



**COMILLAS**  
UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

# **IMPACTO DE LA INTELIGENCIA ARTIFICIAL EN EL DERECHO ¿Peligro u oportunidad?**

Marta Esteban Sánchez

5º E-3 A

MADRID | Marzo 2025



## RESUMEN

Este Trabajo de Fin de Grado examina la compleja interrelación entre la inteligencia artificial y el ordenamiento jurídico, analizando los riesgos y oportunidades que estas tecnologías representan para el sistema judicial. La investigación contextualiza la evolución de la IA desde sus fundamentos hasta su acelerado desarrollo en la Revolución 4.0, destacando el impacto de innovaciones como ChatGPT.

El estudio profundiza en aplicaciones concretas de la IA en el ámbito judicial: la automatización en la gestión de expedientes, sistemas predictivos como Jurimetría y vLex Analytics para la toma de decisiones, y herramientas para Mecanismos Alternativos de Solución de Controversias. Se analiza el fenómeno de los "jueces robots", evaluando experiencias internacionales y limitaciones del razonamiento artificial frente al judicial humano, junto con los desafíos para el futuro del empleo jurídico.

La investigación examina el marco regulatorio emergente, con énfasis en el Reglamento Europeo de IA (UE) 2024/1689 y su paradigma basado en niveles de riesgo. Se analiza la aproximación europea como "tercera vía" que prioriza la protección de derechos fundamentales, y la creación de la Agencia Española de Supervisión de Inteligencia Artificial.

Finalmente, se aborda la problemática ética de la IA jurídica y se concluye que la respuesta a la pregunta "¿Peligro u oportunidad?" trasciende dicotomías simplistas. La verdadera cuestión no es si debemos implementar estas tecnologías, sino cómo hacerlo preservando valores esenciales de justicia y dignidad humana. El éxito de la IA jurídica debe medirse no solo por indicadores de eficiencia, sino por su capacidad para fortalecer, nunca debilitar, los principios fundamentales del ordenamiento jurídico.

**Palabras clave:** Jueces robots, derechos fundamentales, automatización, impacto disruptivo, razonamiento artificial, simbiosis, democratización tecnológica.

## ABSTRACT

This Final Degree Project examines the complex interrelationship between artificial intelligence and the legal system, analyzing the risks and opportunities that these technologies represent for the judicial system. The research contextualizes the evolution of AI from its foundations to its accelerated development in the 4.0 Revolution, highlighting the impact of innovations such as ChatGPT. The study delves into concrete applications of AI in the judicial field: automation in case management, predictive systems such as Jurimetry and vLex Analytics for decision making, and tools for Alternative Dispute Resolution Mechanisms. The phenomenon of "robot judges" is analyzed, evaluating international experiences and limitations of artificial reasoning compared to human judicial reasoning, along with the challenges for the future of legal employment. The research examines the emerging regulatory framework, with emphasis on the European AI Regulation (EU) 2024/1689 and its paradigm based on risk levels. The European approach is analyzed as a "third way" that prioritizes the protection of fundamental rights, and the creation of the Spanish Agency for the Supervision of Artificial Intelligence. Finally, the ethical problems of legal AI are addressed and it is concluded that the answer to the question "Danger or opportunity?" transcends simplistic dichotomies. The real question is not whether we should implement these technologies, but how to do so while preserving essential values of justice and human dignity. The success of legal AI should be measured not only by efficiency indicators, but by its ability to strengthen, never weaken, the fundamental principles of the legal system.

**Keywords:** Robot judges, and fundamental rights, automation, disruptive impact, artificial reasoning, symbiosis, technological democratization.

# ÍNDICE

1.	INTRODUCCIÓN.....	8
1.1.	Justificación .....	8
1.2.	Metodología .....	9
1.3.	Objetivos .....	10
1.4.	Estructura .....	11
2.	GÉNESIS Y EVOLUCIÓN DE LA IA.....	13
2.1.	¿Qué es la inteligencia artificial?.....	13
2.2.	Antecedentes .....	15
2.3.	La revolución 4.0 .....	18
2.4.	La irrupción de ChatGPT en noviembre de 2022 .....	21
2.5.	Prospectiva de la IA .....	22
3.	APLICACIONES DE LA IA EN EL PROCESO JUDICIAL .....	24
3.1.	Gestión de expedientes judiciales .....	24
3.2.	Toma de decisiones judiciales .....	25
3.3.	Resolución alternativa de conflictos (MASC).....	26
3.4.	Jueces robots .....	27
3.5.	Desafíos para los juristas y el futuro del empleo en derecho.....	29
4.	MARCO JURÍDICO .....	31
4.1.	El reglamento europeo de IA .....	31
4.2.	Iniciativas estratégicas europeas para el desarrollo de la IA .....	32
4.3.	Protección de los derechos fundamentales y prohibiciones clave .....	34
4.4.	Marco normativo español complementario .....	38
5.	PROBLEMÁTICA ÉTICA DE LA IA JURÍDICA .....	40
5.1.	Transparencia y explicabilidad de los sistemas de IA en contextos jurídicos.....	40
5.2.	Sesgos algorítmicos y discriminación en la IA jurídica.....	41
5.3.	Gobernanza ética de la IA jurídica.....	42
6.	CONCLUSIONES .....	45
7.	BIBLIOGRAFÍA .....	51
7.1.	Legislación.....	51
7.2.	Obras doctrinales .....	51
7.3.	Recursos de internet.....	53

## ABREVIATURAS

- IA - Inteligencia Artificial
- TFG - Trabajo de Fin de Grado
- BOE- Boletín Oficial del Estado
- COMPAS - Correctional Offender Management Profiling for Alternative Sanctions
- BOSCO - (Sistema mencionado como ejemplo en el contexto español)
- MASC - Mecanismos Adecuados de Solución de Controversias
- AESIA - Agencia Española de Supervisión de Inteligencia Artificial
- UE - Unión Europea
- PIB - Producto Interior Bruto
- GAFA - Google, Amazon, Facebook, Apple
- SEDIA - Secretaría de Estado de Digitalización e Inteligencia Artificial
- WEF - Foro Económico Mundial (World Economic Forum)
- EuroHPC - Empresa Común de Informática de Alto Rendimiento Europea
- PYMEs - Pequeñas y Medianas Empresas
- IoT - Internet de las Cosas (Internet of Things)
- 4RI / 4.0 - Cuarta Revolución Industrial
- LexNET - Sistema de comunicaciones electrónicas judiciales en España
- ChatGPT - Modelo de lenguaje desarrollado por OpenAI
- GPT-3.5 - Generative Pre-trained Transformer 3.5
- OpenAI - Empresa desarrolladora de ChatGPT

- APA - American Psychological Association (referencia al formato de citación)
- vLex Analytics - Herramienta de análisis legal mencionada en el TFG
- REEM - Robot judicial utilizado en Emiratos Árabes Unidos
- IJB-A - (Base de datos de imágenes mencionada en la sección de sesgos algorítmicos)
- RGPD - Reglamento General de Protección de Datos
- Face++ - Sistema de reconocimiento facial mencionado en la sección sobre sesgos algorítmicos
- Next Generation EU - Fondos de recuperación de la Unión Europea

# 1. INTRODUCCIÓN

El ámbito jurídico se encuentra en un punto de inflexión ante el avance de la inteligencia artificial. Este capítulo introductorio establece las bases conceptuales y metodológicas que guiarán nuestra investigación sobre el impacto de la IA en el derecho. Comenzaremos exponiendo la justificación que motiva este trabajo, destacando cómo la automatización de funciones tradicionalmente humanas en la justicia plantea tanto oportunidades como desafíos fundamentales. Posteriormente, detallaremos la metodología empleada, que combina revisión bibliográfica con análisis de normativa reciente. Seguidamente, definiremos los tres objetivos fundamentales que persigue esta investigación: analizar la transformación de la administración de justicia mediante IA, examinar el marco normativo emergente, y formular principios deontológicos para su implementación responsable. Finalmente, presentaremos la estructura organizativa que seguirá el trabajo para abordar sistemáticamente esta compleja interrelación entre tecnología y derecho.

## 1.1. JUSTIFICACIÓN

La inteligencia artificial (IA) se ha convertido en una de las tecnologías más disruptivas en el ámbito jurídico. Ha transformado la manera en que se administran los procesos judiciales y generado nuevos interrogantes sobre su impacto en el derecho. El hecho de que cada vez se automaticen más funciones que siempre han sido humanas en la justicia, plantea oportunidades, como la optimización del tiempo y la reducción de gastos, pero también desafíos que pueden comprometer derechos fundamentales como el debido proceso y la igualdad ante la ley (de Luis García, 2023).

Uno de los principales problemas que provocan estos sistemas al aplicarse al derecho es la falta de transparencia en la toma de decisiones automatizadas. La opacidad de los algoritmos puede dificultar la verificación de criterios utilizados en la resolución de casos, lo que pone en riesgo la garantía de imparcialidad judicial y la confianza en el sistema de justicia (de Luis García, 2023). Asimismo, la implementación de herramientas como el software COMPAS en Estados Unidos ha demostrado que los sistemas de IA pueden mantener sesgos discriminatorios, afectando a la equidad de los fallos judiciales y generando preocupaciones sobre la confiabilidad de estas tecnologías (de Luis García, 2023).

Desde una perspectiva regulatoria, la ausencia de un marco normativo claro en muchos países deja a los operadores jurídicos en una situación de incertidumbre. Aunque iniciativas como el Reglamento Europeo de IA intentan establecer directrices para su aplicación en el sistema de justicia, aún existen vacíos legales que necesitan un análisis detallado para asegurar que la incorporación de la IA en el derecho respete los principios fundamentales del ordenamiento jurídico (de Luis García, 2023).

## 1.2. METODOLOGÍA

Este Trabajo de Fin de Grado utiliza un método de investigación basado principalmente en la revisión y análisis de fuentes de información relevantes sobre la inteligencia artificial y su impacto en el campo del derecho.

Este estudio se basa en un enfoque metodológico que combina la revisión de la literatura existente con el análisis de documentos normativos y propuestas regulatorias recientes, llevando a cabo un examen de estudios académicos y legislación relevante para comprender la relación entre el derecho y la inteligencia artificial.

La metodología adoptada también contempla un enfoque multidisciplinario, integrando perspectivas jurídicas, éticas y tecnológicas. Esto es fundamental para abordar la complejidad del tema, ya que la IA no solo afecta el ámbito legal, sino que también plantea cuestiones éticas sobre la autonomía de las máquinas y la responsabilidad de los diseñadores y usuarios de estos sistemas. A través de este enfoque, se busca proporcionar una visión integral que permita entender las implicaciones de la IA en el derecho y la administración de justicia.

El trabajo se ha desarrollado en varias etapas:

1. Búsqueda inicial: Se realizó una primera exploración en bases de datos académicas y fuentes especializadas para identificar los temas y debates más importantes sobre IA y derecho.
2. Selección de fuentes: Se escogieron principalmente artículos de revistas académicas, informes oficiales y textos legales actuales. Se dio prioridad a publicaciones recientes (2018-2025) para asegurar la actualidad de la información, aunque también se incluyeron algunas referencias más antiguas cuando resultaban importantes para entender la evolución histórica del tema.

3. Análisis de la información: Se examinó cuidadosamente toda la literatura seleccionada para identificar ideas comunes, y tendencias emergentes en el campo de la IA aplicada al derecho.
4. Estudio de la normativa: Se analizaron las regulaciones europeas y españolas relacionadas con la IA, evaluando cómo responden a los desafíos planteados por los expertos en la materia.
5. Desarrollo de conclusiones: Finalmente, se sintetizaron los hallazgos principales para responder a los objetivos del trabajo, señalando tanto las oportunidades como los riesgos que presenta la IA para el ámbito jurídico.

