

La Inteligencia artificial en la justicia penal: Límites del operador jurídico en la resolución de problemas penales

Resumen

El avance de la inteligencia artificial (IA) ha impactado de manera significativa el ámbito jurídico, particularmente en la administración de justicia penal, donde su incorporación plantea importantes desafíos en términos de garantías procesales, autonomía judicial y legitimidad de las decisiones. El presente artículo tiene como objetivo determinar los límites del operador de justicia frente al uso de la inteligencia artificial durante la resolución de procesos penales. Para ello, se analizan las herramientas tecnológicas disponibles, se examina el valor jurídico-crítico de la IA en la toma de decisiones y se proponen criterios que delimiten su uso adecuado. La investigación se desarrolla bajo un enfoque cualitativo, con un diseño de tipo documental, basado en la revisión sistemática de fuentes doctrinales, normativas y académicas. Se concluye que, si bien la IA constituye una herramienta útil para optimizar la gestión judicial, su utilización debe estar sujeta a límites claros que garanticen el respeto por los derechos fundamentales y la preservación del juicio humano como eje central del proceso penal.

Palabras clave: Inteligencia artificial, justicia penal, operador de justicia, debido proceso, toma de decisiones judiciales, límites jurídicos.

Abstract

The advancement of artificial intelligence (AI) has significantly impacted the legal field, particularly in the administration of criminal justice, where its incorporation poses important challenges in terms of procedural safeguards, judicial autonomy, and the

legitimacy of decisions. This article aims to determine the limits of the justice system's role in the use of artificial intelligence during criminal proceedings. To this end, the available technological tools are analyzed, the legal and critical value of AI in decision-making is examined, and criteria are proposed to define its appropriate use. The research is conducted using a qualitative approach, with a documentary design based on a systematic review of doctrinal, regulatory, and academic sources. It concludes that, while AI is a useful tool for optimizing judicial management, its use must be subject to clear limits that guarantee respect for fundamental rights and the preservation of human judgment as the central axis of the criminal process.

Keywords: Artificial intelligence, criminal justice, justice system, due process, judicial decision-making, legal limits.

Sumario

1. Herramientas de inteligencia artificial en la justicia penal: aplicaciones, alcances y riesgos para el operador jurídico. 2. Valor jurídico-crítico de la inteligencia artificial en la toma de decisiones penales: análisis de confiabilidad, legitimidad y debido proceso. 3. Límites del uso de la inteligencia artificial en la resolución de procesos penales: criterios para una implementación garantista.

1. Introducción

La incorporación de sistemas de inteligencia artificial (IA) en la cotidianidad humana representa una ampliación de las posibilidades en el actuar delictivo, toda vez que facilita la automatización de conductas, la sofisticación de modalidades sumada a la dificultad en la identificación de los responsables. De este modo, modalidades como los fraudes automatizados, la suplantación de identidad mediante tecnología de generación de contenido sintético y la manipulación de información son la muestra de que la IA si bien representa un instrumento que puede asistir de manera eficiente las actividades cotidianas, profesionales o académicas, también puede ser implementada como instrumento eficaz para la criminalidad contemporánea (Morales Moreno, 2021).

Desde este punto de vista, la inteligencia artificial representa tensiones significativas relacionadas con los elementos estructurales del delito. Así, la noción de acción comprendida como un comportamiento humano voluntario, se ve atravesada por la intervención de sistemas autónomos con la capacidad de toma de decisiones sin intervención directa e inmediata de una persona. Este panorama conlleva a repensar los fundamentos de la imputación penal, específicamente en lo que respecta a la autoría y la participación (Oieda, 2026)

Así pues, la concepción de imputación objetiva atraviesa complicaciones cuando se trata de determinar el nexo entre la conducta humana y el resultado lesivo siempre que la IA interviene como intermediaria o ejecutora del hecho. La opacidad algorítmica, como característica de la mayoría de IA, complica la previsibilidad del resultado y, en consecuencia, la determinación del riesgo jurídicamente desaprobado. Lo anterior, conlleva incertidumbre en relación con los criterios tradicionales de atribuciones de responsabilidad (Gallegos, 2022).

Ahora bien, en relación con la culpabilidad los criterios resultan igualmente atravesados, la exigencia de dolo o culpa como presupuestos de la responsabilidad penal se ve tensionada cuando los resultados lesivos derivan de decisiones adoptadas por sistemas de inteligencia artificial, cuya programación puede incluir márgenes de aprendizaje autónomo. En estos casos, surge la interrogante sobre si es posible atribuir responsabilidad al programador, al usuario o incluso considerar nuevas formas de responsabilidad jurídica (Puig, 2003).

A nivel normativo, tanto Colombia como México presentan avances limitados en la regulación específica del uso de la inteligencia artificial en contextos delictivos. En ambos países, la respuesta del derecho penal ha consistido principalmente en la adaptación de tipos penales tradicionales a nuevas formas de criminalidad digital, lo que evidencia la existencia de vacíos normativos frente a las particularidades de la IA. Esta situación plantea la necesidad de evaluar la suficiencia de los marcos legales actuales (Gélvez et al, 2021).

1.1. Planteamiento del problema

En Colombia, si bien se han realizado esfuerzos recientes para incorporar conductas relacionadas con tecnologías emergentes dentro del ámbito penal, la regulación de la inteligencia artificial sigue siendo incipiente. Por su parte, en México, aunque existen normas orientadas a combatir la violencia digital y otros delitos tecnológicos, no se observa una regulación integral que aborde de manera específica los riesgos asociados al uso delictivo de la IA (Quiñones et al, 2025).

