de la IA (en adelante FRIA) delimita la regulación en base al siguiente concepto:

"Sistema de IA como el software que se desarrolla siguiendo una o varias técnicas o enfoques específicos[1] *«y que puede, para un conjunto determinado de objetivos definidos por el ser humano, generar resultados tales como contenidos, predicciones, recomendaciones o decisiones que influyen en entornos con los que se interactúa"*

Dada la importancia que tendrá esta regulación, no solo para los territorios comunitarios sino también para el resto de los territorios que comercializan con la Unión, esta definición será probablemente el punto de referencia desde el que se regulará la IA y, por tanto, actualmente se presenta como la principal delimitación jurídica de la materia.

2. LEGISLACIÓN EUROPEA DE LA IA

Los numerosos beneficios y cambios que traerá consigo el uso de la IA han obligado a la UE a delimitar el uso de este tipo de técnicas. En este sentido, el ordenamiento jurídico europeo pretende adoptar una posición favorable a la IA que genere en la sociedad un sentimiento de confianza y seguridad respecto a este medio tecnológico.

2.1.Primera Etapa: Informes y guías de actuación hasta el 2020

Con el objetivo de regular la IA, en esta primera etapa desde la UE se desarrollaron los informes y guías que delimitarían el desarrollo legislativo posterior.

De esta forma, las comunicaciones que se presentaron en 2017 y 2018 (y que ya fueron mencionadas en el anterior capítulo) culminaron con la publicación de las Directrices Éticas para una IA confiable.

[1] Dichas técnicas a las que se hacen referencia se encuentran situadas en el Anexo I de la Propuesta del Futuro Reglamento Europea de IA. Siendo las siguientes:

- Estrategias de aprendizaje automático, incluidos el aprendizaje supervisado, el no supervisado y el realizado por refuerzo, que emplean una variedad de métodos, entre ellos el aprendizaje profundo.
- Estrategias basadas en la lógica y el conocimiento, especialmente la representación del conocimiento, la programación (lógica) inductiva, las bases de conocimiento, los motores de inferencia y deducción, los sistemas expertos y de razonamiento (simbólico).
- Estrategias estadísticas, estimación bayesiana, métodos de búsqueda y optimización.

Este informe se enfocó en cuatro principios básicos sobre los que se detallaban las prácticas que debían seguir las soluciones IA: respeto por la autonomía humana, prevención del daño, justicia y explicabilidad del sistema (Comisión Europea, 2019a).

Sobres estos principios se erigieron varios requisitos: la supervisión humana, la solidez y seguridad técnica, la privacidad y gobernanza de datos, la transparencia, la diversidad, no discriminación y justicia, el bienestar social y medioambiental y la responsabilidad que serán expuestos a continuación.

La supervisión humana enfatizaba la importancia de que los sistemas contasen con controles humanos independientes que habilitasen una intervención humana en cualquier etapa del proceso de la solución IA (Comisión Europea, 2019a)

Por su parte, la solidez y seguridad técnica destacaba la importancia de que los sistemas IA contasen con la capacidad para adaptarse a los cambios del entorno y, sobre todo, de prevenir los potenciales daños a los que pudiese verse afectada una solución IA. En suma, en esta sección del informe se reflejaba la necesidad de que todo sistema debía contar con procesos que mostraran a los usuarios la resistencia, seguridad y confianza de la IA (Comisión Europea, 2019a).

Continuando con los principios básicos enunciados en el informe, la privacidad y gobernanza de los datos obligaba a los sistemas IA a adaptarse a la legislación de protección de datos europea. Dada la especial vulnerabilidad que puede derivarse de un mal uso de los datos en este tipo de procesos, los sistemas IA debían presentar un modelo de gobernanza de datos robusto que abarcase la calidad e integridad de los datos y la necesidad de acceso a los mismos (minimizando la recogida indiscriminada de información)[2] (Comisión Europea, 2019a).

[2] La protección a la gobernanza de datos puede ser observada en sentencias como la emitida por el Tribunal Rechtspraak en una sentencia de 2020. En dicha sentencia por primera vez en Europa se prohibía un sistema IA por la desproporcionalidad de sus algoritmos. En dicha sentencia se declaraba la ilegalidad del sistema algorítmico utilizado por el Gobierno de los Países Bajos para evaluar el riesgo de fraude a la seguridad social o a hacienda, concretamente la recolección de datos se consideraba desproporcionada y el proceso es sí mismo carente de la transparencia debida. Todo ello, vulneraba directamente los requisitos de proporcionalidad expuestos en el artículo 8 del CEDH (Rechtspraak, 2020).

Otra sentencia europea que merece la pena mencionar, es la sentencia 2465/2017 del Tribunale Amministrativo Regionale. En dicha resolución se juzgaba la accesibilidad al procedimiento empleado por el Ministerio de Educación Italiano para automatizar las soluciones de movilidad de los docentes. En concreto, el tribunal consideró que el algoritmo empleado era un acto administrativo y, en consecuencia,

Por otra parte, la transparencia hacía referencia al deber de explicabilidad que debían cumplir los sistemas IA. Específicamente, esta obligación incluía la transparencia tanto en los datos empleados en el sistema como en el proceso de construcción y funcionamiento del modelo entrenado. Al mismo tiempo, se requería a los proveedores de sistema IA, que informasen a los usuarios que interactuasen con la solución de las capacidades y limitaciones asociadas a la IA (Comisión Europea, 2019a).

Seguidamente, la diversidad, no discriminación y justicia con la que debía contar el sistema IA forzaba a la construcción de un modelo de diseño accesible, universal y libre de sesgos. Para ello, se debía fijar el usuario objetivo que iba a utilizar el sistema y adoptar medidas consecuentes al mismo que minimizasen posibles discriminaciones futuras (Comisión Europea, 2019a).

Por su lado, el bienestar social y medioambiental perseguido en los sistemas IA se debía entender como una extensión directa de los objetivos proyectados en el Pacto Verde Europeo (Comisión Europea, 2019b). En los mismos, se abogaba por el desarrollo de tecnologías sostenibles y respetuosas con el medio ambiente, y, en el caso de no poder ser así, establecer el traslado de la responsabilidad de los daños provocados a las empresas proveedoras de la tecnología responsable (Comisión Europea, 2019b).

Finalmente, la responsabilidad mencionada en el documento pretendía asegurar la rendición de cuentas y la reparación de los sistemas IA que se podían ver afectados. Asimismo, desde este requisito se pretendía cuantificar los impactos negativos derivados del uso de los sistemas IA, para ello la correcta localización de la responsabilidad requeriría de la identificación, evaluación, documentación de las soluciones IA (Comisión Europea, 2019b).

Tras la emisión de este informe, la UE fue consciente de la necesidad de imponer el cumplimiento de características y límites a los sistemas IA. En concreto, se comenzó a valorar el desarrollo de una legislación comunitaria en torno a la materia. Con este fin en mente, la UE desarrollaría el Libro Blanco antes mencionado, este informe marcaría la emisión de recomendaciones e informes relativos a la IA y, sobre todo, se convertiría en la piedra angular sobre la que se desarrollaría la Propuesta del Reglamento IA.

---

se debía permitir el acceso al funcionamiento y creación de los sistemas utilizados en el procedimiento (Tribunale Amministrativo Regionale (TAR), 2017).

El objetivo último del Libro Blanco era reflejar las necesidades normativas de la IA. En concreto, pretendía impulsar la intervención de las instituciones europeas en el sector a partir de la presentación de una guía de principios y recomendaciones (De Hoyos Sancho, 2021, pp 9-44). La propuesta del estudio consistía en poner de relieve los requisitos que un sistema IA debía demostrar tener para que ser considerado confiable.

En el Libro Blanco se urgía a los máximos responsables europeos a liderar la construcción de un marco normativo que garantizase la seguridad de los sistemas en Europa, especialmente de los derechos de los datos personales de los ciudadanos. De esta forma, el Libro Blanco actúa en consonancia con la línea doctrinal comunitaria encuadrada en la estrategia de protección de datos, con el RGPD[3] como máximo exponente, que pretendía garantizar la confianza de los usuarios en la IA. Todo ello, convergía en el objetivo último de convertir a la UE en un referente legislativo de primer nivel (Comisión Europea, 2020b).

Los principios enunciados por el Libro Blanco se verían desarrollados en el Primer Plan Estratégico de Horizonte Europa 2021-2024[4]. En este informe, se expusieron tanto las características de sostenibilidad con las que tenían que contar los sistemas IA como los requerimientos mínimos que debían cumplimentarse para garantizar la seguridad y transparencia de las herramientas. Los puntos más importantes que se desprendían del documento fueron los siguientes (Comisión Europea, 2021c):

a) El informe demandaba que los sistemas IA debían desarrollarse en base a aprendizajes abiertos que favoreciesen el progreso del sector tecnológico
b) En el documento se demandaba el aprovechamiento de las soluciones propias de la IA para mejorar los ecosistemas y la biodiversidad en Europa a través de una gestión eficiente y sostenible de los recursos.
c) El uso de la IA ha de estar orientado a convertir a la UE en una referencia global. En particular, se debía posibilitar que este tipo de soluciones fueran empleadas de

[3] Los sistemas IA se nutren de una colección de datos que son tratados para proveer soluciones automatizadas, la conexión directa con la recolección directa se encuentra directamente protegida por los contenidos dispuestos en el RGPD y, por tanto, su análisis es indispensable a la hora de tratar la IA.

[4] Comisión Europea (2021c). *Primer plan estratégico de Horizonte Europa 2021-2024: la Comisión establece prioridades de investigación e innovación para un futuro sostenible*. Disponible en: https://ec.europa.eu/commission/presscorner/detail/es/ip_21_1122

forma que se habilite a la UE a convertirse en la primera economía digital, circular, sostenible y climáticamente neutra.

d) Los sistemas IA tenían que emplearse para crear una sociedad europea más sostenible, integradora y democrática.

Todos estos informes y guías permitieron a la Unión establecer los primeros pasos para desarrollar en la segunda etapa una regulación de la IA.

2.2. Segunda etapa: Situación actual de la futura legislación Europea

Actualmente, la legislación europea en torno a la IA se encuentra en construcción. Específicamente, la UE ha desarrollado una propuesta de reglamento en la que ha enunciado el marco legal en el que se regirán los sistemas IA que operen en el territorio comunitario[5].

En su exposición de motivos, el futuro FRIA expone su voluntad de garantizar la seguridad y legalidad de los sistemas IA introducidos en el mercado comunitario respecto a los derechos fundamentales y los valores de la Unión. En esta línea, se incide en la necesidad de garantizar la seguridad jurídica para impulsar la inversión e innovación en IA, y, al mismo tiempo, de facilitar el desarrollo de un mercado único que permita un uso legal, seguro y fiable de las aplicaciones de IA, evitando de esta forma la fragmentación del mercado (Comisión Europea, 2021a).

Esta futura normativa, se ha creado bajo un enfoque basado en el riesgo desde donde se ha clasificado los sistemas IA en función de los efectos que se desprenden de los mismos. En particular, los riesgos deben calcularse valorando el impacto que pueden tener las soluciones tanto en los derechos como en la seguridad de las personas (Comisión Europea, 2021a).

En base a este planteamiento, la norma incluye una estructura regulatoria con cuatro niveles diferenciados: IA prohibida, IA de alto riesgo, IA de riesgo bajo e IA de riesgo

[5] Comisión Europea (2021a). *Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de IA (Ley de IA) y se modifican determinados actos legislativos de la Unión* (FRIA). COM/2021/206 final. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:52021PC0206

mínimo[6]. En función del nivel de riesgo se incluyen unas obligaciones u otras para los proveedores y distribuidores.

Adicionalmente a las obligaciones expuestas en el Reglamento, en septiembre de 2022 se presentó una Propuesta de Directiva relativa a la responsabilidad de la IA complementaria a la FRIA (Comisión Europea, 2022a). En dicha Directiva se establecieron dos principales novedades relativas a la responsabilidad extracontractual.

En primer lugar, en referencia a la solicitud de información de los sistemas, se permite a los tribunales ordenar la revelación de pruebas relativas a los daños causados por los sistemas IA de alto riesgo. Como límite a esta práctica, el demandante deberá demostrar haber tratado de obtener la información a través de los canales disponibles, y, apoyar la petición con pruebas y hechos capaces de sustentar la reclamación de daños (Pina, 2022).

En segundo lugar, la directiva establece una presunción 'iuris tantum' para aquellas situaciones en las que el demandante demuestre que existe un nexo razonablemente causal entre el daño y uso del sistema. Esta posibilidad solo aplica a los sistemas de alto riesgo cuando la capacidad de probar el nexo sea excesivamente complicada; los sistemas que no sean de alto riesgo cuando no sea posible acceder a las pruebas; y los sistemas de IA para actividades no profesionales cuando el demandante haya obstaculizado el correcto funcionamiento u opte por no aclarar el funcionamiento de la solución (Pina 2022).

Respecto a la situación actual del Reglamento[7], tras la inclusión de esta última directiva, todavía se espera que sea aprobado por el Consejo y el Parlamento Europeo. La fecha más cercana en la se prevé que entre vigor es el segundo semestre del 2024. (Comisión Europea, 2022b).

[6] La clasificación será analizada en el capítulo III.

[7] Adicionalmente, se ha incluido la construcción de un Sandbox regulatorio para testar la eficacia del futuro Reglamento. Un Sandbox regulatorio es un entorno controlado que puede ser utilizado para desarrollar e impulsar proyectos con bases tecnológicas La especialidad que se desprende de estos proyectos reside en la relajación de los requisitos legislativos, de esta forma, a través de una autoridad reguladora se crea un marco temporal para que las empresas puedan explorar las potenciales soluciones tecnológicas, sin dañar al mercado en su conjunto (Marín Moreno, 2020).

En el caso concreto de la Unión, ha sido España el estado elegido para poner en marcha el primer Sandbox regulatorio de la UE en materia de IA. De esta forma, se pretende controlar la eficacia del Reglamento de la IA a partir de la monitorización de proyectos desarrollados por pymes, startups y en empresas en el entorno controlado. Mediante los resultados que se obtengan se matizarán los controles y se podrán, si fuera necesario, redefinir los procesos de supervisión y control.

A expensas de que se apruebe esta futura legislación, la Unión ha emitido el informe "Identificación y evaluación de la legislación comunitaria vigente y en preparación en el ámbito digital"[8] para aclarar aquellos aspectos que pudieran haberse quedado inconclusos en la FRIA.

En dicho informe además de exponer el plan de acción comunitario relativo a la IA, se localizaron las lagunas legislativas existentes. En este sentido, el estudio localiza aquellos aspectos que no han sido incluidos en el Reglamento y que podrían provocar conflictos legislativos en un futuro. Entre las materias que no fueron incluidos en el informe se encuentran aquellos sistemas de IA con objetivos militares o aquellos relativos a la identificación biométrica. Adicionalmente, el informe demostró que los criterios de identificación de los sistemas IA de alto riesgo podían no ser lo suficientemente eficientes (Parlamento Europeo, 2022).

Asimismo, a través de este informe se reflejó la necesidad de emitir normativas de IA humanistas que situasen al ser humano en el centro y persiguiesen los valores de la UE. Entre las diferentes propuestas incluidas en el escrito, se muestra la posibilidad de crear bases de datos públicas que favorezcan el desarrollo de la IA, o, de rebajar las medidas de control a los sistemas que no sean de alto riesgo. Ello con la idea de favorecer que Europa pueda ser competitiva en el sector (Parlamento Europeo, 2022).