Esta metodología ha permitido obtener una visión completa de cómo la inteligencia artificial está transformando el derecho, considerando aspectos tecnológicos, legales y éticos de forma integrada.

### 1.3. OBJETIVOS

El propósito central de esta investigación consiste en examinar la compleja interrelación entre el ordenamiento jurídico y los sistemas de inteligencia artificial, buscando reconocer tanto los retos como las posibilidades que estas tecnologías emergentes representan para nuestras instituciones judiciales. Este trabajo persigue tres propósitos fundamentales:

En primer lugar, analizaremos detalladamente cómo las herramientas basadas en IA están transformando diversas facetas de la administración de justicia. Estudiaremos su aplicación desde las etapas iniciales de investigación hasta la emisión de sentencias, prestando especial atención a su influencia en los mecanismos decisorios de los operadores jurídicos.

Seguidamente, examinaremos la creciente necesidad de desarrollar un marco normativo adecuado que permita garantizar la responsabilidad en la implementación de estos sistemas. Este marco deberá asegurar tanto la debida rendición de cuentas en el uso de la tecnología como la salvaguarda efectiva de los derechos fundamentales de todas las personas.

Finalmente, nuestra investigación pretende formular un conjunto coherente de principios deontológicos que sirvan como guía para la concepción y puesta en práctica de soluciones tecnológicas en el contexto jurídico. Estos principios deberán asegurar que la incorporación de la IA al ámbito judicial se realice preservando inalterablemente los valores democráticos y el respeto por los derechos humanos.

## 1.4. ESTRUCTURA

Se estructura en seis capítulos principales, diseñados para proporcionar una comprensión progresiva y sistemática del impacto de la inteligencia artificial en el ámbito jurídico:

El primer capítulo introduce el trabajo, presentando la justificación del tema elegido, la metodología empleada, los objetivos perseguidos y la estructura del documento.

El segundo capítulo aborda los fundamentos conceptuales e históricos de la inteligencia artificial. Comienza definiendo qué es la IA, explora sus antecedentes históricos, contextualiza su desarrollo en el marco de la cuarta revolución industrial, analiza el impacto disruptivo de ChatGPT desde su lanzamiento en 2022, y finaliza con un análisis prospectivo sobre la evolución futura de estas tecnologías.

El tercer capítulo examina las aplicaciones concretas de la IA en el proceso judicial. Se analizan cinco áreas principales: la gestión de expedientes judiciales, la toma de decisiones judiciales, la resolución alternativa de conflictos, el concepto de "jueces robots", y los desafíos que estas tecnologías plantean para los profesionales del derecho y el futuro del empleo en este sector.

El cuarto capítulo se centra en el marco regulatorio de la IA, con especial atención al Reglamento Europeo de IA de 2024, las iniciativas estratégicas europeas para el desarrollo de estas tecnologías, la protección de derechos fundamentales y el marco normativo español complementario.

El quinto capítulo aborda la problemática ética asociada a la implementación de la IA en contextos jurídicos. Se analizan tres aspectos fundamentales: la transparencia y explicabilidad de los sistemas de IA, los sesgos algorítmicos y la discriminación potencial, y los modelos de gobernanza ética necesarios.

El sexto capítulo presenta las conclusiones del trabajo, sintetizando los hallazgos principales y ofreciendo una reflexión sobre el balance entre oportunidades y riesgos que presenta la IA para el derecho.

Finalmente, se incluye la bibliografía completa que ha servido de base para la investigación, ordenada alfabéticamente y siguiendo las normas APA 7ª edición.

Esta estructura permite abordar el tema desde múltiples perspectivas (técnica, jurídica y ética), facilitando la comprensión integral de un fenómeno complejo y en constante evolución como es la integración de la IA en el ámbito jurídico.

## 2. GÉNESIS Y EVOLUCIÓN DE LA IA

La comprensión del impacto de la inteligencia artificial en el ámbito jurídico requiere conocer sus fundamentos teóricos y desarrollo histórico. Este capítulo ofrece una panorámica evolutiva de la IA desde sus orígenes conceptuales hasta su actual protagonismo tecnológico. Comenzaremos definiendo qué es exactamente la inteligencia artificial, explorando sus distintas categorías y potencialidades. Seguidamente, examinaremos sus antecedentes históricos, desde las contribuciones pioneras de Alan Turing hasta la histórica Conferencia de Dartmouth que acuñó el término en 1956. Continuaremos analizando cómo la IA se ha convertido en elemento central de la llamada Revolución 4.0, transformando estructuras económicas y sociales. Posteriormente, estudiaremos el impacto disruptivo que supuso la irrupción de ChatGPT en noviembre de 2022, difundiendo el acceso a sistemas de IA avanzados. Finalmente, presentaremos una prospectiva sobre el futuro desarrollo de estas tecnologías y su proyección económica hacia 2030, con especial atención a sus implicaciones para España y el contexto europeo.

### 2.1. ¿QUÉ ES LA INTELIGENCIA ARTIFICIAL?

La inteligencia artificial (IA) puede definirse como la habilidad de un sistema para analizar con precisión datos del entorno, extraer conocimientos de ellos y aplicar lo aprendido para cumplir objetivos y realizar tareas de manera adaptable (Kaplan & Haenlein, 2019). De forma complementaria, Jarrahi (2018) la caracteriza como sistemas avanzados capaces de razonar y adquirir conocimientos basándose en el trabajo fundamental de Russell, Norvig e Intelligence (1995).

La IA, a pesar de su importancia en casi todas las salas de juntas empresariales e incluso en muchas conversaciones del día a día, sigue siendo un concepto con muchas cuestiones por resolver aún. No debe considerarse un término tan cerrado, si no que requiere una versión más detallada. (Kaplan & Haenlein, 2019).

Según Kaplan y Haenlein (2019), los sistemas de IA pueden clasificarse en tres categorías fundamentales. En primer lugar, la IA analítica, centrada en las características cognitivas y el razonamiento lógico. En segundo lugar, la IA inspirada en humanos, la cual incorpora elementos cognitivos y emocionales, y por último la IA humanizada que integra capacidades cognitivas, emocionales y sociales.

La IA al influir en tantos procesos organizacionales, ha generado el miedo de que las máquinas inteligentes pronto pueden llegar a reemplazar a muchos humanos en la toma de decisiones. Sin embargo, Jarrahi (2018) establece una perspectiva más proactiva y pragmática que defiende que los humanos y la IA no tienen porque ser excluyentes si no complementarios, dejando a un lado esta mentalidad negativa y analizando cómo tanto uno como otro puede aportar sus propias fortalezas y ventajas en los procesos de toma de decisiones organizacionales.

Con una mayor capacidad tecnológica de procesamiento de información y un enfoque analítico, la inteligencia artificial puede ampliar las capacidades cognitivas humanas al enfrentar desafíos complejos (Jarrahi, 2018). Al mismo tiempo, los humanos tienen esa capacidad de ofrecer un enfoque más intuitivo para lidiar con la incertidumbre y los posibles errores que puedan darse en la toma de decisiones organizacionales. El objetivo de esta idea es que los sistemas de IA deberían diseñarse con la intención de sacar provecho o explotar las contribuciones humanas, no de reemplazarlas.

La inteligencia artificial impulsa una transformación interna que permite completar diversas actividades de forma más ágil, eficiente y con menor inversión de recursos (Kaplan & Haenlein, 2019), afectando incluso a industrias intensivas en conocimiento. El crecimiento de la inteligencia artificial exige una nueva relación simbiótica entre humanos y máquinas, en la que la distribución del trabajo evoluciona constantemente (Jarrahi, 2018). Las perspectivas predominantes sobre esta colaboración sugieren que las máquinas deberían encargarse de las tareas repetitivas, permitiendo a los humanos enfocarse en actividades más creativas.

## 2.2. ANTECEDENTES

La IA tuvo su origen en el proyecto histórico de Conferencia de Verano de Dartmouth en 1956. Como señalan McCarthy et al. (1955), este proyecto se fundamentó en "la conjetura de que cada aspecto del aprendizaje o cualquier otra característica de la inteligencia puede, en principio, ser descrito con tanta precisión que se puede crear una máquina para simularlo" (p. 12). Esta propuesta, elaborada por John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Shannon, estableció las bases para investigar cómo hacer que las máquinas "utilicen el lenguaje, formen abstracciones y conceptos, resuelvan tipos de problemas actualmente reservados para los humanos y se mejoren a sí mismas" (McCarthy et al., 1955, p. 12).

No obstante, las bases teóricas de la IA tienen su origen aproximadamente veinte años antes, gracias a las contribuciones pioneras de Alan Turing, quien durante la década de 1930 formuló conceptos matemáticos que posteriormente serían esenciales para el desarrollo de esta disciplina

En 1936, Turing publicó su influyente artículo "On Computable Numbers", donde introdujo el concepto de lo que posteriormente se conocería como la "Máquina de Turing" (Hodges, 2006). Este trabajo no solo representó un avance crucial en el campo de la lógica matemática, sino que estableció las bases conceptuales para el desarrollo de la tecnología moderna y, como consecuencia, de la inteligencia artificial.

La Máquina de Turing fue concebida como un modelo teórico que permitía definir con precisión el concepto de "proceso mecánico" o "método", utilizando operaciones simples y fundamentales. Como señala Hodges (2006), "Turing definió el concepto de 'lo mecánico' en términos de operaciones atómicas simples" (p. 2). Esto fue crucial, ya que permitió establecer los límites de lo que podía ser calculado mediante procedimientos mecánicos definidos.

Un concepto particularmente relevante para el posterior desarrollo de la IA fue el de la "máquina universal", un tipo especial de Máquina de Turing capaz de simular el comportamiento de cualquier otra Máquina de Turing. Como explica Hodges (2006), "existen infinitas máquinas universales de Turing... son aquellas máquinas con 'tablas de comportamiento' lo suficientemente complejas como para leer las tablas de otras Máquinas de Turing, y luego hacer lo que esas máquinas habrían hecho" (p. 3). Este concepto de universalidad sentó las bases teóricas para lo que eventualmente se convertiría en la computadora digital moderna.

Es importante destacar que Turing desarrolló estos conceptos mucho antes de que existieran los ordenadores tal como los conocemos hoy. Como señala Hodges (2006), "en 1936 no existían tales ordenadores; de hecho, el ordenador moderno surgió de la formulación de 'comportamiento mecánico' que Turing encontró en este trabajo" (p. 3).

Tras sus contribuciones fundamentales en la Segunda Guerra Mundial, donde Turing jugó un papel crucial en el desciframiento de los códigos alemanes en Bletchley Park, su atención se dirigió hacia la posibilidad de crear máquinas inteligentes. Como indica Hodges (2006), "desde 1945, Turing intentó utilizar la electrónica para convertir su máquina universal en una realidad práctica" (p. 5). Este período marcó una transición en su pensamiento de la teoría abstracta de la tecnología a la aplicación práctica en forma de ordenadores.

En 1946, Turing ya anticipaba el futuro de la programación de ordenadores, o como él lo llamaba, "la construcción de tablas de instrucciones". Según Hodges (2006), Turing predijo que "habrá probablemente mucho trabajo por hacer, ya que cada proceso conocido tendrá que ser traducido en forma de tabla de instrucciones en algún momento" (p. 5). Esta visión anticipaba refleja el desarrollo futuro de la industria informática.

A finales de la década de 1940, Turing empezó a expresar sus ideas acerca del potencial de las máquinas para manifestar comportamiento inteligente. Su enfoque suponía una transformación importante frente a las creencias predominantes en aquel momento. En una conferencia de 1947, Turing abordó la aparente contradicción entre infalibilidad e inteligencia: "Si se espera que una máquina sea infalible, no puede ser también inteligente... Pero estos teoremas no dicen nada sobre cuánta inteligencia puede mostrarse si una máquina no pretende ser infalible" (Hodges, 2006, p. 5).

