

SISTEMAS DE INTELIGENCIA ARTIFICIAL PARA LA TOMA AUTOMATIZADA DE DECISIONES: ALGUNOS PROBLEMAS DE LOS CONTRATOS ALGORÍTMICOS CONCLUIDOS CON ASISTENTE DIGITAL

Artificial Intelligence systems for automated decision-making: Some issues with algorithmic contracts concluded with digital assistants

M^a NIEVES PACHECO JIMÉNEZ¹

marianieves.pacheco@uclm.es

Profesora Titular de Derecho Civil

Universidad de Castilla-La Mancha

Recibido: 05.06.2024

Cómo citar / Citation

Pacheco Jiménez, M^a N. (2024).

Sistemas de inteligencia artificial para la toma automatizada de decisiones: algunos problemas de los contratos algorítmicos concluidos con asistente digital.

Cuadernos de Derecho Privado, 9, pp. 140-171

DOI: <https://doi.org/10.62158/cdp.64>

(Recepción: 05/06/2024; aceptación: 02/09/2024; publicación: 02/09/2024)

Resumen

Es innegable el potencial de la Inteligencia Artificial (IA), que permite la automatización de procesos y la toma de decisiones basada en algoritmos avanzados, con importantes ventajas de precisión y personalización, pero también con riesgos en la transparencia y en la privacidad, así como en la perpetuación de sesgos discriminatorios. Uno de los escenarios donde se muestra el auge de la IA es en el de las decisiones automatizadas en la ejecución de contratos de consumo a través de asistentes digitales, que, *a priori*, mejoran la eficiencia en la gestión de la relación contractual. En el primer semestre de 2024 han proliferado recomendaciones y regulaciones sobre IA, manifestando el gran desafío al que se enfrentan todos los agentes implicados para generar un marco de referencia adecuado para el óptimo y seguro desarrollo de la IA.

Palabras clave

Inteligencia artificial; toma de decisiones automatizada; contratos de consumo; garantías; protección de datos.

¹ ORCID ID: <https://orcid.org/0000-0002-9062-2342>

Trabajo realizado en el marco del Proyecto de I+D+i PID2021-128913NB-I00, titulado “Protección de consumidores y riesgo de exclusión social: seguimiento y avances”, financiado/a por MICIU/AEI/10.13039/501100011033/ y “FEDER Una manera de hacer Europa”, dirigido por Ángel Carrasco Perera y Encarna Cordero Lobato; en el marco del Proyecto de Investigación SBPLY/23/180225/000242 “El reto de la sostenibilidad en la cadena de suministros y la defensa del consumidor final”, cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER), para el Programa Operativo de Castilla-La Mancha 2021-2027, dirigido por Ángel Carrasco Perera y Ana Carretero García; y en el marco de las Ayudas para la realización de proyectos de investigación aplicada, dentro del Plan Propio de investigación UCLM, cofinanciadas en un 85% por el FEDER, para el proyecto titulado “Modelos jurídicos eficientes de consumo sostenible”, con Ref.: 2022-GRIN-34487 dirigido por Ángel Carrasco Perera y Ana I. Mendoza Losana.

Abstract

The potential of Artificial Intelligence (AI) is undeniable, enabling the automation of processes and decision-making based on advanced algorithms, with significant advantages in terms of precision and personalization. However, it also presents risks concerning transparency and privacy, as well as the perpetuation of discriminatory biases. One of the scenarios where the rise of AI is evident is in automated decision-making (ADM) in the execution of consumer contracts through digital assistants, which, in principle, improve the efficiency of managing contractual relationships. In the first half of 2024, recommendations and regulations on AI have proliferated, highlighting the great challenge faced by all stakeholders in creating an adequate framework for the optimal and safe development of AI.

Keywords

Artificial intelligence; automated decision-making (ADM); consumer contracts; guarantees; data protection.

SUMARIO:

I. *CAN MACHINES THINK? APROXIMACIÓN A LA INTELIGENCIA ARTIFICIAL.* I.1. *Can machines think?* I.2. Elementos clave y regulación de la inteligencia artificial. II. *SISTEMAS DE DECISIONES AUTOMATIZADAS.* II.1. Auge de los sistemas de decisiones automatizadas. II.2. Fases del sistema de decisiones automatizadas. III. *SISTEMAS DE DECISIONES AUTOMATIZADAS EN LA EJECUCIÓN DE CONTRATOS DE CONSUMO A LA LUZ DE LOS PRINCIPIOS DEL INSTITUTO DE DERECHO EUROPEO.* III.1. Contextualización: automatización y asistentes digitales. III.2. Principios generales del Informe del Instituto de Derecho Europeo sobre el derecho del consumidor de la UE y la toma de decisiones automatizada. III.3. Directivas “listas para ADM”. IV. *GARANTÍAS FRENTE A LAS DECISIONES AUTOMATIZADAS Y REGLAMENTO DE PROTECCIÓN DE DATOS.* V. *CONSIDERACIONES FINALES. Bibliografía.*

I. CAN MACHINES THINK? APROXIMACIÓN A LA INTELIGENCIA ARTIFICIAL**I.1. *Can machines think?***

El primer acercamiento a la denominada Inteligencia Artificial (IA en lo sucesivo) es, para muchos autores, el estudio del matemático inglés Alan Turing en 1950 titulado “*Computing Machinery and Intelligence*”², que, curiosamente, comenzaba con la frase *can machines think*”³. Para medir la inteligencia de la máquina, el *test de Turing* consistía en evaluar la capacidad de aquella para mostrar un comportamiento inteligente similar al de un ser humano, de manera que este no pudiese distinguir que las respuestas las proporcionaba una

² Turing (1950).

³ *Ibidem*, 433.

máquina⁴. En el año 1956 se empleó por primera vez el término “inteligencia artificial”, cuando varios científicos sostuvieron la tesis de que una máquina podía llegar a desempeñar tareas pensadas tradicionalmente para las personas atendiendo al grado de inteligencia que suponía⁵; entre ellos, el profesor emérito de Stanford, John McCarthy, definiéndolo como “la ciencia y la ingeniería de crear máquinas inteligentes”⁶. A finales del siglo XX muchos componentes de la IA comenzaron a utilizarse en la industria, pero no es hasta hace tan solo unos años cuando ha marcado un punto de inflexión en nuestro día a día.

Así las cosas, la IA, que implica el uso de algoritmos⁷ avanzados y ordenadores para simular las capacidades humanas de razonamiento, aprendizaje, creatividad y aptitud de planear, ha revolucionado la forma de tomar decisiones en la sociedad actual. La IA procesa inmensas cantidades de datos, extrae patrones y conocimientos útiles, de manera que los sistemas de aprendizaje automático⁸ pueden realizar análisis complejos, identificar tendencias ocultas y proporcionar información con gran precisión.

⁴ El denominado “imitation game”. *Vid.* Turing (1950: 435-436).

⁵ Palma Ortigosa (2022.b: 30).

⁶ McCarthy (1956: 2)

⁷ Un algoritmo es una estructura de control finita, abstracta y compuesta para cumplir un propósito, resolver un problema, hacer predicciones o guiar la ejecución de una tarea, siendo capaz de aprender con experiencia de entrenamiento. *Vid.* Mittelstadt *et al* (2016: 2). Consta de tres partes: una entrada o *input* (datos sobre los que se aplican las instrucciones); procesamiento (con lo recibido en el *input*, el algoritmo elabora una serie de cálculos lógicos para resolver el problema); salida (*output*) o resultado obtenido. Un sistema de IA puede estructurarse sobre algoritmos cuyas reglas que procesan la información se presentan ya definidas (algoritmos deterministas), que son más interpretables pero que muestran dificultades en entornos complejos con escenarios diversos; o sobre algoritmos cuyas reglas no están totalmente marcadas y, por tanto, gozan de autonomía y realizan aprendizaje automático con cantidades ingentes de datos para procesar la información (algoritmos no deterministas), por lo que, teniendo en cuenta más variables, se adecúan a contextos diferentes. *Vid.* Palma Ortigosa (2022.b:32).

⁸ *Machine learning*, entendida como rama de la IA y la informática que se basa en el uso de datos y algoritmos para imitar la forma en la que aprenden los seres humanos, mejorando gradualmente su precisión. Así, mediante métodos estadísticos, extrae patrones desde los datos previamente procesados para crear estructuras y reglas de conocimiento cuya forma sintética obedece a “si aparece tal patrón entonces la decisión es”. El modelo planteado va siendo reentrenado con nuevos datos y nuevas fuentes de información, de modo que el modelo cada vez aprende más, siendo capaz de emitir sus propios juicios basándose en datos similares. La introducción de técnicas de aprendizaje automático permite mejorar el proceso de estimación de modelos y la mejora de la eficiencia de los procesos a través de la automatización. *Vid.* Semmler & Rose (2017: 86-87); Calvo, Guzmán y Ramos (2018: 6-11 y 23). En función de cómo esos algoritmos sean capaces de aprender del conjunto de datos, el aprendizaje automático podrá ser supervisado (el algoritmo se entrena con un conjunto de datos de entradas y salidas previamente etiquetados, con un resultado predictivo); no supervisado (los datos utilizados no son etiquetados previamente, sino que es el propio algoritmo el que trata de descubrir patrones y conexiones en el conjunto de datos, con un resultado descriptivo); o por refuerzo (al algoritmo no se le proporciona un conjunto de datos de entrenamiento, sino que el sistema aprende a través de la interacción con el entorno a través del modelo prueba-error). *Vid.* Palma Ortigosa (2022.b: 33); en sentido similar, *vid.* Gil Membrado (2024: 505-508).

Ha de añadirse que en los últimos años está cobrando relevancia un nuevo enfoque de aprendizaje automático, concretamente el aprendizaje profundo o *deep learning*, inspirado en el comportamiento de las neuronas del cerebro humano a través de múltiples capas de procesamiento. Este sistema logra potencia y flexibilidad al aprender por cuenta propia (requiriéndose menos intervención humana) a representar el mundo como una jerarquía anidada de conceptos y representaciones, donde cada concepto se define en relación con conceptos más simples, y donde las representaciones más abstractas se calculan en función de las menos abstractas. *Vid.* Bengio, Goodfellow & Courville (2017: 8).

Lo cierto es que la IA ha encontrado su momento de eclosión gracias a diversos factores, tales como: la revolución digital y la interconexión derivada de ella; la posibilidad de almacenar grandes cantidades de datos en la nube, que posteriormente se utilizan para entrenar los sistemas de IA; el *Big Data*, como herramienta de análisis masivo de datos, que permite a la IA un preprocesamiento de grandes cantidades de datos (volumen), con diferentes formatos y estructuras (variedad) y de forma muy rápida (velocidad)⁹.

La IA posibilita la automatización de procesos y la toma de decisiones basada en algoritmos avanzados, con evidentes beneficios, entre ellos: precisión mejorada, al examinar grandes cantidades de datos con rapidez y meticulosidad; análisis profundo, al identificar patrones ocultos en conjuntos de datos complejos; velocidad y eficiencia ya que los algoritmos de IA procesan datos en tiempo real; personalización, en tanto en cuanto estos sistemas se adaptan a las preferencias y necesidades individuales; mejor detección de anomalías porque los sistemas de IA buscan patrones anormales y riesgos potenciales; gran capacidad de aprendizaje debido a que la IA puede instruirse de experiencias anteriores y mejorar atendiendo a ellas; y reducción de errores humanos puesto que la IA automatiza ciertos aspectos de la toma de decisiones¹⁰.

Pero también comporta importantes riesgos de diversa índole¹¹: en la transparencia, siendo decisiva la información facilitada para no generar asimetría informativa¹²; en la privacidad de las personas por el hecho de que se pueden elaborar perfiles detallados que invadan dicho ámbito personal reservado sin el consentimiento oportuno, o se empleen sistemas de vigilancia masiva (v. gr., reconocimiento biométrico); así como en la perpetuación de sesgos discriminatorios ya que los algoritmos predictivos pueden discriminar a personas según su género, origen étnico o nivel socioeconómico, entre otras variables.