En suma, la legislación europea relativa a la IA está en construcción y se espera que sea definida en los próximos meses. No obstante, si bien la UE muestra una intención de convertirse en una potencia en IA, del FRIA se desprende una contradicción evidente. La innovación en materias tan novedosas requiere de una libertad de creación suficiente para desarrollar nuevas soluciones, lo cual indirectamente va a requerir del uso de un gran número de datos. La férrea protección que se desprende tanto del RGPD como del FRIA pueden ser obstáculos para los propósitos innovadores de la UE.

Así las cosas, la UE está tratando de encontrar un equilibrio entre el impulso tecnológico de la región y la protección de los derechos fundamentales de los ciudadanos, y, desde nuestro punto de vista puede estar fallando a la hora de llevar a cabo este equilibrio. De esta forma, la gran cantidad de obligaciones impuestas denotan la dificultad

[8] Parlamento Europeo (2022). *Identificación y evaluación de la legislación comunitaria vigente y en preparación en el ámbito digital*. Disponible en: https://www.europarl.europa.eu/RegData/etudes/STUD/2022/703345/IPOL_STU(2022)703345_EN.pdf

de la UE para competir con regiones con regulaciones más laxas que facilitan el desarrollo de soluciones IA.

Adicionalmente, la inversión de la UE en la materia no se ajusta con sus objetivos, la organización se ha quedado rezagada en comparación con países como EE. UU. donde la inversión pública y privada en 2022 alcanzó los 3300 y 47400 millones de dólares respectivamente, cifras alejadas de las manejadas en la UE (Maslej et al., 2023, pp 14 y 288). A esto se le añade el hecho de que la mayor parte de la inversión en IA proviene del sector privado, lo que provoca una dependencia de la Administración Pública que entorpece la posibilidad de impulsar la IA desde el sector público.

Finalmente, el rápido avance de la IA está provocando que la regulación no encuadre todos los sistemas que se están desarrollando. Partiendo de la base que la propia definición de IA ofrecida en el FRIA puede haberse quedado obsoleta, soluciones como Chat GPT están reflejando la incapacidad del actual borrador para legislar todas las consecuencias que se van a desprender de soluciones que ni siquiera están concebidas actualmente. Si bien siempre se ha considerado que los procedimientos legislativos de la UE podían demorarse en exceso, en materias como la IA dicha laxitud en los tiempos puede provocar graves perjuicios para los ciudadanos.

En consecuencia, la UE debe plantear métodos para acelerar los cambios que sean demandados en el FRIA de forma que puedan hacer frente a las situaciones provocadas por el uso de la IA en el menor tiempo posible. Adicionalmente, ha de asumir que no todo puede ser regulado y que, una normativa excesiva, puede provocar un desaliento de los proveedores de la materia y una fuga de talento.

3. LEGISLACIÓN NACIONAL DE LA IA

Desde el plano nacional surgen diversas manifestaciones de la regulación de la IA. En ese sentido, España sigue el enfoque europeo definiendo un futuro marco legislativo en el que establecer las condiciones y limitaciones para el desarrollo, uso y aplicación de sistemas de IA en el territorio nacional.

3.1. Legislación Española previa al 2020

Previamente al 2020, si bien no existía una ley como tal, se localizaron algunas menciones a las decisiones automatizadas de la Administración Pública que podrían ser extensivas al uso de la IA. Dichas referencias son las siguientes:

a) La LGT en su artículo 96 prevé el uso de procedimientos y actuaciones automatizadas (como puede ser, aunque en su momento no se concibiese la posibilidad, el empleo de la IA). En concreto, se obliga a los proveedores de dichos sistemas a contar con la identificación de los órganos competentes para la programación y supervisión del sistema de información para resolver los recursos que puedan interponerse. Asimismo, la LGT incluye la necesidad de que los programas o soluciones empleadas para la toma de dichas decisiones deberán haber sido previamente aprobados por la Administración de la forma reglamentaria correspondiente (artículo 96 Ley 58/2003).

b) En segundo lugar, cabe citar la Ley 11/2007, de 22 de junio, de acceso electrónico de los ciudadanos a los Servicios Públicos. Del mismo modo que la anterior normativa, en esta regulación se menciona la actividad administrativa automatizada en los artículos 38 y 39 de la misma.

   En el artículo 38 se considera la posibilidad de adoptar y notificar resoluciones en aquellos procedimientos previstos (artículo 38 Ley 11/2007). En dicho artículo si bien no se menciona la IA de una forma específica (probablemente por la fecha de la normativa y la falta de conocimiento de dicha tecnología) podría ser de aplicación directa a la materia en caso de que la Administración optase por incluir la tecnología en el proceso administrativo.

   Por su parte, respecto al artículo 39 de la misma ley, se incluye una mención expresa a la actuación administrativa automatizada. En la misma se establece que la actuación deberá contar con el nombramiento de un órgano u órganos competentes que definirán las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente. Adicionalmente, se deberá incluir el órgano que será responsable a efectos de impugnación (artículo 39 Ley 11/2007).

c) La Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público también regula actuación administrativa automatizada en su artículo 41. Dicho precepto

sigue la línea del artículo 39 de la Ley 11/2007 y busca establecer un procedimiento de control de las especificaciones técnicas del sistema de información y del código fuente (artículo 41 Ley 40/2015).

d) Singularmente relevante para este trabajo son la Ley Orgánica de Protección de Datos (LOPD) y su Reglamento de Desarrollo (RLOPD) y el Reglamento General de Protección de datos (RGPD). Estas regulaciones relativas a la protección de datos afectan de forma directa a la IA por la forma en la que estas soluciones tratan los datos de los usuarios que hacen uso de las mismas. No obstante, este ámbito se expondrá con mayor detalle en el capítulo V de este trabajo.

La legislación anteriormente expuesta sin bien puede afectar a la IA de una forma más o menos directa, no señala de forma explícita una regulación concreta sobre la IA. Por ello, a falta de una legislación concreta y a la espera del desarrollo del reglamento IA en España, se han elaborado importantes informes que pueden servir de guías para la Administración y las empresas mientras se desarrolla un marco normativo completo.

3.2. Informes, guías y primera normativa desde el 2020

En esta segunda etapa, España, motivada por el impulso legislativo comunitario, fue consciente de la necesidad de implementar un marco regulatorio completo desde el que impulsar y salvaguardar las soluciones IA en el territorio nacional.

El primer informe al que hay que hacer referencia es la Estrategia Nacional de IA[9]. En este documento se sitúan siete objetivos estratégicos para impulsar el desarrollo de la IA en España: la excelencia científica e innovación en IA; la creación de empleo cualificado; la transformación del tejido productivo; la creación de un entorno de confianza en relación a la IA; la incorporación de valores humanistas en la IA y el desarrollo de una IA inclusiva y sostenible, todo ello por medio de seis ejes estratégicos (Gobierno de la España, 2020).

[9] Gobierno de España. (2019). *Estrategia Nacional de IA*. Disponible en: https://www.ciencia.gob.es/stfls/MICINN/Ayudas/Convocatorias/Documents/20191030_EstrategiaNacionalInteligenciaArtificial.pdf

Dentro de la estrategia y centrándose en las medidas que afectan a la Administración Pública, se pueden identificar cinco medidas concretas (Vestri, 2022, pp 38-43):

a) Incorporar la IA en la Administración para mejorar la eficiencia y reducir los retrasos propios de los organismos.
b) Puesta en marcha de laboratorios de innovación para descubrir nuevas oportunidades y aplicaciones de la IA en la administración.
c) Desarrollo de un programa de IA para la gestión pública de los datos.
d) Promoción de proyectos estratégicos en el ámbito de la administración pública donde la IA puede tener impacto (especialmente en los ámbitos de salud, justicia y empleo

Continuando con los informes emitidos por los organismos oficiales, tiene una gran relevancia la Carta de Derechos Digitales[10]. El documento, si bien no tiene carácter normativo, se presenta como una oportunidad para identificar los potenciales retos de aplicación e interpretación que la adaptación de los derechos al entorno digital puede ocasionar. Asimismo, se pretenden reconocer principios y políticas asociadas a la materia y que han podido pasar desapercibidas en el pasado, en concreto (Ministerio de Asuntos Económicos y Transformación Digital, 2021):

a) La IA debe tener un enfoque holístico con el que encontrar el bien común y garantizar el cumplimiento del principio de no maleficencia.
b) Durante el diseño y el mantenimiento del ciclo de vida de los sistemas de IA, se deberá comprobar que se cumple el derecho a la no discriminación, independientemente de su origen, respecto a las decisiones, uso de datos y procesos basados en IA. Asimismo, se deberán desarrollar condiciones de transparencia, auditabilidad, explicabilidad, trazabilidad, supervisión humana y gobernanza mediante procesos accesibles, comprensibles y fiables.
c) Los usuarios deben tener la posibilidad de solicitar una supervisión e intervención humana que permita reevaluar las decisiones automatizadas adoptadas por la IA que causen efectos en su esfera personal y patrimonial.

[10]Ministerio de Asuntos Económicos y Transformación Digital. (2021). *Carta de Derechos Digitales*. Disponible en: https://www.mincotur.gob.es/es-es/GabinetePrensa/NotasPrensa/2021/Documents/20210127_Carta_Derechos_Digitales.pdf

En la Carta se incluyen una serie de derechos en relación con la IA en el marco de la actuación administrativa. En concreto, se especifica que: las decisiones y actividades deberán respetar los principios de buen gobierno y el derecho a una buena Administración digital; el uso de instrumentos IA tendrán que garantizar la transparencia en el funcionamiento y el alcance de cada procedimiento (especialmente en aquellos sistemas empleados para tomar decisiones) a través de la regulación legal de las condiciones de transparencia y el acceso al código fuente (evitando resultados discriminatorios); se deberá contar con una motivación razonable en lenguaje natural de las decisiones que se tomen en el entorno digital (teniendo los usuarios derecho a conocer la justificación administrativa detrás de las decisiones adoptadas por la IA); la adopción de decisiones discrecionales deberá quedar reservada a las personas salvo que normativamente se prevea otra opción (en cualquier caso se deberá contar con una evaluación de impacto de los derechos digitales en el diseño de los algoritmos en el caso de adoptar decisiones automatizadas o semiautomatizadas) (Vestri, 2022, pp 38-43).

Concretamente este deber de transparencia se ha garantizado a través de órganos públicos como la Comisión de Garantía del Derecho de Acceso a la Información Pública de Cataluña en Resolución de la Reclamación 123/2016[11], desde la que se ha promovido la posibilidad de declarar los algoritmos como información pública.

Aunque como se comentó previamente todavía no exista una normativa concreta de IA, en julio de 2022 se emitió la primera ley en la que se hacía referencia a una regulación directa de la IA. Dicha remisión se encuadra en la Ley 15/2022 integral para la igualdad de trato y la no discriminación, que busca convertirse en un marco común normativo que proteja contra cualquier tipo de discriminación presente y futura. Con esta vocación en mente, la regulación hace referencia directa al uso de la IA por las administraciones públicas y las empresas en nuestro país tratando de proteger a los usuarios de los potenciales riesgos adversos asociados a dicha tecnología (Fernández, 2022b)

En la ley se subraya la forma en la que las Administraciones Públicas deben diseñar los algoritmos que se utilicen en los mecanismos de toma de decisiones. Específicamente es necesario remitirse al artículo 23 de la norma, en dicho precepto se continua en la ley de la Estrategia Nacional de la IA, de la Carta de Derechos digitales y las iniciativas

[11] Comisión de Garantía del Derecho de Acceso a la Información Pública de Cataluña. (2016). Resolución de la Reclamación 123/2016. Disponible en: https://www.accesoinfo.cat/sites/default/files/arxius/resolucio_cgaip_123_2016.pdf

europeas de la IA indicando que las administraciones públicas deberán favorecer la puesta en marcha de procesos que garanticen que los algoritmos cuenten con criterios de minimización de sesgos, transparencia y rendición de cuentas cuando sea posible (Artículo 23 de la Ley 15/2022).

Para lograrlo se incluirán evaluaciones de impacto capaces de determinar el sesgo discriminatorio. Asimismo, se pide tanto a la administración como a las empresas promover el uso de una IA ética, confiable y respetuosa con los derechos fundamentales [12] de acuerdo con las indicaciones comunitarias (Artículo 23 de la Ley 15/2022). El mayor problema que se desprende del artículo es su carácter progmático y poco normativo, de esta forma, consideramos que la norma no llega a introducir ningún cambio tangible y se encuentra muy lejos de una verdadera referencia legislativa en la materia.

En definitiva, la futura legislación española sigue la línea legislativa europea y se espera que pivote sobre la protección de los derechos fundamentales de los usuarios.

[12] En esta línea, se ha de hacer referencia a la STC 76/2019 primera en fallar en contra del uso de la IA en el sector público. Dicha sentencia surge en respuesta a un recurso de inconstitucionalidad presentado en contra del antiguo artículo 58 bis 1 de la LOREG. En la sentencia se hace mención expresa a la IA y su función en el tratamiento de datos y procedimientos complejos orientados a modificar, forzar o desviar la voluntad de los electores. El tribunal sostiene que estas técnicas pueden comprometer directamente el derecho a la participación política en los asuntos públicos garantizado en el artículo 23 CE (STC 76/2019)

CAPÍTULO III. CONFLICTOS DERIVADOS DE LA CLASIFICACIÓN DE LOS SISTEMAS IA CON UNA MIRADA A LA PROTECCIÓN DE DATOS

1. CLASIFICACIÓN PRELIMINAR DE LA IA EN EL FRIA

La futura normativa europea ha decido optar por una división basada en el riesgo, dividiendo los sistemas IA en cuatro grupos diferenciados: IA de riesgo inaceptable, IA de alto riesgo, IA de riesgo bajo e IA de riesgo mínimo (Exposición de motivos FRIA).

En función del nivel de clasificación se impondrán a los sistemas unas obligaciones acordes a los riesgos propios de la categoría. Esta última idea denota la importancia de dicha delimitación, puesto que los medios de control y supervisión que deben presentarse oscilan en gran medida dependiendo de la clase de IA empleada.

1.1. Sistema IA de riesgo inaceptable

Los sistemas IA prohibidos se encuentra recogidos dentro del Título II en el artículo 5 del FRIA. La razón subyacente a su prohibición recae en los perjuicios que pueden provocar a los usuarios. En concreto, el uso de estos sistemas supone un atentado a los valores de la UE, ocasionando daños inaceptables que no pueden ser suplidos[13] por los beneficios que se pudiesen derivar del uso estos sistemas IA (Exposición de motivos FRIA).

En base a lo anterior, entre las prácticas prohibidas se encuadrarían los sistemas IA que empleen técnicas de manipulación subliminal que pudiesen ocasionarle a sí misma o a un tercero daños físicos o psicológicos, especialmente, si la manipulación estuviese dirigida hacia sectores específicos de la población en base a su edad o discapacidad (artículo 5.1. FRIA).

[13] De forma genérica, como veremos más adelante podría haber ocasiones en las que desde el plano Administrativo el uso de estos sistemas estaría justificado

Asimismo, se prohíbe la elaboración de perfiles en los que las personas fueran clasificadas atendiendo a sus características personales o a sus comportamientos, cuando estos fueran a ser empleados por las autoridades públicas para tomar decisiones administrativas, siempre que, dicha práctica provoque un trato perjudicial o desfavorable o bien desproporcionado para las personas físicas de las que se hayan tomado los datos, o bien por tratarse de usuarios que no guarden relación con el contexto en el que fueron recogidos los datos (artículo 5.1. FRIA).