En sus trabajos posteriores, Turing exploró cómo las máquinas podrían realizar operaciones que no parecían ser "mecánicas" en el sentido común del término. Su principal objetivo era demostrar que los procesos del cerebro humano podían ser replicados mediante operaciones computables. Como señala Hodges (2006), Turing pensaba que "'la iniciativa' podría surgir de sistemas donde el algoritmo aplicado no está diseñado conscientemente, sino que se llega a él por otros medios" (p. 5). Esto lo llevó a proponer ideas precursoras de las redes neuronales y los algoritmos genéticos.

El momento definitivo en la conceptualización de la inteligencia artificial llegó con la publicación en 1950 del artículo "Computing Machinery and Intelligence" en la revista Mind. En este influyente trabajo, Turing propuso lo que más tarde se conocería como el "Test de Turing", un criterio para determinar si una máquina podía exhibir un comportamiento inteligente que no pudiese distinguirse del de un ser humano (Hodges, 2006).

Este artículo no solo difundió la idea de la inteligencia artificial a más gente, sino que también estableció un programa de investigación concreto. Según Hodges (2006), Turing culminó su trabajo con "ideas para máquinas de aprendizaje y la educación de máquinas-niños" (p. 5), anticipando líneas de investigación que serían fundamentales en el desarrollo posterior del campo.

A lo largo de las décadas, la definición de inteligencia artificial ha evolucionado significativamente. Según Kim (2016), Goodfellow et al. (2016), conceptualizan el aprendizaje profundo como una modalidad de aprendizaje automático que faculta a los sistemas computacionales para adquirir conocimiento a través de la experiencia y comprender la realidad mediante una estructura jerárquica de conceptos. Esta perspectiva destaca la habilidad de los sistemas para obtener conocimiento mediante la experiencia, eliminando la necesidad de que un humano defina explícitamente toda la información necesaria para la máquina.

En tiempos más recientes, Kaplan y Haenlein (2019) definen la IA como la habilidad de un sistema tecnológico para comprender información proveniente del entorno, extraer conocimientos a partir de esos datos y emplear tales conocimientos para alcanzar metas concretas mediante una adaptación dinámica y versátil. Esta definición refleja una comprensión más contemporánea que integra la interpretación de datos, el aprendizaje y la adaptación como componentes esenciales de los sistemas de IA.

Kaplan y Haenlein (2019) proponen que la IA no debe considerarse un término monolítico, sino que requiere una visión más matizada, clasificando los sistemas de IA en tres categorías fundamentales:

1. IA analítica: Centrada en las características cognitivas y el razonamiento lógico.
2. IA inspirada en humanos: Incorpora elementos cognitivos y emocionales.
3. IA humanizada: Integra capacidades cognitivas, emocionales y sociales.

Un aspecto crucial en la evolución de la IA ha sido la reconsideración de la relación entre humanos y máquinas. Jarrahi (2018) propone una aproximación innovadora que resalta la colaboración entre personas y tecnología inteligente, analizando cómo sus capacidades distintivas pueden integrarse para mejorar la toma de decisiones en el contexto organizacional. Esta visión refleja un cambio de paradigma desde la competencia hacia la colaboración, donde los sistemas de inteligencia artificial expanden las capacidades mentales humanas al abordar problemas de gran complejidad, en tanto que las personas conservan su habilidad para ofrecer perspectivas integradoras e intuitivas ante escenarios de ambigüedad e incertidumbre (Jarrahi, 2018).

Esta evolución conceptual ilustra cómo la IA ha pasado de ser una aspiración teórica en sus inicios, con las contribuciones fundamentales de Turing, a convertirse en una realidad tecnológica con profundas implicaciones para prácticamente todos los ámbitos de la actividad humana.

### 2.3. LA REVOLUCIÓN 4.0

La Cuarta Revolución Industrial (4RI) se presenta como un modelo disruptivo que está cambiando radicalmente las estructuras económicas, sociales y formativas en todo el mundo. Como señala Penprase (2018), la 4RI se caracteriza por la convergencia e integración de diversas tecnologías que se desarrollan exponencialmente, como la inteligencia artificial, las biotecnologías y los nanomateriales. Este modelo revolucionario de industria va más allá de nuestras ideas convencionales sobre producción industrial, ya que elementos distintivos de etapas previas, como las fábricas fijas y centralizadas o las grandes plantillas de trabajadores en empresas de gran tamaño, podrían desaparecer en este nuevo escenario.

La 4RI representa una evolución natural después de las tres revoluciones industriales previas: la Primera, caracterizada por la máquina de vapor y la mecanización; la Segunda, impulsada por la electricidad y la producción en masa; y la Tercera, definida por la computarización y la conectividad basada en la web (Penprase, 2018). Sin embargo, a diferencia de sus predecesoras, la 4RI se distingue por la convergencia e integración de múltiples tecnologías que se desarrollan exponencialmente, creando un efecto multiplicador que acelera notablemente el ritmo del cambio.

El Foro Económico Mundial (WEF) ha identificado varios puntos de inflexión en los que las tecnologías de la 4RI alcanzarán una adopción suficientemente amplia como para generar cambios sociales significativos. Penprase (2018) menciona que una encuesta a expertos en tecnología y ejecutivos proyectó fechas para estos puntos críticos, anticipando desarrollos como teléfonos implantables, amplia presencia digital de la población mundial, dispositivos de lectura conectados a internet, indumentaria con conectividad, y acceso generalizado a internet en los próximos años.

En el campo de la inteligencia artificial, Penprase (2018) señala que los expertos prevén una incorporación profunda de esta tecnología en diversos sectores, incluyendo su participación en juntas directivas corporativas, funciones de auditoría y servicios farmacéuticos, lo que refleja la penetración que se espera que tenga en diversos ámbitos económicos y profesionales.

Un elemento fundamental de la 4RI es el Internet de las Cosas (IoT). Penprase (2018) explica que este nuevo paradigma de manufactura permitirá diseñar objetos digitalmente para luego materializarlos mediante impresoras 3D capaces de utilizar una amplia gama de materiales, incluyendo tejidos biológicos. Esta tecnología facilitará una transformación bidireccional entre el mundo digital y el físico, atenuando las fronteras entre ambos.

En cuanto a los materiales para impresión 3D, Penprase (2018) describe un espectro que abarca desde termoplásticos convencionales hasta materiales de construcción a gran escala e incluso estructuras atómicas de precisión nanométrica. Estas capacidades permitirán construir edificios completos, crear microestructuras con tolerancias de alta precisión, o desarrollar estructuras biológicas para implantes médicos y trasplantes de órganos.

Otra dimensión crucial de la 4RI es el desarrollo de una bioeconomía impulsada por avances biotecnológicos. Penprase (2018) advierte que el crecimiento poblacional y la reducción de tierras cultivables debido al cambio climático exigirán un incremento considerable en la eficiencia de producción alimentaria en las próximas décadas, lo que demandará soluciones revolucionarias basadas en las nuevas tecnologías.

Penprase (2018) señala que el desarrollo de “biorefinerías” que utilicen organismos genéticamente modificados podría ser fundamental en el panorama de la 4RI. Estas instalaciones podrían aprovechar diversas materias primas como celulosa, biomasa y azúcares simples para producir masivamente combustibles, productos farmacéuticos y alimentos, contribuyendo a reducir la dependencia de combustibles fósiles.

Las implicaciones de la 4RI para la educación superior son profundas y demandan atención urgente. Penprase (2018) advierte sobre el potencial de estas tecnologías para generar tanto impactos sociales positivos como daños ambientales significativos, así como los riesgos asociados a la posible pérdida de control sobre sistemas de inteligencia artificial cada vez más autónomos en sectores financieros e infraestructuras urbanas.

Para enfrentar estos desafíos, Penprase (2018) enfatiza la necesidad de transformar sustancialmente los planes de estudio en ciencia y tecnología para desarrollar competencias en áreas emergentes como genómica, ciencia de datos, inteligencia artificial, robótica y nanomateriales. Esto implica repensar la enseñanza de las ciencias fundamentales tradicionales e incorporar una mayor formación en informática como alfabetización esencial para la era de la 4RI.

La 4RI exige especial énfasis en la adaptabilidad y el aprendizaje autodirigido. Penprase (2018) observa que la rápida obsolescencia de las habilidades en el entorno actual requiere que los trabajadores actualicen constantemente sus competencias y aprendan por sí mismos sobre tecnologías e industrias emergentes, incluso aquellas que no existían durante su formación inicial.

Además, Penprase (2018) sostiene que cualquier modelo educativo efectivo para la 4RI debe incorporar formación ética, sensibilidad intercultural y pensamiento crítico para garantizar una aplicación responsable e informada de las tecnologías emergentes. Los graduados deberían estar preparados para contribuir al avance material del mundo futuro mientras promueven un desarrollo tecnológico sostenible y ético.

La 4RI y sus tecnologías asociadas cuestionan nuestras concepciones fundamentales sobre la condición humana y nuestra relación con el entorno natural. Ante esto, Penprase (2018) propone que más allá de contenidos específicos, el currículum debe fomentar en los estudiantes capacidades de razonamiento ético, conciencia sobre impactos sociales y humanos, y comprensión de cómo las nuevas tecnologías afectan a las personas.

Un currículum de artes liberales adaptado a la 4RI debe responder a las transformaciones sociales provocadas por estas tecnologías. Penprase (2018) observa que ya se percibe una desconexión entre el crecimiento económico corporativo, los aumentos de productividad y los incrementos salariales, tendencia que probablemente se intensificará con la adopción de sistemas inteligentes automatizados en el ámbito empresarial.

La 4RI, con sus tecnologías exponenciales convergentes, está transformando rápidamente nuestras sociedades y economías de maneras que apenas comenzamos a comprender. A diferencia de las revoluciones industriales anteriores, la velocidad y la escala de estos cambios exigen una respuesta más proactiva del sector educativo.

Penprase (2018) argumenta que los estudiantes formados en entornos colaborativos con diversidad de perspectivas, donde desarrollen soluciones en equipo, estarán mejor preparados para los desafíos laborales de la 4RI. Por tanto, las instituciones de educación superior deben implementar pedagogías más interactivas y currículums que integren múltiples enfoques disciplinarios y culturales.

Finalmente, educar en tiempos de la 4RI requiere más que simplemente actualizar contenidos para incluir las últimas innovaciones tecnológicas. Necesitamos una transformación fundamental en cómo concebimos la educación, con el objetivo de formar individuos que no solo comprendan las tecnologías emergentes, sino que también sean capaces de orientar su desarrollo en contextos de rápida transformación. Esto implica cultivar profesionales con sensibilidad ética que puedan anticipar y evaluar las consecuencias sociales de sus decisiones en un mundo cada vez más interconectado y tecnológicamente complejo

## 2.4. LA IRRUPCIÓN DE CHAGPT EN NOVIEMBRE DE 2022

El 30 de noviembre de 2022, *OpenAI* lanzó *ChatGPT*, un *chatbot* de inteligencia artificial que revolucionaría la tecnología a nivel global. Aunque inicialmente fue concebido como "una demostración de la tecnología misma" y "un avance de investigación" para "recopilar opiniones sobre sus defectos abriéndolo al público" (Osornio, 2023), el impacto de esta herramienta fue inmediato y sin precedentes. En solo seis días alcanzó el millón de usuarios y en febrero de 2023 se convertiría en "la herramienta de consumo de mayor crecimiento de todos los tiempos" (Osornio, 2023). Basado en la tecnología GPT-3.5, este *chatbot* ofrecía capacidades conversacionales sorprendentes, como por ejemplo, era capaz de leer preguntas y respuestas de una conversación y entender el contexto en el que se estaba dando, capaz de componer poesía y de resolver los problemas que le plantearas. (Osornio, 2023). A diferencia de otros *chatbots* existentes, *ChatGPT* se distinguió por el enfoque "modesto y cauteloso" de *OpenAI*, que implementó restricciones para "mantener a *ChatGPT* a raya" (Osornio, 2023), una estrategia que permitió su rápida aceptación y expansión.