Este panorama normativo fragmentado refleja una problemática más amplia relacionada con la capacidad del derecho penal para adaptarse a los avances tecnológicos. La velocidad con la que evoluciona la inteligencia artificial contrasta con la rigidez

inherente a los sistemas jurídicos, generando una brecha que puede ser aprovechada por actores delictivos. En este sentido, resulta fundamental analizar si las respuestas actuales son suficientes o si se requiere una reformulación más profunda del marco penal.

Desde la perspectiva de la política criminal, la inteligencia artificial también plantea interrogantes sobre las estrategias más adecuadas para prevenir y sancionar su uso indebido. La adopción de medidas exclusivamente punitivas podría resultar insuficiente frente a un fenómeno de carácter global y altamente dinámico. Por ello, se hace necesario considerar enfoques integrales que incluyan prevención, regulación tecnológica y cooperación internacional (Jímenez, 2024).

En este contexto, surge el problema de investigación que orienta el presente estudio: la insuficiencia de los marcos dogmáticos y normativos tradicionales del derecho penal para abordar adecuadamente las conductas delictivas resueltas con mediación de inteligencia artificial en Colombia y México. Este problema se manifiesta en la dificultad para imputar responsabilidad penal, la existencia de vacíos legales y la ausencia de políticas criminales articuladas frente a este fenómeno.

1.2. Objetivos

1.2.1. General

Determinar los límites del operador de justicia frente al uso de la inteligencia artificial durante la resolución de procesos penales.

1.2.2. Específicos

- Analizar las herramientas que puede emplear el operador de justicia durante la resolución de procesos penales, características, aplicaciones y posibles riesgos dentro del sistema jurídico

- Identificar el valor jurídico-crítico admisible respecto de la inteligencia artificial durante la solución de problemas jurídicos penales, evaluando su confiabilidad, legitimidad y compatibilidad con los principios del debido proceso.
- Proponer los límites aceptables para el uso de la inteligencia artificial durante la resolución de procesos penales, estableciendo criterios que orienten su implementación de manera responsable y acorde con el marco normativo vigente.

1.3. Diseño Metodológico

Desde el punto de vista metodológico, la investigación se desarrolla bajo un enfoque cualitativo, el cual resulta pertinente para analizar fenómenos complejos relacionados con la interacción entre derecho y tecnología. Este enfoque permite una comprensión profunda del problema a partir de la interpretación de fuentes teóricas y normativas. En cuanto a la naturaleza de la investigación, esta se enmarca dentro del tipo de investigación básica jurídica, en la medida en que su propósito no radica en la aplicación inmediata de soluciones prácticas, sino en la construcción y sistematización de conocimiento teórico orientado a la comprensión de los límites del operador jurídico frente al uso de la inteligencia artificial en el ámbito penal.

En desarrollo de lo anterior, se emplean fuentes de carácter primario y secundario, dentro de las cuales se incluyen normas jurídicas vigentes en Colombia y México, jurisprudencia relevante, doctrina especializada en derecho penal y tecnología, así como artículos científicos indexados y documentos institucionales. Desde el punto de vista lógico, la investigación se apoya principalmente en el método deductivo, al partir de postulados generales del derecho penal y principios constitucionales para analizar su

aplicabilidad frente a casos mediados por inteligencia artificial; sin perjuicio de la utilización complementaria de elementos del método inductivo, en la medida en que se examinan situaciones particulares y desarrollos tecnológicos específicos para derivar conclusiones generales sobre la necesidad de delimitación jurídica de su uso.

El tipo de investigación es documental, dado que se fundamenta en la revisión y análisis de literatura especializada, incluyendo artículos académicos, doctrina jurídica, informes institucionales y normativas relacionadas con el uso de la inteligencia artificial en la justicia. El alcance del estudio es descriptivo y analítico, en tanto no solo se busca caracterizar el fenómeno objeto de estudio, sino también examinar sus implicaciones y proponer soluciones teóricas que contribuyan al debate académico y jurídico.

Así pues, la técnica principal de recolección de información es la revisión bibliográfica, mediante la cual se seleccionan y analizan fuentes relevantes para la construcción del marco conceptual. Con ello, se pretende aportar elementos que permitan delimitar el uso de la inteligencia artificial en la justicia penal, garantizando el respeto por los derechos fundamentales y la centralidad del operador jurídico en la toma de decisiones.

La revisión documental se centra en la recopilación y análisis de legislación vigente en Colombia y México, así como en estudios académicos, artículos científicos y documentos institucionales que abordan la relación entre inteligencia artificial y derecho penal. Este proceso permite construir un marco teórico sólido y fundamentado que sustenta el análisis comparado.

Finalmente, el presente estudio busca contribuir al debate académico y jurídico sobre la necesidad de adaptar el derecho penal a los desafíos que plantea la inteligencia artificial. A través del análisis comparado entre Colombia y México, se pretende ofrecer

elementos que orienten la formulación de respuestas normativas y de política criminal más eficaces, acordes con las transformaciones tecnológicas contemporáneas.

2. Herramientas de inteligencia artificial en la justicia penal: aplicaciones, alcances y riesgos para el operador jurídico

La integración de la inteligencia artificial en la administración de justicia representa un cambio paradigmático que busca transformar la eficiencia, celeridad y precisión de los procesos judiciales. Definida como una disciplina científica dedicada a la creación de programas informáticos que ejecutan operaciones comparables a las de la mente humana, como el aprendizaje y el razonamiento lógico, la IA ha dejado de ser una promesa futurista para convertirse en una herramienta de uso común (Serna, 2025). En el ámbito penal, este fenómeno se manifiesta a través de algoritmos que procesan grandes volúmenes de información, permitiendo a los operadores jurídicos gestionar casos con una rapidez antes inimaginable, aunque su implementación plantea debates profundos sobre la ética y la legalidad del sistema (Román & Domínguez, 2025).