De esta última prohibición, se desprende que la elaboración de perfiles por parte de las autoridades, independientemente de los efectos que este tipo de actividades puede provocar en los usuarios, está permitida siempre que no provoque efectos secundarios en los grupos anteriormente mencionados (Cotino et al., 2021). Si bien el tratamiento de estos datos se encuentra salvaguardado por las leyes de protección de datos, la falta de una prohibición explícita lleva a cuestionarse si la UE debiera instar a regular este tipo de soluciones IA que están provocando graves perjuicios en países como China[14].

Por último, gran parte de los sistemas IA prohibidos hacen referencia a aquellos empleados en la identificación biométrica en tiempo real. Este asunto ha sido incluido también en los sistemas IA de alto riesgo, provocando una dificultad para discernir cuándo una solución IA se encuentra prohibida y cuándo no.

En conclusión, la clasificación de los sistemas de IA dentro de la categoría de riesgo inaceptable se construye sobre la idea de proteger los derechos fundamentales de los ciudadanos. No obstante, pese a los efectos claramente perjudiciales que pueden causar a los usuarios, se habilita el empleo de estas técnicas en el caso de contar con una autorización judicial o administrativa. Ello provoca que sea necesario que las autoridades lleven a cabo una evaluación cuidadosa de los riesgos, y al mismo tiempo, monitoricen que la IA cuenta con una estructura de gobernanza de la información capaz de proteger los intereses de los usuarios. El no llevar a cabo una correcta supervisión produciría que casos sustancialmente iguales se tratasen de manera distinta, autorizando situaciones con un nivel de riesgo similar a otras prohibidas por los mismos organismos.

[14] A finales del año pasado, China aprobó la creación de una clasificación social a los ciudadanos en base a cinco principios principales: la solvencia financiera, la ejecución judicial, la confianza comercial, la confianza social y la integridad gubernamental. En caso de obtener una mala puntuación podría derivar en consecuencias directas en los viajes, el empleo o el acceso a la financiación (Fernández, 2022b).

1.2. Sistema de IA de alto riesgo

Los sistemas IA de alto riesgo se encuentran recogidos dentro del Título III en el artículo 6 del FRIA. En esta categoría se regulan aquellos sistemas que pueden producir un daño potencial a la salud o los derechos fundamentales de los usuarios (Exposición de motivos FRIA).

El primer grupo de soluciones reguladas hacen referencia a aquellos sistemas IA que actúen como componentes de seguridad de los productos contemplados en el Anexo II del FRIA[15] (artículo 6.1 FRIA).

Por su parte, en el segundo grupo de soluciones encuadran aquellos sistemas IA que por sus características pueden producir graves efectos en los usuarios que hagan uso de los mismos.

Ello incluye sistemas de identificación biométrica, de gestión y funcionamiento de infraestructuras esenciales[16], de aquellos empleados en la educación y formación profesional[17], de los utilizados para gestionar el empleo, a los trabajadores y el acceso al autoempleo[18], de los que se hagan uso para acceder y disfrutar de los servicios esenciales ya sean públicos y privados[19], de gestión de la migración, el asilo y el control de fronteras, de los que faciliten la aplicación de la ley[20], y finalmente de los que favorezcan la administración de justicia y los procesos democráticos (artículo 6.2 FRIA).

---

[15] Los productos contemplados en el Anexo son los siguientes: Máquinas, juguetes, embarcaciones de recreo y motos acuáticas, ascensores, aparatos y sistemas de protección para uso en atmósferas potencialmente explosivas, comercialización de equipos radioeléctricos, comercialización de equipos a presión, instalaciones de transporte por cable, equipos de protección individual, aparatos que queman combustibles gaseosos, productos sanitarios, aviación civil, vehículos de dos o tres ruedas y los cuatriciclos, vehículos agrícolas o forestales, equipos marinos, interoperabilidad del sistema ferroviario y vehículos de motor y sus remolques (Anexo II, FRIA). Asimismo, también serán considerados de alto riesgo los propios productos contemplados en el Anexo II (artículo 6.1 FRIA).

[16] Dentro de este apartado entrarían aquellos sistemas relativos al tráfico, suministro de agua, gas calefacción y electricidad (Anexo III FRIA).

[17] En concreto, en los procesos de admisión de candidatos o de evaluación de los mismos (Anexo III FRIA).

[18] Específicamente, el uso de la IA en procesos de contratación, selección de personal, ascensos o resoluciones contractuales (Anexo III FRIA).

[19] Dentro de este tipo de soluciones entrarían los análisis de acceso a prestaciones y servicios de asistencia pública, las evaluaciones de solvencia crediticia, o los envíos de servicios de intervención en situaciones de emergencia (Anexo III FRIA).

[20] Dentro de este grupo entrarían las evaluaciones de riesgos cuyo fin último sea la determinación de la probabilidad de que una persona cometa o reincida en una infracción penal, los sistemas capaces de

Esta categoría se erige como el principal núcleo del FRIA, precisamente por ello son estos sistemas los que incluyen el mayor número de obligaciones. La gran variedad que conforman esta clase de IA deja intuir la problemática que se deriva de la misma. Por un lado, algunas de las soluciones deberían considerarse IA de riesgo inaceptable (solo aceptables con la correspondiente autorización) y, por otro, no todos los sistemas han de requerir el mismo nivel de obligaciones, debería graduarse en función de la herramienta IA concreta.

En conclusión, si bien la clasificación de sistemas IA de riesgo inaceptable se sustentaba en la protección y salvaguarda de los valores de la Unión, en el caso de los sistemas IA de alto riesgo, el objetivo último de la clasificación reside en la necesidad de proteger a la sociedad de los daños que puede provocar el uso de estos sistemas en la sociedad. En última instancia, esta sección de la clasificación es un paso fundamental para garantizar que la IA se desarrolle y utilice de una forma responsable y segura.

1.3. Sistema de IA de riesgo limitado

Los sistemas IA de riesgo limitado se encuentran recogidos tanto en la exposición de motivos como en el Título IV del FRIA. En esta parte de la clasificación se integran aquellos sistemas que interactúan con personas físicas y, que, debido a su interfaz y procesamiento del lenguaje, pueden provocar que el usuario no sea consciente de que está tratando con una IA. Asimismo, entrarían dentro de esta categoría aquellos sistemas IA capaces de detectar emociones (artículo 52 FRIA). Ejemplos de este tipo de sistemas serían los chatbot o "*Deep fake"* entre otras posibles soluciones.

detectar el estado emocional de las personas físicas (polígrafos o similares), los sistemas destinados a detectar falsificaciones, las evaluaciones destinadas a comprobar la veracidad de las pruebas empleadas durante la investigación o enjuiciamiento de una infracción penal, los sistemas que hagan uso de extensas bases de datos que permitan detectar modelos o encontrar relaciones en los datos que de otra forma podrían haber pasado desapercibidas. Asimismo, serán considerados sistemas de alto riesgo aquellos empleados para elaborar perfiles. En consecuencia, entrarían las soluciones IA que creen perfiles en base a características individuales o propias de un grupo de personas que revelen la frecuencia o reiteración de una infracción y aquellas técnicas IA destinadas a establecer perfiles asociados a la detección, investigación o enjuiciamiento de infracciones penales. Adicionalmente, serán sistemas de alto riesgo aquellos utilizados para verificar la autenticidad de los documentos de viaje y los documentos justificativos, así como aquellos empleados para detectar falsificaciones documentales (Anexo III FRIA).

Las obligaciones en esta sección están orientadas a la transparencia del sistema y a la advertencia a los usuarios que hagan uso de la IA.

En primer lugar, se deberá garantizar que las personas sean conscientes en todo momento que están interactuando con un sistema IA (excepto en caso de que por las características propias de la IA o por el contexto en el que es utilizada, hagan evidente que se trata de una IA) (artículo 52.1 FRIA).

En segundo lugar, cuando el sistema IA sea empleado para crear o manipular contenido audiovisual similar a personas, objetos, lugares o sucesos reales, que puedan llevar a una persona física a pensar que es real, el sistema IA deberá contar con un sistema de publicidad que haga público el contenido generado por la solución. Estarán exentos de cumplir dicha obligación los sistemas que sean vehículos de la libertad de expresión o el derecho a libertad de las artes y de las ciencias (artículo 52.3 FRIA).

Como se intuye la mayor problemática recae en el hecho de que las obligaciones son muy limitadas. De esta manera, se trata de requisitos orientados a la transparencia y a requisitos de publicidad. Las rápidas evoluciones de la IA han provocado que la falta de una regulación más férrea pueda ser insuficiente para controlar las consecuencias de estos sistemas (el mayor exponente de ello es Chat GPT que se valorará más adelante en el trabajo).

En conclusión, en esta categoría el objetivo es salvaguardar la transparencia de la interacción humana con los sistemas IA. El avance constante de este tipo de soluciones puede provocar que sea difícil discernir para el usuario si se está tratando con una IA, ello puede provocar graves consecuencias para la persona y por ello la Unión ha establecido límites y garantías para este tipo de sistemas que, con las nuevas soluciones, pueden requerir ser actualizadas.

1.4. Sistema de IA de riesgo mínimo

En esta categoría entrarían el resto de los sistemas IA comercializados en la Unión cuyo riesgo es considerado mínimo. La falta de una regulación más extensa en esta sección es derivada de la necesidad de permitir la innovación e impulso de la IA en Europa, que podría verse afectada en caso de incluir una sobrerregulación en sistemas cuyos efectos en la sociedad son mínimos (Comisión Europea, 2021d).

La ausencia de un riesgo cuantificable los convierte en soluciones que pueden desarrollarse y utilizarse sin mayores obligaciones jurídicas a las expuestas por la legislación vigente. No obstante, la Comisión anima a los proveedores de este tipo de sistemas a adherirse de forma voluntaria a las obligaciones incluidas en el resto de las categorías (Comisión Europea, 2021d).

2. IDENTIFICACIÓN BIOMÉTRICA: CLASIFICACIÓN MÚLTIPLE DE LA MATERIA Y CONSECUENCIAS DERIVADAS DE LA MISMA

2.1. La identificación biométrica en la protección de datos

A la hora de la clasificar la identificación biométrica, la UE ha optado por una asignación dual del riesgo en función del uso que se vaya a llevar a cabo de la solución. El problema de dicha distinción es que existe un abanico de soluciones que, por su nivel de riesgo, deberían encuadrarse como IA prohibida, y, sin embargo, se encuentran configuradas como IA de alto riesgo o no clasificadas en ninguna categoría en concreto.

Es preciso hacer una mención al derecho configurado por la normativa vigente en protección de datos, y en especial, a su finalidad última. En este sentido, el RGPD se configura como el epítome legislativo de este derecho, y su objetivo no es otro que proteger mediante la regulación los datos personales. Por datos personales, el RGPD entiende cualquier tipo de datos que se pueden utilizar para identificar directa o indirectamente a una persona. La IA, alimentada primordialmente por una base de datos, tendrá en su configuración una base de datos que la alimente, por lo que la incidencia de la normativa aplicable en protección de datos es de importante calado.

Especialmente relevante en este contexto resulta el artículo 9.1 del RGPD, que establece como tratamiento de categorías especiales de datos, entre otros, los «datos biométricos dirigidos a identificar de manera unívoca a una persona física». Ello hace que los datos biométricos de este calado resulten merecedores de una protección especial, con un haz de obligaciones para los responsables y encargados del tratamiento y consecuentes garantías para los interesados.

Ahora bien, para poder entender el riesgo que se desprende de la materia es necesario definir en primera instancia que son los datos biométricos. Para ello, se ha acudido a la delimitación expuesta en el artículo 4 del RGPD que ha servido de base para configurar el FRIA:

"Los datos biométricos son aquellos datos personales obtenidos a partir de un tratamiento técnico especifico, 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"

La capacidad de identificar a una persona de forma inequívoca supone un riesgo directo a los derechos fundamentales de las personas físicas. En concreto, las particularidades de los datos biométricos provocan que su tratamiento pueda vulnerar tanto el derecho a la intimidad como el derecho a la protección de datos personales (Orrico, 2021, pp 107-113).

Los datos biométricos ofrecen una información más precisa sobre la persona a la que pertenecen. De esta forma, pueden permitir conocer la raza o el género, su estado emocional, las enfermedades, discapacidades y características generales propias del usuario, entre otros muchos aspectos. El problema es que este conocimiento es subsidiario de la recogida inicial de los datos y, por ende, la obtención de la información está incluida en el consentimiento inicial al ser inseparable del dato originario. Ello implica que el usuario no puede impedir el tratamiento de dicha información subsidiaria (AEPD, 2020c).

En base a ello, como dato personal el tratamiento de los datos biométricos requiere de la adecuación de los mismos a una serie de principios fundamentales. Ello incluiría el principio de licitud, lealtad y transparencia, el principio de limitación de la finalidad, el principio de exactitud, el principio del plazo de conservación (los datos solo podrán ser almacenados mientras dure el fin del tratamiento), el principio de integridad y seguridad, el principio de responsabilidad proactiva (AEPD, 2020f). Adicionalmente, conviene indicar que la obtención lícita de los datos es tan relevante que ha sido acuñada por el RGPD como un principio *sine qua non* para el tratamiento legítimo de los datos de los interesados: el principio de licitud (artículo 5.1.a del RGPD).

Estos principios regirán independientemente de la categoría en la que se incluya la IA respecto del FRIA.

Se presenta necesario acotar la diferencia que se hace en la doctrina entre la identificación y la verificación biométrica de las personas. Mientras que la verificación supone la comparación entre datos en dos plantillas concretas (que pertenecen a una persona), la identificación supone el tratamiento de datos biométricos sobre bases de datos (siendo el riesgo hacia la intimidad personal mucho mayor). Dicha distinción tendrá importancia para entender la razón por la que la gestión transfronteriza, que entraría dentro de una verificación biométrica, se encuadra dentro de la IA de alto riesgo y no de la IA de riesgo inaceptable. Asimismo, aunque el artículo 4 del RGPD engloba ambos supuestos, de la propia AEPD se alude a la necesidad de otorgar únicamente a los datos biométricos tratados en la identificación biométrica como dato personal especial (AEPD, 2020a).

Es importante la distinción pues un dato personal especial requiere un mayor número de obligaciones que uno de carácter general. Como se ha mencionado anteriormente, al ser los datos biométricos considerado un dato de categoría especial, su obtención y consiguiente tratamiento sólo será lícito cuando concurra una de las bases legitimadoras del artículo 9.2 del RGPD. En el mencionado artículo 9 del RGPD, se prohíbe el tratamiento de este tipo de datos, habilitando el mismo únicamente cuando se cuente con alguna de las siguientes bases legitimadoras: un consentimiento explícito para uno o más fines específicos, el tratamiento es necesario para el cumplimiento contractual, el tratamiento es imprescindible para proteger los intereses vitales de las personas físicas, el tratamiento es llevado a cabo por una fundación u organismo sin ánimo de lucro[21] cuya finalidad sea política, filosófica, religiosa o sindical, los datos tratados son públicos, el tratamiento es necesario para la formulación, ejercicio o la defensa de reclamaciones, y, finalmente y probablemente el que más va a importar para habilitar la IA de riesgo inaceptable, el tratamiento por razones de un interés público esencial (AEPD, 2018a).