## 2.5. PROSPECTIVA DE LA IA

La inteligencia artificial se proyecta como una tecnología transformadora que impactará de manera significativa en el plano económico en los próximos años. Según datos citados por Alonso (2023), esta tecnología generó alrededor de 1700 millones de euros al producto interior bruto global en 2018, y las proyecciones sugieren que hacia 2030 su contribución alcanzará aproximadamente 14 mil millones de euros, magnitud comparable al PIB contemporáneo del conjunto de naciones que conforman la Unión Europea. El desarrollo de la IA se desenvuelve en un contexto de rivalidad entre potencias, donde China ha establecido como meta estratégica convertirse en el primer país del mundo en esta tecnología para 2030, mientras Estados Unidos lidera actualmente el sector mediante el dominio de las empresas conocidas como plataforma GAFA (Google, Amazon, Facebook y Apple).

En este contexto, la Unión Europea ha optado por una "tercera vía", buscando "promover asimismo una IA robusta, transparente, auditable y al servicio de la Humanidad, con las personas en el centro" (Alonso, 2023, p. 2).

Las tendencias futuras indican una transformación radical del mercado laboral, donde "3 de cada 4 trabajadores tendrán que cambiar o adaptar su trabajo hacia 2030" (Alonso, 2023, p. 4). El avance tecnológico generará una demanda creciente de nuevas especialidades laborales, particularmente en sectores vinculados a la tecnología. Según revela el Informe LinkedIn de empleos emergentes 2020, los profesionales en Inteligencia Artificial encabezan la lista de perfiles más solicitados por las empresas, registrando un incremento del 76% en comparación con el período previo.

Para España, la prospectiva indica que la adopción de la IA en las empresas crecerá "enormemente en los próximos años hasta poder alcanzar el 80%, previsiblemente alrededor de 2025" (Alonso, 2023, p. 3), lo que requerirá inversiones sostenidas y planificación a largo plazo. Como señala la autora, "es necesario comprometerse en inversiones a largo plazo, planificarlas y darles estabilidad contra cambios políticos, recesiones, etc." (Alonso, 2023, p. 5).

Impulsar proyectos clave con impacto a largo plazo, que integren toda la cadena industrial y fortalezcan el desarrollo regional, es una estrategia fundamental para aprovechar los fondos de recuperación Next Generation EU, de los cuales el 20% se destina al sector digital. En este contexto, las colaboraciones entre empresas, centros de investigación y academias serán fundamentales para "conseguir un salto de calidad en la transferencia y la innovación" (Alonso, 2023, p. 7).

### **3. APLICACIONES DE LA IA EN EL PROCESO JUDICIAL**

El ámbito de la justicia siempre ha sido el sector más complejo y tradicionalmente más lento en cuanto a su modernización. Sin embargo, gracias a la inteligencia artificial, el sistema judicial se ha agilizado, permitiendo este avance que sea mucho más eficiente a la hora de optimizar tareas que históricamente han sido manuales, repetitivas y propensas a errores humanos. En este apartado hablaremos de la gestión de expedientes judiciales, la toma de decisiones judiciales y la resolución alternativa de conflictos (MASC). Por otro lado, también se hablará sobre los jueces robots así como los desafíos con los que se encuentran los juristas hoy en día y los riesgos de su empleo.

#### **3.1. GESTIÓN DE EXPEDIENTES JUDICIALES**

Como señala Galindo Ayuda (2019), la evolución hacia el "expediente judicial electrónico" representa un hito fundamental en la modernización de la administración de justicia, permitiendo "la tramitación de los procedimientos judiciales en formato electrónico, sin el uso del papel" (p. 50). Esta transformación digital implica que los documentos iniciales del proceso estén en formato digital, así como que todo el procedimiento, desde la presentación de escritos hasta las resoluciones judiciales, se tramite electrónicamente. La implantación de sistemas como LexNET, que desde enero de 2017 hizo "obligatorio el envío de escritos a Juzgados y Tribunales electrónicamente" (Galindo Ayuda, 2019, p. 50), ha sentado las bases para una gestión automatizada que puede ser potenciada mediante programas de inteligencia artificial.

### 3.2. TOMA DE DECISIONES JUDICIALES

La integración de la inteligencia artificial en los procesos judiciales plantea importantes consideraciones éticas y legales, especialmente cuando se trata de la toma de decisiones automatizadas. Galindo Ayuda (2019) analiza este fenómeno destacando que "todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado" (p. 54), un principio recogido en el Reglamento General de Protección de Datos de la Unión Europea. El autor analiza programas como COMPAS, utilizado por las autoridades judiciales en Estados Unidos, y señala que generan preocupaciones legítimas cuando las decisiones automatizadas individuales, aparentemente son tomadas en cuenta no solo por las fuerzas policiales sino también por jueces que fundamentan sus fallos en las evaluaciones del programa. Esta realidad refleja la tensión entre la innovación tecnológica y los principios fundamentales del Estado de Derecho, donde se establece que "en el diseño e implantación de los programas se respeten los derechos y deberes establecidos por las leyes y las declaraciones de derechos" (p. 37).

En el contexto del sistema judicial moderno, han surgido diversas herramientas de análisis predictivo que buscan auxiliar a los profesionales del derecho. Según señala Galindo Ayuda (2019), programas como "Jurimetría" "ayuda a diseñar la mejor línea argumental y las posibilidades de éxito en relación a un caso" (p. 40), mientras que sistemas como *vLex Analytics* permiten "tener estimaciones de plazos y probabilidades de éxito en virtud tanto del juzgado y de los recursos como de los asuntos" (p. 42). Sin embargo, estas aplicaciones deben operar dentro de un marco normativo que proteja los derechos individuales. El autor enfatiza que la regulación europea establece que "en ambos casos, la decisión adoptada debe garantizar los derechos y libertades de la persona aplicando las garantías adecuadas" (p. 54), incluyendo "como mínimo el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión" (p. 55). Estas salvaguardas son fundamentales para mantener la integridad del sistema judicial en una era de creciente automatización.

La complejidad de implementar estos sistemas se acrecienta cuando se consideran sus implicaciones prácticas. El autor advierte sobre los riesgos de que el "acceso quede afectado por intereses" (p. 52), limitando "el ámbito dejado a la libertad de elección de los mismos que, en conformidad con las normas del Estado de Derecho, irremediablemente han de tener los juristas" (p. 52). Así, mientras la inteligencia artificial promete mejorar la eficiencia judicial, su aplicación debe equilibrarse cuidadosamente con los principios de independencia judicial, transparencia y debido proceso.

### 3.3. RESOLUCIÓN ALTERNATIVA DE CONFLICTOS (MASC)

La reforma introducida por la Ley Orgánica 1/2025 de medidas en materia de eficiencia del Servicio Público de Justicia implica una transformación fundamental en la manera de resolver conflictos dentro del sistema jurídico español, abarcando también las disputas societarias. Como señala García-Villarrubia (2025), esta reforma puede considerarse "la más importante en el campo de la resolución de disputas entre particulares desde que, allá por el 2000, se promulgó la vigente Ley de Enjuiciamiento Civil". El núcleo de esta transformación radica en la potenciación de los Mecanismos Adecuados de Solución de Controversias (MASC), convertidos ahora en requisito de procedibilidad para acceder a la jurisdicción. Estos MASC, definidos como "cualquier tipo de actividad negociadora, reconocida en esta u otras leyes, a la que las partes de un conflicto acuden de buena fe para encontrar una solución extrajudicial" (García-Villarrubia, 2025), incluyen diversas modalidades como la conciliación privada, la oferta vinculante confidencial, la opinión de persona experta independiente y el proceso de derecho colaborativo, además de la negociación directa y la mediación.

En el contexto societario, esta reforma plantea numerosos interrogantes y desafíos prácticos. García-Villarrubia (2025) observa que "en este tipo de litigios es obligatorio acudir a un MASC antes de formular la acción o acciones de que se trate", aunque reconoce que en casos como la impugnación de acuerdos sociales, "no parece que este tipo de procesos sea un campo especialmente propicio para que las partes alcancen una solución transaccional". Resulta particularmente complejo el caso de las acciones de responsabilidad contra administradores, donde el sistema de legitimación se enfrenta a dudas sobre quién debe conducir el MASC y cómo afecta a los plazos establecidos en la ley, pues como señala el autor, "la sociedad dispone de un plazo de un mes para entablar la acción, a partir del cual podrá hacerlo la minoría cualificada. Pero ese plazo no es de caducidad ni de prescripción", lo que genera incertidumbre sobre su interacción con la suspensión de plazos que conlleva el inicio de un MASC.

### 3.4. JUECES ROBOTS

La inclusión de inteligencia artificial (IA) en el ámbito judicial es uno de los temas más controvertidos y cautivadores en nuestra era tecnológica. Según Barona Vilar (2024), el principal desafío es si los sistemas computacionales podrían reemplazar a los jueces humanos en la toma de decisiones legales.

A nivel global, se están realizando varios experimentos para explorar este avance. En Estonia, por ejemplo, se ha creado un robot judicial que resuelve casos de disputas contractuales por menos de 7.000 euros, aunque siempre se puede apelar a un juez humano. En China, desde 2017, existen "tribunales de internet" que se encargan de casos relacionados con contratos, derechos de autor y consumo. Por su parte, en los Emiratos Árabes, el robot judicial REEM se ocupa de cuestiones como accidentes de tráfico y sanciones por infracciones (Barona Vilar, 2024, p.112).

Sin embargo, la autora señala que la verdadera dificultad radica en que las máquinas no son capaces de "pensar" como los humanos. Se basan en datos y estadísticas, pero carecen de memoria, de la capacidad de percibir el paso del tiempo, de ser creativas y de comprender el contexto emocional. A pesar de su capacidad para procesar grandes cantidades de información rápidamente, las máquinas no pueden generar un "diálogo interior" ni entender la continuidad de la consciencia como un ser humano (Barona Vilar, 2024).

El proceso judicial requiere un tipo de razonamiento que va más allá de la simple lógica de las máquinas, ya que un juez humano interpreta normas y situaciones con emociones e intuiciones. Por otro lado, un "juez robot" solo realiza operaciones basadas en datos (Barona Vilar, 2024).

La principal preocupación, como subraya Barona Vilar (2024), es que la IA no simula el pensamiento humano, sino que ofrece una versión limitada de este que podría poner en riesgo los derechos fundamentales de los involucrados en el proceso judicial.

Los sistemas basados en inteligencia artificial podrían ofrecer varias ventajas, como una mayor rapidez, un manejo más exhaustivo de la jurisprudencia, un acceso más fácil a las leyes, decisiones más consistentes y la eliminación de posibles sesgos o influencias corruptas, ya sean de carácter político, económico o social. Sin embargo, como señala Ríos (2020, citado en Cárdenas Krenz, 2021), una pregunta importante es si la sociedad estaría dispuesta a confiar aspectos tan cruciales de su vida, como su libertad, patrimonio, familia o empleo, a un robot.

El debate sobre la conveniencia de los jueces robots invita a reflexionar sobre las diferencias entre la inteligencia artificial y la humana. Según Searle (citado en Lacruz, 2018, referido por Cárdenas Krenz, 2021), la función de los ordenadores se limita a manipular símbolos, prestando atención solo a su estructura formal y sintáctica, sin considerar su significado, a diferencia de la mente humana, que opera en función del significado de los símbolos que procesa.