Desde una perspectiva técnica, es fundamental distinguir entre los diferentes tipos de IA que impactan el sistema penal. La IA "estrecha" o especializada está diseñada para tareas específicas, como el reconocimiento facial o la detección de fraudes, mientras que la IA "general" aspira a replicar la flexibilidad del pensamiento humano, un nivel que aún se encuentra en desarrollo (Guerrero Gualdrón, 2024). En la práctica criminal, la mayoría de las herramientas actuales operan bajo modelos matemáticos de procesamiento de datos, como el machine learning y el deep learning, que permiten a las máquinas aprender patrones a partir de la información suministrada, mejorando su capacidad predictiva con el tiempo (Lara Novoa, 2024).

El alcance de estas herramientas comienza en la fase de prevención e investigación criminal, donde la denominada "policía predictiva" o predictive policing ha cobrado relevancia. Al utilizar técnicas cuantitativas para identificar objetivos y lugares con mayor riesgo delictivo, las agencias de seguridad pueden optimizar el despliegue de sus recursos (Miró Llinares, 2018). Estas aplicaciones, basadas en teorías de criminología ambiental, permiten pasar de una postura reactiva a una proactiva, identificando concentraciones desproporcionadas de criminalidad, conocidas como hot spots, mediante el análisis de datos históricos (Polo Herazo & Padilla Maussa, 2025).

En el contexto colombiano, herramientas como Watson de IBM han sido fundamentales para la Fiscalía General de la Nación. Esta plataforma funciona como un sistema de Big Data que logra hallar relaciones entre noticias criminales y bases de datos históricas, facilitando el esclarecimiento de patrones delictivos complejos que podrían pasar inadvertidos para un analista humano (Serna, 2025). La capacidad de Watson para procesar información no estructurada, como documentos en PDF o redes sociales, permite a los investigadores asociar casos y realizar análisis de contexto que fortalecen la persecución penal (Espinoza et al, 2024).

Otra aplicación significativa es el sistema Prisma (Perfil de Riesgo de Reincidencia para la Solicitud de Medidas de Aseguramiento), utilizado por los fiscales para evaluar la probabilidad de que un individuo vuelva a delinquir. A través de algoritmos que analizan el historial criminal y las capturas previas, Prisma genera un documento digital que sirve de apoyo en las audiencias de solicitud de medida de aseguramiento (Serna, 2025). No obstante, su uso se concibe como una herramienta de apoyo y no como una base única para la decisión, buscando que el juez cuente con información objetiva y racional antes de decretar una detención preventiva (Polo Herazo & Padilla Maussa, 2025).

En el ámbito de las altas cortes, el sistema PretorIA ha revolucionado la selección de expedientes de tutela en la Corte Constitucional de Colombia. Dada la inmensa cantidad de acciones de amparo que recibe esta corporación, PretorIA utiliza algoritmos para clasificar, priorizar y determinar hechos jurídicamente relevantes, simplificando un proceso que antes requería múltiples etapas manuales (Morales et al, 2025). Al reducir los tiempos de gestión y facilitar la detección de casos de interés nacional, esta herramienta mejora la eficiencia del órgano de cierre, aunque siempre bajo la supervisión humana de los magistrados (Polo Herazo & Padilla Maussa, 2025).

Internacionalmente, herramientas como COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) han marcado el camino y la polémica en la valoración del riesgo de reincidencia. Utilizada en Estados Unidos, COMPAS informa decisiones sobre libertad condicional y regímenes penitenciarios basándose en modelos actuariales que agrupan datos sociodemográficos y comportamentales (Miró Llinares, 2018). A pesar de su utilidad declarada para reducir sesgos subjetivos de los jueces, ha sido objeto de críticas por su opacidad y la falta de transparencia en sus algoritmos, protegidos por leyes de propiedad intelectual (Acosta, 2025).

La aplicación de la IA generativa, como ChatGPT, también ha comenzado a filtrarse en la labor de los operadores jurídicos para la investigación legal y la redacción de documentos. Su capacidad para generar resúmenes, redactar proyectos de fallos y ofrecer explicaciones sobre conceptos legales complejos la convierte en un asistente virtual potente (Serna, 2025). Sin embargo, su uso por parte de jueces en Colombia, como en el caso de la Sentencia T 323 de 2024, ha generado alertas sobre la necesidad de establecer criterios claros que impidan que estas herramientas reemplacen la racionalidad humana (Cruz & León, 2023).

El alcance legal de estas tecnologías en Colombia fue abordado precisamente por la Corte Constitucional en la Sentencia T 323 de 2024, la cual establece que el uso de la IA es aceptable siempre que se respeten los derechos fundamentales y el debido proceso. La Corte enfatiza que la IA debe ser un apoyo para la investigación y la consulta de precedentes, pero la interpretación de la ley y la toma de decisiones finales deben permanecer exclusivamente en manos del juez (Valeria, 2022). Este pronunciamiento subraya que el uso indiscriminado de la IA podría comprometer la legitimidad de las decisiones judiciales si se delegan funciones racionales indelegables (Díaz, 2024).

Uno de los riesgos más críticos identificados en el uso de la IA penal es el denominado "sesgo algorítmico". Este fenómeno ocurre cuando los algoritmos, nutridos con datos que reflejan prejuicios históricos, producen resultados discriminatorios contra grupos específicos por razones de raza, género o condición socioeconómica (Serna, 2025). En la justicia penal, un algoritmo sesgado podría recomendar penas más severas o medidas cautelares innecesarias para minorías, perpetuando injusticias estructurales bajo una apariencia de objetividad matemática (Miró Llinares, 2018).

Asimismo, surge la preocupación por la privacidad y la protección de datos personales. El uso de IA en la investigación criminal requiere la gestión masiva de información, lo que pone a prueba los límites del derecho fundamental a la intimidad y al habeas data (Serna, 2025). La recolección indiscriminada de metadatos y la vigilancia sistemática mediante algoritmos pueden derivar en una "tecnovigilancia" que, si no cuenta con un marco legal robusto y supervisión judicial, vulnera las garantías constitucionales de los ciudadanos (Miró Llinares, 2018).