Una vez comprendida la importancia de estos datos en la protección de datos y en los derechos fundamentales de los usuarios, se ha de acudir al FRIA para determinar cómo se ha dividido la identificación biométrica en las categorías de riesgo.

[21] El tratamiento es efectuado, en el ámbito de sus actividades legítimas y con las debidas garantías, por una fundación, una asociación o cualquier otro organismo sin ánimo de lucro, cuya finalidad sea política, filosófica, religiosa o sindical, siempre que el tratamiento se refiera exclusivamente a los miembros actuales o antiguos de tales organismos o a personas que mantengan contactos regulares con ellos en relación con sus fines y siempre que los datos personales no se comuniquen fuera de ellos sin el consentimiento de los interesados (artículo 9 RGPD)

2.2.La identificación biométrica encuadrada en la IA prohibida

Respecto a la IA prohibida, se encuadrará dentro de la categoría la identificación biométrica remota que sea llevada a cabo "en tiempo real en zonas de acceso público", a expensas de que sea necesaria para identificar víctimas de un delito, prevenir amenazas a la vida o seguridad de las personas físicas, y detener criminales por delitos específicamente regulados[22] (artículo 5.1. FRIA).

El primer problema que plantea esta delimitación es el concepto de "tiempo real", esta definición se encuentra recogida en el artículo 3 del FRIA y puede no ser lo suficientemente completa como para encuadrar todos los supuestos que implican un riesgo inaceptable.

De esta manera, dicho concepto se limita a marcar la identificación en tiempo real como aquella que se produce de forma instantánea o sin demoras significativas.

No obstante, se ha de ponderar aquellos sistemas que si bien no realizan el tratamiento de forma directa, llevan a cabo grabaciones constantes que son examinadas al día siguiente (evitando la inmediatez del tiempo real). Se puede convenir que el daño a los derechos fundamentales en estas circunstancias es casi tan elevado como si hubiesen sido analizados directamente. De esta forma, el punto de inflexión debería recaer en una identificación biométrica continua o en una que no se enfoca en un periodo concreto (Parlamento Europeo, 2021).

[22] "Pertenencia a organización delictiva, terrorismo, trata de seres humanos, explotación sexual de los niños y pornografía infantil, tráfico ilícito de estupefacientes y sustancias psicotrópicas, tráfico ilícito de armas, municiones y explosivos, corrupción, fraude, incluido el que afecte a los intereses financieros de las Comunidades Europeas con arreglo al Convenio de 26 de julio de 1995 relativo a la protección de los intereses financieros de las Comunidades Europeas, blanqueo del producto del delito, falsificación de moneda, incluida la falsificación del euro, delitos de alta tecnología, en particular delito informático, delitos contra el medio ambiente, incluido el tráfico ilícito de especies animales protegidas y de especies y variedades vegetales protegidas, ayuda a la entrada y residencia en situación ilegal, homicidio voluntario, agresión con lesiones graves, tráfico ilícito de órganos y tejidos humanos, secuestro, detención ilegal y toma de rehenes, racismo y xenofobia, robos organizados o a mano armada, tráfico ilícito de bienes culturales, incluidas las antigüedades y las obras de arte, estafa, chantaje y extorsión de fondos, violación de derechos de propiedad industrial y falsificación de mercancías, falsificación de documentos administrativos y tráfico de documentos falsos, falsificación de medios de pago, tráfico ilícito de sustancias hormonales y otros factores de crecimiento, tráfico ilícito de materiales radiactivos o sustancias nucleares, tráfico de vehículos robados, violación, incendio voluntario, delitos incluidos en la jurisdicción de la Corte Penal Internacional, secuestro de aeronaves y buques, sabotaje." (Artículo 2 apartado 2 de la Decisión Marco 2002/584/Jai del Consejo)

En segundo lugar, también presenta problemas el papel que el FRIA le otorga a la "identificación biométrica remota" siendo esta aquellas basadas en la comparación con los datos biométricos contenidos en una base de datos de referencia, sin que el sistema de IA sepa previamente si la persona estará incluida en la misma (artículo 3 FRIA).

Ello implica que el reglamento sitúe la falta de conocimiento sobre si la persona estará presente o no para diferencia entre los métodos de verificación e identificación. No obstante, la identificación también puede darse cuando los proveedores del sistema utilicen la solución para localizar a personas que conformen la base de datos. Por ejemplo, en el ámbito laboral para controlar las jornadas laborales de los empleados, en este caso el sistema IA conocerá de forma previa que la persona estará incluida en la base de datos (Parlamento Europeo, 2021).

Estas disyuntivas que se plantean, tanto la distinción de identificación con la autentificación, como los problemas que se derivan del concepto de "tiempo real y remoto" muestran la necesidad de redefinir la categoría. Las consecuencias que se derivan de no hacerlo pueden suponer un grave atentado contra los derechos fundamentales de los usuarios. Aunque gran parte de los problemas entran dentro de la protección del RGPD, no se puede dejar al arbitrio de dicha regulación el control de la IA, de esta forma, se presenta necesario el estudio completo de este tipo de actividades.

Por otro lado, en caso de que se autorizase a las autoridades para hacer uso de estos sistemas prohibidos, se ve necesario hacer mención de las obligaciones que los proveedores habrán de cumplir.

De esta forma, de acuerdo con el artículo 5.2 del FRIA el uso de la identificación biométrica requerirá que el sistema IA empleado pondere la situación de uso y, en concreto, la gravedad, probabilidad e impacto del daño que podría desprenderse si no se hiciera uso de la IA (por ejemplo, si el no aprovechar la identificación biométrica para identificar un terrorista pudiese derivar en un mayor número de atentados).

Del mismo modo, en el artículo 5.2 del FRIA se incluye la obligación de analizar los efectos que se derivarían del uso de la identificación biométrica en los derechos y libertades de aquellos involucrados en el proceso. Igualmente, al tratarse de sistemas de identificación remota "en tiempo real" en zonas de acceso público, se deberán cumplir con una serie de garantías y condiciones adecuadas al uso.

Asimismo, en el artículo 5.2 del FRIA también se regula el empleo de la identificación biométrica en zonas de acceso público, imponiendo la adquisición de una autorización previa por parte de una autoridad judicial o administrativa del EM donde se vaya a aplicar el sistema. Dicho permiso deberá sustentarse en una solicitud motivada y adaptada a las normas de Derecho interno[23]. Cabe mencionar, que en caso de que la situación donde se fuera a utilizar el sistema presentase un carácter de urgencia lo suficientemente justificado, la autorización requerida podría aplazarse.

Las obligaciones impuestas para este tipo de sistemas se muestran coherentes y adecuadas para el tipo de IA en cuestión. La ponderación de los efectos de la IA, tanto si se opta por utilizar el sistema como en el caso contrario, implica una protección clara al bien jurídico protegido (el derecho a la intimidad y a la protección de datos personales). De esta forma, los requisitos son factibles para los proveedores y suficientemente fuertes como para que en caso de ser necesario el uso de la IA prohibida se desarrolle de la mejor forma.

En conclusión, los sistemas IA de identificación biométrica considerados de riesgo inaceptable pueden no ser suficientes para proteger los derechos de los ciudadanos, al no incluir todos los casos de uso con riesgos similares. En particular, se han de ponderar las situaciones que, si bien no son a tiempo real, se pueden asemejar a las mismas. Una laguna en esta sección de la regulación puede traducirse en graves perjuicios para los usuarios, sus derechos a la intimidad y a la protección de datos sin duda se ven directamente atacados. Por su parte, las obligaciones impuestas para acceder a esta categoría se presentan suficientes para proteger las situaciones actualmente incluidas en la misma.

[23] "Los Estados Miembros podrán decidir contemplar la posibilidad de autorizar, ya sea total o parcialmente, el uso de sistemas de identificación biométrica remota «en tiempo real» en espacios de acceso público con fines de aplicación de la ley dentro de los límites y en las condiciones que se indican en el apartado 1, letra d), y los apartados 2 y 3. A tal fin, tendrán que establecer en sus respectivos Derechos internos las normas detalladas necesarias aplicables a la solicitud, la concesión y el ejercicio de las autorizaciones a que se refiere el apartado 3, así como a la supervisión de estas. Dichas normas especificarán también para cuáles de los objetivos enumerados en el apartado 1, letra d), y en su caso en relación con cuáles de los delitos indicados en su inciso iii), se podrá autorizar que las autoridades competentes utilicen esos sistemas con fines de aplicación de la ley" (artículo 5.4 FRIA)

2.3.La identificación biométrica encuadrada en la IA de alto riesgo

Por su lado, respecto a los sistemas IA de identificación biométrica localizados en la categoría de alto riesgo, se han incluido las soluciones IA utilizadas para gestionar la migración, el asilo y el control transfronterizo. De esta forma, al igual que en los sistemas IA de aplicación legal se encuadrarían técnicas destinadas a detectar el estado emocional de las personas físicas (polígrafos o similares) durante la gestión de la inmigración. Similarmente, se considerará IA de alto riesgo aquella que se emplearía para llevar a cabo evaluaciones de riesgo para la seguridad, salud o similares relativas a la inmigración ilegal en la UE (Anexo III FRIA).

En esta categoría entrarían los sistemas de verificación biométrica, ya antes explicados, los sistemas de identificación biométrica remota en los que no se sepa previamente si la persona va a ofrecer una coincidencia con la base de datos (aquellos empleados por ejemplo para identificar personas físicas en un ambiente cerrado).

El mayor problema de esta clasificación es que algunos de los asuntos que se incluyen en este sector del FRIA son discutiblemente menos peligrosos que los de la anterior clase. De esta forma, herramientas como los utilizados en la vigilancia policial predictiva o la gestión la migración son soluciones IA que pueden incidir directamente con el derecho a la intimidad y la protección de datos (EDRi, 2021)

En nuestra opinión, si bien se ha de permitir la innovación y habilitar la creación de nuevos sistemas, no se puede negar que los riesgos asociados a estos usos son considerables y que, por ello, deberían contar con alguno de los requisitos impuestos en la categoría de riesgo inaceptable. En concreto, considero que se debería requerir a los proveedores de este tipo de sistemas una autorización administrativa o judicial pues los derechos fundamentales de los usuarios deben contar con la más alta protección.

Entre las obligaciones que deben cumplir los sistemas de identificación biométrica en esta categoría entraría la creación de sistemas de gestión de riesgo y gobernanza, la recogida de la documentación técnica asociada a la solución, una trazabilidad y transparencia del sistema, la garantía de una vigilancia humana y unas medidas de precisión, solidez y ciberseguridad adecuadas. Asimismo, antes de poder introducirlos en el mercado deberán cumplir con una evaluación de la conformidad.

Los requisitos impuestos a las soluciones IA de esta categoría son los más férreos del reglamento, y sobre ellos, rige el control de las autoridades sobre el funcionamiento de los sistemas. El objetivo último de esta monitorización es proteger a los ciudadanos de los riesgos asociados a esta categoría. Se profundizará en dichos requisitos en el capítulo IV de este trabajo.

En conclusión, se debe ponderar la situación de la identificación biométrica en la categoría de IA de alto riesgo. Las consecuencias fruto del uso de este tipo de datos son sin duda cuantiosas y no se pueden obviar. Debido a ello, se debe considerar tanto la posibilidad de incorporar en la categoría de riesgo inaceptable usos que se han incluido en esta categoría, como, por otro lado, se debe valorar la posibilidad de graduar esta categoría en función de las características propias de cada actividad.

2.4. Sistemas de identificación biométrica no incorporados en ninguna categoría en concreto

Pese a la gravedad que se desprende del tratamiento de los datos biométricos, se ha de exponer que no todos los usos de la identificación biométrica han sido regulados por el FRIA. En consecuencia, con el ánimo de no impedir la innovación han surgidos lagunas en la futura legislación.

De esta forma, no se encontrarían regulados en ninguna categoría tanto la identificación biométrica que no se llevase a cabo en remoto, como la verificación mediante los datos asociados a la propia identidad (por ejemplo, la comparación entre una imagen facial con otra previamente registrada) (Veridas, 2022, p 7).

En este caso conviene separar ambos conceptos, si bien la verificación antes mencionada no supone un peligro real contra los intereses de los usuarios (actúa como un mero mecanismo de prueba) y por ello no debería contar con mayores requisitos, en el caso de la identificación biométrica los riesgos son mayores. Aunque no se aboga por incorporarlos en la categoría de riesgo alto (pues requiere del consentimiento activo del afectado) sí debería incluirse algún tipo de requisito de transparencia similar a las obligaciones de la categoría de riesgo limitado.

Por otra parte, los sistemas biométricos de emociones y categorización no solo no están considerados sistemas de riesgo inaceptable, sino que, en general, ni siquiera se encuentran encuadrados en los sistemas IA de alto riesgo. La mención a los mismos se encuentra limitada a lo dispuesto en el artículo 52 del FRIA donde se obliga a los proveedores a ofrecer al usuario la información relativa al funcionamiento del sistema (Cotino, 2021).

Dado el carácter especial de esta información cabe preguntarse si es coherente no incluir este uso en ninguna categoría en específica. De esta forma, se deben valorar los riesgos que se pueden derivar del empleo de esta información, el avance constante de la IA supone sin duda un factor considerable. Por todo ello, considero que debería incluirse la identificación biométrica de emociones y categorización en el sector de IA de alto riesgo.

En conclusión, las lagunas surgidas en el FRIA relativas a la identificación biométrica deben resolverse ya sea incorporando las materias más controvertidas en IA de alto riesgo, o creando nuevas secciones que tengan en cuenta los riesgos asociados a estas herramientas.

3. SOBREDIMENSIÓN DE LA CLASIFICACIÓN DE LOS SISTEMAS IA DE ALTO RIESGO

Como se comentó en el apartado de clasificación, la IA de alto riesgo engloba el mayor número de supuestos del reglamento. En base a ello, surge el debate sobre si la categoría es demasiada amplia provocando que no se esté ponderando el riesgo correctamente.

De esta forma, es necesario considerar que se pueden estar imponiendo unas mismas obligaciones a situaciones con implicaciones muy distintas. En consecuencia, debería plantearse la posibilidad de incorporar subniveles con unos requisitos y obligaciones acordes al nivel de riesgo de cada situación (Cotino et al., 2021).

Así las cosas, se deben repasar las materias asignadas a este sector del FRIA. No volveremos sobre los casos de identificación biométrica, la gestión de la migración, el

asilo, el control de las fronteras y las consecuencias de la clasificación que han sido suficientemente desarrollados en el anterior apartado.

En primer lugar, respecto a las materias de gestión y funcionamiento de las infraestructuras esenciales, se hace referencia a las aplicaciones IA empleadas en el funcionamiento de dichas actividades. La aplicación de la IA sin duda ofrece grandes beneficios a esta materia, su uso puede permitir una gestión más eficiente de los recursos y de los procesos, no obstante, al tratarse de elementos críticos para la sociedad los elementos de seguridad han de ser rigurosos y resistentes (Laplante et al, 2020, pp 45-52). Esto último sustenta la incorporación de esta materia en la clasificación de IA de alto riesgo, no suponen una amenaza tan elevada como la IA de riesgo inaceptable pero sus consecuencias pueden provocar graves efectos en los ciudadanos.

En segundo lugar, en referencia a la educación y la formación profesional, la materia es controvertida al poder ser la IA empleada para ponderar los resultados de los alumnos provocando un efecto directo en los usuarios. Por ello, se requiere discernir entre las múltiples aplicaciones y valorar en función de las mismas.