En lugar de reemplazar a los jueces por completo, podría ser más sensato ver a los robots como colaboradores judiciales. Ejemplos como los sistemas Prometea en Argentina, Siarelis y Pretoria en Colombia muestran cómo la tecnología puede asistir a los jueces en la aceleración de los procesos, el análisis de jurisprudencia y la preparación de borradores de decisiones (Cárdenas Krenz, 2021).

Como concluye Cárdenas Krenz (2021), "la inteligencia artificial puede ser un valioso soporte para la labor del juez, mas no su sustituto, en una relación de colaboración que recién empieza" (p. 9). El autor subraya que "la noble y delicada tarea de administrar justicia no es una actividad meramente mecánica" y que la justicia "no solo trata de la manera debida de distribuir las cosas. Trata también de la manera debida de valorarlas" (p. 6), citando a Sandel (2018).

### 3.5. DESAFÍOS PARA LOS JURISTAS Y EL FUTURO DEL EMPLEO EN DERECHO

La incorporación progresiva de sistemas de inteligencia artificial al ámbito legal está generando retos significativos que los profesionales jurídicos deben afrontar. McGinnis y Pearce (2014) sostienen que la profesión legal enfrenta una "gran disrupción" debido al desarrollo exponencial de las tecnologías de la información. Esta transformación no es un evento aislado que los abogados simplemente deberán acomodar, sino "una fuerza aceleradora que invadirá un territorio cada vez mayor y ejercerá un dominio más firme sobre esta área más amplia" (McGinnis & Pearce, 2014, p. 3043).

El impacto de esta disrupción en el empleo jurídico será diferenciado según el tipo de profesional. Los "abogados superestrella" se beneficiarán de esta revolución tecnológica, ya que "con grandes métricas de comparación, discernir quiénes son las superestrellas será más fácil" y podrán "ampliar su alcance a través de la tecnología: entregan sus soluciones innovadoras a los problemas más rápidamente y a una gama más amplia de clientes" (McGinnis & Pearce, 2014, p. 3054). También se verán favorecidos aquellos abogados que logren adaptar su práctica o su organización para aprovechar los aportes de menor costo proporcionados por las máquinas.

Por otro lado, los "abogados artesanos", aquellos que redactan testamentos rutinarios, gestionan cierres de casas, escriben contratos estándar y revisan documentos, enfrentan "un futuro mucho más sombrío, porque las máquinas realizarán muchas de estas tareas legales rutinarias" (McGinnis & Pearce, 2014, p. 3042). Estas tareas, que representan gran parte del trabajo de muchos despachos jurídicos, son precisamente las más propensas a ser automatizadas.

Existe un sector de abogados que permanecerá relativamente menos afectado. McGinnis y Pearce (2014) identifican tres categorías:

1. Los abogados orales, ya que "las máquinas no hablarán en los tribunales en el futuro previsible" (p. 3042).
2. Los abogados especialistas en áreas de cambio legal rápido, como la regulación Dodd-Frank, donde "las máquinas funcionarán mejor en áreas más rutinizadas y establecidas" (p. 3042).

3. Los consejeros que deben persuadir a clientes reacios a hacer lo que es de su interés, puesto que "las máquinas no podrán crear los lazos emocionales necesarios con los clientes" (p. 3042).

Es importante señalar que la integración de la inteligencia artificial en el ámbito jurídico no solo transformará la profesión, sino que también podría mejorar el acceso a la justicia. Las tecnologías inteligentes pueden ayudar a "los abogados, mediante habilidad o mejor organización, a aumentar la prestación de servicios a precios muy bajos", satisfaciendo "necesidades legales insatisfechas que existen en todo el país, generalmente para personas de ingresos bajos y medios que no pueden permitirse los precios que cobran los abogados" (McGinnis & Pearce, 2014, p. 3054).

El futuro de la profesión jurídica dependerá, por tanto, de la capacidad de adaptación de los profesionales a un nuevo ecosistema donde las tareas rutinarias y mecánicas serán progresivamente asumidas por sistemas de inteligencia artificial, mientras que aquellas funciones que requieren un alto grado de especialización, creatividad, empatía o interacción humana seguirán a cargo de los abogados humanos.

## 4. MARCO JURÍDICO

### 4.1. EL REGLAMENTO EUROPEO DE IA

El avance de la inteligencia artificial (IA) ha llevado a la Unión Europea (UE) a desarrollar un marco normativo integral que garantice su aplicación de manera segura, ética y alineada con los derechos fundamentales. Como respuesta a esta necesidad, el Parlamento Europeo y el Consejo aprobaron el Reglamento (UE) 2024/1689 el 13 de junio de 2024. Su objetivo principal es establecer normas comunes para evitar la fragmentación legislativa entre los Estados miembros y facilitar la libre circulación de sistemas de IA dentro del territorio de la UE. No obstante, se prevé la posibilidad de imponer restricciones cuando sea necesario para la protección de valores esenciales como la democracia, el Estado de Derecho y los derechos humanos (Parlamento Europeo y Consejo de la UE, 2024).

Este reglamento introduce un modelo de regulación basado en el nivel de riesgo, permitiendo categorizar los sistemas de IA según su impacto potencial. Los principios fundamentales que rigen esta normativa son los siguientes:

#### 1. **Inteligencia Artificial centrada en el ser humano y fiable:**

El reglamento establece que todo sistema de IA debe garantizar seguridad, transparencia y supervisión humana, asegurando su uso responsable en beneficio de la sociedad y el respeto a los derechos fundamentales (Parlamento Europeo y Consejo de la UE, 2024).

#### 2. **Clasificación según el nivel de riesgo:**

- 1) IA de riesgo inaceptable: Se prohíben aquellos sistemas de IA que representen una amenaza directa para los derechos fundamentales, como:
  - A. Tecnologías de manipulación subliminal que alteren el comportamiento humano sin el consentimiento de la persona.
  - B. Sistemas que exploten vulnerabilidades de individuos en función de su edad, discapacidad u otras condiciones de vulnerabilidad.

- C. Modelos de puntuación social similares a los aplicados en algunos países, que puedan afectar el acceso a servicios o derechos (Parlamento Europeo y Consejo de la UE, 2024).
- 2) IA de alto riesgo: Esta categoría abarca los sistemas utilizados en sectores críticos como la salud, la educación, las infraestructuras esenciales, el empleo y la justicia. Para estos casos, se exige el cumplimiento de estrictos requisitos en materia de transparencia, supervisión humana y evaluaciones de impacto para mitigar posibles riesgos (Parlamento Europeo y Consejo de la UE, 2024).
  - 3) IA de riesgo limitado: Los sistemas que presentan un impacto moderado en los derechos fundamentales se incluyen en esta categoría. Se requiere que estos cumplan con normas de transparencia específicas, como la obligación de informar a los usuarios cuando interactúan con *Chatbots*, *Deepfakes* y sistemas de recomendación. De esta manera, se garantiza que las personas sean conscientes de que están tratando con tecnologías de IA (Parlamento Europeo y Consejo de la UE, 2024).
  - 4) IA de riesgo mínimo: incluye aplicaciones de uso cotidiano que presentan un riesgo bajo, como filtros de spam y sistemas de recomendación en plataformas digitales. Estos sistemas no están sujetos a regulaciones estrictas debido a su impacto limitado en la sociedad (Parlamento Europeo y Consejo de la UE, 2024).

Para garantizar el cumplimiento del reglamento, se han establecido mecanismos de supervisión y gobernanza, incluyendo la creación de la Oficina Europea de Inteligencia Artificial y el Consejo de IA. Estas entidades serán responsables de la aplicación de la normativa y de la revisión periódica de la clasificación de riesgos, así como de las prácticas prohibidas (Parlamento Europeo y Consejo de la UE, 2024).

El reglamento entrará en vigor de manera progresiva. Algunas disposiciones se comenzaron a aplicar a partir del 2 de febrero de 2025, mientras que la normativa completa será obligatoria a partir del 2 de agosto de 2026 (Parlamento Europeo y Consejo de la UE, 2024).

## 4.2. INICIATIVAS ESTRATÉGICAS EUROPEAS PARA EL DESARROLLO DE LA IA

Junto a la evolución del marco normativo fundamental, la UE ha puesto en marcha varias acciones estratégicas orientadas a potenciar el entorno de inteligencia artificial en Europa. Estas medidas complementan la aplicación del Reglamento de IA y buscan garantizar que Europa mantenga la competitividad necesaria mientras implementa un enfoque ético y centrado en el ser humano.

Entre las iniciativas más destacadas se encuentra la ampliación de los objetivos de la Empresa Común de Informática de Alto Rendimiento Europea (EuroHPC). En mayo de 2024, el Consejo de la UE alcanzó un acuerdo político sobre un Reglamento para expandir las capacidades de EuroHPC hacia el desarrollo de la inteligencia artificial, incorporando un objetivo adicional: "desarrollar y explotar factorías de IA en apoyo de un ecosistema de inteligencia artificial en la Unión" (Consejo de la Unión Europea, 2024, p.1).

Este consenso político constituye un momento decisivo en la planificación europea para el fortalecimiento de competencias en inteligencia artificial, pues facilita que la actual estructura de supercomputación se ajuste particularmente a los requerimientos de las nuevas empresas europeas del sector. En palabras del vicepresidente de Valonia, Willy Borsus, la colaboración ha sido fundamental para Europa, permitiendo alcanzar importantes logros como el desarrollo de superordenadores de nivel mundial, y ahora el objetivo es aprovechar la capacidad de estas avanzadas tecnologías mediante sistemas de inteligencia artificial confiables (Consejo de la Unión Europea, 2024)

Los organismos conocidos como "factorías de IA" funcionarán como entidades proveedoras de infraestructura para servicios de supercomputación diseñados expresamente para aplicaciones de inteligencia artificial. Esta potencia de cálculo es determinante para desarrollar y entrenar modelos de IA a gran escala, un elemento que hasta ahora ha situado a las empresas europeas en desventaja frente a competidores internacionales.

La iniciativa introduce modificaciones importantes para garantizar que las nuevas capacidades tecnológicas sean accesibles para el tejido empresarial europeo. El acuerdo del Consejo "garantiza que las actividades cubiertas por las factorías de IA ofrezcan oportunidades de acceso justo a los superordenadores optimizados para la IA, abriéndolos a un mayor número de usuarios públicos y privados" (Consejo de la Unión Europea, 2024, p.1).

Un elemento notable es la referencia directa a *startups* y *PYMEs* como beneficiarios prioritarios, creando un punto centralizado de acceso para simplificar la utilización de los servicios. Esta apertura de acceso a tecnologías esenciales muestra un plan estratégico que busca impulsar a las empresas nuevas y pequeñas en el campo de la inteligencia artificial, asegurando que no solo las grandes empresas tecnológicas puedan innovar y desarrollar nuevas soluciones.

El Reglamento establece también mecanismos de financiación específicos, permitiendo que "las entidades anfitrionas puedan recibir una contribución financiera de la Unión que cubra hasta el 50% de los costes de adquisición de los superordenadores de IA y hasta el 50% de sus costes de funcionamiento" (Consejo de la Unión Europea, 2024, pp. 1-2). Este modelo de cofinanciación busca maximizar el impacto de la inversión pública y crear un efecto multiplicador a través de colaboraciones público-privadas.

La transferencia de la propiedad de los superordenadores a las instituciones que los albergan después de cinco años de su aceptación constituye un estímulo importante para que más entidades participen en el programa, estableciendo así un modelo viable para su administración en el largo plazo.

Esta acción forma parte de una estrategia más amplia que la presidenta de la Comisión Europea, Ursula von der Leyen, dio a conocer durante su discurso sobre el estado de la Unión en 2023. La iniciativa marca un avance importante hacia la independencia tecnológica de Europa en un campo que se considera fundamental para el futuro económico y la posición geopolítica del continente.