En este punto es fundamental señalar la trascendencia de la elección de los algoritmos que se utilizarán para entrenar los datos; elección que corresponde a quienes diseñan el sistema. Entre los factores que influyen en ese diseño se encuentran: el tipo de datos presentes (si es posible o no su etiquetado para un aprendizaje supervisado); y el contexto donde se pretende que tendrá sus efectos el sistema de IA (necesitándose algoritmos más o menos complejos)¹³.

⁹ Palma Ortigosa (2022.a: 44).

¹⁰ Calvo, Guzmán y Ramos (2018: 13-29).

¹¹ Todos ellos alineados con los retos de conseguir confianza y seguridad para las personas usuarias, máxime en el ámbito contractual. *Vid.* Sáenz de Jubera Higuero (2023: 50-51).

¹² *Ibidem*, 50.

¹³ Palma Ortigosa (2022.b: 40).

Retomando aquí los sesgos en los algoritmos, estos pueden ser conscientes o inconscientes. Los primeros se producen intencionadamente por la persona¹⁴ que selecciona los datos y desarrolla el algoritmo; los segundos se originan sin que quien diseña el algoritmo tenga intención de introducirlos, pudiendo concernir: 1) al contexto cultural geográfico y cultural (los datos recogidos no representan con precisión el entorno en el que ha de ejecutarse el programa, produciéndose un sesgo de muestra), o el diseño elimina algunas características pensando que son irrelevantes basándose en creencias preexistentes (sesgo de exclusión); 2) al perfil de la persona que desarrolla los algoritmos, que puede llegar a ellos con prejuicios (sesgo de prejuicios); 3) a la recogida de propios datos, que puede verse alterada o distorsionada por el propio instrumento de recogida de datos (sesgos de medición y recogida de datos)¹⁵. Sea como sea, y a pesar de la apariencia de objetividad de estos sistemas, debe tenerse presente que son diseñados y desarrollados en un contexto social basado en estructuras de desigualdad¹⁶.

I.2. Elementos clave y regulación de la inteligencia artificial

Ante los posibles errores y consecuencias negativas que se han reflejado en líneas anteriores, la Comisión Europea presentó en abril de 2019 las líneas maestras para el desarrollo de una IA fiable y segura¹⁷, a saber: necesidad de supervisión humana; robustez y seguridad; privacidad y gobernanza de datos; transparencia; diversidad, no discriminación y equidad; bienestar social y ambiental; y rendición de cuentas.

Por su parte y también en 2019, el Consejo de la Organización para la Cooperación y el Desarrollo Económicos (OCDE¹⁸) suscribió el 22 de mayo los denominados *Principios de la OCDE sobre la IA*¹⁹, con el objetivo de fomentar la innovación y la confianza en la IA promoviendo una IA confiable, que garantizase los derechos humanos y los valores democráticos. Para ello se centraba en cuestiones políticas específicas de la IA, esforzándose

¹⁴ Aunque no es objeto de este estudio, sí quisiera señalar una preocupación quizás menos analizada, que no es otra que la de la labor humana que hay detrás de la IA, concretamente el ámbito del entrenamiento supervisado de algoritmos, donde los datos son etiquetados y tratados por seres humanos. Se trata del proceso de *microworking* o “microtrabajo”, como tipo de trabajo de plataforma, que contempla tareas de transcripción de archivos de sonido o de imágenes a texto, descripción de imágenes, etc. El trabajo de plataformas, extrapolado al ámbito de la labor humana tras la IA, plantea la cuestión de la necesaria regulación y protección de los y las trabajadoras promoviendo un trabajo digno.

¹⁵ Salazar García (2022: 60).

¹⁶ Soriano Arnanz (2021: 93).

¹⁷ European Commission (2019). Disponible en: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (22/03/2024)

¹⁸ OECD por sus siglas en inglés: Organisation for Economic Cooperation and Development.

¹⁹ Disponible en: [https://one.oecd.org/document/C/MIN\(2019\)3/FINAL/en/pdf](https://one.oecd.org/document/C/MIN(2019)3/FINAL/en/pdf)

por establecer un estándar que fuese aplicable y lo suficientemente flexible para resistir el paso del tiempo en un escenario en continua evolución²⁰.

Esta Recomendación de principios fue revisada por el Consejo de la OCDE en noviembre de 2023 para actualizar su definición de “sistema de IA” con el objeto de que fuese técnicamente precisa y acorde con los avances tecnológicos. En este sentido, la actualización clarificó los objetivos de un sistema de IA, incluyó su aplicación a sistemas de IA generativos (que producen contenido) y matizó que algunos sistemas de IA pueden seguir evolucionando después de su diseño e implementación²¹.

Su última modificación (*Recomendación del Consejo sobre Inteligencia Artificial*)²², más extensa, data de principios de mayo de 2024 (aprovechando la aprobación dos meses antes del Reglamento de Inteligencia Artificial²³ -en adelante RIA-), y pone el acento en las siguientes cuestiones²⁴: 1) Reflejar la creciente importancia de afrontar la desinformación y la información falsa, sobre todo atendiendo al contexto de la IA generativa. 2) Acometer usos (intencionales o no intencionales) fuera del propósito previsto. 3) Aclarar la información que los actores de la IA deben proporcionar sobre los sistemas para garantizar la transparencia y la divulgación responsable. 4) Abordar preocupaciones de seguridad en tanto en cuanto los sistemas de IA causen un daño indebido, pudiendo entonces ser anulados, reparados y/o desactivados de manera segura mediante la interacción humana. 5) Enfatizar la conducta empresarial responsable a lo largo del ciclo de vida de la IA. 6) Subrayar la necesidad de que las jurisdicciones trabajen juntas para promover entornos de gobernanza y políticas interoperables para la IA. 7) Introducir una referencia expresa a la sostenibilidad ambiental, teniendo en cuenta que su importancia ha aumentado desde la Recomendación de 2019. Todo ello fundamentándose en el enorme e incuestionable potencial que tiene la IA, pero también en sus efectos dispares en relación con los cambios económicos, la competencia, el mercado laboral, las desigualdades, los derechos humanos, la protección de datos y la seguridad digital²⁵.

²⁰ Vid. OECD (2019).

²¹ Estas divergencias en los posicionamientos de cada institución comunitaria sobre la definición de IA se justifican por la importancia de su concreción de cara a establecer el ámbito de aplicación del RIA. Vid. Sáenz de Jubera Higuero (2023: 46).

²² Vid. OECD (2024). Disponible en: <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449?s=03#backgroundInformation>

²³ Resolución legislativa del Parlamento Europeo, de 13 de marzo de 2024, sobre la propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial y se modifican determinados actos legislativos de la Unión (COM (2021)0206 – C9-0146/2021 – 2021/0106 (COD)). Disponible en: https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_ES.pdf

²⁴ OECD (2024).

²⁵ Y en aras a esa consecución de una IA fiable, la Recomendación aconseja una serie de principios de administración responsable: 1) “Crecimiento inclusivo, desarrollo sostenible y bienestar”, con la finalidad de buscar resultados beneficiosos para las personas y el planeta. 2) “Respetar el Estado de Derecho, los derechos

Asimismo, se acogen las siguientes definiciones para comprender el sentido de las recomendaciones que posteriormente se realizan: A) “Sistema de IA”, como “sistema basado en máquinas que, para objetivos explícitos o implícitos, infiere, a partir de los datos de entrada que recibe, cómo generar información de salida como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos reales o virtuales; y presentando diversos niveles de autonomía y capacidad de adaptación”. B) “Ciclo de vida del sistema de IA”, que “comprenden generalmente varias fases: planificación y diseño; recopilación y tratamiento de datos; creación de modelo(s) y/ adaptación de modelo(s) existente(s) a tareas específicas; prueba, evaluación, verificación y validación; entrada en servicio/despliegue; explotación y supervisión; y retirada/desmantelamiento”. C) “Actores de la IA”, quienes “desempeñan un papel activo en el ciclo de vida del sistema de IA, como pueden ser entidades y personas que despliegan y explotan la IA”²⁶.

Coincidiendo con el desarrollo experimentado en los últimos meses del marco de referencia de la IA, se aprueba en España el pasado 14 de mayo por el Consejo de Ministros, a propuesta del Ministerio para la Transformación Digital y de la Función Pública, la *Estrategia de Inteligencia Artificial 2024*²⁷. Su propio Prólogo señala que la Estrategia “persigue acelerar, facilitar, acompañar y promover el desarrollo y la expansión de la IA en nuestro país”; con el firme “compromiso de garantizar que esta transición se lleve a cabo respetando los derechos humanos, promoviendo la seguridad tecnológica, la inclusión social y económica, y observando el cumplimiento del Reglamento Europeo de IA”. Su despliegue se realizará en 2024 y 2025, y se estructura en tres ejes: Eje 1) Refuerzo de las capacidades para el desarrollo de la IA, atendiendo al soporte de la supercomputación, la capacidad de almacenamiento sostenible, los modelos de lenguaje en castellano y lenguas cooficiales, y la necesidad de talento especializado en IA a través de creación de redes y búsqueda de sinergias entre proyectos²⁸. Eje 2) Facilitar la expansión de la IA en el sector público y privado fomentando la innovación y la

humanos y los valores democráticos, incluidas la equidad y la privacidad”, como obligación para los actores de la IA a lo largo del ciclo de vida de los sistemas de IA. 3) “Transparencia y explicabilidad”, debiendo los actores proporcionar información significativa, transparente y comprensible “sobre las fuentes de datos, los factores, los procesos y/o el razonamiento que subyace a las predicciones, contenidos, recomendaciones y decisiones”. 4) “Solidez, seguridad y protección” de los sistemas de IA para que “funcionen adecuadamente y no planteen riesgos excesivos en materia de seguridad y/o protección”. 5) “Responsabilidad” por parte de los actores de la IA en su correcto funcionamiento y en el respeto de los principios enumerados. *Vid.* OECD (2024).

²⁶ *Ibidem*.

²⁷ Ministerio para la Transformación Digital y de la Función Pública (2024). Disponible en: https://portal.mineco.gob.es/es-es/digitalizacionIA/Documents/Estrategia_IA_2024.pdf. Esta Estrategia es el resultado de la evolución de la publicada en 2020, que necesitaba un claro refuerzo y aceleramiento a tenor de la evolución tecnológica de los últimos años.

²⁸ Ministerio para la Transformación Digital y de la Función Pública (2024: 10-35).

ciberseguridad, complementándose con la futura Ley de Ciberseguridad, en tanto en cuanto esta es un componente intrínseco de la propia digitalización y del desarrollo de nuevas tecnologías²⁹. Eje 3) Desarrollar una IA transparente, responsable y humanística, con intervención de la Agencia Española de Supervisión de la IA³⁰ (AESIA)³¹.

Solo tres días después de la publicación de la antedicha Estrategia, el 17 de mayo, el Consejo de Europa adopta el *Convenio Marco sobre Inteligencia Artificial y Derechos Humanos, Democracia y Estado de Derecho*³² (que se abrirá a su firma el próximo mes de septiembre), como primer Tratado internacional jurídicamente vinculante que abarca todo el ciclo de vida de los sistemas de IA en aras de promover una innovación responsable y respetuosa de los derechos de las personas³³.

Su artículo 1 indica como objetivo del Convenio “asegurar que las actividades dentro del ciclo de vida de los sistemas de inteligencia artificial³⁴ sean totalmente coherentes con los derechos humanos, la democracia y el Estado de Derecho”. Para ello “cada Parte deberá adoptar o mantener las medidas legislativas, administrativas u otras apropiadas para dar efecto a las disposiciones establecidas en este Convenio”, debiendo graduarse y diferenciarse “teniendo en cuenta la gravedad y la probabilidad de que ocurran impactos adversos sobre los derechos humanos, la democracia y el Estado de Derecho a lo largo del ciclo de vida de los sistemas de IA”. Así, “para asegurar la implementación efectiva de sus disposiciones por las Partes, este Convenio establece un mecanismo de seguimiento y prevé la cooperación internacional”³⁵.