Otro riesgo notable es el de las "alucinaciones" de la IA, que se refiere a la generación de información falsa o inexacta por parte de los modelos generativos. En un

proceso judicial, una alucinación sobre un precedente inexistente o una interpretación errónea de los hechos podría viciar la motivación de una sentencia, comprometiendo el derecho de defensa y la seguridad jurídica (Serna, 2025). Por ello, la verificación rigurosa de la fiabilidad de la información arrojada por la IA es un deber funcional ineludible para cualquier operador jurídico (Díaz, 2024).

La opacidad o el problema de la "caja negra" representa otro desafío ético y procedimental. Muchos sistemas de IA operan de una forma que impide conocer exactamente cómo se llegó a una conclusión determinada, lo que choca con la obligación judicial de motivar las decisiones de manera transparente y comprensible (Guerrero Gualdrón, 2024). Si el operador jurídico no puede explicar la lógica detrás de una recomendación algorítmica, se debilita la posibilidad de que las partes ejerzan un derecho de contradicción efectivo (Moreno et al, 2024).

La garantía del juez natural e imparcial también se ve amenazada si se permite que la IA influya de manera desproporcionada en la mente del juzgador. La independencia judicial exige que las sentencias sean el resultado de un razonamiento autónomo, libre de influencias externas no fundadas en el derecho (Polo Herazo & Padilla Maussa, 2025). Una dependencia excesiva de las herramientas predictivas podría reducir la supervisión humana y la discrecionalidad necesaria para atender a la singularidad y humanidad de cada caso concreto (Serna, 2025).

Ante estos riesgos, el papel del juez penal debe evolucionar hacia un perfil con nuevas competencias técnicas y éticas. No basta con el dominio tradicional de la norma; el juez del siglo XXI debe poseer "discernimiento tecnológico" para comprender los alcances y límites de los sistemas automatizados que utiliza (Guerrero Gualdrón, 2024). Esto incluye el desarrollo de competencias cognitivas para identificar riesgos epistémicos y

competencias instrumentales para manejar herramientas digitales de manera crítica (Guerrero Gualdrón, 2024).

La formación judicial debe, por tanto, incorporar el componente de la "alfabetización digital" de forma transversal. Los operadores deben estar capacitados para evaluar la fiabilidad de las pruebas generadas por algoritmos y para leer técnicamente informes periciales automatizados (Guerrero Gualdrón, 2024). Sin esta preparación, el funcionario judicial corre el riesgo de convertirse en un operador acrítico de herramientas que no comprende totalmente, lo que pone en peligro la equidad del sistema (Polo Herazo & Padilla Maussa, 2025).

El principio de transparencia cobra una relevancia renovada en este ecosistema digital. Los jueces tienen el deber de revelar a las partes si han utilizado herramientas de IA en la construcción de sus decisiones, permitiendo que estas sean objeto de debate y control (Polo Herazo & Padilla Maussa, 2025). La transparencia no solo fortalece la confianza ciudadana, sino que asegura que los beneficios de la tecnología se compartan de forma equilibrada en la sociedad, sin abusos ni exclusiones (Serna, 2025).

A pesar de los desafíos, los beneficios en términos de eficiencia son innegables. Herramientas como Prometea han demostrado que tareas que tradicionalmente tomaban días pueden reducirse a minutos, permitiendo que el talento humano se enfoque en los aspectos más sustanciales y complejos del derecho (Polo Herazo & Padilla Maussa, 2025). Esta agilización contribuye a combatir la congestión judicial, un problema crónico que afecta el acceso a una tutela judicial efectiva en muchos países (Serna, 2025).

El uso de la IA también permite una mayor objetividad en áreas como la valoración del riesgo, siempre que los algoritmos se diseñen y auditen correctamente. Al basarse en grandes volúmenes de datos empíricos, la IA puede ayudar a superar prejuicios subjetivos e

inconscientes que a menudo afectan a los tomadores de decisiones humanos (Miró Llinares, 2018). El reto consiste en encontrar un equilibrio donde la máquina potencie la capacidad humana sin deshumanizar el proceso penal (Guerrero Gualdrón, 2024).

Desde el punto de vista de la infraestructura, la implementación de la IA en la justicia penal enfrenta retos logísticos significativos. Se requiere una inversión sostenida en redes de internet estables y dispositivos electrónicos modernos que sean compatibles con los nuevos programas (Polo Herazo & Padilla Maussa, 2025). Además, la brecha tecnológica entre diferentes despachos judiciales puede generar desigualdades en la calidad de la administración de justicia, un aspecto que debe ser gestionado mediante políticas públicas inclusivas (Guerrero Gualdrón, 2024).

La regulación ética de los algoritmos es una necesidad imperativa que precede a su uso extensivo. No se debe avanzar en la automatización sin un marco normativo que establezca criterios de responsabilidad, rendición de cuentas y respeto a la dignidad humana (Serna, 2025). Los legisladores y organismos internacionales deben trabajar en auditorías periódicas que aseguren que la IA no se convierta en un instrumento de manipulación o engaño (Serna, 2025).

El futuro de la profesión jurídica no es la desaparición ante la IA, sino su transformación. Los abogados y jueces seguirán siendo esenciales para interpretar normas, argumentar en juicios y tomar decisiones éticas fundamentadas en el juicio crítico y la empatía (Serna, 2025). La IA será un socio estratégico que automatizará lo repetitivo, permitiendo que el operador jurídico recupere su rol creativo y deliberativo frente a la innovación (Guerrero Gualdrón, 2024).