El Reglamento establece claramente que los superordenadores de IA deben utilizarse "principalmente para crear, ensayar, evaluar y validar modelos de formación de IA de uso general a gran escala y aplicaciones emergentes de IA, así como para seguir desarrollando soluciones de IA en la Unión" (Consejo de la Unión Europea, 2024, p.2), alineando así los recursos tecnológicos con las prioridades definidas en la estrategia europea para la IA.

#### 4.3. PROTECCIÓN DE LOS DERECHOS FUNDAMENTALES Y PROHIBICIONES CLAVE

El Reglamento (UE) 2024/1689 instaura una serie de limitaciones y prohibiciones rigurosas con el objetivo de salvaguardar los derechos fundamentales y combatir aplicaciones incorrectas de la IA. Estas disposiciones están diseñadas para contrarrestar amenazas vinculadas con la manipulación del comportamiento, prácticas discriminatorias, sistemas de vigilancia generalizada y otras acciones que puedan comprometer la dignidad de las personas y su derecho a la privacidad (Parlamento Europeo y Consejo de la Unión Europea, 2024).

#### **4.3.1. Manipulación subliminal y explotación de vulnerabilidades**

El Reglamento prohíbe los sistemas de IA diseñados para manipular el comportamiento humano de manera subliminal o engañosa, especialmente cuando puedan causar daños físicos, psicológicos o económicos a los individuos. Según el considerando 29, estos sistemas pueden utilizar "estímulos de audio, imagen o vídeo que las personas no pueden percibir, u otras técnicas manipulativas que socaven su autonomía y capacidad de elección" (Parlamento Europeo y Consejo de la Unión Europea, 2024, p. 8).

Además, se prohíbe la explotación de vulnerabilidades de ciertos grupos debido a su edad, discapacidad o situación socioeconómica. El reglamento señala que "estos sistemas pueden utilizarse para alterar sustancialmente el comportamiento de una persona de manera que le cause un perjuicio considerable" (Parlamento Europeo y Consejo de la Unión Europea, 2024, p. 8).

### **4.3.2. Identificación biométrica y vigilancia masiva**

El empleo de tecnologías que permiten reconocer personas a distancia mediante características físicas en entornos públicos y en tiempo real está sujeto a estrictas normativas por su considerable afectación a la privacidad. La legislación no permite su implementación generalizada, si bien contempla situaciones excepcionales como la localización de personas extraviadas o la anticipación de actos terroristas, requiriendo en todos los casos una aprobación previa por parte de las autoridades judiciales. (Parlamento Europeo y Consejo de la Unión Europea, 2024).

Además, se prohíbe que las autoridades adopten decisiones con efectos jurídicos adversos basadas exclusivamente en los resultados de salida de un sistema de identificación biométrica, evitando así la posibilidad de errores o sesgos que puedan perjudicar a las personas identificadas (Parlamento Europeo y Consejo de la Unión Europea, 2024).

### **4.3.3. Categorización biométrica basada en datos sensibles**

La normativa prohíbe la implementación de tecnologías de inteligencia artificial que agrupen o etiqueten a individuos basándose en atributos protegidos y sensibles, tales como origen étnico, creencias religiosas, preferencias sexuales o tendencias políticas, a través de la utilización de información biométrica como impresiones digitales o sistemas de identificación facial. (Parlamento Europeo y Consejo de la Unión Europea, 2024).

Esta prohibición no afecta a procesos lícitos de etiquetado de conjuntos de datos que están en conformidad con las leyes actuales y persiguen objetivos concretos, como la clasificación de imágenes según el color de ojos o de cabello en investigaciones policiales (Parlamento Europeo y Consejo de la Unión Europea, 2024).

### **4.3.4. Evaluación de impacto sobre los derechos fundamentales**

El Reglamento impone la obligación de realizar una evaluación de impacto sobre los derechos fundamentales antes de desplegar sistemas de IA de alto riesgo en ámbitos sensibles como la sanidad, la educación o los servicios públicos. Según el considerando 96, el objetivo de la evaluación es determinar los riesgos específicos que puedan afectar a los derechos de las personas y definir las medidas necesarias para mitigarlos (Parlamento Europeo y Consejo de la Unión Europea, 2024).

Esta evaluación se aplica especialmente a entidades públicas y empresas privadas que prestan servicios de interés público, como educación, salud y asistencia social, garantizando que el uso de IA en estos sectores no genere discriminación o vulneraciones de derechos (Parlamento Europeo y Consejo de la Unión Europea, 2024).

#### **4.3.5. Uso de IA en la toma de decisiones sobre servicios esenciales**

Cuando la IA ayuda a decidir quién recibe beneficios importantes como seguridad social, atención médica, ayudas para vivienda o préstamos, la ley europea considera estos sistemas como "de alto riesgo". Esto significa que deben seguir reglas estrictas de transparencia y justicia, ya que pueden afectar seriamente la vida de las personas (Parlamento Europeo y Consejo de la Unión Europea, 2024).

Existe preocupación porque estos sistemas podrían repetir formas antiguas de discriminación. Por ejemplo, un programa de IA podría rechazar préstamos a ciertos grupos de personas injustamente. Por eso, la ley exige que estos sistemas pasen por revisiones cuidadosas antes de usarse (Parlamento Europeo y Consejo de la Unión Europea, 2024).

El Reglamento 2024/1689 crea límites claros para asegurar que la IA se use de forma ética y respete los derechos básicos de todos. Prohíbe la manipulación oculta, controla estrictamente el uso de identificación biométrica (como reconocimiento facial) y protege contra la discriminación.

También obliga a evaluar el impacto de los sistemas de alto riesgo antes de usarlos y a realizar auditorías para verificar que funcionen correctamente.

Para vigilar que todos cumplan estas reglas, se creará un Consejo Europeo de Inteligencia Artificial que supervisará cómo se aplican las normas en todos los países miembros.

El reglamento también establece sanciones económicas severas en caso de incumplimiento (Parlamento Europeo y Consejo de la Unión Europea, 2024):

- a) Hasta 35 millones de euros o el 7% del volumen de negocio global en infracciones muy graves.
- b) Hasta 15 millones de euros o el 3% del volumen de negocio por incumplimientos menos severos.

- c) Hasta 7,5 millones de euros o el 1% del volumen de negocio por información incorrecta o engañosa sobre los sistemas de IA (Parlamento Europeo y Consejo de la Unión Europea, 2024).

#### 4.4. MARCO NORMATIVO ESPAÑOL COMPLEMENTARIO

España ha desarrollado un marco regulatorio específico para la supervisión y el desarrollo responsable de la inteligencia artificial, anticipándose a la plena implementación de la normativa europea en esta materia. El principal instrumento normativo es el Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial (AESIA).

La creación de la AESIA responde a un mandato establecido previamente en la Ley 22/2021, de 28 de diciembre, de Presupuestos Generales del Estado para el año 2022, que en su Disposición adicional centésimo trigésima autorizaba al Gobierno a "impulsar una Ley para la creación de la Agencia Española de Supervisión de Inteligencia Artificial" (Real Decreto 729/2023, p. 5). Esta iniciativa fue reforzada posteriormente por la Ley 28/2022, de 21 de diciembre, de fomento del ecosistema de las empresas emergentes, que preveía específicamente la creación de esta entidad.

La AESIA se configura como una entidad de derecho público "con personalidad jurídica pública, patrimonio propio y autonomía en su gestión" (Real Decreto 729/2023, p. 9). Su establecimiento está fundamentado en la obligación que, según el preámbulo del Real Decreto, deriva de "lo dispuesto en la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial)" (Real Decreto 729/2023, p. 5), que establece para los Estados miembros la necesidad de seleccionar una "autoridad nacional de supervisión".

Entre los objetivos fundamentales de la AESIA, según el artículo 4 de su Estatuto, se encuentran:

- a) "La minimización de los riesgos que puede suponer el uso de esta nueva tecnología, el adecuado desarrollo y potenciación de los sistemas de inteligencia artificial" (Real Decreto 729/2023, p. 10).

- b) La eliminación o reducción de "los riesgos para la integridad, la intimidad, la igualdad de trato y la no discriminación, en particular entre mujeres y hombres, y demás derechos fundamentales que pueden verse afectados por el mal uso de los sistemas" (Real Decreto 729/2023, p. 10).
- c) "La concienciación, divulgación y promoción de la formación, y del desarrollo y uso responsable, sostenible y confiable de la inteligencia artificial" (Real Decreto 729/2023, p. 10).
- d) "El fomento de entornos reales de prueba de los sistemas de inteligencia artificial, para reforzar la protección de los usuarios y evitar sesgos discriminatorios" (Real Decreto 729/2023, p. 10).

Un aspecto destacable del marco normativo español es la competencia otorgada a la AESIA para realizar tareas de supervisión, asesoramiento, concienciación y formación dirigidas a entidades tanto públicas como privadas, lo que refleja un enfoque holístico en la gestión de la IA (Real Decreto 729/2023). Adicionalmente, la Agencia tiene la capacidad para "emitir informes sobre el impacto generado por un sistema de inteligencia artificial" y para realizar "evaluaciones de impacto para que los algoritmos involucrados en la toma de decisiones (...) tengan en cuenta criterios de minimización de sesgos, transparencia y rendición de cuentas" (Real Decreto 729/2023, p. 20).

La regulación española muestra una especial preocupación por la perspectiva de género en el desarrollo de la IA, estableciendo en su artículo 10.1.c) el compromiso de incorporar "el principio de igualdad de oportunidades entre mujeres y hombres" y de promover "la realización de evaluaciones de impacto capaces de identificar posibles sesgos discriminatorios" (Real Decreto 729/2023, p. 12).

Este marco normativo se complementa con mecanismos de certificación voluntaria para sistemas de IA, establecidos en el artículo 26, que permiten a la AESIA "diseñar y publicar el programa o esquema de certificación de los sistemas de inteligencia artificial" (Real Decreto 729/2023, p. 20), adelantándose así a los requisitos que eventualmente establecerá la normativa europea.

## 5. PROBLEMÁTICA ÉTICA DE LA IA JURÍDICA

### 5.1. TRANSPARENCIA Y EXPLICABILIDAD DE LOS SISTEMAS DE IA EN CONTEXTOS JURÍDICOS

La implementación de sistemas de inteligencia artificial en entornos jurídicos plantea desafíos particulares en cuanto a transparencia y explicabilidad. Según el análisis de Pere Simón Castellano en su trabajo "Taxonomía de las Garantías Jurídicas en el Empleo de los Sistemas de Inteligencia Artificial", estos conceptos constituyen elementos fundamentales, pero no suficientes por sí solos para garantizar la protección de derechos.

En contextos jurídicos, la transparencia algorítmica se refiere a "la publicidad activa y al derecho de acceso a la información relativa a los algoritmos y, más concretamente, la que hace referencia a los propósitos y fines, a la estructura y diseño, a las acciones subyacentes, a las bases de datos empleadas y a disponibilidad de la información relativa a los mecanismos de participación humana" (Simón Castellano, 2023, p. 180). Esta información permite mantener la confianza en el sistema y capacita a los afectados para impugnar sus resultados.

El documento establece una diferencia conceptual clave entre transparencia y explicabilidad. Mientras la primera se centra en hacer accesible información sobre el sistema en general, la explicabilidad busca hacer "inteligible los resultados de los sistemas de inteligencia artificial en un caso concreto, así como la comprensibilidad de los datos, procesos y comportamientos asociados a la decisión específica que se proyecta sobre los individuos" (Simón Castellano, 2023, p. 184).

La problemática se hace evidente en casos como COMPAS y BOSCO mencionados en el documento, donde "la imposibilidad de impugnar una decisión automatizada derivaba de una falta de transparencia y explicabilidad que, a su vez, tenía su origen en un problema mayor" (Simón Castellano, 2023, p. 187).