El Capítulo II del Convenio aborda las obligaciones generales de las Partes³⁶, a saber: protección de los derechos humanos, según lo consagrado en el Derecho Internacional y en el

²⁹ *Ibidem*, 36-51.

³⁰ Que actuará en una triple dirección: como centro de pensamiento y análisis de la IA, como supervisor de un despliegue responsable de la IA, y como referente internacional.

³¹ Ministerio para la Transformación Digital y de la Función Pública (2024: 52)

³² Consejo de Europa (2024). Disponible en:

https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680afb11f (20/05/2024)

³³ Este Convenio es el resultado de dos años de trabajo del Comité sobre Inteligencia Artificial, que reunió a los 46 Estados miembros del Consejo de Europa, la Unión Europea y 11 Estados no miembros (Argentina, Australia, Canadá, Costa Rica, Estados Unidos, Israel, Japón, México, Perú, la Santa Sede y Uruguay), así como a representantes, en calidad de observadores del sector privado, la sociedad civil y el mundo académico. *Vid.* <https://www.coe.int/es/web/portal/-/council-of-europe-adopts-first-international-treaty-on-artificial-intelligence?s=03> (20/05/2024)

³⁴ Según el art. 2 del Convenio, es aquel “sistema basado en máquinas que, con objetivos explícitos o implícitos, infiere, a partir de la entrada que recibe, cómo generar salidas tales como predicciones, contenido, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales”.

³⁵ Consejo de Europa (2024).

³⁶ Ha de reseñarse que las Partes del Convenio no estarán obligadas a aplicar las disposiciones de este Tratado a las actividades relacionadas con la protección de los intereses de la seguridad nacional, pero sí a garantizar que dichas actividades respetan el Derecho Internacional, los derechos humanos y las instituciones y procesos democráticos. En cuanto a la regulación del sector privado, las Partes pueden optar por obligarse directamente por las disposiciones pertinentes o, atendiendo a la heterogeneidad de los sistemas jurídicos nacionales, adoptar otras

Derecho interno (*ex art. 4*); integridad de los procesos democráticos y respeto al Estado de Derecho, evitando socavar la integridad de las instituciones e incluyendo el acceso justo de los individuos a la participación en el debate público, protegiendo a su vez la capacidad para formar opiniones libremente (*ex art. 5*); respeto de la dignidad humana y la autonomía individual (*ex art. 7*); implementación de requisitos de transparencia y supervisión adaptados a los contextos y riesgos específicos (*ex art. 8*); rendición de cuentas y responsabilidad por los impactos adversos en los derechos humanos, la democracia y el Estado de Derecho (*ex art. 9*); igualdad (incluida la igualdad de género) y no discriminación, estableciendo medidas dirigidas a superar las desigualdades y lograr resultados justos, equitativos e imparciales (*ex art. 10*); privacidad y protección de datos personales (*ex art. 11*); fiabilidad de los sistemas de IA y confianza en sus resultados (*ex art. 12*); innovación segura a través de entornos controlados (*sandboxes*) para probar y desarrollar sistemas de IA bajo la supervisión de sus autoridades competentes (*ex art. 13*).

Atendiendo a la evolución cronológica, en este punto resulta imprescindible mencionar la reciente publicación en el Diario Oficial de la Unión Europea (12 de julio de 2024) del Reglamento de Inteligencia Artificial³⁷, tras años de arduas negociaciones y modificaciones de la Propuesta planteada en 2021 por la Comisión Europea. La norma entrará en vigor el 1 de agosto, siendo plenamente aplicable 24 meses después, exceptuando algunas de sus obligaciones, que irán cobrando vigencia paulatinamente: prohibiciones de prácticas prohibidas (6 meses después); códigos de práctica (9 meses después); normas de IA de uso general, incluida la gobernanza (12 meses después); obligaciones para sistemas de alto riesgo (36 meses después).

El objetivo principal del RIA, atendiendo a su Considerando 1, es “mejorar el funcionamiento del mercado interior mediante el establecimiento de un marco jurídico uniforme, en particular para el desarrollo, la introducción en el mercado, la puesta en servicio y la utilización de sistemas de IA³⁸ en la Unión, de conformidad con los valores de la Unión, a

medidas que permitan cumplir con las disposiciones del Tratado a la vez que se respetan plenamente sus obligaciones en materia de derechos humanos, democracia y Estado de Derecho (*ex arts. 23-25*). Además, para velar por el adecuado cumplimiento del Convenio, cada Parte establecerá un mecanismo de supervisión independiente (*ex art. 26*).

³⁷ Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) núm. 300/2008, (UE), núm.167/2013, (UE) núm. 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial). Disponible en <https://www.boe.es/doue/2024/1689/L00001-00144.pdf>

³⁸ Sistema de IA *ex art. 3* RIA: “un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos

fin de promover la adopción de una IA centrada en el ser humano y fiable, garantizando al mismo tiempo un elevado nivel de protección de la salud, la seguridad y los derechos fundamentales consagrados en la Carta de los Derechos Fundamentales de la Unión Europea”.

Teniendo como objetivos específicos, basados en la implementación de un marco jurídico armonizado para la IA, los que prosiguen: a) garantizar que los sistemas de IA comercializados en la UE sean seguros y respeten la legislación de la UE vigente; b) garantizar la seguridad jurídica para facilitar la inversión y la innovación en la IA; c) mejorar la gobernanza y la aplicación efectiva de la legislación de la UE sobre derechos fundamentales y requisitos de seguridad aplicables a los sistemas de IA; d) facilitar el desarrollo de un mercado único de aplicaciones de IA lícitas, seguras y fiables, y evitar la fragmentación del mercado.

El RIA es consciente de las repercusiones que la inteligencia artificial puede tener en la sociedad; de ahí la necesidad de generar confianza y de que su marco reglamentario sea acorde con los valores de la Unión consagrados en el artículo 2 del Tratado de la Unión Europea, los derechos y libertades fundamentales preceptuados en los diferentes Tratados (*ex* Considerando 6). Por los derechos e intereses en juego, la Comisión consideró que la regulación debía seguir un enfoque basado en el riesgo³⁹.

Así pues, el RIA define cuatro niveles de riesgo para los sistemas de IA, esto es: riesgo mínimo o nulo⁴⁰, riesgo limitado⁴¹, alto riesgo⁴² y riesgo inaceptable. Este último tipo quedaría terminantemente prohibido (*ex* art. 5 Capítulo II) porque supondría riesgos inadmisibles, en

explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales”.

³⁹ Muñoz García (2023: 55-56).

⁴⁰ Se trataría de aquellos irrelevantes para el Derecho por entender que no generan riesgos (*v. gr.*, aplicaciones como videojuegos habilitados para IA o filtros de *spam*). Para estos riesgos se exigen requisitos de documentación, información y registro, trazabilidad y explicabilidad, además de gobernanza de datos en función de la probabilidad del riesgo. No obstante, algún autor aperece de que ningún sistema de IA puede entenderse “inocente”, aconsejando la inclusión de alguna cláusula genérica que recomiende la supervisión humana de los resultados generados por el sistema ya que puede contener errores significativos. *Vid.* Vestri (2024).

⁴¹ Este riesgo estaría asociado a la falta de transparencia, en la medida en que sus usuarios/as deben estar debidamente informados/as fomentado así su confianza. Sería el caso de *chatbots* o asistentes digitales. Es reseñable el caso de la IA generativa, productora de contenido, en el sentido de que sistemas *a priori* irrelevantes podrían crear contenido creíble y vulnerar derechos de las personas. (*Vid.* Muñoz García (2023: 61).

⁴² Los sistemas de IA identificados como de alto riesgo (*ex* art. 6 Capítulo II) incluyen la tecnología de IA utilizada en: a) infraestructuras críticas (por ejemplo, el transporte), que podrían poner en peligro la vida y la salud de la ciudadanía; b) formación educativa o profesional, que puede determinar el acceso a la educación y el curso profesional de la vida de una persona (por ejemplo, puntuación de los exámenes); c) componentes de seguridad de los productos (por ejemplo, aplicación de IA en cirugía asistida por robot); d) empleo, gestión de trabajadores y acceso al trabajo por cuenta propia (por ejemplo, software de clasificación de CV para procedimientos de contratación); e) servicios públicos y privados esenciales (por ejemplo, calificación crediticia que dificulta a la ciudadanía la obtención de un préstamo); f) aplicación de la ley que pueda interferir con los derechos fundamentales de las personas (por ejemplo, evaluación de la fiabilidad de las pruebas); g) gestión de la migración, el asilo y el control fronterizo (por ejemplo, examen automatizado de las solicitudes de visado); h) administración de justicia y procesos democráticos (por ejemplo, soluciones de IA para buscar resoluciones judiciales). (*Vid.* <https://digital-strategy.ec.europa.eu/es/policies/regulatory-framework-ai> -15/05/2024-).

tanto en cuanto se trataría de sistemas dirigidos a implantar ideas en los seres humanos, de manera subliminal o aprovechándose de condiciones de especial vulnerabilidad, y que alteren de forma sustancial su comportamiento o sean susceptibles de provocarles perjuicios físicos o psicológicos; o porque pudieran ser empleados por las autoridades con el fin de perfilar a los seres humanos atendiendo a su conducta social o a características personales o de su personalidad (conocidas o predichas) cuando ello cause un trato perjudicial o desfavorable a determinados grupos de personas o colectivos; o porque se usen sistemas de identificación biométrica⁴³ remota⁴⁴ en tiempo real⁴⁵ en espacios de acceso público (salvo en situaciones excepcionales relacionadas con la prevención o persecución de actividades delictivas, estando sujetas a la autorización de un órgano judicial u otro organismo independiente). Este último supuesto, de identificación biométrica, es especialmente preocupante por la posible vulneración de derechos fundamentales.

Evidentemente, la protección debe ser más acusada cuando se trate de sistemas de IA de alto riesgo por la magnitud de las consecuencias adversas para los derechos fundamentales. Entre dichos derechos “se incluyen el derecho a la dignidad humana, el respeto de la vida privada y familiar, la protección de datos de carácter personal, la libertad de expresión y de información, la libertad de reunión y de asociación, la no discriminación, el derecho a la educación, la protección de los consumidores, los derechos de los trabajadores, los derechos de las personas discapacitadas, la igualdad entre hombres y mujeres, los derechos de propiedad intelectual, el derecho a la tutela judicial efectiva y a un juez imparcial, los derechos de la defensa y la presunción de inocencia, y el derecho a una buena administración” (*ex* Considerando 48). De ahí la necesidad de implementar medidas de gestión de riesgos adecuadas y específicas que eliminen o mitiguen los riesgos detectados y evaluados, o que se implanten medidas de reducción y control frente a los riesgos que no puedan eliminarse (*ex* art. 9 Capítulo II); sin olvidar medidas *ex post* por parte de los proveedores e implementadores de sistemas de IA de alto riesgo. Cobra especial importancia aquí el requisito de documentación técnica⁴⁶ que

⁴³ Entendida como “el reconocimiento automatizado de características humanas de tipo físico, fisiológico, conductual o psicológico para determinar la identidad de una persona física comparando sus datos biométricos con los datos biométricos de personas almacenados en una base de datos” (*ex* Considerando 35).

⁴⁴ A través de la cual un sistema de IA identifica “a las personas físicas sin su participación activa y generalmente a distancia comparando sus datos biométricos con los que figuran en una base de datos de referencia” (*ex* Considerando 41).

⁴⁵ Aquí “la recogida de datos biométricos, la comparación y la identificación se producen sin una demora significativa” (*ex* Considerando 42).