En esta línea, en caso de la IA fuera empleada como un método de monitorización del proceso de evaluación, sería necesario distinguir soluciones dedicadas a detectar objetos (apuntes, libros…) o soluciones enfocadas al reconocimiento facial (por ejemplo, para detectar suplantaciones). En el primer caso estaríamos ante un supuesto que ya ha sido convalidado por la AEPD (AEPD, 2022b), mientras que en el segundo se estaría ante datos especialmente protegidos y ante un proceso de identificación biométrica que ya ha sido valorado en este trabajo.

Por su lado, en caso de que la IA fuera directamente empleada para llevar a cabo la evaluación de los exámenes o trabajos. Siempre que se evite la toma de decisiones automatizadas, la creación de perfiles y se mantenga una vigilancia humana que supervise el correcto funcionamiento de la herramienta localizando sesgos si los hubiera, no parece necesario incluir este método dentro de la clasificación de IA de alto riesgo.

En tercer lugar, los sistemas IA relativos al empleo y a los trabajadores (actuaciones como la contratación o la promoción) requieren un mayor control contra las decisiones automatizadas pues afectan directamente a la posición social de los usuarios. Por ello,

aunque si sigue las indicaciones del RGPD[24] se protege en gran medida los derechos de los usuarios, incluso con dicha protección las consecuencias del uso de estas herramientas justifican el mantenimiento de esta categoría como IA de alto riesgo (AEPD, 2020b).

En cuarto lugar, los sistemas IA que sean empleados para acceder y disfrutar de los servicios esenciales siguen un razonamiento análogo a las infraestructuras esenciales. Debido a ello, no requieren mayor explicación para comprender su inclusión en la IA de alto riesgo.

En quinto lugar, aquellos sistemas utilizados para aplicar la ley. Dentro de este grupo de sistemas entrarían aquellos empleados para valorar la probabilidad de reincidencia penal, los polígrafos, la detección de falsificaciones o las evaluaciones de pruebas en una investigación. Todas estas soluciones van a provocar consecuencias de gran calado en los ciudadanos, entra en juego el derecho de los acusados de contar con un proceso judicial con todas las garantías. Por ello, el uso de estas herramientas debe apoyarse en un sistema de obligaciones completo capaz de demostrar el correcto funcionamiento de la IA, así como se ofrece la presencia humana en todo el ciclo de vida de la IA.

En sexto y último lugar, los sistemas empleados en los procesos democráticos, al igual que en el caso anterior se presenta un uso que requiere del sometimiento a un proceso capaz de garantizar el correcto funcionamiento de la IA, por ello, es indudable que ha de mantenerse la clasificación de IA de alto riesgo para estos sistemas.

En conclusión, la vasta extensión de la categoría IA de alto riesgo es reflejo de la conversión del sector en una especie de cajón de sastre para el legislador, absorbiendo en la mayor parte de los usos que pueden desarrollarse con los sistemas IA. Urge por tanto una revisión de la categoría y una potencial redistribución de los sistemas IA, incorporados incluyendo nuevas categorías para las particularidades de ciertos usos de la IA.

[24] Artículo 22 del RGPD "derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar.")

4. PROCESADORES DEL LENGUAJE (CHAT GPT) CLASIFICACIÓN Y CONSECUENCIAS CON UNA ESPECIAL MIRADA A LA PROTECCIÓN DE DATOS

En el momento en el que se desarrolló el FRIA, no se podía siquiera imaginar las posibilidades que los procesadores del lenguaje, comúnmente denominados Chatbot, iban a desarrollas. Sin duda entre las múltiples soluciones de este tipo destaca Chat GPT por el impacto que ha tenido en la sociedad en tan poco tiempo.

En una aproximación general, este tipo de soluciones entrarían dentro de la clasificación de riesgo limitado, cuyas obligaciones se limitan a requisitos de transparencia e información. Sin embargo, Chat GPT ha mostrado una gran cantidad de situaciones que no han sido incluidas en el FRIA y que pueden afectar los derechos de los usuarios.

En primer lugar, como modelo de lenguaje los Chatbot como Chat GPT se construyen a través de la información publicada en internet, lo cual implica directamente un riesgo de infringir los derechos de autor de las obras recopiladas por la IA (art 10 LPI). La capacidad de resumir o plagiar la información desplegada en Internet junto con el desplazamiento de la responsabilidad hacia el usuario, provocan que las personas físicas puedan infringir derechos de propiedad intelectual sin ser plenamente conscientes de ello.

Se ha de tener en cuenta que lo que protege la regulación de la propiedad intelectual no son simplemente las ideas más la forma en la que estas se han desplegadas, el objeto jurídicamente protegido es la forma de expresar las ideas por parte de los autores (OMPI, 2016). Debido a los derechos en juego se debería promover la inclusión de unas obligaciones más estrictas en el funcionamiento de la IA. Por ejemplo, además del requisito de transparencia que se impone a la IA se debería valorar la incorporación de citas como ya se está haciendo en el Chatbot de Bing que emplea tecnología de Chat GPT. Al comprobar que es factible incorporar fuentes preliminares a las soluciones IA, se debe extender el requisito al resto de procesadores pues dotaría a los derechos de propiedad intelectual de un mecanismo de protección.

Adicionalmente, cobra especial relevancia la inclusión de mecanismos de vigilancia humana en los sistemas IA puesto que, como se ha comprobado, los resultados emitidos por Chat GPT pueden ser incorrectos o contener contenido ofensivo. La capacidad de usar

la IA como un vehículo de desinformación provoca que sea necesaria establecer técnicas para controlar los sesgos, así como la posibilidad de incluir descargos de responsabilidad en los textos que permitan que los usuarios sean conscientes en todo momento de la potencial falta de precisión que ofrece la respuesta otorgada por el Chatbot.

En deferencia a los resultados ofrecidos por las soluciones de esta categoría, se ha de tener en cuenta que de acuerdo con el artículo 5 de la LPI, los derechos de propiedad intelectual protegen las obras asociadas a una persona física (no entrando dentro del ámbito de protección aquellas creadas por la IA). Dicha idea es sustentada por el TJUE en el asunto C-145/10, dicha resolución dictamina que serán protegidas por derechos de autor solo aquellas obras en las que el autor refleje su personalidad. Al no tener una personalidad jurídica ningún texto producido por una IA podrá ser protegidos por la regulación vigente.

Una vez expuesto el funcionamiento y los derechos asociados a los resultados de la IA, conviene dar un paso atrás y volver al origen de la recopilación de la información empleada para crear las respuestas. Como se puede intuir para llevar a cabo la finalidad de la IA es necesario una recogida masiva de datos, lo cual puede afectar directamente con la regulación de protección de datos en Europa.

Italia ha sido el primer país europeo en prohibir la actuación de Chat GPT hasta que no cumpla con las obligaciones impuestas en el RGPD. De esta forma, en el procedimiento de medidas cautelares emitido por el organismo de protección de datos italiano[25], se alude a la falta de información relativa a la recogida de datos y a la ausencia de una base de legitimación adecuada para llevar a cabo el tratamiento. Asimismo, la resolución alude al tratamiento defectuoso de los datos de los menores de 13 años, por un lado, no se verifica que sean mayores de dicha edad para acceder al servicio, y, por otro, Chat GPT no incorpora los filtros requeridos a las respuestas que pueden obtener los menores al hacer uso de la aplicación (AIPD, 2023).

Si bien Italia es el primero en tomar esta decisión, Alemania o Francia también se plantea revisar la actuación de Chat GPT en relación con la protección de datos (Galán, 2023).

[25] Se han infringido de acuerdo con la AIPD los artículos 5,6,8,13 y 25 del RGPD

De cualquier forma, los estados han de tener en cuenta que el hecho de prohibir la solución no va a suponer que los ciudadanos no puedan hacer uso de la misma. Actualmente, herramientas digitales como las extensiones VPN permiten a los usuarios establecer conexiones de red privadas sobre las que ocultar la ubicación geográfica. Ello permite al usuario establecer una dirección IP en un país diferente al de origen, así el usuario podrá cambiar su dirección IP a la de un Estado donde Chat GPT esté habilitado. Dicha situación lleva a cuestionarse la utilidad de prohibir en su totalidad la aplicación de la solución.

No obstante, aunque la prohibición completa pueda ser exagerada por las múltiples posibilidades para sortearlas, los problemas que se están derivando de esta actuación requieren de la inclusión de la materia en el FRIA. Cuando se emitió el borrador no se contaba con la información adecuada, pero actualmente las consecuencias prácticas que se están produciendo por uso de la IA requieren de la incorporación de nuevas categorías de riesgo o del reforzamiento de la categoría de riesgo limitado. No es concebible que lo mayores requisitos impuestos a dicho sector de la IA se limiten a obligaciones de transparencia.

En línea con dicha idea, el Parlamento Europeo ha emitido una declaración sobre la actividad derivada de procesadores del lenguaje como Chat GPT. Lo verdaderamente novedoso de la declaración es el reconocimiento de que la clasificación de IA de alto riesgo en base a fines concretos, obvia los mecanismos como Chat GPT orientados a un propósito general (y por tanto libres de dicha clasificación y de las obligaciones asociadas a la misma) (Parlamento Europeo, 2023).

En base a lo anterior, como se ha comentado en este trabajo, la IA de alto riesgo ya se encuentra sobredimensionada, por ello, estos sistemas no deberían incluirse en dicha categoría sino en una nueva que tenga en cuenta las particularidades de la IA cuyo propósito sea de carácter general.

La última precisión que se debe hacer de este tipo de sistemas es la posición del derecho al olvido, regulado en el artículo 17 del RGPD (también denominado derecho de supresión). Se trata de la potestad de una persona a impedir la difusión de información

universal e indiscriminada de los datos personales[26], cuando la misma este desfasada o haya perdido la relevancia e interés público (AEPD, 2022a).

El mayor problema del ejercicio de este derecho en soluciones de Chat GPT es la imposibilidad de eliminar los datos recopilados. El ejercicio, por tanto, se limita enfocar la atención de la base de datos a la nueva información recopilada. En consecuencia, se debe realizar una valoración completa sobre los mecanismos de cumplimiento de la normativa para garantizar la protección de los derechos de los usuarios.

En conclusión, el avance continuo de la IA requiere de un análisis del sistema de clasificación de riesgos que se presenta actualmente en el FRIA. Es innegable que las consecuencias de la IA son imprevisibles, por ende, la regulación debe ser más abierta para incluir nuevos casos, limitando las prohibiciones a las situaciones realmente contrarias a los valores de la UE.

[26] En los motores de búsqueda tradicionales dicho derecho ya ha sido avalado por sentencias como la STJUE C-131/12 donde se garantiza el derecho de los usuarios a demandar la supresión de información personal cuando se cumplen unos requisitos

CAPÍTULO IV. MATERIALIZACIÓN DE LAS OBLIGACIONES DE LOS SISTEMAS IA DE RIESGO ALTO CON ESPECIAL ATENCIÓN A LA PROTECCIÓN DE DATOS

Una vez expuesta la clasificación y los problemas asociados a la misma, se presenta necesario ponderar las obligaciones impuestas a la categoría de riesgo alto.

El mayor problema de sobredimensión de este grupo radica precisamente en el alto número de obligaciones que se imponen, provocando que los proveedores deban llevar a cabo un esfuerzo añadido para poder comercializar con estos sistemas.

En consecuencia, este análisis podrá mostrar las posibilidades de materializar las obligaciones y pondrá en perspectiva la necesidad o no de imponer las mismas a todos los sistemas contenidos en el FRIA.

1. OBLIGACIONES GENERALES PARA TODOS LOS SISTEMAS IA DE ALTO RIESGO

1.1. Exposición de motivos y sistemas de gestión de riesgo y gobernanza

Debido al carácter de los sistemas IA de este grupo, las obligaciones que se imponen son variadas y afectan a todos los grupos sociales que interaccionan con el sistema. Dentro del FRIA se han establecido requisitos principalmente a los proveedores, a los distribuidores y a los usuarios, pero antes de profundizar en los mismos, se han establecido una serie de indicaciones generales para todo sistema IA de alto riesgo, que serán capaces de facilitar a las autoridades un control general sobre la adecuación del sistema al FRIA.

En primer lugar, para poder verificar el cumplimiento de la solución IA se acudirá tanto al objetivo perseguido por el sistema como a la existencia de un sistema de gestión de riesgos (artículo 8 FRIA).

Esta obligación sigue la estela marcada en el RGPD respecto al tratamiento de datos, el establecimiento de un fin concreto para el uso de los datos forma parte del Registro de Actividades de Tratamiento, conformando la información mínima que han de cumplir los encargados del tratamiento. Este proceso permite a las autoridades conocer la legitimidad del proceso desde una perspectiva general (AEPD, 2022c).

Como se expuso anteriormente, definir el objetivo perseguido es clave para conocer el nivel de riesgo del sistema IA, lo que consecuentemente permite una clasificación adecuada de la solución en el FRIA. Esta obligación se presenta esencial para cualquier IA, independientemente de la categoría en la que se encuadre, siendo absolutamente necesario que cualquier proveedor marque la finalidad perseguida por su herramienta.

Por su parte, la obligación referida al sistema de gestión se encuadra en el artículo 9 del FRIA debiendo consistir en un proceso reiterativo que será ejecutado durante toda la vida útil de la solución. Dicho sistema tendrá que demostrar capacidad para identificar tanto los riesgos conocidos como los previsibles, por el uso directo o indirecto del sistema IA, antes y después de la comercialización. Asimismo, en función de los riesgos localizados se deberá establecer un sistema de medidas de seguimiento posterior a la entrada en mercado de la IA que garanticen la correcta gestión de los riesgos.

En suma, las medidas implementadas deben ser capaces de minimizar o eliminar los riesgos, controlar aquellos que no puedan ser eliminados y considerar los conocimientos técnicos, la experiencia, la educación y la formación de los usuarios y del entorno donde se aplicará la IA. Además, se requiere que la IA actúe de manera coherente con su finalidad, que se realicen pruebas adecuadas antes de la comercialización o puesta en servicio, y que se definan métricas y umbrales previamente. Especialmente se deberá prestar atención a aquellos riesgos que pudiesen afectar a menores (artículo 9 FRIA).

Dicha obligación ya ha sido establecida para el tratamiento de los datos y cuenta con numerosas herramientas para llevarlo a cabo[27], en el caso de la IA los riesgos se presentan mayores pues las aplicaciones de la misma parecen ilimitadas. Dentro de esta categoría, se encuadran materias tan sensibles como la gestión migratoria, un mal uso de las herramientas de la IA podría incluso derivar en la inadmisión de los nacionales de un país determinado. Los proveedores de estas herramientas deben ser conscientes de estos

[27] Véase *"Gestiona EIDP"*, herramienta gratuita ofrecía por la AEPD para llevar a cabo la gestión de riesgos propios del tratamiento de datos personales (AEDP, 2020d)

riesgos, de los inherentes al sistema y los provocados por un mal uso. Paralelamente, se ha de vigilar los sesgos y los ataques, por todo ello esta obligación debe ser ineludible para aquellos sistemas con un riesgo alto.

No obstante, es cierto que los requisitos que se plantean no son sencillos, para facilitar el trabajo de los proveedores se recomienda la aplicación análoga de los sistemas de gestión de riesgos de protección de datos, y, a partir de ellos construir un sistema de gestión del riesgo referente a la IA que tenga en cuenta las particularidades de la misma. De esta forma, esta obligación, aunque extensa, es básica para garantizar el correcto funcionamiento del sistema.