El autor señala que "la explicabilidad es una propiedad que contribuye a crear una falsa sensación de seguridad que actúa a modo de lastre contra la innovación tecnológica" (Simón Castellano, 2023, p. 185), especialmente cuando se prioriza la comprensibilidad sobre la efectividad.

Finalmente, Simón Castellano (2023) concluye que para contextos jurídicos, "el conjunto de las garantías jurídicas estudiadas pretende, en definitiva, un objetivo mayor: asegurar que ningún ciudadano sea objeto de una decisión automatizada sin poder recurrir, replicar o defenderse frente a tal decisión" (p. 193).

## 5.2. SEGOS ALGORÍTMICOS Y DISCRIMINACIÓN EN LA IA JURÍDICA

La investigación de Buolamwini y Gebru (2018) mostró que la tecnología de reconocimiento facial no funciona igual para todos. Su estudio "*Gender Shades*" encontró que estos sistemas cometen muchos más errores cuando analizan mujeres de piel oscura que cuando identifican hombres de piel clara.

Al probar varios sistemas comerciales en 2017, descubrieron que el sistema de IBM se equivocaba hasta en un 34.7% de los casos al identificar mujeres de piel oscura. En cambio, el sistema Face++ solo fallaba un 0.8% de las veces con hombres de piel clara (Buolamwini & Gebru, 2018).

Esto es especialmente preocupante porque la policía usa esta tecnología para identificar sospechosos. Aunque estos sistemas no deciden quién es culpable, ayudan a determinar a quién investigar (Buolamwini & Gebru, 2018). Además, según un estudio citado por las autoras, los afroamericanos ya tienen más probabilidades de ser sometidos a este tipo de vigilancia (Garvie et al., 2016, como se citó en Buolamwini & Gebru, 2018).

El problema empieza con los datos usados para crear estos sistemas. Las investigadoras encontraron que las bases de imágenes estándar contienen principalmente personas de piel clara: 79.6% en la base IJB-A y 86.2% en Adience (Buolamwini & Gebru, 2018).

Para mejorar esta situación, Buolamwini y Gebru crearon un nuevo conjunto de datos llamado "*Pilot Parliaments Benchmark*" con una representación más equilibrada de diferentes tonos de piel, usando la escala *Fitzpatrick* que utilizan los dermatólogos (Buolamwini & Gebru, 2018).

Las autoras concluyen que las empresas deben evaluar sus sistemas considerando tanto el género como el tono de piel simultáneamente. Solo así podrán crear tecnología de reconocimiento facial que sea justa y confiable para todos (Buolamwini & Gebru, 2018).

### 5.3. GOBERNANZA ÉTICA DE LA IA JURÍDICA

La supervisión ética de los sistemas de inteligencia artificial en el campo legal es muy importante para asegurar que estas tecnologías se desarrollen de manera equilibrada, respetando los derechos básicos de las personas y los valores democráticos. Según lo que establece la Agencia Española de Supervisión de la Inteligencia Artificial (AESIA), hay varios elementos principales que forman parte de esta supervisión.

La visión española y europea sobre la IA se caracteriza por un enfoque humanista que coloca "a las personas y sus derechos en el centro" (Artero Muñoz et al., 2023, p. 33). Esta aproximación constituye una "posición de privilegio al potenciar a nivel global esta visión humanística, al margen de las tendencias y corrientes coyunturales" (Artero Muñoz et al., 2023, p. 33). El desarrollo de tecnologías de IA debe ser compatible "con los derechos de las personas, ética y respetuosa con el medio ambiente" (Artero Muñoz et al., 2023, p. 35).

La AESIA establece principios fundamentales que orientan cualquier desarrollo de IA, incluyendo aplicaciones jurídicas:

- a) Principio de no discriminación: Considerado "el origen y fundamento de la propia creación de la Agencia, en cuanto pretende eliminar o luchar contra los sesgos que, inevitablemente, están inmersos en los sistemas de IA desarrollados por los humanos" (Artero Muñoz et al., 2023, p. 41).
- b) Principio de transparencia: se basa en "la presentación de la información sobre las cuentas de manera precisa y completa, además de las obligaciones de buen gobierno" (Artero Muñoz et al., 2023, p. 41).
- c) Principio de autonomía: Orientado a "alcanzar los objetivos, gestionando sus propios medios" (Artero Muñoz et al., 2023, p. 41).
- d) Principio de independencia técnica: Fundamentado en "la alta capacitación y profesionalidad del personal al servicio de la Agencia, adaptado a la ética profesional" (Artero Muñoz et al., 2023, p. 41).

- e) Principio de cooperación interinstitucional: Entendido como "la búsqueda de sinergias en la colaboración con otras Administraciones Públicas, agentes e instituciones, públicas o privadas, nacionales e internacionales para el fomento del conocimiento en todos sus ámbitos" (Artero Muñoz et al., 2023, pp. 41-42).

Además, se establece una clasificación de los sistemas de IA según su nivel de riesgo, que resulta crucial para determinar el nivel de supervisión necesaria:

- a) Sistemas inaceptables: "Aquel que manipulara el comportamiento humano de forma subliminal o que explotara a las personas vulnerables, como un juego de azar que indujera a la adicción" (Artero Muñoz et al., 2023, p. 40).
- b) Sistemas de alto riesgo: "Aquel que tuviera un impacto directo sobre la vida o la salud de las personas, como un robot cirujano que operara a un paciente o un sistema de conducción autónoma que controlara un vehículo" (Artero Muñoz et al., 2023, p. 40).
- c) Sistemas de riesgo limitado: "Aquel que generara o modificara contenido de forma automática, como una aplicación de reconocimiento facial que creara un avatar personalizado o un asistente virtual que respondiera a las consultas de los usuarios" (Artero Muñoz et al., 2023, p. 40).
- d) Sistemas de riesgo mínimo: "Aquel que tuviera una finalidad lúdica o recreativa, como un juego de mesa que usara IA para crear desafíos o un filtro de realidad aumentada que cambiara el aspecto de una foto" (Artero Muñoz et al., 2023, p. 40).

Por otro lado, la gobernanza ética de la IA jurídica requiere de instrumentos concretos. La AESIA ha desarrollado diversos mecanismos de implementación:

- a) Entornos controlados de prueba (sandbox): "Como un entorno controlado de pruebas del Reglamento europeo IA, que se está desarrollando desde la SEDIA con objeto de operacionalizar los principios del Reglamento aplicables a los sistemas de inteligencia artificial de alto riesgo" (Artero Muñoz et al., 2023, p. 42).
- b) Marco de certificación voluntaria: "Para entidades privadas y fomentará la dinamización de casos de uso de IA para desarrollar soluciones que permitan transitar hacia modelos productivos y sociales sostenibles" (Artero Muñoz et al., 2023, p. 42).

- c) Observatorio de Impacto Social: Para analizar y determinar "el impacto social en la IA y creando de esta manera un observatorio de tendencias, a través del Observatorio de Impacto Social de los Algoritmos" (Artero Muñoz et al., 2023, p. 42).
- d) Think tank y centro de análisis: La AESIA prevé "convertirse en un potente think tank, centro de análisis, aportando su visión y su experiencia a los debates y las políticas públicas sobre esta materia" (Artero Muñoz et al., 2023, p. 43).

## 6. CONCLUSIONES

Al término de esta investigación sobre la interrelación entre inteligencia artificial y derecho, podemos formular un conjunto de reflexiones que sintetizan los hallazgos principales y ofrecen respuestas a los objetivos planteados inicialmente. Estas conclusiones abordan la transformación que la IA está produciendo en el ámbito jurídico, el marco normativo emergente, y los principios deontológicos necesarios para su implementación responsable.

Nuestro análisis ha revelado que la inteligencia artificial está modificando profundamente múltiples facetas de la administración de justicia. Desde la digitalización de expedientes hasta sistemas predictivos y herramientas decisorias, la tecnología redefine procesos que históricamente han dependido exclusivamente del juicio humano.

La gestión documental automatizada, implementada a través de plataformas como LexNET, ha significado un primer paso fundamental hacia la modernización judicial. El expediente electrónico no solo reduce la carga administrativa, sino que sienta las bases para análisis de datos más sofisticados que pueden identificar patrones relevantes en la información jurídica.

En las etapas de investigación, sistemas basados en IA facilitan la búsqueda, clasificación y análisis de evidencias, permitiendo a los operadores jurídicos procesar volúmenes de información que serían inabarcables mediante métodos tradicionales. Herramientas como Jurimetría y vLex Analytics proporcionan análisis predictivos que orientan estrategias procesales basadas en tendencias jurisprudenciales, alterando fundamentalmente la forma en que abogados evalúan casos y formulan argumentaciones.

La aplicación de la IA en los mecanismos decisorios presenta sus mayores potencialidades y riesgos. Por un lado, ofrece oportunidades para decisiones más consistentes y fundamentadas en análisis exhaustivos de precedentes. Por otro, casos como COMPAS en Estados Unidos evidencian que estos sistemas pueden perpetuar sesgos sociales preexistentes bajo una aparente objetividad técnica. La opacidad algorítmica plantea interrogantes fundamentales sobre garantías procesales como el derecho de defensa y la motivación de resoluciones.

Respecto a los Mecanismos Alternativos de Solución de Controversias, la reforma introducida por la Ley Orgánica 1/2025 encuentra en la IA una potencial aliada para su implementación efectiva. Sin embargo, como observamos en nuestra investigación, existen complejas cuestiones sobre legitimación, plazos e intervención humana que requieren soluciones específicas, particularmente en disputas societarias.

La figura del "juez robot" merece consideración especial por sus profundas implicaciones. Mientras experiencias internacionales como las de Estonia, China o Emiratos Árabes demuestran la viabilidad técnica de automatizar decisiones en casos específicos, nuestra investigación ha confirmado que la administración de justicia trasciende la mera aplicación lógica de normas. El juicio humano integra valoraciones contextuales, consideraciones éticas y comprensión de realidades sociales que difícilmente pueden codificarse algorítmicamente.

En cuanto al futuro profesional jurídico, concluimos que la IA reconfigurará dramáticamente el mercado laboral legal, pero no necesariamente mediante una sustitución total. Siguiendo el análisis de McGinnis y Pearce, identificamos una polarización donde tareas rutinarias serán progresivamente automatizadas, mientras funciones que requieren empatía, creatividad y juicio ético permanecerán predominantemente humanas. El modelo más prometedor parece ser una colaboración sinérgica entre profesionales y sistemas inteligentes.

Nuestra investigación ha verificado la existencia de un emergente marco regulatorio que busca equilibrar innovación tecnológica con protección de derechos fundamentales. El Reglamento Europeo (UE) 2024/1689 establece un paradigma basado en niveles de riesgo que proporciona tanto seguridad jurídica como flexibilidad adaptativa.

La clasificación de sistemas jurídicos como aplicaciones "de alto riesgo" reconoce acertadamente su impacto potencial sobre derechos esenciales, estableciendo requisitos estrictos de transparencia, supervisión humana y evaluaciones de impacto. Las prohibiciones absolutas para tecnologías de manipulación subliminal, explotación de vulnerabilidades y sistemas de puntuación social reflejan una aproximación principialista que prioriza dignidad humana y autonomía personal.

En el contexto español, la creación de la Agencia Española de Supervisión de Inteligencia Artificial (AESIA) mediante el Real Decreto 729/2023 representa un avance institucional significativo. Sus competencias para supervisión, asesoramiento y certificación constituyen un modelo interesante de gobernanza que podría inspirar a otras jurisdicciones. Particularmente innovadores resultan los "sandboxes" regulatorios que permiten experimentación controlada con nuevas aplicaciones.