⁴⁶ Regulada en el Anexo IV del RIA, debiendo incluir, entre otros: descripción detallada del sistema de IA y de su proceso de desarrollo; información acerca de la supervisión, el funcionamiento y el control del sistema de IA (atendiendo a fuentes de riesgo para la salud y la seguridad, los derechos fundamentales y la discriminación); descripción pormenorizada del sistema de gestión de riesgos; descripción de los cambios pertinentes realizados

ha de elaborarse para un sistema de IA de alto riesgo antes de su introducción en el mercado o puesta en servicio.

El RIA prevé un sistema de sanciones si se produce la vulneración de las disposiciones previstas en él (*ex arts. 99 y siguientes*), pudiendo variar según la gravedad y la naturaleza de la infracción, y yendo desde meras amonestaciones por parte de las autoridades nacionales competentes a multas comprendidas entre 7,5 y 35 millones de euros⁴⁷.

De la lectura sosegada de todo este abordaje sistemático (nótese la proliferación de recomendaciones y regulaciones durante el pasado mes de mayo, solo un par de meses después de la aprobación del RIA), se infiere el gran desafío al que se enfrentan todos los agentes implicados en la generación de un marco de referencia adecuado para el óptimo y seguro desarrollo de la IA, presentando las mayores controversias aquellas cuestiones de responsabilidad para con los derechos humanos y de uso malicioso o doloso.

II. SISTEMAS DE DECISIONES AUTOMATIZADAS

II.1. Auge de los sistemas de decisiones automatizadas

Tras una necesaria contextualización y revisión del trazado regulatorio de los sistemas de IA, han de plantearse los motivos por los que el uso de sistemas automatizados basados en tecnologías de IA está experimentando su auge actualmente. Las más relevantes serían: la mayor disponibilidad de datos; el desarrollo exponencial del procesamiento de datos; y la mejora del rendimiento de los sistemas computacionales⁴⁸.

La mayor disponibilidad de datos está estrechamente relacionada con la revolución digital y, más concretamente, con el Internet de las cosas (*IoT* o *Internet of Things*). Se trataría de sistemas de dispositivos físicos que reciben y transfieren datos a través de redes inalámbricas, con poca o nula intervención humana, ya que esos dispositivos están debidamente integrados en todo tipo de objetos. Los datos, al ser generados en línea, permiten a las diversas organizaciones recopilarlos y procesarlos, ofreciendo servicios y publicidad personalizados⁴⁹. Esa disponibilidad masiva de datos está inexorablemente conectada a su tratamiento y

por el proveedor en el sistema a lo largo de su ciclo de vida; un listado de normas armonizadas, aplicadas total o parcialmente; una copia de la declaración de la UE de conformidad.

⁴⁷ Multas de hasta 35 millones de euros o el 7% del volumen de negocios total a nivel mundial del ejercicio financiero anterior para vulneraciones de prácticas prohibidas o protección de datos; multas de hasta 15 millones de euros o el 3% del volumen de negocios anual total a nivel mundial del ejercicio financiero anterior en caso de incumplimiento de cualquier otra obligación del RIA; multas de hasta 7,5 millones de euros o el 1,5% del volumen de negocios anual total a nivel mundial del ejercicio financiero anterior si se proporciona información inexacta, incompleto o engañosa a organismos y autoridades nacionales competentes.

⁴⁸ Palma Ortigosa (2022.a: 45). En sentido similar: Coglianese & Lehr (2017: 1162); Soriano Arnanz (2021: 90).

⁴⁹ Palma Ortigosa (2022.a: 45).

almacenamiento, realizados en muchas ocasiones por los denominados *data centers* (de naturaleza pública o privada), gestionando los sistemas de información de datos. Asimismo, el almacenamiento de estos datos en entornos virtuales o nubes (*clouds*) supone un abaratamiento de costes⁵⁰. Por último, la mejora en el rendimiento (v. gr., velocidad de proceso y capacidad de almacenamiento) de las computadoras que procesan esas cantidades ingentes de datos permite que se puedan analizar escenarios cambiantes y complejos con precisión adecuada⁵¹.

Este contexto ha propiciado que técnicas como *Big Data* o IA puedan desarrollarse óptimamente, mostrando todo su potencial, y estando entre sus funcionalidades la toma de decisiones automatizadas, con unos claros beneficios de eficiencia y ahorro de recursos. De hecho, las empresas y los poderes públicos emplean cada vez más herramientas de toma automatizada de decisiones, en ámbitos tales como la salud, las resoluciones jurídicas, la educación, los servicios financieros, el comercio, el mercado laboral, la publicidad o la contratación⁵². En estos sectores las máquinas vienen a sustituir tareas que tradicionalmente eran desempeñadas por personas. Como contrapunto, el uso de algoritmos a la hora de adoptar decisiones conlleva problemas tales como discriminación, decisiones injustas, privacidad o denegaciones de un servicio o producto por parte de una máquina⁵³. Y ello porque las decisiones tomadas por los algoritmos basados en correlaciones se consideran suficientemente fiables, aun sin prueba de causalidad, generándose el riesgo de inferencias erróneas⁵⁴.

II.2. Fases del sistema de decisiones automatizadas

La expresión “decisiones automatizadas” procede del anglicismo *Automated Decision-Making* (en adelante ADM), significando un proceso computacional, incluidas técnicas y enfoques de IA, que, alimentado por entradas y datos recibidos o recogidos del entorno, pueden generar, a partir de un conjunto de objetivos predefinidos, resultados en una amplia variedad de formas (v. gr, contenidos, valoraciones, recomendaciones, decisiones, predicciones...), logrando, pues, la automatización de la toma de decisiones, que, de otro modo, sería llevada a cabo por humanos.

Para que esas decisiones automatizadas acaben aplicándose en la vida real, que es el objetivo de las organizaciones que implementan este tipo de sistemas, es preciso que se superen las fases de diseño y de despliegue, tal y como se exponen a continuación.

⁵⁰ *Ibidem*, 46.

⁵¹ *Ibidem*.

⁵² Roig (2020: 17).

⁵³ Roberto Granero (2022: 108).

⁵⁴ Soriano Arnanz (2021: 91).

La fase de diseño, que es fundamental porque influirá en el ulterior despliegue, consiste en elaborar un modelo que posteriormente será implantado en un sistema de toma de decisiones. Esta etapa abarca múltiples subfases, a saber: A) Planificación del proyecto por parte de los desarrolladores, fijando los objetivos que se pretenden con el sistema. B) Recopilación y obtención de datos representativos de calidad, eliminando los datos “sucios” porque su utilización generaría decisiones inadecuadas con respecto a la realidad donde han de adoptarse dichas decisiones. C) Preprocesamiento de datos, como tarea esencial, cuyo objetivo es depurar y transformar los datos a un formato reconocible al algoritmo y que le permita trabajar adecuadamente para alcanzar el objetivo previsto, siendo de gran utilidad aquí las técnicas basadas en *Big Data*. D) Desarrollo del modelo a través del entrenamiento del algoritmo que cuenta con los datos previamente procesados. E) Prueba del modelo mediante la comprobación del rendimiento de los distintos modelos creados. F) Elección y validación del modelo, teniendo en cuenta los parámetros de interpretabilidad⁵⁵ y exactitud para optar por un modelo u otro⁵⁶.

Por su parte, la fase de despliegue, basada en la integración e implementación de los sistemas en el proceso organizativo y su puesta en marcha en la realidad donde se adoptarán las decisiones, cuenta con las subfases que prosiguen: A) Puesta en marcha del sistema de IA en el contexto específico donde va a desplegar sus efectos, con la recomendación de realización de pruebas de forma local y limitadas a un grupo de personas. B) Procesamiento de datos, que son transformados en un resultado y la valoración de si se convertirá inmediatamente en una decisión plenamente automatizada o formará parte de otra serie de factores que fundamentarán la decisión final (decisión parcialmente automatizada). C) Evolución del modelo desde el momento en que comienza a interactuar con el entorno para el que fue diseñado, teniendo en cuenta si estamos ante sistemas dinámicos (más prácticos en entornos cambiantes que requieren una continua adaptación porque aprovechan los datos de entrada para mejorar y ajustarse) o estáticos (aconsejables en escenarios no excesivamente cambiantes ya que dejan de aprender una vez que han sido diseñados y su programación no se altera)⁵⁷.

III. SISTEMAS DE DECISIONES AUTOMATIZADAS EN LA EJECUCIÓN DE CONTRATOS DE CONSUMO A LA LUZ DE LOS PRINCIPIOS DEL INSTITUTO DE DERECHO EUROPEO

⁵⁵ Aunque va existiendo un mayor interés en la creación de sistemas interpretables, hoy por hoy se prefieren los modelos que puedan generar predicciones o resultados más precisos.

⁵⁶ Palma Ortigosa (2022.a: 48-52).

⁵⁷ *Ibidem*, 52-57.

III.1. Contextualización: automatización y asistentes digitales

Como se ha señalado en líneas anteriores, uno de los ámbitos donde se viene observando una creciente adopción de decisiones automatizadas es el de la ejecución de contratos de consumo. La IA, gracias a su capacidad predictiva derivada del análisis de datos, puede realizar diversas tareas en la negociación, formación y celebración de un contrato. Reflejo de la automatización en este contexto son las obligaciones de pago a través de débitos directos, la suscripción para la entrega de bienes a intervalos determinados, la renovación automática de los contratos al final del período contractual acordado, o la utilización de motores de búsqueda para comparar precios. Sin embargo, existen otros escenarios donde la automatización es la de la propia decisión de celebrar un contrato; estaríamos hablando de la ejecución a través de un asistente digital⁵⁸. Así las cosas, la automatización de la toma de decisiones se basaría en el uso de algoritmos y sistemas informáticos para analizar datos y adoptar determinaciones en nombre de una de las partes contratantes, sin intervención humana directa.

Estos asistentes digitales⁵⁹ son sistemas de IA que utilizan algoritmos y técnicas de aprendizaje automático para procesar información y generar respuestas adecuadas en función de criterios predefinidos. Por tanto, dependen en gran medida de la programación del inicial algoritmo, los datos de entrenamiento y cualquier capacidad de “autoaprendizaje” codificada en aquel. En el concreto ámbito de contratos de consumo, la finalidad del asistente digital es mejorar la eficiencia y la precisión en la gestión de la relación contractual entre sujeto proveedor y consumidor/a encaminándose a la negociación, conclusión o ejecución de un contrato. Este, en consecuencia, será un contrato algorítmico, entendido como aquel en el que una o ambas partes utilizan un asistente digital para llevar a cabo uno más aspectos del proceso contractual, delegando, por tanto, el uso de su libertad contractual⁶⁰.

El funcionamiento del asistente digital podría resumirse como prosigue: Recopila datos (procedentes de registros de transacciones pasadas, preferencias, historial crediticio, etc.) relevantes sobre el consumidor o la consumidora, el producto o servicios objeto de contrato y

⁵⁸ Algunos/as autores/as manifiestan su reticencia a la automatización de la negociación en contratos complejos, donde dicha automatización debería configurarse más bien como instrumento de apoyo. *Vid.* Navas Navarro (2024: 65).

⁵⁹ *A priori*, y según el RIA, encajarían en la clasificación de riesgo limitado, poniéndose el énfasis en la transparencia en el uso del sistema de IA si se quiere fomentar la confianza de sus usuarios/as.

⁶⁰ En sede de contratos algorítmicos es preciso matizar que los sistemas de inteligencia artificial predeterminada (ofrecen siempre el mismo resultado para unos mismos datos) no tienen capacidad de desarrollar sus propias respuestas, por lo que la emisión de la oferta o la aceptación será en realidad la ejecución de una voluntad previamente programada. Sin embargo, los sistemas de inteligencia artificial autónomos sí tienen capacidad de aprendizaje y de adaptación a nuevas circunstancias, lo que pone en entredicho el consentimiento libre y consciente de la persona que puso en marcha tal contrato. *Vid.* Pertíñez Vílchez (2024: 106-108).

cualquier otra información pertinente para la toma de decisiones. Una vez recopilados, los datos son procesados por el sistema empleando algoritmos específicos, pudiéndose aplicar técnicas de análisis de datos para identificar patrones o tendencias que ayuden a informar la toma de decisiones. El asistente digital genera una decisión concreta respecto al contrato de consumo en cuestión (v. gr., elección de ciertos productos o servicios, fijación de precios, gestión de riesgos...). Una vez tomada la determinación, el asistente digital puede implementarla a través de acciones como envío de notificaciones, ajuste de términos del contrato o realización de transacción, pudiendo desplegar un seguimiento y realizar ajustes según sea o no necesario.