La justicia penal basada en IA debe ser una "justicia centrada en la persona". Esto implica que cada avance tecnológico debe evaluarse bajo la lupa de los derechos

fundamentales y el bienestar social (Guerrero Gualdrón, 2024). La tecnología no es un fin en sí misma, sino un medio para alcanzar un sistema penal más justo, equitativo y transparente (Serna, 2025).

Es fundamental que la integración de estas herramientas cuente con la participación activa de los propios funcionarios judiciales y de la academia. Solo a través de un diálogo interdisciplinario entre científicos de datos y juristas se podrán diseñar arquitecturas de IA que sean éticamente robustas y legalmente legítimas (Miró Llinares, 2018). El monitoreo constante y la capacidad de rectificación ante errores algorítmicos son garantías esenciales para este tránsito tecnológico (Guerrero Gualdrón, 2024).

En conclusión, las herramientas de IA ofrecen un potencial inmenso para fortalecer la justicia penal, aportando precisión y celeridad a la labor del operador jurídico. Sin embargo, su éxito dependerá de la capacidad del sistema para mitigar riesgos como el sesgo, la opacidad y la deshumanización. El juez penal debe erigirse como el guardián de la racionalidad humana y los principios constitucionales, asegurando que, en la era de los algoritmos, la justicia siga siendo una virtud humana al servicio de la dignidad de las personas (Serna, 2025).

3. Valor jurídico-crítico de la inteligencia artificial en la toma de decisiones penales: análisis de confiabilidad, legitimidad y debido proceso

Incorporar la inteligencia artificial en la toma de decisiones penales, eccede la mera adopción tecnológica con el fin de situarse en el centro de un debate jurídico crítico en relacion con la existencia de la función judicial. Dicho fenómeno plantea un cambio paradigmático donde la eficiencia procesal presenta tensiones con las garantías fundamentales, obligando a reevaluar si un sistema fundamentado en algoritmos puede

cumplir con los estándares de justicia humana (Beltran & Quesada, 2025). La riqueza crítica de esta tecnología reside en la forma en que su implementación repercute en la relación entre el Estado, el juez y el ciudadano procesado (Miró Linares, 2018).

En concordancia, los criterios de confiabilidad de los sistemas de IA representan uno de los cuestionamientos fundamentales, en razón a que estos modelos operan fundamentalmente como procesadores matemáticos de datos y no como entidades dotadas de comprensión jurídica. Se entiende que las máquinas, a diferencia de la mente humana no poseen conciencia ni sensibilidad ética, lo que limita su capacidad para la interpretación en el contexto social y humano de un delito (Serna, 2025). Así pues, la confiabilidad de un resultado del algoritmo se encuentra ligada completamente tanto a la calidad de los datos entregados como a la supervisión permanente del operario jurídico (Miró Linares, 2018).

Ahora bien, un riesgo latente para la confiabilidad técnica son las denominadas “alucinaciones” de la IA, fenómeno en el que los sistemas generativos producen respuestas equivocadas, incoherentes o sin sentido con una aparente veracidad. En un contexto penal, una alucinación en relación con un precedente o una prueba podría viciar la motivación de una providencia, comprometiendo la seguridad jurídica (Basquez, 2020). La Corte Constitucional de Colombia ha advertido que, frente a la incapacidad de la IA para presentar respuestas acertadas, pueden presentarse soluciones descabelladas que el juez debe filtrar rigurosamente (García et al, 2024).

Así, la legitimidad de las decisiones apoyadas en IA se sostiene en la premisa de que la tecnología debe ser un auxiliar y nunca implementarse a modo de reemplazo de la racionalidad del ser humano. Al respecto, el sistema jurídico colombiano ha prohibido a partir de la Sentencia T-323 (2024), la sustitución del juez por la máquina, en tanto la función de administración de justicia es una labor jurisdiccional que no puede ser delegada

(Serna, 2025). De este modo, la legitimidad se pierde en el momento en que el juez cede su capacidad de juicio y discreción frente a la comodidad del resultado automatizado (Polo Herazo & Padilla Maussa, 2025).

Ahora bien, desde el punto de vista epistemológico, la legitimidad del uso de la IA depende, además, del *discernimiento tecnológico* del magistrado. El juez penal moderno debe contar con competencias que le posibiliten la comprensión de los alcances y las limitaciones de los sistemas automatizados para la garantía de un ejercicio ético de la función judicial (Guerrero Gualdrón, 2024). Sin esta preparación, la judicatura incurre en la posibilidad de validar ciegamente procesos de *caja negra* que son imposibles para el escrutinio de las partes (Miró Linares, 2018).

De este modo, la opacidad algorítmica puede amenazar las condiciones de debido proceso, cuando los criterios que utiliza la IA para llegar a una conclusión o son públicos es decir corresponden a propiedad intelectual de empresas privadas. Esta ausencia de transparencia imposibilita que el procesado tenga conocimiento de la lógica de la decisión, lo que anula la garantía de contradicción y el derecho de defensa (Miró Linares, 2018). Si las partes no tienen la disposición de debatir la base del razonamiento algorítmico, el proceso se puede tornar arbitrario y carente de las garantías mínimas exigidas por el Estado Social de Derecho (Polo Herazo & Padilla Maussa, 2025).

Así pues, la garantía de que el juez actúe de forma natural e imparcial es otro factor importante para analizar. La independencia judicial requiere que las sentencias correspondan al resultado de un análisis humano libre de influencias externas no fundamentadas en el derecho (Polo Herazo & Padilla Maussa, 2025). Una dependencia excesiva de herramientas predictivas de riesgo podría sesgar la imparcialidad del juzgador, que podría verse condicionado por la presunta autoridad técnica del software cuando sea

necesario decidir, teniendo en cuenta que se trata de asuntos relacionados con la integridad y libertad de los individuos (Miró Linares, 2018).