En segundo lugar, los sistemas IA de alto riesgo deberán garantizar que cuentan con un sistema de gobernanza y de recopilación de datos robusto y transparente. Para ello, los conjuntos de datos deberán adecuarse a una serie de criterios mínimos de calidad encuadrados en el artículo 10 del FRIA.

En concreto, en el artículo se impone el establecimiento de un diseño de gobernanza adecuado, capaz de llevar a cabo un tratamiento de datos correcto durante todas las etapas de preparación y análisis de la solución. Similarmente, el sistema de gobernanza deberá evaluar la accesibilidad, cantidad e idoneidad de los datos, del mismo modo deberá detectar los posibles sesgos asociados con el objetivo de corregirlos antes de hacer uso activo de la información.

Al igual que en la gestión del riesgo, para que los proveedores de IA cuenten con un correcto sistema de gobernanza se debe acudir al RGPD para completar las obligaciones del FRIA. La calidad de los datos es fundamental para que la información obtenida sea útil para la toma de decisiones en una organización. Debido a ello, el RGPD indica que los objetivos de la gobernanza de datos deberán estar alineados con los objetivos de la organización, garantizando la responsabilidad estratégica, la definición de estándares de datos, la alineación de objetivos de la organización y la participación integral de todos los niveles de la organización (AEPD, 2020e).

Un sistema de gobernanza de datos robusto permite una mayor protección de los datos de los usuarios y, consecuentemente, una cobertura reforzada de los derechos fundamentales de los usuarios (en concreto, a su derecho a la intimidad y a la protección de los datos personales). Al igual que con el sistema de gestión del riesgo, las obligaciones impuestas al sistema de gobernanza son requisitos mínimos que permiten garantizar el

buen funcionamiento del sistema. El correcto desarrollo de estos sistemas se ha visto desde la emisión del RGPD, de hecho, los beneficios que procura la incursión de estos métodos denotan la proporcionalidad de la obligación.

En esta primera categoría de requisitos se busca conocer el objetivo último de la IA para comprobar la clasificación correcta del sistema como herramienta de alto riesgo. En base a esta verificación se instauran las principales obligaciones que han de cumplir este tipo de soluciones: la gestión de riesgo y la gobernanza de datos. No debemos olvidar que los sistemas IA de alto riesgo tratan de salvaguardar a los usuarios de los potenciales daños que pueden derivarse del uso de la IA. En consecuencia, el establecimiento de estos sistemas se erige como un contrapeso contra las consecuencias de una mala gestión de los sistemas IA y se presumen obligaciones necesarias para los proveedores de estos sistemas.

1.2. Documentación técnica de la solución IA

En esta sección, el legislador busca las pruebas que sustenten la clasificación de la IA como sistema de alto riesgo. En base a esta documentación, se podrá comprobar tanto el proceso de construcción de la solución como el funcionamiento técnico de la herramienta. Todo ello permitirá la construcción de una visión completa del sistema.

Dicha documentación deberá presentarse antes de introducir la IA en el mercado y mantenerse constantemente actualizada a los cambios que puedan incorporarse a la herramienta (artículo 11 FRIA).

La información mínima que deberá presentarse incluirá una descripción detallada del sistema de IA, su relación con el hardware y el software, y las instrucciones de uso e instalación. También el proveedor deberá proporcionar información sobre la lógica general, el diseño, el funcionamiento y los resultados obtenidos por la IA, así como los recursos informáticos y datos utilizados en la construcción y entrenamiento de la IA. Además, se debe incluir un análisis completo en caso de cambios en la IA, los procedimientos para asegurar el cumplimiento constante de las obligaciones y un informe sobre la monitorización, funcionamiento y control del sistema de IA. El informe deberá examinar las normas asociadas a la IA que se puedan referenciar en la FRIA, y también

tendrá que incluir las compensaciones en caso de que el sistema de IA no cumpla con las obligaciones impuestas en el futuro reglamento (Anexo IV FRIA).

Similarmente, la documentación deberá incluir las categorías utilizadas para clasificar los resultados, así como las razones inherentes al perfilamiento de las categorías, por ejemplo, en el caso de valorar la idoneidad para un crédito, las opciones serían "apto" o "no apto" y la justificación subyacente serían la salud financiera de la persona. Paralelamente, la documentación deberá proporcionar una valoración exhaustiva de la vigilancia humana incluida en el funcionamiento normal de la IA, los métodos de validación y ensayo, y finalmente los indicadores empleados para evaluar la precisión, la solidez, la ciberseguridad y el cumplimiento de los requisitos establecidos en el FRIA. Especial atención recibirán los posibles daños discriminatorios derivados del uso de la IA que deberán ser examinados para conocer sus efectos sobre la salud y seguridad y los derechos fundamentales (Anexo IV FRIA).

En el RGPD esta obligación tiene una importancia considerable, en concreto, se encuentra regulada en los artículos 74 y 78. En este ámbito, la documentación técnica debe describir cómo se procesan los datos personales, quién es responsable del tratamiento de los mismos, quién tiene acceso a los datos y cómo se protegen. La elaboración de documentación técnica en la protección de datos es un proceso continuo, que exige que los responsables del tratamiento de datos realicen evaluaciones de impacto sobre la protección de datos de forma regular (CEPD, 2021).

En conclusión, el proceso de elaboración de la documentación técnica en el RGPD y en el FRIA es un elemento básico para conocer los entresijos de la solución. A partir de él, se podrán establecer las cargas de responsabilidad en caso de que el sistema falle. En conjunto, se trata de una obligación con carácter probatorio que protege tanto a los proveedores como a los usuarios, no se trata de un requisito especialmente difícil técnicamente por lo que su realización podría ser extensible al resto de categorías (que ya por tratar datos deben llevar a cabo una documentación en el ámbito del RGPD).

1.3.Trazabilidad y transparencia de los sistemas IA

Tanto la trazabilidad como la transparencia son mecanismos esenciales para conocer la forma en la que se están utilizando los sistemas IA una vez puesto en funcionamiento. Estas obligaciones son herramientas auxiliares que permiten a los usuarios y a las autoridades protegerse de las consecuencias de los resultados previstos por la IA.

En primer lugar, respecto a la obligación de trazabilidad para los sistemas IA de alto riesgo, las soluciones deberán contar con técnicas capaces de registrar los eventos que surjan automáticamente por el uso de la IA, y, al mismo tiempo, tendrán que garantizar una monitorización adecuada del funcionamiento durante todo el ciclo de vida de la IA, ajustándose a las normas y especificaciones comunes (artículo 12 FRIA).

El requisito de trazabilidad es al final una actividad conexa al establecimiento de los sistemas de riesgos y gobernanza, se trata por tanto de una obligación que permite observar el funcionamiento de la IA tras la comercialización de la solución. Por ello, se trata de una referencia para el resto de las obligaciones generales y, dado que no supone una gran complejidad, se ve necesario extender dicha obligación al resto de categoría del FRIA.

En segundo lugar, las soluciones IA deberán demostrar un grado de transparencia capaz de habilitar a los usuarios para interpretar y utilizar los resultados emitidos por la IA.

Con dicho objetivo en mente, la IA de alto riesgo deberá llevar aparejada las instrucciones de uso del sistema. Dichas indicaciones deberán incluir información sobre la identidad y datos del proveedor o representante de la IA, así como una descripción completa de las capacidades y límites del sistema, incluyendo la finalidad prevista, el grado de precisión, solidez y ciberseguridad esperable y los posibles riesgos asociados a su mal uso. Paralelamente, se habrá de incluir detalles sobre el funcionamiento general que los usuarios promedio llevarán a cabo y, si es necesario, características de los conjuntos de datos de entrenamiento. Por último, los sistemas IA deberán demostrar ser transparentes, revelando cualquier cambio en el funcionamiento previsto en la evaluación de conformidad inicial, ofreciendo las medidas de supervisión humana y de conservación y protección necesarias para garantizar el correcto funcionamiento de la solución (artículo 13 FRIA).

Dentro de la protección de datos, el principio de transparencia tiene una importancia trascendental sobre el tratamiento de los datos. El propósito último del principio es proporcionar a los interesados la información necesaria sobre los tratamientos que les afecten y sus derechos, siendo necesaria que ésta sea concisa, transparente, fácil de entender y accesible. Para completarlo, se debe evitar el uso de lenguaje legal complejo y las cláusulas informativas deben explicar claramente el contenido al que se refieren, de manera que cualquier persona pueda entenderlo, independientemente de su nivel de conocimiento en la materia (AEPD, 2018b). En concreto dicha obligación se encuentra recogida en el artículo 12 del RGPD.

En conclusión, la obligación de trazabilidad y transparencia surge con la finalidad de llevar a cabo un seguimiento completo de los usos que se llevan a cabo por los usuarios, y al mismo tiempo, demostrar un conocimiento completo de la forma en la que funciona la solución. Todo ello con el objetivo de aportar herramientas al usuario y a los organismos de responsables para defender los intereses y derechos de los ciudadanos ante las consecuencias de un mal uso del sistema IA.

1.4. Vigilancia humana y medidas de precisión, solidez y ciberseguridad

En este apartado, se busca garantizar la seguridad del sistema a través de la incorporación de un elemento humano capaz de intervenir en cualquier momento del ciclo de vida de la solución IA. Por su parte, las medidas de precisión, solidez y ciberseguridad buscan una protección general sobre el sistema frente a ataques externos y a los posibles fallos derivados del uso de la IA.

Entrando en materia, la supervisión humana tiene como finalidad evitar o minimizar los riesgos asociados a la IA en relación con la salud, seguridad o los derechos fundamentales de los usuarios (artículo 14 FRIA).

De esta forma, se ve necesario establecer una correcta supervisión de los sistemas IA, deben definirse medidas concretas orientadas al usuario garantizando en todo momento el correcto funcionamiento de la herramienta. Para ello, se debe designar una persona capaz de comprender las capacidades y limitaciones del sistema IA y supervisar su funcionamiento, recordando en todo momento el riesgo que supone depositar una

confianza ciega o excesiva en los resultados emitidos por la IA. Específicamente, en caso de una IA destinada a la identificación biométrica remota de personas, ninguna acción o decisión debe ser tomada por el usuario basándose únicamente en la identificación resultante del sistema, a menos que haya sido verificada y confirmada por al menos dos personas físicas (artículo 14 FRIA).

Respecto a la protección de datos, la presencia de esta obligación se encuentra referida a las decisiones automatizadas, como se ha mencionado en otras partes del trabajo, el artículo 22.1 del RGPD incluye el derecho de las personas a no ser objeto de decisiones automatizadas. En esta línea, el artículo 22.3 del RGPD reconoce el derecho del particular a obtener una intervención humana, a expresar su opinión y a impugnar las decisiones automatizadas. No obstante, es necesario mencionar que existe una excepción a este requisito en el artículo 22.2 RGPD, de esta forma cuando el tratamiento se haya amparado en el consentimiento explícito de los interesados o, el mismo está autorizado por la normativa española o europea no cabre un derecho a la oposición (Palma, 2019).

Si bien la vigilancia humana en la protección de datos es importante, en la IA el calibre de esta obligación adquiere un cariz aun más relevante. Las decisiones que toman los sistemas de alto riesgo, como se ha expuesto anteriormente, pueden derivar en graves consecuencias para los usuarios, por ejemplo, un error en el otorgamiento de un crédito puede propiciar la pérdida de financiación de una persona solvente por motivos ajenos a la misma. En consecuencia, la vigilancia humana puede ayudar a detectar errores y sesgos, así como a garantizar que se toman decisiones éticas y responsables.

Por otro lado, respecto a las medidas de precisión, solidez y ciberseguridad, los sistemas IA deberán garantizar una serie de procedimientos técnicos que salvaguarden el correcto funcionamiento de la herramienta.

De esta forma, la IA deberá asegurar que su diseño y funcionamiento está orientado a cumplimentar su finalidad de forma acorde con los requisitos de precisión, solidez y ciberseguridad necesarios durante todo su ciclo de vida. En consonancia, dichos requisitos habrán de demostrar capacidad para resistir a los problemas, fallas e incongruencias derivadas del uso de la IA o del entorno en el que se emplea el sistema. Para lograr esto, se deben incluir soluciones técnicas específicas para el tipo de IA y métodos para resolver vulnerabilidades y prevenir ataques que busquen manipular el

conjunto de datos de entrenamiento, los datos de entrada o los defectos del modelo (artículo 15 FRIA).

Ejemplos de implementaciones de estas obligaciones serían evaluaciones a los conjuntos de datos, testeos a los modelos IA, establecimiento de mecanismos de detección y corrección de errores, pruebas de resistencias o técnicas de encriptación y autenticación para proteger los datos de entrada y salida del sistema IA.

En conclusión, en esta sección se busca la protección de los sistemas IA frente a los daños producidos por ataques externos o por el propio funcionamiento de la IA. Esta protección es realizada desde dos frentes diferenciados, la vigilancia humana y la incorporación de medidas técnicas al modelo. Mientras que las medidas tratan de impedir que el sistema IA derive en un mal uso, la vigilancia humana va más allá no solo cumpliendo este objetivo, sino que además actúa como el último dique de protección del usuario siendo capaz de revertir los resultados de la IA cuando los mismos sean fruto de un proceso defectuoso.

2. OBLIGACIONES A LOS PROVEEDORES Y FABRICANTES DE SISTEMAS IA DE ALTO RIESGO

Los proveedores de IA de alto riesgo, además de contar con un deber de cumplimiento de los requisitos antes mencionados, tienen una serie de obligaciones fruto de su posición clave para el desarrollo de la IA. Asimismo, se ha de notar que las obligaciones que se exponen a continuación serán aplicables de forma idéntica a los fabricantes de los productos del Anexo 2 A (artículo 24 FRIA).

2.1. Sistema de gestión de la calidad

Como proveedores, las obligaciones que se imponen en esta sección deben permitir garantizar que el producto que ofrecen cumple con las garantías adecuadas para ser comercializado, por ello, un sistema de gestión de la calidad es fundamental para lograr dicho objetivo.

Para ello se ha de acudir al artículo 17 del FRIA, en el mismo se establece que los proveedores de IA deben contar con un sistema de gestión de la calidad que incluya políticas, procedimientos e instrucciones documentadas para demostrar la calidad de la solución. El sistema de gestión deberá incluir una estrategia completa para cumplir con la normativa y los procedimientos de evaluación de la conformidad y gestión de las posibles modificaciones de la herramienta.

Un buen sistema de gestión de la calidad se va a basar en gran medida en cumplimentar las obligaciones de documentación y gestión de riesgos. No obstante, otras formas con las que configurar este sistema sería incluir herramientas para monitorizar el cumplimiento normativo, definir un sistema de notificación relativo a las fallas combinado con una gestión continuada de la comunicación con las autoridades competentes. Adicionalmente, sería conveniente establecer de forma clara el marco de responsabilidad respecto a la dirección y miembros del personal del proveedor de la IA, lo que permitirá al usuario defenderse ante los efectos negativos que se pudiesen derivar del uso de la solución.

En conclusión, los proveedores deberán demostrar que cuentan con un sistema de gestión de la calidad adecuado a los requisitos dispuestos en la FRIA, y que permitirán a los usuarios tener la confianza de que están haciendo uso de un sistema con las máximas garantías de calidad y, al mismo tiempo, ofrecerán al usuario una persona hacia la que tomar acciones legales en caso de que el sistema les produjesen un daño.