Este fenómeno plantea desafíos significativos desde una perspectiva de Derecho de Consumo, concretamente en lo que concierne a la protección de la persona consumidora, que, como se ha apuntado previamente, delega su libertad contractual en el asistente. De ahí que haya que analizar si el marco normativo actual podría dar adecuada respuesta a los nuevos retos que surgen en el entorno digital. El Informe del Instituto de Derecho Europeo (ELI por sus siglas en inglés *-European Law Institute-*) sobre el Derecho del consumidor de la Unión Europea y la toma de decisiones automatizadas, como resultado provisional del proyecto sobre Principios Rectores y Reglas Modelo en Contratos Algorítmicos, consciente de la creciente presencia de la toma de decisiones automatizada, ofrece una interesante guía con recomendaciones para adaptar el marco legal existente y abordar de manera efectiva las implicaciones de la automatización en la ejecución de contratos de consumo.

En cuanto a los asistentes digitales en el escenario de contratos de consumo, el Informe del Instituto de Derecho Europeo revela una serie de formas típicas en los que pueden intervenir⁶¹: a) un producto inteligente con un sistema integrado de asistente digital⁶²; b) una aplicación de asistente digital independiente⁶³ que se configura para gestionar ciertos tipos de compras para un consumidor o una consumidora; c) una plataforma que proporciona un servicio integrado de asistente digital para realizar compras a comerciantes en la misma plataforma; d) un asistente digital con un tipo de *software* “hecho en casa” o de código abierto⁶⁴. Además, y atendiendo a la necesaria sostenibilidad, se plantea la oportunidad de diseñarlos de manera que hicieran del consumo sostenible una prioridad, como serían los casos de ralentizar el proceso de compra o de operar con valores predeterminados verdes.

⁶¹ European Law Institute (2023: 19).

⁶² Caso de una máquina de café con opciones de reordenamiento automático basadas en datos sobre la persona usuaria de la máquina.

⁶³ Por ejemplo, para un teléfono inteligente o para un ordenador.

⁶⁴ Su código fuente está disponible públicamente, por lo que cualquier persona puede utilizarlo de acuerdo con los términos de una licencia de código abierto, añadiendo nuevas funcionalidades, corrigiendo errores o adaptándolo a diversas plataformas.

Para los fines del citado Informe, el debate se centra en los asistentes digitales tipo a), tipo b) y tipo c). Y se darían en tres posibles escenarios: 1) consumidor/a implementa un asistente digital a través del cual firma un contrato con el comerciante; 2) consumidor/a concluye un contrato con comerciante que utiliza un asistente digital para la celebración del contrato; 3) tanto consumidor/a como comerciante se sirven de asistentes digitales para la conclusión del contrato.

Sea el escenario que sea, se deben aplicar los Principios comunes que indica el Informe del Instituto de Derecho Europeo, y que podrían constituir una parte general de una futura ley europea sobre contratación algorítmica.

III.2. Principios generales del Informe del Instituto de Derecho Europeo sobre el derecho del consumidor de la UE y la toma de decisiones automatizada

El meritado Informe utiliza como clave de bóveda los diferentes Principios generales que deben informar la legislación de la UE en materia de Derecho de Consumo para dar una respuesta adecuada a las ADM. A partir de ellos se construyen las recomendaciones normativas que pueden contribuir a un óptimo desarrollo de las decisiones automatizadas en los contratos de consumo.

Se analizan⁶⁵ a continuación cada uno de los Principios referenciados con la finalidad de entender cuál es su repercusión en las ADM contractuales y, por ende, si implicaría la necesidad de reforma legislativa.

A) Principio 1: Atribución de las acciones del asistente digital al consumidor⁶⁶

Si el consumidor o la consumidora emplea un asistente digital para concluir un contrato, el Informe acoge la tesis de que el despliegue de dicho asistente implica que todas las acciones de este se traten legalmente como acciones de quien consume, atribuyéndoselas a efectos jurídicos. Así, el Principio de atribución debe combinarse con requisitos que permitan a la persona consumidora mantener suficiente control sobre su asistente digital. De ahí que el Informe sugiera los siguientes mecanismos de control: 1) Posibilidad de aprobar u oponerse (a través de una suerte de derecho a objetar⁶⁷) a la celebración de un contrato concertado a través de un asistente digital antes de que sea jurídicamente vinculante. 2) Establecer y revisar los

⁶⁵ *Vid.* como trabajo preliminar al respecto, Pacheco Jiménez (2024).

⁶⁶ European Law Institute (2023: 23-25).

⁶⁷ Más acertado que introducir un “período de gracia” para retirarse una vez notificado al/a la consumidor/a que el contrato se va a celebrar y que este no continúe; y esto porque muchos contratos pueden tener excluido el derecho de desistimiento.

parámetros que utiliza un asistente digital para llegar a sus decisiones⁶⁸. 3) Facultad de suspender o desconectar un asistente digital, si así lo precisan las personas consumidoras ya que no están obligadas a utilizar un asistente digital. 4) Derecho a no conformidad y, por ende, invalidación del contrato, por decisiones del asistente digital derivadas del autoaprendizaje, excediéndose del rango de toma de decisiones.

Estos requisitos se podrían implementar a través de un “deber de diseño”, de manera que el/la consumidor/a pueda tener el control último sobre las decisiones adoptadas por su asistente digital. Es más, ha de añadirse en este punto el temor a que el asistente digital tome una decisión completamente imprevista, esto es, fuera del rango de decisiones que un algoritmo de autoaprendizaje podría tomar⁶⁹.

Todo lo expuesto se alinea con las pautas del nuevo Reglamento de Inteligencia Artificial, en tanto en cuanto propone que el sistema de IA cumpla con la exigencia de supervisión y control humanos.

B) Principio 2: Aplicación de la legislación sobre consumidores a los contratos algorítmicos⁷⁰

En tanto en cuanto el contrato algorítmico celebrado a través de un asistente digital entre consumidor/a y comerciante entra en el ámbito de aplicación de la legislación de la UE en materia de Derecho de Consumo, la utilización del asistente digital por la persona consumidora no debería alterar las obligaciones legales del/la comerciante para el tipo específico de contrato en virtud de la legislación aludida, sino actuar como si el contrato se hubiera negociado o concluido el sin el uso de dicho asistente.

No puede obviarse que el asistente digital no siempre ha de conducir a una mejor toma de decisiones. Piénsese en el supuesto de que el sistema de IA haya sido diseñado para promover los intereses de quien comercia, siendo en este caso más pertinente aún la protección ofrecida a los y las consumidores/as desde la normativa europea. Ante un contenido sesgado proporcionado por el asistente, se podría materializar el derecho a compensación por pérdidas debido a información insuficiente o engañosa. Según el Informe, y para mayor garantía, los asistentes digitales podrían diseñarse sobre la base del marco jurídico pertinente. Es más, el/la comerciante no puede intentar evadir su responsabilidad por las acciones derivadas de su asistente digital, infringiendo derechos de la persona consumidora basándose que fue resultado de las decisiones tomadas por un asistente digital.

⁶⁸ De conformidad con el Principio 7.

⁶⁹ La solución que propone el Informe sería la no conformidad por parte de la persona consumidora, desplegándose responsabilidad por parte del proveedor del asistente digital. *Vid.* European Law Institute (2023: 24-25).

⁷⁰ European Law Institute (2023: 26-27).

C) Principio 3: Obligaciones de información precontractual⁷¹

Partiendo de la premisa de que el/la consumidor/a informado/a toma decisiones informadas (racionales), se colige que la tarea de reporte previo es parte determinante en el proceso de contratación. Pues bien, dependiendo de cómo evolucionen la tecnología de la IA y su aplicación, podría ser necesario recalibrar el modelo actual sobre información precontractual.

Podemos cuestionarnos si es factible que un asistente digital gestione los datos de forma totalmente fiable y más eficaz que la propia persona consumidora o, por el contrario, la desinformación o el sesgo del asistente puede llegar a ser un riesgo real. En este sentido, no todos los asistentes digitales tienen la misma capacidad predeterminada para llenar vacíos de información; comparar, combinar y verificar datos; o expresar se manera comprensible y personalizada; ni tampoco están diseñados con las mismas funcionalidades. Por tanto, el riesgo de desinformación o información sesgada procedente de un asistente digital es más que posible.

Dado que, hoy por hoy, la trayectoria de la evolución de los asistentes digitales es demasiado incierta, dudándose de su completa fiabilidad, en el Informe se parte de la base de que la lógica en la que se sustenta la obligación de facilitar información precontractual al sujeto más débil de la relación contractual no se vería afectada por el hecho de que utilice un asistente digital (de hecho, los deberes de información precontractual se aplicarían igual que para cualquier otro contrato para el suministro de contenido/servicios digitales), incluso cuando este disponga de funciones independientes de recopilación de datos y capacidad de procesamiento. En definitiva, la obligación de proporcionar información previa al contrato por parte de quien comercia es un deber general que ha de persistir, independientemente de que la persona consumidora se sirva de un asistente digital.

D) Principio 4: Principios de no discriminación/no barreras⁷²

Los/las consumidores/as deben tener libertad para utilizar asistentes digitales, no pudiendo quienes comercian prohibirlo de manera general, arbitraria o injustificada, a no ser que existan excepciones claramente definidas (*v. gr.*, evitar la manipulación o la saturación del sistema por un alto volumen de interacciones de asistentes digitales). Por tanto, permitir la prohibición injustificada de los asistentes digitales o el trato diferenciado de las personas consumidoras que hacen uso de ellos podría conllevar una discriminación en el mercado. De ahí que estos Principios de no discriminación y de no barreras impliquen tanto la prevención de medidas técnicas que impidan el empleo de asistentes digitales (*v. gr.*, a través de “bloqueadores

⁷¹ *Ibidem*, 27-29.

⁷² *Ibidem*, 29.

ADM”), como la necesidad de un diseño de interfaces en línea “amigable” (es decir, formato legible por máquina⁷³) para las ADM.

E) Principio 5: Divulgación del uso del asistente digital⁷⁴

El Informe propone que se divulgue el uso, si es el caso, de un asistente digital, normalmente a través de medios tecnológicos adecuados, permitiendo que se aplique tanto al uso de asistentes digitales por parte de personas consumidoras como de comerciantes, alertando así tanto a unos como a otros de la utilización de los mismos en la relación contractual. Y máxime en aquellas circunstancias en las que consumidores/as pueden tener dudas sobre si están interactuando con un humano o con una máquina. Es más, en aras a esta transparencia, el asistente debería programarse para revelarse a sí mismo.

F) Principio 6: Protección del asistente digital frente a manipulaciones⁷⁵

La actual legislación de la UE en materia de Derecho de Consumo ya trata de impedir la manipulación de las personas consumidoras cuando deciden celebrar un contrato, así como a la hora de ejercer sus derechos. Además de la tradicional protección contra prácticas engañosas y agresivas, regulaciones más recientes en la Directiva de Servicios Digitales⁷⁶, en el Reglamento de Mercados Digitales⁷⁷ y en la Directiva sobre contratos de servicios financieros celebrados a distancia⁷⁸, se centran en los “patrones oscuros” que explotan deliberadamente las debilidades de comportamiento de los consumidores. De ahí que la protección jurídica ya existente deba ampliarse a los intentos de manipulación dirigida de un

⁷³ El conjunto de datos está estructurado de una manera que permita el procesamiento automatizado, es decir, esos datos se pueden importar o leer mediante un sistema informático sin intervención humana.