Las iniciativas estratégicas europeas para fortalecer infraestructuras tecnológicas, como la ampliación de competencias de la Empresa Común de Informática de Alto Rendimiento, demuestran conciencia sobre la dimensión geopolítica de la IA. La soberanía tecnológica y el acceso equitativo a capacidades computacionales avanzadas emergen como condiciones necesarias para desarrollar soluciones alineadas con valores democráticos europeos.

Sin embargo, nuestro análisis también identifica desafíos significativos en la implementación efectiva de estos marcos normativos. La velocidad de evolución tecnológica contrasta con los ritmos legislativos tradicionales, creando potenciales desajustes. Asimismo, la complejidad técnica de sistemas avanzados de IA plantea retos importantes para verificación de cumplimiento y auditoría efectiva.

La perspectiva comparada revela aproximaciones regulatorias divergentes entre regiones. Mientras el modelo europeo prioriza salvaguardas preventivas y enfoque principialista, otras jurisdicciones favorecen aproximaciones más permisivas orientadas hacia innovación acelerada. Esta fragmentación podría generar desafíos para interoperabilidad de sistemas y protección transfronteriza de derechos.

Concluimos que un marco normativo efectivo requiere equilibrar precisión técnica con fundamentación ética, capacidad adaptativa con seguridad jurídica, y mecanismos vinculantes con instrumentos de soft law. La colaboración internacional y participación multisectorial resultan indispensables para desarrollar estándares que trasciendan fronteras nacionales, reflejando la naturaleza global de los retos planteados.

Nuestra investigación ha identificado un conjunto coherente de principios deontológicos que deben orientar el desarrollo e implementación de sistemas de IA en contextos jurídicos. Estos principios, lejos de constituir meras aspiraciones teóricas, pueden operacionalizarse en prácticas concretas que preserven valores democráticos y derechos humanos.

El principio de transparencia y explicabilidad emerge como fundamento ineludible. Los sistemas implementados en contextos jurídicos deben proporcionar explicaciones comprensibles sobre sus procesos decisorios, permitiendo a los afectados entender y potencialmente impugnar resultados. Esta transparencia debe extenderse tanto a los datos utilizados como a la lógica algorítmica, reconociendo que la opacidad socava la legitimidad del sistema judicial.

El principio de no discriminación adquiere particular relevancia ante la evidencia de sesgos algorítmicos documentada por investigadores como Buolamwini y Gebru. Los sistemas jurídicos deben incorporar mecanismos de detección y mitigación de sesgos, evaluaciones de impacto diferencial en poblaciones vulnerables, y auditorías periódicas que verifiquen resultados equitativos independientemente de características protegidas.

El principio de supervisión humana significativa constituye una salvaguarda esencial, especialmente en decisiones con impacto directo sobre derechos fundamentales. Nuestro análisis ha confirmado que la intervención humana no debe limitarse a una validación formal, sino que debe garantizar capacidad efectiva para comprender, revisar y eventualmente anular determinaciones algorítmicas. La rendición de cuentas requiere asignación clara de responsabilidades entre desarrolladores, implementadores y operadores de sistemas.

El principio de proporcionalidad tecnológica implica evaluar cuidadosamente qué procesos son susceptibles de automatización, considerando no solo eficiencia sino también valores procesales fundamentales. Concluimos que ciertos ámbitos jurídicos, por su complejidad ética o impacto potencial, deberían mantener mayor prevalencia de juicio humano, mientras otros procedimientos más estandarizados pueden beneficiarse de mayor automatización sin comprometer garantías esenciales.

El principio de inclusión y accesibilidad exige que la implementación de IA jurídica no genere exclusiones o barreras adicionales. Nuestro estudio ha identificado el riesgo de crear una "brecha digital jurídica" donde el acceso efectivo a la justicia dependa de alfabetización tecnológica o recursos económicos. Los sistemas diseñados deben considerar necesidades de grupos diversos, incluyendo personas con discapacidad, limitaciones tecnológicas o pertenecientes a comunidades marginadas.

Finalmente, el principio de sostenibilidad institucional requiere que la incorporación de IA no socave fundamentos democráticos del sistema judicial. Las tecnologías deben implementarse preservando valores como independencia judicial, contradicción procesal, y acceso a recursos efectivos, considerando impactos a largo plazo sobre confianza ciudadana e integridad institucional.

La operacionalización de estos principios requiere mecanismos concretos como certificaciones éticas, evaluaciones de impacto en derechos humanos, programas de alfabetización tecnológica para operadores jurídicos, y procesos participativos que incorporen perspectivas diversas en diseño e implementación. La ética no debe considerarse un componente separado o posterior al desarrollo tecnológico, sino integrada desde las etapas iniciales de concepción.

La respuesta a la pregunta central de nuestra investigación "¿Peligro u oportunidad?" trasciende dicotomías simplistas. La inteligencia artificial representa simultáneamente ambas realidades para el derecho, con potencial tanto para transformar positivamente la administración de justicia como para comprometer principios fundamentales del Estado de Derecho.

El análisis desarrollado confirma que nos encontramos en un momento definitorio donde las decisiones actuales sobre regulación, gobernanza e implementación determinarán la trayectoria futura de estos sistemas. La prospectiva hacia 2030 sugiere una integración creciente de tecnologías inteligentes en procesos jurídicos, exigiendo adaptaciones profundas en formación profesional, estructuras institucionales y marcos conceptuales.

Esta investigación ha cumplido sus objetivos fundamentales, ofreciendo un análisis detallado de las transformaciones en curso, examinando marcos normativos emergentes, y formulando principios deontológicos coherentes. Sin embargo, reconocemos que el carácter dinámico de estas tecnologías exigirá actualizaciones continuas y nuevos estudios que aborden desafíos específicos.

Concluimos que el camino responsable implica adoptar un enfoque reflexivo pero no tecnofóbico, que reconozca tanto oportunidades transformadoras como riesgos sustanciales. El éxito de la IA jurídica no debería medirse exclusivamente por indicadores de eficiencia o precisión técnica, sino fundamentalmente por su capacidad para fortalecer —nunca debilitar— los valores de justicia, igualdad y dignidad humana que constituyen la razón de ser del ordenamiento jurídico.

La integración de sistemas de inteligencia artificial en el derecho no representa meramente una evolución tecnológica, sino una transformación que interpela fundamentos filosóficos de nuestras instituciones jurídicas y concepciones de justicia. El verdadero desafío consiste en desarrollar un enfoque que aproveche el potencial de estas herramientas sin comprometer los principios humanistas que han guiado el desarrollo del derecho moderno, asegurando que la tecnología permanezca siempre al servicio de la persona y sus derechos fundamentales.

## 7. BIBLIOGRAFÍA

### 7.1. LEGISLACIÓN

España. (2025). *Ley Orgánica 1/2025, de 2 de enero, de medidas en materia de eficiencia del Servicio Público de Justicia*. Boletín Oficial del Estado. <https://www.boe.es/eli/es/lo/2025/01/02/1/con>

Parlamento Europeo y Consejo de la Unión Europea. (2024). *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 varios reglamentos y directivas*. Diario Oficial de la Unión Europea.

España. (2023). *Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial*. Boletín Oficial del Estado, 210, 18911. <https://www.boe.es/eli/es/rd/2023/08/22/729>

### 7.2. OBRAS DOCTRINALES

Alonso, A. (2023). La inteligencia artificial y la prospectiva. *AROEC: 11 Jornadas de Prospectiva Económica, Volumen Especial*, 1-8. <https://dialnet.unirioja.es/servlet/articulo?codigo=8789522>

Barona Vilar, S. (2024). Justicia con algoritmos e inteligencia artificial, ¿acuercando garantías y derechos procesales o liquidándolos? *Derechos y Libertades*, 51, 83-115. <https://doi.org/10.20318/dyl.2024.8584>

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1-15. <https://www.media.mit.edu/publications/gender-shades-intersectional-accuracy-disparities-in-commercial-gender-classification/>

Cárdenas Krenz, R. (2021). ¿Jueces robots? Inteligencia artificial y derecho. *Justicia & Derecho*, 4(2), 1-10. <https://doi.org/10.32457/rjyd.v4i2.1345>

de Luis García, E. (2023). *Justicia, inteligencia artificial y derecho de defensa*. IDP. *Revista d'Internet, Dret i Política*, 39, 1-12. <https://doi.org/10.7238/idp.v0i39.417164>

Galindo, F. (2019). ¿Inteligencia artificial y derecho? Sí, pero ¿cómo? *Revista Democracia Digital e Governo Eletrónico*, 2(18), 36-57. [https://zaguan.unizar.es/record/79533/files/texto\\_completo.pdf](https://zaguan.unizar.es/record/79533/files/texto_completo.pdf)

García-Villarrubia, M. (2025). Los mecanismos alternativos de resolución de controversias (MASC) y la litigación societaria. *Almacén de Derecho*. <https://almacenederecho.org/los-mecanismos-alternativos-de-resolucion-de-controversias-masc-y-la-litigacion-societaria>

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. [http://alvarestech.com/temp/deep/Deep%20Learning%20by%20Ian%20Goodfellow,%20Yosua%20Bengio,%20Aaron%20Courville%20\(z-lib.org\).pdf](http://alvarestech.com/temp/deep/Deep%20Learning%20by%20Ian%20Goodfellow,%20Yosua%20Bengio,%20Aaron%20Courville%20(z-lib.org).pdf)

Hodges, A. (2006). *Alan Turing: The logical and physical basis of computing*. ResearchGate. [https://www.researchgate.net/publication/336539669\\_Alan\\_Turing\\_the\\_logical\\_and\\_physical\\_basis\\_of\\_computing/link/5da51b5045851553ff920021/download?tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19](https://www.researchgate.net/publication/336539669_Alan_Turing_the_logical_and_physical_basis_of_computing/link/5da51b5045851553ff920021/download?tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19)

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577-586. <https://doi.org/10.1016/j.bushor.2018.03.007>

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15-25. <https://doi.org/10.1016/j.bushor.2018.08.004>

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, 27(4), 12-12. <https://doi.org/10.1609/aimag.v27i4.1904>

McGinnis, J. O., & Pearce, R. G. (2014). The Great Disruption: How machine intelligence will transform the role of lawyers in the delivery of legal services. *Fordham Law Review*, 82, 3041-3066. <https://ssrn.com/abstract=2436937>

Muñoz, A. A., de Toledo Rodríguez, C. F. R., & Medina, P. M. (2023). Agencia Española de Supervisión de la Inteligencia Artificial, la clave para un desarrollo tecnológico ético, justo y sostenible. *Revista española de control externo*, (7475), 32-45. <https://hdl.handle.net/20.500.14352/117670>

Penprase, B. E. (2018). The fourth industrial revolution and higher education. *Higher Education in the Era of the Fourth Industrial Revolution*, 10(1), 207-229. <https://doi.org/10.1007/978-981-13-0194-0>

Russell, S., & Norvig, P. (1995). *Artificial intelligence: A modern approach*. Prentice Hall.

Simón Castellano, P. (2023). Taxonomía de las garantías jurídicas en el empleo de los sistemas de inteligencia artificial. <https://doi.org/10.5944/rdp.117.2023.37929>

### 7.3. RECURSOS DE INTERNET

Consejo de la Unión Europea. (2024, 23 de mayo). *Inteligencia artificial: el Consejo alcanza un acuerdo político sobre el uso de la supercomputación para el desarrollo de la IA* [Comunicado de prensa 432/24]. <https://www.consilium.europa.eu/es/press/press-releases/2024/05/23/ai-council-reaches-political-agreement-on-the-use-of-super-computing-for-ai-development/>

Osornio, A. (2023, noviembre 30). *Año uno de ChatGPT: así ha cambiado el mundo desde la llegada del chatbot de OpenAI*. *Wired en Español*. <https://es.wired.com/articulos/ano-uno-de-chatgpt-asi-ha-cambiado-el-mundo-desde-la-llegada-del-chatbot-de-openai>