En consecuencia, la discriminación algorítmica comporta un desafío mayúsculo para la equidad del sistema penal. Los algoritmos de IA tienen la capacidad de reproducir o ampliar consideraciones de origen prejuicioso presentes en los datos con los que fueron entrenados, lo que trae consigo resultados injustos contra grupos vulnerables por razones de raza, género o condición socioeconómica. En este sentido, la IA puede recomendar medidas cautelares o penas más severas basándose en patrones de discriminación estructural, lo que volaría frontalmente el principio de igualdad (Serna, 2025).

En relación con el debido proceso probatorio exige que sea el juez quien revise la legalidad, utilidad y pertinencia de cada medio de convicción. Delegar la valoración de las pruebas a un sistema automatizado compromete esta garantía, toda vez que la IA carece de la capacidad de realizar un razonamiento lógico-humano para interpretar los hechos en su singularidad (Polo Herazo & Padilla Maussa, 2025). El juez debe actuar como un filtro crítico frente a las pruebas generadas por IA, asegurando que estas cumplan con estándares de fiabilidad exigidos por la ley penal (Guerrero Guadrón, 2024).

La tensión entre la exactitud (accuracy) y la equidad (fairness) es un desencuentro ético central en la IA penal. Si bien incluir variables como el género o la étnia pueden incrementar la precisión predictiva de una herramienta de riesgo, su uso plantea dudas sobre si resulta justo responsabilizar un individuo por características que no dependen de su voluntad (Miró Linares, 2018). Esta práctica puede desdibujar el principio de culpabilidad, que exige que la sanción se base en el hecho cometido y no en la pertenencia a un grupo demográfico específico (Miró Linares, 2018).

De cualquier modo, la transparencia representa un deber funcional que la Corte Constitucional ha sido aumentando a criterio de validez procesal. El juez que utiliza IA debe revelar que herramientas ha empleado y en qué medida su motivación se compone de aportes de la IA (Polo Herazo & Padilla Maussa, 2025). Por tanto, ocultar el uso de la tecnología no solo es un acto de deslealtad procesal, sino que compromete la integridad de la justicia al impedir que las partes ejerzan un control de veracidad sobre los resultados obtenidos (Serna, 2025).

El valor jurídico crítico de la IA pasa, además, por la protección del derecho a la intimidad y el habeas data. El uso masivo de metadatos e información personal para nutrir sistemas de predicción pone a prueba los límites del poder de vigilancia del Estado (Serna, 2025). Es imperativo que el proceso de recopilación de datos para fines preventivos criminal no se torne en una tecnovigilancia indiscriminada que afecte el núcleo esencial de la privacidad de los ciudadanos (Miró Linares, 2018).

Así, la legitimidad social del sistema penal se sostiene en la confianza ciudadana en que las decisiones son tomadas por humanos con responsabilidad ética. Si la percepción pública es que la justicia ha sido relegada a las máquinas frías e impersonales, se afecta la función resocializadora y pedagógica de la pena (Guerrero Gualdrón, 2024). La justicia, como fenómeno cultural y normativo, requiere que el infractor sea consciente del reproche social a través de una autoridad legítima y humana.

En cuanto a la automatización, se entiende que el juez debe contar con la capacidad de identificar riesgos epistémicos, lo que incluye la capacidad para la evaluación de la robustez metodológica y la presentación de riesgos que puedan invalidar el resultado. Sin que medie este escrutinio de carácter técnico y jurídico el operador judicial se transforma en

un espectador de la función de la IA, por lo que pierde su rol de guardián de las garantías constitucionales (Guerrero Gualdrón, 2024).

El debido proceso, por su parte, implica el derecho a una justicia flexible que atienda a la excepcionalidad de cada caso. La automatización del derecho tiende a expandir la regla general en detrimento de la excepción, lo que puede perjudicar los intereses de las minorías en relación con los criterios de diversidad (Serna, 2025). Una aplicación de carácter rígido y estandarizado de la ley fundamentada en algoritmos eliminaría la discrecionalidad necesaria para que el juez administre justicia con equidad y humanidad (Serna, 2025).

Además, la rendición de cuentas representa también un reto jurídico de carácter fundamental, en caso de que un sistema de IA induzca a un error judicial con consecuencias lesivas, debe haber claridad sobre la responsabilidad tanto del operario como del desarrollador (Serna, 2025). La justicia no debe ampararse en el *fallo técnico* para la justificación de la vulneración de derechos fundamentales, por lo que el control humanitario debe ser siempre el eje central de la decisión (Polo Herazo & Padilla Maussa, 2025).

De otro lado, la confiabilidad de la IA penal se ve reforzada cuando el diseño de las herramientas cuenta con la participación de juristas y científicos sociales, y no solo de ingenieros. Solo un enfoque de carácter interdisciplinario permite que los algoritmos repseten principios como la proporcionalidad y la mínima intervención. La tencología, debe ser, por tanto, diseñada desde el comienzo bajo principios de *privacidad por diseño* y *ética por diseño* para que sea considerada legítima en un entorno judicial (Guerrero Gualdrón, 2024).

Así, en el marco de la Sentencia T-323 de 2024 se concluye que el uso de la IA representa una realidad irreversible, pero condicionada al repeto irrestricto de la dignidad humana. La Corte destaca que la tecnología debe entenderse y usarse como socio estratégico que aumente la eficiencia sin que se vea comprometida la esencia racional del fallo judicial (Serna, 2025). El valor de la IA, por tanto, resulta instrumental y nunca sustantivo en la arquitectura del debido proceso (Polo herazo & Padilla Maussa, 2025).