⁷⁴ European Law Institute (2023: 30).

⁷⁵ *Ibidem*, 31.

⁷⁶ Directiva 2000/31/CE del Parlamento Europeo y del Consejo, de 8 de junio de 2000 (DOCE núm. 178, de 17 de julio de 2000). Su objetivo era crear un espacio digital más seguro en el que se protejan adecuadamente los derechos fundamentales de todos los usuarios de servicios digitales, que incluyen un gran número de servicios y plataformas en línea que conectan, transmiten y alojan la información que fluye por Internet. Modificada por el Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo, de 19 de octubre de 2022, relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE (DOUE núm. 277, de 27 de octubre de 2022). Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32022R2065>

⁷⁷ Establece unas condiciones de competencia equitativas para fomentar la innovación el crecimiento y la competitividad, tanto en el mercado único europeo como a escala mundial. Actualmente regulación por el Reglamento de Mercados Digitales (UE) 2022/1925, del Parlamento Europeo y del Consejo, de 14 de septiembre de 2022, sobre mercados disputables y equitativos en el sector digital y por el que se modifican las Directivas (UE) 2019/1937 y 2022/1828 (DOUE núm. 265, de 12 de octubre de 2022). Entró en vigor el 2 de mayo de 2023. Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32022R1925>

⁷⁸ Directiva (UE) 2023/2673, del Parlamento Europeo y del Consejo, de 22 de noviembre de 2023, por la que se modifica la Directiva 2011/83/UE en lo relativo a los contratos de servicios financieros celebrados a distancia y se deroga la Directiva 2002/65/CE (DOUE núm. 2673, de 28 de noviembre de 2023). Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202302673

asistente digital en los contratos algorítmicos⁷⁹. Para ello el Informe indica la posibilidad de que las reglas de diseño de los sistemas de IA en general o los asistentes digitales en concreto incorporen medidas tecnológicas para minimizar dichos riesgos de manipulación.

G) Principio 7: Capacidad para determinar parámetros/divulgación de parámetros preestablecidos⁸⁰

El funcionamiento de los asistentes digitales precisará la inclusión de diversos parámetros, aunque algunos ya vendrán preestablecidos por el proveedor de aquellos. Si quien contrata no puede conocer dichos parámetros empleados por el asistente digital, el principio de atribución (*ex* Principio 1) de los efectos jurídicos de las acciones y decisiones adoptadas por dicho asistente puede verse socavado y, consecuentemente, propiciar que quien provee el asistente sea responsable por el daño irrogado. Asimismo, cuando se proporciona el asistente digital operado por quien comercia (por ejemplo, porque es una característica integrada de un producto inteligente o de una plataforma en línea), sería adecuado que el/la consumidor/a pudiera ser capaz de determinar (a través de una función de selección) los parámetros para el uso del asistente al tratar con quien comercia. Por tanto, atendiendo a la necesidad de que las personas consumidoras conserven el control general, habrá de divulgarse la existencia de parámetros preestablecidos y preverse, asimismo, la posibilidad de revisarlos⁸¹.

H) Principio 8: Asistentes digitales y conflictos de intereses⁸²

Un problema concreto que se plantea con los asistentes digitales integrados en productos inteligentes (*v. gr.*, vehículos, dispositivos médicos, conectados a Internet), o aquellos que se proporcionan en plataformas en línea, es el posible conflicto que surgiría entre, por un lado, los intereses comerciales de quien provee el producto inteligente o de quien opera la plataforma en línea, y aquellos/as comerciantes que venden a través de la plataforma, y, por otro, los intereses de la persona consumidora. Aunque se plantea la opción de exigir que un asistente digital se centre exclusivamente en asistir al/a la consumidor/a, quizás la solución más viable sería revelar claramente ese posible conflicto de intereses, haciendo que quien consume sea plenamente

⁷⁹ Para ello las plataformas deben ser más transparentes sobre sus algoritmos y sobre la publicidad (evitando la elaboración de perfiles utilizando categorías especiales de datos personales); y han de crearse guardianes de acceso para mitigar las discriminaciones algorítmicas. *Vid.* Espín Alba (2024: 171-172).

⁸⁰ European Law Institute (2023: 31-32).

⁸¹ De hecho, el propio Informe sugiere un borrador para esa divulgación: “Cuando un comerciante utilice un asistente digital para la celebración de contratos con consumidores, los parámetros en base a los cuales actúe el asistente digital cuando trata con un consumidor deben serle comunicados, incluidos los parámetros preestablecidos por el comerciante y los principales criterios para tomar decisiones. El consumidor debería tener la opción de seleccionar los tipos de parámetros que se utilizarán y establecer valores para dichos parámetros”. Asimismo, y alternativamente, el Informe sostiene que esta disposición podría formularse como un deber de diseño en el asistente digital.

⁸² European Law Institute (2023: 32).

consciente de que el asistente digital únicamente considerará celebrar contratos con los/las comerciantes de la plataforma en cuestión o con los/las comerciantes seleccionados por quien fabrique el producto inteligente, pero que otros/as comerciantes que no se hayan tomado en consideración por el asistente también estarían disponibles. Igualmente, una solución alternativa sería, de nuevo, un requisito de diseño basado en que el asistente digital solo pueda ayudar a una de las partes, aunque esta medida podría afectar en la práctica a la provisión de asistentes digitales de plataformas.

Atendiendo a los Principios desarrollados, el Informe recomienda la adopción de un pequeño número de nuevas disposiciones para alinear la legislación europea existente, pero manteniendo que todas las Directivas implicadas hoy por hoy están casi “listas para ADM” (“ADMready”). De hecho, en la fase 2 del Proyecto sobre Principios Rectores y Reglas Modelo en Contratos Algorítmicos se abordarán cuestiones más amplias de cara a un marco jurídico óptimo, como es el derecho de protección de datos o la implementación de la Inteligencia Artificial tras la aprobación de su norma europea definitiva.

III.3. Directivas “listas para ADM”

Me permito mostrar a continuación algunas de esas Directivas aludidas, que, siguiendo lo observado por los Principios ELI en el marco de la contratación con asistentes digitales, necesitarían unos meros ajustes para su correcta implementación en el ámbito de ADM⁸³.

En primer lugar, la Directiva 93/13/CEE, del Consejo, de 5 de abril de 1993, sobre las cláusulas abusivas en los contratos celebrados con consumidores⁸⁴, cuyos términos deben extrapolarse indefectiblemente a los contratos algorítmicos, manteniéndose las mismas garantías, entre ellas: que los términos estén redactados de forma clara y lenguaje inteligible, así como en un formato legible por máquina permitiendo un procesamiento digital; que no se puedan bloquear los asistentes digitales, siendo abusiva la restricción del uso de estos por parte del consumidor; y que se puedan rechazar términos ya declarados como abusivos a través de una función tecnológica⁸⁵.

En segundo lugar, la Directiva 2005/29/CE, del Parlamento Europeo y del Consejo, de 11 de mayo de 2005, relativa a las prácticas comerciales desleales de las empresas en sus relaciones con los consumidores en el mercado interior⁸⁶, cuyo estándar para discernir si la

⁸³ Vid. como trabajo preliminar al respecto, Pacheco Jiménez (2024).

⁸⁴ DOCE núm. 95, de 21 de abril de 1993. Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:01993L0013-20220528&from=EN>

⁸⁵ European Law Institute (2023: 33-35).

⁸⁶ DOUE núm. 149, de 11 de junio de 2005. Disponible en: <https://www.boe.es/doue/2005/149/L00022-00039.pdf>

información es engañosa se conservaría, máxime en el seno de los contratos propiciados por un asistente digital, que puede encontrar nuevas formas de prácticas manipulativas, aportando información incorrecta o sesgada y distorsionando la capacidad de la persona consumidora para tomar una decisión informada de manera óptima. El listado de prácticas prohibidas que observa el Anexo I de la Directiva no se ajustaría, por cuestiones evidentes de evolución tecnológica desde la fecha de su entrada en vigor, al escenario de los asistentes digitales en sede de contratación. Para solventarlo, la evaluación de la práctica dudosa debería efectuarse caso por caso, conforme a los requisitos específicos para identificar prácticas engañosas (*ex arts. 6 y 7*) y agresivas (*ex arts. 8 y 9*); y, en último término, atendiendo a la definición general de práctica comercial desleal del art. 5, que contempla la violación de la diligencia profesional del comerciante y el menoscabo apreciable de la capacidad para tomar una decisión informada⁸⁷. Por último, podría hacerse una remisión al RIA en la medida en que los sistemas de IA pueden, por un lado, implementar técnicas subliminales que consigan dirigir a la persona consumidora hacia una determinada decisión sin ser plenamente consciente de ello; y por otro, elaborar perfiles, lo que podría redundar en la actual definición de “consumidor vulnerable”⁸⁸.

En tercer lugar, la Directiva 2011/83/UE, del Parlamento Europeo y del Consejo, de 25 de octubre de 2011, sobre los derechos de los consumidores⁸⁹, que podría complementarse con alguna disposición adicional coherente con los Principios ELI para dar cabida al uso de asistentes digitales y la debida información sobre ello. Esas posibles modificaciones tendrían que ver con los criterios utilizados por el asistente digital al seleccionar las contrapartes de la persona consumidora; con la obligación de información sobre los intereses que representa el asistente digital (*v. gr.*, interés de consumidor/a o económico o del mercado); o con la necesidad de revelar que se están empleando asistentes digitales para la negociación o celebración de un contrato. En este punto ha de traerse a colación el RIA ya que, en sus arts. 50 y siguientes, se contemplan las obligaciones de transparencia de los proveedores y responsables del despliegue

⁸⁷ European Law Institute (2023: 39).

⁸⁸ Como concepto dinámico, cambiaría en función del contexto existente, pero siempre teniendo presente la situación de manifiesta desigualdad en el ejercicio de sus derechos, que podría verse obligado a tomar decisiones o aceptar condiciones contractuales que, de no encontrarse en situación de vulnerabilidad, no aceptaría.

⁸⁹ DOUE núm. 304, de 22 de noviembre de 2011. Disponible en: <https://www.boe.es/doue/2011/304/L00064-00088.pdf>. Modificada por la Directiva (UE) 2019/2161 del Parlamento Europeo y del Consejo de 27 de noviembre de 2019 por la que se modifica la Directiva 93/13/CEE del Consejo y las Directivas 98/6/CE, 2005/29/CE y 2011/83/UE del Parlamento Europeo y del Consejo, en lo que atañe a la mejora de la aplicación y la modernización de las normas de protección de los consumidores de la Unión (DOUE núm. 328, de 18 de diciembre de 2019). Disponible en: <https://www.boe.es/doue/2019/328/L00007-00028.pdf>

de determinados sistemas de IA. Es más, la falta de información podría implicar el incumplimiento del deber precontractual de informar⁹⁰.

En cuarto lugar, la Directiva (UE) 2019/771, del Parlamento Europeo y del Consejo, de 20 de mayo de 2019, relativa a determinados aspectos de los contratos de compraventa de bienes⁹¹ y Directiva (UE) 2019/770, del Parlamento Europeo y del Consejo, de 20 de mayo de 2019, relativa a determinados aspectos de los contratos de suministro de contenidos y servicios digitales⁹², que constituyen una sólida base para la persona consumidora que busque solución en aquellos supuestos en los que un asistente digital actúa de una manera diferente a la razonablemente esperada y ello conlleve una falta de conformidad. En este sentido, el asistente digital puede tomar una decisión que exceda del rango de determinaciones que tiene diseñadas aquel; o no se desempeña en la forma preestablecida de conformidad con los términos del contrato de suministro del asistente digital. Sin olvidar que los procesos de aprendizaje automático derivados de la IA pueden implicar que las decisiones adoptadas por el asistente digital superen lo previsto por quién crea o provee el asistente. La referida falta de conformidad debería solventarse atendiendo a las soluciones que contemplan sendas Directivas (*ex* artículos 7 y 8 de la Directiva de compraventa de bienes, y artículos 6 y 7 de la Directiva de servicios digitales), esto es, reparación, sustitución, reducción de precio o rescisión del contrato. A lo que habría de añadirse el derecho a una indemnización de daños y perjuicios para compensar las posibles pérdidas ocasionadas por la falta de conformidad derivada de la actuación de un asistente digital⁹³.