2.2. Evaluación de conformidad

Esta obligación surge como un deber para los proveedores de demostrar que están cumpliendo con los requisitos asociados con los sistemas IA de alto riesgo. Asimismo, este deber permite a los usuarios comprobar que están haciendo uso de un sistema IA adecuado que cumple con la legislación europea.

De esta forma, los proveedores de IA de alto riesgo han de garantizar que las soluciones que se desarrollen cumplen con un proceso de evaluación de la conformidad aceptable, antes de la comercialización o puesta en servicio de la herramienta. En concreto, los proveedores deberán redactar una declaración UE de conformidad y colocar

un marcado CE (artículo 18 FRIA)[28]. Dicha declaración deberá mantener accesible durante al menos diez años tras la comercialización del sistema IA, de igual modo, se deberá entregar una copia del documento a las autoridades nacionales pertinentes en caso de que se solicite (artículo 48 FRIA)[29].

Por su parte, en referencia con el marcado CE, éste será incluido de forma visible, accesible y comprensible en todos los sistemas IA de alto riesgo[30]. Además, tendrá que incorporar el número de identificación del organismo responsable de supervisar los procesos de evaluación de la conformidad (artículo 49 FRIA) [31].

Retomando la obligación de llevar a cabo una evaluación de conformidad, en el caso de que el proveedor se hubiese adherido a normas armonizadas (artículo 40 FRIA) o especificaciones comunes (artículo 41 FRIA), la evaluación podrá basarse en un control interno[32], o, en una evaluación completa del sistema de gestión de la calidad (artículo 43 FRIA).

En caso de que el sistema IA hiciese referencia a una solución de identificación biométrica, y no se hubiesen aplicado correctamente las normas armonizadas o no se hubiesen aplicado las especificaciones comunes, el proveedor tendrá que demostrar haber seguido el procedimiento de evaluación de la conformidad, identificando la autoridad competente[33] que va a llevar a cabo la vigilancia del mercado y estar en posición de una certificación vigente[34] producida por el organismo incluido en el Anexo VII.

---

[28] En el caso de los sistemas de IA de alto riesgo mencionados en el punto 5, letra b), del anexo III, introducidos en el mercado o puestos en servicio por proveedores que sean entidades de crédito reguladas por la Directiva 2013/36/UE, la evaluación de la conformidad se llevará a cabo como parte del procedimiento a que se refieren los artículos 97 a 101 de la mencionada Directiva.

[29] No será necesario redactar más de una declaración cuando se incluyan en la declaración original todas las legislaciones comunitarias aplicables al sistema IA (Artículo 48 FRIA)

[30] Si dicha incorporación no fuese factible o no pudiese avalarse en base al origen del sistema IA de alto riesgo, el marcado CE deberá incorporarse al embalaje o en los documentos adjuntos (Artículo 49 FRIA)

[31] El marcado CE estará sujeto a los principios generales contemplados en el artículo 30 del Reglamento 768/2008: "Cuando un producto plantea un riesgo grave, es necesario intervenir rápidamente, lo que puede implicar que el producto sea retirado del mercado o recuperado, o que se prohíba su comercialización. En esas situaciones, es necesario poder recurrir a un sistema de intercambio rápido de información entre los Estados miembros y la L 218/32 ES Diario Oficial de la Unión Europea 13.8.2008 Comisión. El sistema previsto en el artículo 12 de la Directiva 2001/95/CE ha demostrado su eficacia y eficiencia en el ámbito de los productos de consumo. Para evitar una duplicación innecesaria, debe utilizarse dicho sistema a efectos del presente Reglamento. Además, la vigilancia del mercado coherente en toda la Comunidad requiere un intercambio completo de información sobre las actividades nacionales en la materia, más allá del presente sistema"

[32] Véase Anexo VI FRIA

[33] Véase artículo 63 FRIA

[34] Véase Artículo 44 FRIA

Finalmente, en los casos excepcionales, se tendrá que llevar a cabo una nueva evaluación de la conformidad en caso de que el sistema IA fuera modificado considerablemente (artículo 43 FRIA).

Con todo lo anterior, se puede concluir que la evaluación de conformidad es una obligación necesaria para garantizar que el sistema IA no solo funciona correctamente, sino que, además, cumple con las regulaciones europeas y nacionales relativas a la IA. Si bien es cierto, que recabar el marcado CE para cualquier IA incluida en la categoría de alto riesgo parece excesivo pues como se vio en el capítulo de clasificación no todas las materias cuentan con el mismo riesgo, y esta obligación puede ser excesiva, por ejemplo para las herramientas de evaluación educativas (con la vigilancia humana podría suplirse este requisito pues no presenta un riesgo realmente elevado).

2.3. Acciones de diligencia

Las acciones de diligencian engloban todas aquellas obligaciones asociadas a un sistema IA que no esté cumpliendo los requisitos del FRIA.

En primer lugar, desde el instante en el que el proveedor sea consciente de que el sistema está incumpliendo alguna cláusula del FRIA, se deberán adoptar las medidas pertinentes para adecuar el sistema a los requisitos solicitados, y, en caso de que no fuese posible incorporar ninguna medida, retirar el sistema IA (artículo 21 FRIA).

En segundo lugar, el proveedor tendría que garantizar que en caso de presentar riesgos que pudiesen derivar en daños potenciales a los usuarios, se han de marcar procesos para notificar a las autoridades nacionales de la disponibilidad de la solución IA, y si fuese el caso, notificar de forma inmediata a los organismos pertinentes (artículo 22 FRIA).

En tercer lugar, el proveedor tendrá que incorporar mecanismos claros de notificación hacia las autoridades nacionales. En esa línea se ha de diseñar un proceso que incluya una solicitud previa y los datos y documentos necesarios para demostrar el cumplimiento del sistema de las obligaciones de la FRIA. Por su parte, en caso de estar incumpliendo algún precepto, establecer un cauce de notificación que prevea la presentación de una solicitud previa motivada a los registros creados por la IA (artículo 22 FRIA).

En conclusión, todas estas acciones se erigen como medidas correctoras ante un incumplimiento de las obligaciones impuestas en el futuro reglamento. Se trata de medidas básicas y coherentes que pueden, y deberían, ser extrapolables a todas las categorías de riesgo puesto que permiten advertir y, en su caso eliminar, soluciones IA cuyos efectos sean dañinos para la sociedad.

2.4. Nombramiento de un representante en la UE

Esta obligación recae directamente sobre aquellos proveedores establecidos fuera de la Unión y que no cuentan con un importador identificable. El objetivo de este requisito es establecer un nexo entre los proveedores de la IA y los distribuidores en la UE de dichos sistemas.

Dicho representante deberá conservar una copia de la declaración conforme de la UE y de la documentación técnica. Además, serán los encargados de facilitar a la autoridad nacional, previa solicitud motivada, todos los datos y documentación que permitirán garantizar el cumplimiento de la IA con el FRIA. Asimismo, los representantes tendrán que cooperar constantemente con las autoridades nacionales (artículo 25 FRIA).

En conclusión, está ultima obligación está orientada a aquellos proveedores extranjeros que no cuentan con un importador reconocible. Este requisito trata de proteger a los usuarios frente a la comercialización de sistemas IA externos.

3. OBLIGACIONES A LOS IMPORTADORES Y DISTRIBUIDORES

El FRIA incorpora una sección destinada a regular las obligaciones específicas que deben cumplimentar los importadores de los sistemas IA.

De esta forma, los importadores deberán garantizar que los proveedores han cumplimentado la evaluación de conformidad, han presentado correctamente la documentación técnica pertinente y, en última instancia, que la solución IA cuenta con el marcado de conformidad correspondiente y con las instrucciones e información necesarias. Del mismo modo, deberán comprobar que se cumplen todas las obligaciones

de los sistemas IA y mantener una estrecha cooperación con las autoridades (artículo 25 y 26 FRIA).

Adicionalmente, en caso de que el importador note que la solución IA no se adecue con el FRIA, no podrá comercializar el sistema hasta que se adapte el mismo a la normativa europea. En esta línea, en caso de que el sistema IA provoque un riesgo para la salud, la seguridad o los derechos fundamentales deberá avisar tanto al proveedor como a las autoridades de vigilancia del mercado. Similarmente, los importadores deberán establecer su nombre, su nombre comercial (o marca registrada) y su dirección de contacto en la herramienta IA (artículo 25 y 26 FRIA).

En conclusión, los importadores y distribuidores deberán servir como un dique entre los usuarios y los proveedores. El objetivo último es que garanticen que se están cumpliendo con las obligaciones vinculadas con los sistemas IA y que, al mismo tiempo, puedan ayudar a las autoridades a identificar aquella IA que provoque riesgos dañinos para las personas físicas que hagan uso de ella.

CAPÍTULO V. CONCLUSIONES

Las principales conclusiones que pueden extraerse de este trabajo son las siguientes:

Primera. Los constantes cambios relativos a la definición de la IA denotan la constante evolución de la tecnología. El concepto que se tiene actualmente de la misma se ha quedado obsoleta con las nuevas soluciones desarrolladas. En consecuencia, la UE debería actualizar la explicación expuesta en el FRIA pues la misma es básica para el posterior desarrollo legislativo.

Segunda. El equilibrio que pretende lograr la UE con el FRIA, proteger los derechos de los ciudadanos y convertirse en una potencia de IA, choca frontalmente con la incorporación de una legislación tan exhaustiva. El gran número de obligaciones impide un desarrollo pleno de los sistemas puesto que este tipo de tecnología depende en gran medida de los datos que se incorporen a la solución (materia altamente regulada tanto por el FRIA como por el RGPD). Asimismo, la falta de una inversión similar a la de estados como EE. UU. impiden a la UE posicionarse como una potencia líder del sector.

Tercera. La rápida evolución de la IA provoca la aparición de ininterrumpida de soluciones IA no previstas en el FRIA, lo cual unido a la considerable demora que suele venir aparejada con la tramitación de normativa comunitaria, puede provocar situaciones de indefensión para los ciudadanos por la falta de legislación. La UE debería incorporar marcos generales en el FRIA (que puedan proteger frente a las soluciones más dañinas mientras dure los cambios legislativos) o establecer mecanismos que agilicen los procesos de modificación de la norma.

Cuarta. La identificación biométrica no se encuentra correctamente regulada en el FRIA. La dispersión de los supuestos en el FRIA y la decisión de no regular otros, esta propiciando una situación de indefensión para los ciudadanos

Quinta. Los conceptos de “tiempo real” y “remoto” relativos a la IA propician la aparición de lagunas legislativas que impiden un correcto tratamiento de los datos en la identificación biométrica.

Sexta. La sobredimensión de la categoría de alto riesgo impide otorgar a la misma el valor que merece. Se están aglutinando materias con un nivel de riesgo muy distinto y, con ello, se están imponiendo obligaciones no acordes con los supuestos encuadrados en

la sección. Este problema ocurre tanto por un nivel de riesgo superior al de la categoría (identificación biométrica, gestión de la inmigración) como por un nivel inferior (evaluaciones en la educación).

Séptima. Los procesadores de lenguaje no se encuentran correctamente regulados en el FRIA. Las consecuencias derivadas del Chat GPT demuestran el riesgo que se deriva de un uso indiscriminado de los datos con un propósito general. Asimismo, la prohibición de la herramienta en Italia, así como las dudas de Alemania y Francia, denotan la necesidad de regular de forma adecuada este tipo de soluciones IA. No obstante, regular no implica prohibir, herramientas como los VPN permitirían eludir la prohibición y no se estaría protegiendo de forma efectiva los derechos de los ciudadanos al optar por no enfrentar las vicisitudes normativas de este tipo de soluciones.

Octava. Las obligaciones de la IA de alto riesgo se muestran no solo factibles sino necesarias para las materias que realmente cuenten con un nivel de riesgo alto. No obstante, aquellas soluciones que no cuentan con un nivel de riesgo tan elevado muestran que las obligaciones pueden ser excesivas para los proveedores (por ejemplo, la obtención de un certificado CE por los problemas administrativos que suele llevar aparejado). Esta idea esta íntimamente relacionada con importancia de que el FRIA pondere la relación entre los bienes jurídicamente protegidos y sus medidas limitativas a su preservación, de forma que sean proporcionales y acordes a la realidad regulada.

FUENTES DE INVESTIGACIÓN

LEGISLACIÓN

Comisión Europea (2022a). *Propuesta de directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsbilidad civil extracontractual a la Inteligencia Artificial*. COM/2022/496 final. Unión Europea. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52022PC0496

Comisión Europea (2021a). *Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de IA) y se modifican determinados actos legislativos de la Unión* (FRIA). COM/2021/206 final. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex:52021PC0206

Consejo de la Unión Europea. (2002). *Decisión Marco 2002/584/JAI del Consejo de 13 de junio de 2002 relativa a la orden de detención europea y a los procedimientos de entrega entre Estados miembros*. Diario Oficial de la Unión Europea L 190/1. Disponible en: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002F0584:ES:HTML

Convenio Europeo de Derechos Humanos. (1950). *Artículo 8: Derecho al respeto de la vida privada y familiar*. Disponible en: https://www.echr.coe.int/Documents/Convention_ENG.pdf

Constitución española de 1978. Boletín Oficial del Estado, número 311, de 29 de diciembre de 1978. Disponible en: https://www.boe.es/buscar/act.php?id=BOE-A-1978-31229

Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación. Boletín Oficial del Estado, 13 de julio de 2022, núm. 167. Disponible en: https://www.boe.es/buscar/pdf/2022/BOE-A-2022-11589-consolidado.pdf

Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público. Boletín Oficial del Estado, núm. 236, de 2 de octubre de 2015. Disponible en: https://www.boe.es/buscar/act.php?id=BOE-A-2015-10566

Ley 11/2007, de 22 de junio, de Acceso Electrónico de los Ciudadanos a los Servicios Públicos. Boletín Oficial del Estado, núm. 151, de 23 de junio de 2007. Disponible en: https://www.boe.es/buscar/act.php?id=BOE-A-2007-12352

Ley 58/2003, de 17 de diciembre, General Tributaria. Boletín Oficial del Estado, núm. 302, de 18 de diciembre de 2003. Disponible en: https://www.boe.es/buscar/act.php?id=BOE-A-2003-23186

Ley 1/1996, de 12 de abril, de Propiedad Intelectual. BOE núm. 97, de 22 de abril de 1996. Disponible en: https://www.boe.es/buscar/act.php?id=BOE-A-1996-8930

Ley Orgánica 5/1985, de 19 de junio, del Régimen Electoral General. Boletín Oficial del Estado número 147, de 20 de junio de 1985, modificación del 6 de diciembre de 2018. Disponible en: https://www.boe.es/buscar/act.php?id=BOE-A-1985-11672&b=82&tn=1&p=20181206#ac

JURISPRUDENCIA

Rechtspraak. (2020). ECLI:NL:RBDHA:2020:1878. Uitspraken.rechtspraak.nl. Dispobible en : https://uitspraken.rechtspraak.nl/#!/details?id=ECLI:NL:RBDHA:2020:1878

Tribunal Constitucional. (2019). Sentencia 76/2019, de 13 de junio de 2019. Recurso de inconstitucionalidad 1674-2014. BOLETÍN OFICIAL DEL ESTADO [BOE], número 163, de 9 de julio de 2019. Disponible en: https://www.boe.es/boe/dias/2019/06/25/pdfs/BOE-A-2019-9548.pdf