En concordancia, la confiabilidad externa depende también de que los Estados establezcan marcos regulatorios sólidos antes de la implementación generalizada de dichas herramientas. La ausencia de norma específica en Colombia, genera un vacío jurídico que dificulta el control gubernamental y judicial de la IA (Serna, 2025). Siempre que no existan auditorias periódicas y estándares de calidad para los algoritmos, su uso en la justicia penal debe ser altamente cauteloso y limitado (Guerrero Gualdrón, 2024).

Críticamente, la IA puede incentivar un derecho penal de autor encubierto en datos actuariales, al predecir futuras conductas en base a estadísticas de grupos, se corre el riesgo de sancionar a la persona por quien es o a que grupo pertenece y no por lo que hizo (Miró Linares, 2018). Este retroceso hacia modelos preventivos desproporcionados debe ser evitado a partir de la reafirmación del principio de culpabilidad por el hecho (Miró Linares, 2018).

El debido proceso, se fortalece cuando la IA se implementa para tareas de tipo administrativo y documental que liberan al juez de cargas repetitivas, permitiéndole centrarse en el análisis sustancial del caso (Guerrero Gualdrón, 2024). La tecnología es más legítima cuando agiliza el acceso a la justicia y reduce la congestión, siempre que el acto creativo y deliberativo de juzgar permanezca en la esfera humana (Sena, 2025).

Así las cosas, el equilibrio entre la eficiencia tecnológica y las garantías humanas es el principal desafío de la justicia contemporánea. El valor jurídico que puede llegar a aportar el uso de la IA proviene de su capacidad para otorgar información con criterios de objetividad que apoye la superación de los posibles sesgos cognitivos del juez, pero solo si esta información se somete al escrutinio permanente del razonamiento humano (miro Linares, 2018). Lo anterior, quiere decir que la maquina debe informar, pero es el humano quien debe tomar las decisiones asumiendo la responsabilidad moral que conlleva la acción de juzgar.

En última instancia, la legitimidad de la IA dentro del espectro de la justicia penal se encuentra intrínsecamente relacionada de forma interdependiente a su alineación con los tratados internacionales de derechos humanos. Por ende, no debe existir algoritmo por avanzado que sea, que supere o esté por encima de la presunción de inocencia o del derecho a un juicio justo (debido proceso) (Guerrero Gualdrón, 2024). El control judicial, debe ser, por tanto, el baluarte que impida que la eficiencia del algoritmo reemplace a la justicia material.

Para cerrar el presente apartado, debe mencionarse que el análisis jurídico crítico de la IA en la toma de decisiones penales evidencia que la confiabilidad y legitimidad que deben asistirle no son automáticas, sino que se construyen a través de un debido proceso de transparencia y que se enfoque en las personas, es decir con carácter humano. El juez penal debe, por tanto, actuar como líder ético de cara a la innovación, para asegurar que la tecnología cumpla con potenciar el funcionamiento burocrático de la justicia sin que esto implique deshumanizarla y que el imperio de la ley siga estando por encima de la función de los algoritmos.

4. Límites del uso de la inteligencia artificial en la resolución de procesos penales: criterios para una implementación garantista

La implementación de la inteligencia artificial en el sistema de justicia penal no puede entenderse como una carta blanca para la automatización, sino que debe estar sujeta a límites infranqueables derivados del respeto a la dignidad humana y las garantías procesales. En el contexto de un Estado de derecho, la búsqueda de la eficiencia judicial a través de algoritmos nunca puede justificar el sacrificio de los derechos fundamentales de los sujetos procesales (Yáñez García-Bernalt, 2024). Por tanto, establecer criterios para una implementación garantista implica reconocer que la tecnología es un medio instrumental y no un fin en sí mismo, cuya validez depende de su subordinación a los mandatos constitucionales (Serna, 2025).

El límite primordial en la resolución de procesos penales es el principio de no sustitución de la racionalidad humana por la máquina. Según lo establecido en la Sentencia T-323 de 2024 de la Corte Constitucional de Colombia, las funciones jurisdiccionales que requieren razonamiento lógico, interpretación de pruebas y motivación de decisiones son estrictamente indelegables (Polo Herazo & Padilla Maussa, 2025). Delegar el acto de juzgar a un sistema automatizado no solo desnaturaliza la función judicial, sino que vulnera la garantía del juez natural, quien debe ser un ser humano capaz de comprender el espíritu de las normas y la singularidad de cada caso concreto (Guerrero Gualdrón, 2024).

Desde un enfoque internacional, el Reglamento Europeo de Inteligencia Artificial (RIA) establece un marco de límites basado en el riesgo, prohibiendo sistemas que atenten

directamente contra la integridad y libertad de las personas. Entre las prácticas prohibidas se encuentran la clasificación biométrica para inferir datos sensibles o la evaluación indiscriminada de individuos basada en comportamientos sociales que conduzcan a tratos perjudiciales (Reglamento UE 2024/1689). En el ámbito penal, este reglamento considera de "alto riesgo" a los sistemas destinados a ayudar a las autoridades judiciales a investigar e interpretar hechos, exigiendo para ellos una supervisión humana obligatoria y estándares rigurosos de ciberseguridad (Yáñez García-Bernalt, 2024).

La transparencia y la explicabilidad se erigen como criterios esenciales para una implementación garantista, permitiendo que las decisiones apoyadas en IA sean rastreables y comprensibles tanto para las partes como para la ciudadanía. La Comisión Iberoamericana de Ética Judicial subraya que el uso de algoritmos de "caja negra", donde el procesamiento de datos es opaco, es incompatible con un servicio público de justicia democrático (Dictamen CIEJ, 2023). Un proceso justo exige que el justiciable sepa cuándo está interactuando con un sistema de IA y cuáles son las bases lógicas que fundamentan una recomendación algorítmica, garantizando así el derecho a la contradicción (Polo Herazo & Padilla Maussa, 2025).