IV. GARANTÍAS FRENTE A LAS DECISIONES AUTOMATIZADAS Y REGLAMENTO DE PROTECCIÓN DE DATOS

Tal y como ya se ha apuntado en este estudio, para que los algoritmos funcionen se precisan datos masivos caracterizados por su volumen, variedad y velocidad. Evidentemente, la cuestión de la protección de datos en el ámbito de las decisiones automatizadas presenta una intrincada problemática.

⁹⁰ Navas Navarro (2024: 56).

⁹¹ DOUE núm. 136, de 22 de mayo de 2019. Disponible en: <https://www.boe.es/doue/2019/136/L00028-00050.pdf>

⁹² DOUE núm. 126, de 22 de mayo de 2019. Disponible en: <https://www.boe.es/doue/2019/136/L00001-00027.pdf>

⁹³ Así se desprende del Informe del European Law Institute (2023: 48-49).

Un caso complejo es el que concierne a la elaboración de perfiles⁹⁴ en función de sus características⁹⁵ (v. gr., género, edad, hábitos y comportamientos). Con el “*profiling*” se generalizan y agrupan personas, pero hay que tener presente que cada persona puede presentar variables concretas que la diferencien del resto de miembros del grupo al que ha sido asignada. En este sentido, los arts. 13.2.f) y 14.2g) del Reglamento Europeo de Protección de Datos (RGPD), obligan al responsable del tratamiento a informar sobre “la existencia de decisiones automatizadas, incluida la elaboración de perfiles, a que se refiere el artículo 22, apartados 1 y 4, y, al menos en tales casos, información significativa sobre la lógica aplicada, así como la importancia y las consecuencias previstas de dicho tratamiento para el interesado”.

Entre las problemáticas posibles en la elaboración de perfiles podemos hallar algoritmos erróneos que conecten a una persona a determinadas particularidades o comportamientos con los que no se identifica, o algoritmos sesgados que crean estereotipos discriminatorios (siendo especialmente preocupante la reconducción forzosa a parámetros comunes de la generalidad de los ciudadanos obviando, por tanto, la diversidad o la discapacidad). Sobre discriminación algorítmica directa se pueden deducir dos tipos: atribución de un valor negativo al grupo desaventajado o inferencia de la pertenencia al grupo protegido por medio de otros datos a los que se atribuye un valor negativo y que no tienen valor predictivo en el marco de la decisión automatizada⁹⁶. En ambos casos, el algoritmo contraviene el tenor del art. 9.1 RGPD: “Quedan prohibidos el tratamiento de datos personales que revelen el origen étnico o racial, las opiniones políticas, las convicciones religiosas o filosóficas, o la afiliación sindical, y el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física, datos relativos a la salud o datos relativos a la vida sexual o la orientación sexual de una persona física”. Puede observarse en esta categorización que no se incluyen datos relativos a la discapacidad, la edad o el estatus socioeconómico; algunos podrían interpretarse a través de las

⁹⁴ El art. 4 del 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 esos datos (disponible en: <https://www.boe.es/doue/2016/119/L00001-00088.pdf>), define la elaboración de perfiles como “toda forma de tratamiento automatizado de datos personales consistente en utilizar datos personales para evaluar determinados aspectos personales de una persona física, en particular para analizar o predecir aspectos relativos al rendimiento profesional, situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, ubicación o movimientos de dicha persona física”.

⁹⁵ Dos tipologías de perfil relevantes serían los perfiles individuales y los perfiles grupales. En los perfiles individuales se agrega información sobre un individuo y se usa para derivar o inferir características desconocidas. Los grupales categorizan a las personas en un grupo que comparte todas o ciertas características. *Vid.* Palma Ortigosa (2022: 84).

⁹⁶ Soriano Arnanz (2022: 149).

categorías existentes, pero otros no terminarían de encajar, por lo que en la práctica podría verse cercenada la protección de ciertos grupos vulnerables⁹⁷.

El RGPD dispone en su art. 22 un régimen específico para aquellos tratamientos donde los sistemas algorítmicos adoptan decisiones plenamente automatizadas y estas generan efectos relevantes en la esfera de los particulares. Así, según el meritado precepto, el derecho a no ser objeto de decisiones automatizadas solo se podrá ejercitar cuando la persona en cuestión sea consciente de que las decisiones que le afectan se han basado únicamente en el tratamiento de sus datos, incluida la elaboración de perfiles, produciendo efectos jurídicos sobre aquella o afectándole de manera significativa⁹⁸. Hay que matizar que deben darse los dos elementos cumulativos, esto es: decisión enteramente basada en un tratamiento automatizado (descartándose los sistemas donde se haya incorporado intervención humana); y decisión que produce efectos jurídicos (*v. gr.*, denegación de una prestación económica, no obtención de un contrato) en la persona afectada o le afecten significativamente de modo similar (*v. gr.*, precio o anuncio personalizado que puede influir en el acceso a determinados bienes o servicios)⁹⁹. No obstante, este derecho es limitado en la medida en que no se reconoce a los particulares una oposición generalizada a tal tratamiento, sino que se podrán negar a verse sometidos a dicho tratamiento solo cuando la base jurídica utilizada por su responsable no sea la permitida por la norma o, a pesar de ser la adecuada, no se hayan cumplido convenientemente todos los requisitos legales¹⁰⁰.

Observa, además, las siguientes excepciones al ejercicio de ese derecho, a saber: que esta automatización sea “necesaria para la celebración o la ejecución de un contrato entre el interesado y el responsable del tratamiento”; que esté “autorizada por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento y que establezca

⁹⁷ La Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación (BOE núm. 167, de 13 de julio de 2022), destina su art. 23 a la IA y los mecanismos de toma de decisiones automatizados, con especial mención la minimización de sesgos en los algoritmos en su art. 23.1: “En el marco de la Estrategia Nacional de Inteligencia Artificial, de la Carta de Derechos Digitales y de las iniciativas europeas en torno a la Inteligencia Artificial, las administraciones públicas favorecerán la puesta en marcha de mecanismos para que los algoritmos involucrados en la toma de decisiones que se utilicen en las administraciones públicas tengan en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas, siempre que sea factible técnicamente. En estos mecanismos se incluirán su diseño y datos de entrenamiento, y abordarán su potencial impacto discriminatorio. Para lograr este fin, se promoverá la realización de evaluaciones de impacto que determinen el posible sesgo discriminatorio”.

⁹⁸ Puede accederse al formulario de derecho de oposición a decisiones individuales automatizadas, incluida la elaboración de perfiles, en este enlace: <https://www.aepd.es/sites/default/files/2019-09/formulario-derecho-de-oposicion-decisiones-automatizadas.pdf> (22/05/2024).

⁹⁹ Palma Ortigosa (2022: 71-72). En sentido similar, *vid.* Sáenz de Jubera Higuero (2023: 61-62).

¹⁰⁰ Roberto Granero (2022: 127).

asimismo medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado”; que se base “en el consentimiento explícito del interesado”.

Lo cierto es que el derecho a una explicación de las decisiones que toma un programa informático de manera automatizada genera muchas controversias. Y ello por dos motivos: primero, el carácter vinculante o no de dicho derecho ya que esta concreta facultad no viene reconocida en ningún precepto del Reglamento, considerándose la posibilidad de integrar este derecho a través de una interpretación sistemática de los arts. 13, 14, 15 y 22 RGPD; y segundo, aun reconociéndose este derecho, su alcance y contenido no quedaría claro debido a que su ejercicio chocaría con otros bienes jurídicos, tales como propiedad intelectual o secretos comerciales. Sea como sea, el responsable del tratamiento, ante una decisión automatizada, tendrá la obligación de explicar al particular, en la medida de lo posible (esto es, propiciando una comprensión adecuada, lo que no significa un entendimiento completo del funcionamiento del sistema), la decisión adoptada por el programa informático, de manera que la persona afectada pueda ejercitar los derechos que el RGPD le concede (*v. gr.*, derecho a impugnar la decisión o derecho a expresar su punto de vista)¹⁰¹.

En este punto resulta interesante la sentencia de la Sala Primera del Tribunal de Justicia de la Unión Europea de 7 de diciembre de 2023¹⁰², como primera sentencia sobre decisiones automatizadas y sus implicaciones jurídicas para la protección de datos y el Reglamento de IA. Adopta un enfoque garantista expandiendo el alcance del art. 22 RGPD más allá del responsable formal de la decisión, abarcando a terceros que procesen datos. Así, el caso involucra a SCHUFA, una empresa que proporciona información de solvencia de consumidores basada en procedimientos matemáticos y estadísticos, realizándose una clasificación de la persona atendiendo a su comportamiento. El demandante, persona física, alega que su solicitud de préstamo fue denegada sobre la base de la información proporcionada por SCHUFA, ante la que ejerció su derecho de acceso de protección de datos. Pero dicha empresa únicamente le facilitó contenido genérico, sin mostrar al demandante los datos sobre sí mismo y sin explicar la ponderación que de esos datos había realizado en forma de valores de probabilidad. Y, para no facilitar más información, SCHUFA alegó que no era ella la responsable de la denegación del crédito, sino sus socios contractuales, a quienes les facilitaba la información de solvencia. La sentencia acomete una interpretación garantista al considerar que la generación del valor de probabilidad por parte de SCHUFA no es un mero acto preparatorio, por lo que cae en el ámbito

¹⁰¹ *Ibidem*, 128-129.

¹⁰² (TJCE 2023, 146). Para un acertado análisis de ella, *vid.* Cotino Hueso (2024) y Espín Alba (2024: 156-160).

de aplicación del art. 22 RGPD (proyectándose más allá del responsable formal que toma la decisión respecto del afectado), en tanto en cuanto este “debe interpretarse en el sentido que de la generación automatizada, por una agencia de información comercial, de un valor de probabilidad a partir de datos personales relativos a una persona y acerca de la capacidad de esta para hacer frente a compromisos de pago en el futuro, constituye una decisión individual automatizada, en el sentido de la mencionada disposición, cuando de ese valor de probabilidad dependa de manera determinante que un tercero, al que se comunica dicho valor, establezca, ejecute o ponga fin a una relación contractual con esa persona” (FD 73).

Y no olvidemos que el régimen del art. 22 RGPD será aplicable cuando haya un tratamiento de datos mediante sistemas de IA, por lo que la conexión con la norma de IA es más que evidente y adquiere, por tanto, una nueva (y más compleja) dimensión. Piénsese en aquellos supuestos donde los sistemas de IA se utilicen para influir sustancialmente en ámbitos específicos como educación, empleo o seguros; sistemas clasificados como de alto riesgo.

V. CONSIDERACIONES FINALES

La implementación de sistemas de IA en sede de contratos es ya una realidad; una realidad no exenta de riesgos de diversa índole. Como se ha desarrollado a lo largo del presente estudio, el abordaje regulatorio de la IA es prolijo y, a tenor de las diferentes situaciones que concurren en la práctica, se irán viendo sus aciertos y carencias, máxime en un escenario tan técnicamente complejo y con repercusiones jurídicas tan relevantes como es el de la Inteligencia Artificial.