Tribunale Amministrativo Regionale (TAR) Lazio-Roma. (2017). Sentenza n. 2465/2017. Disponible en: http://www.dirittomedicinasport.it/wp-content/uploads/2017/04/TARLazioRoma_2017_3372s1ter.pdf

TJUE. (2014). Sentencia del Tribunal de Justicia (Gran Sala) en el asunto C-131/12. Disponible en: https://curia.europa.eu/juris/document/document.jsf?docid=152065&doclang=ES

TJUE. (2011). Sentencia del Tribunal de Justicia (Sala Tercera) en el asunto C-145/10. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:62010CJ0145&from=ES

BIBLIOGRAFÍA

Agencia Española de Protección de Datos (AEPD). (2022a). *Derecho de supresión ("al olvido"): buscadores de internet*. Disponible en: https://www.aepd.es/es/areas-de-actuacion/internet-y-redes-sociales/derecho-al-olvido

Agencia Española de Protección de Datos (AEPD). (2022b). *Expediente N.º: EXP202200367*. Disponible en: https://www.aepd.es/es/documento/ai-00086-2022.pdf

Agencia Española de Protección de Datos (AEPD). (2022c). *Registro de actividades del tratamiento*. Disponible en: https://www.aepd.es/es/derechos-y-deberes/cumple-tus-deberes/medidas-de-cumplimiento/actividades-tratamiento

Agencia Española de Protección de Datos (AEPD). (2020a). *Consulta 0036/2020*. Disponible en: https://www.aepd.es/es/documento/2020-0036.pdf

Agencia Española de Protección de Datos (AEPD). (2020b). *La adecuación al Reglamento General de Protección de Datos de la IA*. Disponible en: https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia.pdf

Agencia Española de Protección de Datos (AEPD). (2020c). *Nota sobre los equívocos en torno a la biometría*. Disponible en: https://www.aepd.es/es/documento/nota-equivocos-biometria.pdf

Agencia Española de Protección de Datos (AEPD). (2020d). *Gestiona EIPD*. Disponible en: https://www.aepd.es/es/guias-y-herramientas/herramientas/gestiona-eipd

Agencia Española de Protección de Datos (AEPD). (2020e). *Gobernanza y política de protección de datos*. Disponible en: https://www.aepd.es/es/prensa-y-comunicacion/blog/gobernanza-y-politica-de-proteccion-de-datos

Agencia Española de Protección de Datos (AEPD). (2020f). *Protección de datos: Guía para el ciudadano*. Disponible en: https://www.aepd.es/sites/default/files/2020-05/guia-ciudadano.pdf

Agencia Española de Protección de Datos (AEPD). (2018a). *¿Cuáles son las bases de legitimación para el tratamiento de las categorías especiales de datos?*. Disponible en: https://www.aepd.es/es/preguntas-frecuentes/2-rgpd/5-bases-juridicas-de-los-tratamientos/FAQ-0215-cuales-son-las-bases-de-legitimacion-para-el-tratamiento-de-las-categorias-especiales-de-datos#:~:text=Las%20categor%C3%ADas%20especiales%20de%20datos%20son%20aquellas%20que%20incluyen%20datos,relativos%20a%20la%20salud%20o%20a

Agencia Española de Protección de Datos (AEPD). (2018b). *¿En qué consiste el principio de transparencia e información a los afectados?*. Disponible en: https://www.aepd.es/es/preguntas-frecuentes/2-rgpd/3-principios-relativos-al-tratamiento/FAQ-0210-en-que-consiste-el-principio-de-transparencia-e-informacion-a-los-afectados

Agencia Italiana de Protección de Datos (AIPD) (2023). *Resolución nº 9870832*. Disponible en:

https://www.gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9870832

Casado, E. G. (2021). *El enfoque europeo de Inteligencia Artificial*. Revista de Derecho Administrativo, 20, 268–289. Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=8510535

Comisión de Garantía del Derecho de Acceso a la Información Pública de Cataluña. (2016). Resolución de la Reclamación 123/2016. Disponible en: https://www.accesoinfo.cat/sites/default/files/arxius/resolucio_cgaip_123_2016.pdf

Comisión Europea (2022b).*Un enfoque europeo de la Inteligencia Artificial*. Estrategia europea de la IA. Disponible en: https://digital-strategy.ec.europa.eu/es/policies/european-approach-artificial-intelligence

Comisión Europea. (2021b). *Preguntas y respuestas: Regulación de IA para aumentar la confianza en una tecnología segura y justa*. Disponible en: https://ec.europa.eu/commission/presscorner/detail/es/qanda_21_1683

Comisión Europea (2021c). *Primer plan estratégico de Horizonte Europa 2021-2024: la Comisión establece prioridades de investigación e innovación para un futuro sostenible*. Disponible en: https://ec.europa.eu/commission/presscorner/detail/es/ip_21_1122

Comisión Europea (2020a). *Comunicación de la comisión al parlamento europeo, al consejo europeo, al consejo, al comité económico y social europeo y al comité de las regiones.* Diario Oficial de la Unión Europea SWD 137, p 1. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52018DC0237&from=ES

Comisión Europea (2020b). *Libro Blanco sobre la Inteligencia Artificial - un enfoque europeo orientado a la excelencia y la confianza*. COM/2020/65 final. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52020DC0065

Comisión Europea (2019a). *Directrices éticas para una IA fiable*, Dirección General de Redes de Comunicación, Contenido y Tecnologías, Oficina de Publicaciones. Disponible en: https://data.europa.eu/doi/10.2759/14078

Comisión Europea. (2019b). *Pacto Verde Europeo: Hacia una Europa climáticamente neutra*. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52019DC0640&from=ES

Comité Económico y Social Europeo (2018). *Dictamen del Comité Económico y Social Europeo sobre «Inteligencia Artificial: anticipar su impacto en el trabajo para garantizar una transición justa»*. Diario Oficial de la Unión Europea C 440, p 3. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52018IE1473&from=ES

Comité Económico y Social Europeo (2017). *Dictamen del Comité Económico y Social Europeo «IA: las consecuencias de la Inteligencia Artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad»*. Diario Oficial de la Unión Europea C 288, pp. 58-65. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52016IE5369&from=ES

Comité Europeo de Protección de Datos. (2021). *Directrices 3/2019 sobre la aplicación del Reglamento General de Protección de Datos (RGPD): Protección de datos por diseño y por defecto. Versión 2.0.* Disponible en: https://edpb.europa.eu/system/files/2021-04/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_es.pdf

Cotino, L., Castillo, J. A., Salazar, I., Benjamins, R., Cumbreras, M., & María Esteban, A. (2021). *Un análisis crítico constructivo de la Propuesta de Reglamento de la Unión Europea por el que se establecen normas armonizadas sobre la Inteligencia Artificial (Artificial Intelligence Act)*. Diario La Ley 7980/2021. Disponible en: https://laleydigital.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAEAE2Qy27CMBBFv6beIFUOIVAWXhBYVxWk3Q_2AKO6dvAjkL_vOFGlWjqyxvO4d3zPGMYOn0lpC5EupEHTS5byUrsFuYSWrug0wQJCKmkCK-LovBt_VBcyigTnqCopXza6UDE1s2IaZsO8MduSmwrWAnTKYA9eq2pVAhqwg7NqhA8GQzsqKZJPYI8YVVOJePOPdxjoCom8ayHMwmSMOnSST73cVs

1WDBgiF6ivYjmhiAhB3z7giqrN8Z7RwCvE_iko7vo--AGNWsplLVc8oa7XwrpvNnWa2maNgfChyBl87iGYnTPlq7j_01FRA_u_eNZrc0ps4pzcnBPa8n2AhHuw6Myffeh7Ox695R2n-OKDxhPGaYXEL7--0qUumwEAAA==WKE#I5

Cotino Hueso, L. (2021). *Sistemas de Inteligencia Artificial con reconocimiento facial y datos biométricos. Mejor regular bien que prohibir mal*. El cronista del estado social y democrático de derecho, (100). Disponible en: https://www.uv.es/cotino/publicaciones/cronistacotinopublicado.pdf

De Hoyos Sancho, M. (2021). *El libro blanco sobre Inteligencia Artificial de la Comisión Europea: reflexiones desde las garantías esenciales del proceso penal como "sector de riesgo". Civitas, 76*, 9-44. Disponible en : http://www.revistasmarcialpons.es/revistaespanoladerechoeuropeo/article/view/534/536

EDRi. (2021). *Llamado de la sociedad civil a establecer límites para la IA en la propuesta de la Unión Europea*. Disponible en: https://edri.org/our-work/civil-society-call-for-ai-red-lines-in-the-european-unions-artificial-intelligence-proposal/

Fernández, C. (2020). *El Libro Blanco de la Comisión Europea sobre Inteligencia Artificial*. Diario La Ley, 37. Disponible en : https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAEAMtMSbF1jTAAAkNTIwMzM7Wy1KLizPw827DM9NS8klS1zGLHgoKi_LLUFFsjAyNjAyNDA0NLE3MAGK89UzgAAAA=WKE

Fernández Hernández, C. (2022a). *China contará con un sistema de crédito social para ciudadanos, administraciones y empresas*. Diario La Ley N.º 67. Disponible en: https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAEAA3BQQqAIBAF0Nu4NjWihYuobhDRVppPDIjKaJ6_3mMiv1_6N0zj4JzqkMo5-ZMfpAbFdSlFcgd5o43VTltj7azuiCBbaFhDRKIg_pAXH9rBj4BPAAAAWKE

Fernández Hernández, C. (2022b). *La Ley 15/2022 introduce la primera regulación positiva de la IA en España*. Diario La Ley. Disponible en: https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAEAE2QQWvDMAyFf019MYwkbdnJl3Y9FMoYXRm7qo6aiLl2JstZ8--npgxmEJJ4T48Pfxfk6YQ3cQewB5xsvV40l6ZqGktROLXFow1gB6YrMljGrgTwtChVdVlGO6RMQiPYdrbpDQbqMHoCCyx0IZ2CxWh3eYD5qgaTp5jidHUnLmgEztktVwa8FAgvybv6PtOIJzjfhcQt8mZylZEkEI6YXWNyn35eYaQOhFLcAD_CqG3d7rPSVz8363plRuSsBvdxpxI0PXX9QUse_ozAvn-DDt0-Kmt6gjzcTIhfCvI-i_-NmyKiYWeJD834oP0FBLcQMLZ_GDAMYTqmoKjzrt-keeWqCPu4BU4lY3DVL7NWWR9-AQAAWKE

Galán Feced, C. (2023). *Alemania debate si seguir los pasos de Italia y prohibir ChatGPT: "Sería concebible"*. Business Insider España. Disponible en: https://www.businessinsider.es/alemania-debate-prohibir-chatgpt-1226000

Gobierno de España. (2019). *Estrategia Nacional de Inteligencia Artificial*. Disponible en:

https://www.ciencia.gob.es/stfls/MICINN/Ayudas/Convocatorias/Documents/20191030_EstrategiaNacionalInteligenciaArtificial.pdf

Laplante P. , Milojicic D. , Serebryakov S. y Bennett. D (2020). IA y Sistemas Críticos: De la Hype a la Realidad. Computer, 53(11), 45-52. Disponible en: https://www.osti.gov/servlets/purl/1713282

Marín Moreno, F. (2020). *¿Sandbox en el sector legal?* Diario La Ley, Nº 38, Sección Ciber derecho, Wolters Kluwers. Disponible en: https://laleydigital.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAEAC2OQU_DMAyFfw2-TEIpG-t2yKXrcUIICnc3sbpIIYHEKe2_x1mJ9Cl28t6zfwqldaCFdcYwxmWXaCoeOSYXdw4hryGG9UsPqRAwjlk3Sj20ptIIe-EgPAtHoa3vtTkJ59rclUdAwwV9H41uau1mGnCsYRCTpdStWgFHRv9GMgPyLf6-4OwmZBdDh2nbwFmr-0HJ2av2_HSCmVIWgf50EwUmyITJ3F5xIn11mVHSime0MT9i_l7A5Y_gqgn9-1265W62rjBL1shh-wPj5e6R6YKegv3f4g8sydRQNAEAAA==WKE#I11

Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J., & Perrault, R. (2023). *The AI Index 2023 Annual Report*. Stanford University, pp 14 y 288. Disponible en: https://aiindex.stanford.edu/report/

Ministerio de Asuntos Económicos y Transformación Digital. (2021). *Carta de Derechos Digitales*. Disponible en: https://www.mincotur.gob.es/es-es/GabinetePrensa/NotasPrensa/2021/Documents/20210127_Carta_Derechos_Digitales.pdf

Ministerio de Economía y Empresa de España. (2022). *El Gobierno de España presenta, en colaboración con la Comisión Europea, el primer piloto del Sandbox de regulación de IA en la UE*. Disponible en: https://portal.mineco.gob.es/es-es/comunicacion/Paginas/20220627-PR_AI_Sandbox.aspx

Organización Mundial de Propiedad Intelectual (OMPI). (2016). Principios básicos del derecho de autor y los derechos conexos. Disponible en: https://www.wipo.int/edocs/pubdocs/es/wipo_pub_909_2016.pdf

Orrico, F. J. F. (2021). *Criterios sobre uso de dispositivos tecnológicos en el ámbito laboral (Enfoque Laboral)* (1.ª ed.) pp 107-113. Tirant Lo Blanch.

Palma Ortigosa, A. (2019). *Decisiones automatizadas en el RGPD. El uso de algoritmos en el contexto de la protección de datos*. Revista General de Derecho Administrativo, 50. Iustel. Disponible en: https://laadministracionaldia.inap.es/noticia.asp?id=1509629

Parlamento Europeo (2023). *Inteligencia Artificial de uso general*. Disponible en:

https://epthinktank.eu/2023/03/31/general-purpose-artificial-intelligence/

Parlamento Europeo (2022). *Identificación y evaluación de la legislación comunitaria vigente y en preparación en el ámbito digital*. Disponible en: https://www.europarl.europa.eu/RegData/etudes/STUD/2022/703345/IPOL_STU(2022)703345_EN.pdf

Parlamento Europeo. (2021). *Reconocimiento biométrico y detección del comportamiento: Evaluación de los aspectos éticos de las técnicas de reconocimiento biométrico y detección del comportamiento, centrándose en su uso actual y futuro en espacios públicos.* Disponible en: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/696968/IPOL_STU(2021)696968_EN.pdf

Pina, C. (2022). *Inteligencia Artificial (IA): así es la Propuesta de Directiva para adaptar las normas de responsabilidad extracontractual*. Garrigues Digital. Disponible en: https://www.garrigues.com/es_ES/garrigues-digital/inteligencia-artificial-ia-asi-es-propuesta-directiva-adaptar-normas

S.J. y Norvig, P. (2008). *Inteligencia Artificial:Un Enfoque Moderno*, Las seis disciplinas que abarcan la mayor parte de la IA (2 ª edición, p.2.). Pearson, pp 2-3.

Veridas (2022). Informe ejecutivo: Datos biométricos y sistemas biométricos: marco jurídico en España y en Europa, p 7. Disponible en: https://das-gate.com/wp-content/uploads/2022/04/dasGate-Datos-biometricos-sistemas-biometricos.pdf

Vestri, G. (2022). *La disrupción tecnológica en la administración pública. Retos y desafíos de la IA.* ARANZADI / CIVITAS 38-43.