En lo que respecta a contratos en los que la persona consumidora emplea un asistente digital, la normativa europea vigente en materia de Derecho de Consumo se adecuaría bastante a ese concreto escenario, precisando unos pequeños ajustes, que posiblemente vendrán de la mano de la implementación del RIA. Y ello en la medida en que los sistemas de IA, tal y como se ha apuntado en líneas anteriores, pueden llegar a manipular o sesgar la información que se proporciona a la persona usuaria o pueden crear una comunicación comercial (véase la incursión del IA generativa) o pueden elaborar perfiles concretos de personas consumidoras ofreciendo un determinado tipo de productos/servicios o incluso evaluar su capacidad crediticia. En principio, y según el RIA, los asistentes digitales encajarían en la clasificación de riesgo limitado, poniéndose el énfasis en la transparencia en el uso del sistema de IA si se quiere fomentar la confianza de sus usuarios/as. De ahí la importancia de ir comprobando si la

evolución de los asistentes digitales en el ámbito de la contratación cumple los requisitos para esta clasificación, a efectos de aplicación del RIA.

Otra de las cuestiones que precisan una regulación jurídica satisfactoria es la procedente de la discriminación algorítmica, perpetuando estructuras sociales de desigualdad. Para solventar la problemática sería recomendable elaborar un marco jurídico¹⁰³ que asegurase el control previo y constante de los algoritmos (incluso en función del nivel de riesgo), partiendo desde su propio diseño¹⁰⁴, basado en la transparencia¹⁰⁵, y sin hacer que queden inservibles por eliminar aquellos datos que revelen la pertenencia a un determinado grupo social¹⁰⁶. No puede obviarse que la opacidad de los algoritmos utilizados y la posible existencia de sesgos ponen en jaque la fiabilidad de los sistemas de IA en sede de contratación ya que pueden llegar a vulnerarse la autonomía de la voluntad y la capacidad electiva de quien contrata¹⁰⁷.

Asimismo, el meritado diseño es fundamental para el futuro ciclo de vida del asistente digital, influyendo en su ulterior despliegue, tal y como se ha explicado en el análisis de los Principios generales del Informe del Instituto de Derecho Europeo en materia de decisiones automatizadas, cuando aluden a ese óptimo diseño para procurar una interfaz en línea amigable para las ADM o para minimizar riesgos de manipulación o para evitar conflictos de intereses.

En materia de protección de datos, a pesar de las previsiones del RGPD, y ante el desarrollo y las especialidades que nos depara la IA (mención especial de la IA generativa de contenidos y los datos sintéticos¹⁰⁸) deberán revisarse las disposiciones normativas tuitivas de los datos de usuarios y usuarias para que no queden obsoletas.

¹⁰³ En este sentido recuérdese el mencionado Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022, relativo a un mercado único de servicios digitales, que obliga a las plataformas a ser más transparentes sobre sus algoritmos y prácticas comerciales, prohibiendo que aquellas presenten anuncios basados en la elaboración de perfiles.

¹⁰⁴ Soriano Aranz (2022: 166).

¹⁰⁵ Sin embargo, los objetivos de transparencia se ven entorpecidos por las denominadas “cajas negras” (situaciones donde se ven los datos de entrada y de salida, pero se desconoce qué sucede en el ínterin, es decir, los criterios empleados por el algoritmo). *Vid.* Espín Alba (2024: 146).

¹⁰⁶ Soriano Aranz (2021: 117).

¹⁰⁷ Sáenz de Jubera Higuero (2023: 59).

¹⁰⁸ Se parecen mucho a los datos reales (de hecho, son generados por la IA para sustituir los datos reales) y pueden tener patrones o detalles que facilitarían la identificación de personas o características sensibles.

A la fecha de finalización de este trabajo se publica la “Declaración Conjunta sobre la Competencia en Modelos Fundamentales de IA Generativa y Productos de IA” (23 de julio de 2024) por parte de la Unión Europea, el Reino Unido y los Estados Unidos de América. Su finalidad es la promoción de mercados justos, abiertos y competitivos en el ámbito de la IA generativa, vigilando los riesgos que puedan obstaculizarlo. Disponible en: https://competition-policy.ec.europa.eu/document/download/79948846-4605-4c3a-94a6-044e344acc33_en?filename=20240723_competition_in_generative_AI_joint_statement_COMP-CMA-DOJ-FTC.pdf

Para finalizar, comparto las palabras del profesor Roberto Granero (2022)¹⁰⁹ cuando afirma que “los algoritmos no son ni serán nunca fuente del Derecho”; sino que, en el mejor de los casos, las tradicionales fuentes del Derecho podrán ser ejecutadas mediante el algoritmo y consecuentemente supervisadas por un ser humano. De ahí que uno de los pilares básicos del Reglamento de Inteligencia Artificial sea el de la supervisión humana para evitar la conculcación de derechos fundamentales, particularmente en los sistemas calificados de alto riesgo.

Bibliografía

- BENGIO, Y., GOODFELLOW, I. & COURVILLE, A. (2017): *Deep learning*, vol. 1. Cambridge, MA, USA: MIT press.
- CALVO, J., GUZMÁN, M. y RAMOS, D. (2018): “Machine Learning, una pieza clave en la transformación de los modelos de negocio”, *Management Solutions* (pp. 1-42).
- COGLIANESE, C. & LEHR, D. (2017): “Regulating by robot: administrative decision making in the machine-learning era”, *The Georgetown Law Journal*, vol. 105, n. 5 (pp. 1147-1223).
- COTINO HUESO, L. (2024): “La primera sentencia del Tribunal de Justicia de la Unión Europea sobre decisiones automatizadas y sus implicaciones para la protección de datos y el Reglamento de inteligencia artificial”, *Diario LA LEY*. Disponible en: <https://diariolaley.laleynext.es/dll/2024/01/17/la-primera-sentencia-del-tribunal-de-justicia-de-la-union-europea-sobre-decisiones-automatizadas-y-sus-implicaciones-para-la-proteccion-de-datos-y-el-reglamento-de-inteligencia-artificial>
- ESPÍN ALBA, I. (2024): “Sesgos discriminatorios en la toma automatizada de decisiones en la contratación y protección de datos”, ÁLVAREZ LATA, N. (Coord.) et al.: *Derecho de contratos, responsabilidad extracontractual e inteligencia artificial*, Aranzadi, Las Rozas (Madrid) (pp. 123-186).
- GIL MEMBRADO, C. (2024): “Daños producidos por la IA: la opacidad del algoritmo y el efecto de caja negra”, ÁLVAREZ LATA, N. (Coord.) et al.: *Derecho de contratos, responsabilidad extracontractual e inteligencia artificial*, Aranzadi, Las Rozas (Madrid) (pp. 501-547).
- MCCARTHY, J. (1956): *What is artificial intelligence?* Disponible en: <https://www-formal.stanford.edu/jmc/whatisai.pdf>.
- MITTELSTADT, B. et al. (2016): “The ethics of algorithms: Mapping the debate”, *Big Data & Society*, 3 (2) (pp. 1-21).
- MUÑOZ GARCÍA, C. (2023): *Regulación de la inteligencia artificial en Europa. Incidencia en los regímenes jurídicos de protección de datos y de responsabilidad por productos*, Tirant lo Blanch, Valencia.
- NAVAS NAVARRO, S. (2024): “Deshumanizando al contrato. Una mirada desde la inteligencia artificial”, ÁLVAREZ LATA, N. (Coord.) et al.: *Derecho de contratos, responsabilidad extracontractual e inteligencia artificial*, Aranzadi, Las Rozas (Madrid) (pp. 31-100).
- PACHECO JIMÉNEZ, M^a N. (2024): “Contratos algorítmicos y toma de decisiones automatizadas. El papel de la normativa europea de Consumo”, *Blog Centro de Estudios de Consumo*. Disponible en:

¹⁰⁹ Roberto Granero (2022:133).

https://centrodeestudiosdeconsumo.com/images/Contratos_algoritmicos_y_toma_de_decisiones_automatizada.pdf

- PALMA ORTIGOSA, A. (2002.a): *Decisiones automatizadas. Y protección de datos. Especial atención a los sistemas de inteligencia artificial*, Dykinson, Madrid.
- (2022.b): “El ciclo de vida de los sistemas de Inteligencia Artificial. Aproximación técnica de las fases presentes durante el diseño y despliegue de los sistemas algorítmicos”, en COTINO HUESO, L. (Dir.) *et al: Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thomson Reuters Aranzadi, Cizur Menor (pp. 29-51).
- PERTIÑEZ VÍLCHEZ, F. (2024): “La contratación entre sistemas de inteligencia artificial”, ÁLVAREZ LATA, N. (Coord.) *et al.: Derecho de contratos, responsabilidad extracontractual e inteligencia artificial*, Aranzadi, Las Rozas (Madrid) (pp. 101-122).
- ROBERTO GRANERO, H. (2022): “Derechos y garantías frente al uso de inteligencia artificial y decisiones automatizadas, especialmente en el ámbito judicial y de aplicación de la ley”, en COTINO HUESO, L. (Dir.) *et al: Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thomson Reuters Aranzadi, Cizur Menor (pp. 107-137).
- ROIG, A. (2020): *Las garantías frente a las decisiones automatizadas. Del Reglamento General de Protección de Datos a la gobernanza algorítmica*, Bosch, Barcelona.
- SÁENZ DE JUBERA HIGUERO, B. (2023): “Retos jurídicos de la inteligencia artificial en el ámbito del consentimiento contractual: una aproximación general”, *Revista de Derecho Civil*, vol. X, núm. 2 (pp. 41-70).
- SALAZAR GARCÍA, I. (2022): “Retos actuales de la ética en la Inteligencia Artificial”, en COTINO HUESO, L. (Dir.) *et al: Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thomson Reuters Aranzadi, Cizur Menor (pp. 53-66).
- SEMMLER, S. & ZEEVE R. (2017): “Artificial Intelligence: Application Today and Implications Tomorrow”, *Duke Law and Technology Review*, 16 (pp. 85-99).
- SORIANO ARNAZ, A. (2021): “Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos”, *Revista de Derecho Público: Teoría y Método*, Marcial Pons, Madrid (pp. 85-127).
- SORIANO ARNAZ, A. (2022): “Discriminación algorítmica: garantías y protección jurídica”, en COTINO HUESO, L. (Dir.) *et al: Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*, Thomson Reuters Aranzadi, Cizur Menor (pp. 139-169).
- TURING, A. M. (1950): “Computing Machinery and Intelligence”, *Mind* 49.
- VESTRI, G. (2024): “La Unión Europea estrena el Reglamento de Inteligencia Artificial (RIA)”, *Diario La Ley*, núm. 10550. Disponible en: <https://diariolaley.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAAEAMtMSbF1jTAAAiNTMzMjY7WylKLizPw8WYMDIxMDc0NTIbz8lNQOF2fb0ryU1LTMvNQUtZzEpNQcZ5BwakWJbXpqXmpRZrJrUVF-EUh3ZlqlS35ySGVBqm1aYk5xqlpqUn5-Nool8TDDAbeRB19-AAAAWKE>

Textos oficiales consultados

- CONSEJO DE EUROPA (2024): *Convenio Marco sobre Inteligencia Artificial y Derechos Humanos, Democracia y Estado de Derecho*. Disponible en: https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680afb11f
- DIRECTORATE-GENERAL FOR COMPETITION, EUROPEAN COMMISSION (2024): “Join Statement on Competition in Generative AI Foundation Models and AI Products”. Disponible en: https://commission.europa.eu/about-european-commission/departments-and-executive-agencies/competition_en
- EUROPEAN COMMISSION (2019): “Ethics guidelines for trustworthy AI”. Disponible en: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

- EUROPEAN LAW INSTITUTE (2023): “Eu Consumer Law and Automated Decision-Making (ADM): Is EU Consumer Law Ready for ADM?”.
- MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA (2024): *Estrategia de Inteligencia Artificial*. Disponible en: https://portal.mineco.gob.es/es-es/digitalizacionIA/Documents/Estrategia_IA_2024.pdf
- OECD (2019): “Recommendation of the Council on Artificial Intelligence”. Disponible en: [https://one.oecd.org/document/C/MIN\(2019\)3/FINAL/en/pdf](https://one.oecd.org/document/C/MIN(2019)3/FINAL/en/pdf)
- OECD (2024): “Recommendation of the Council on Artificial Intelligence”. Disponible en: <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449?s=03#backgroundInformation>