El debido proceso se ve comprometido cuando el "secreto empresarial" de las desarrolladoras de software impide el acceso al código fuente del algoritmo utilizado en el juicio. En sistemas como el estadounidense, el caso *State v. Loomis* evidenció la tensión entre la propiedad intelectual y el derecho del acusado a conocer la información que sustenta su nivel de riesgo (Nounou, 2024). No obstante, la doctrina garantista sostiene que en el proceso penal debe primar la publicidad de las herramientas utilizadas, ya que sin transparencia absoluta es imposible que la defensa cuestione la fiabilidad del sistema o detecte errores metodológicos (Serventich, 2022).

Un criterio garantista fundamental es la mitigación de los sesgos algorítmicos para proteger el derecho a la igualdad y la no discriminación. Los algoritmos pueden reproducir y amplificar prejuicios históricos presentes en los datos de entrenamiento, castigando a grupos vulnerables por factores sociodemográficos ajenos a su conducta individual (Miró Llinares, 2018). Para evitar un "derecho penal de autor" encubierto por la tecnología, los sistemas de IA no deben basar sus predicciones únicamente en rasgos de personalidad o perfiles grupales, sino que deben estar limitados por el principio de culpabilidad por el hecho cometido (Reglamento UE 2024/1689; Miró Llinares, 2018).

La protección del derecho a la intimidad y el habeas data impone límites estrictos al tratamiento masivo de datos personales en la investigación criminal. La denominada "tecnovigilancia", que incluye el reconocimiento facial masivo y el análisis predictivo de comportamientos en espacios públicos, debe ser excepcional y responder a los principios de necesidad y proporcionalidad (Yáñez García-Bernalt, 2024). La recolección indiscriminada de metadatos para alimentar sistemas de IA puede vulnerar el núcleo esencial de la privacidad si no existe un control judicial previo que legitime la intromisión del Estado en la vida digital de los ciudadanos (Miró Llinares, 2018).

El respeto a la presunción de inocencia exige que las herramientas predictivas de riesgo no sean utilizadas como prueba de culpabilidad ni para predeterminar la mente del juez. Instrumentos como COMPAS o Prisma, que evalúan la probabilidad de reincidencia, corren el riesgo de ser percibidos como infalibles, generando un "prejuicio automatizado" que sitúa al investigado en una posición de desventaja (Nounou, 2024). Una implementación garantista requiere que estos resultados sean considerados únicamente como informes auxiliares de apoyo y que el juez realice una valoración crítica independiente basada en las pruebas del proceso (Yáñez García-Bernalt, 2024).

La supervisión humana efectiva es el baluarte que impide que la eficiencia técnica desplace a la justicia material. No basta con una revisión formal de los resultados de la IA; el operador jurídico debe ejercer un liderazgo ético y deliberativo, siendo capaz de apartarse de la sugerencia algorítmica si esta choca con los principios constitucionales (Guerrero Gualdrón, 2024). La justicia penal debe seguir siendo una actividad de origen humano, donde la empatía, el juicio moral y la comprensión del contexto social guíen la aplicación de la ley, facultades que las máquinas, por definición, no poseen (Dictamen CIEJ, 2023; Serna, 2025).

La fiabilidad técnica es un requisito de validez procesal que debe ser verificado mediante auditorías externas e independientes. El uso de IA en la valoración probatoria, como el análisis de microexpresiones faciales para detectar mentiras, ha sido cuestionado por carecer en muchos casos de consenso científico y por el riesgo de "alucinaciones" tecnológicas (Cantos Pardo, 2024). Para que un sistema sea legítimo en el ámbito judicial, debe someterse a pruebas de "frugabilidad" y validación que certifiquen su precisión, evitando que pseudociencias algorítmicas ingresen al proceso bajo una apariencia de tecnicismo avanzado (Cantos Pardo, 2024; Serna, 2025).

El acceso a la justicia también se ve limitado por la brecha digital, lo que exige criterios de inclusión para los colectivos vulnerables. La automatización de trámites y el uso de asistentes virtuales pueden generar barreras para personas con bajos recursos o nula formación tecnológica, profundizando la desigualdad ante la ley (Nounou, 2024). Por tanto, la digitalización de la justicia penal debe ser acompañada de mecanismos que garanticen la asistencia letrada y la orientación judicial humana para quienes no puedan acceder de forma equitativa a las herramientas digitales (Yáñez García-Bernalt, 2024).

La rendición de cuentas (accountability) es un límite ético que obliga a identificar responsables ante fallos técnicos que vulneren derechos. En un sistema garantista, no es aceptable escudarse en el error del software para justificar una detención arbitraria o una sentencia errónea; el Estado y los desarrolladores deben responder por la integridad de los sistemas empleados (Gutiérrez, 2020). La creación de agencias reguladoras, similares a las de protección de datos, podría ser una solución para supervisar y suspender el uso de algoritmos defectuosos en la administración de justicia (Yáñez García-Bernalt, 2024).

El principio de especialidad en las diligencias de investigación tecnológica debe aplicarse rigurosamente al uso de la IA. La intromisión en la privacidad mediante algoritmos no puede ser genérica ni "a ciegas", sino que debe estar vinculada a la investigación de un delito concreto y debidamente justificada ante la autoridad judicial (Yáñez García-Bernalt, 2024). El uso de la IA para una vigilancia sistemática preventiva sin una sospecha fundada transgrede los límites de la proporcionalidad y convierte al Estado en un vigilante omnipresente que afecta el libre desarrollo de la personalidad (Miró Llinares, 2018).

La motivación de las sentencias constituye una garantía sustancial del debido proceso que no puede ser simplificada por la IA. Aunque la tecnología puede redactar borradores o resumir expedientes, el juez tiene el deber funcional de expresar de manera clara y lógica las razones de su decisión, revelando qué parte de su razonamiento se apoyó en IA y verificando la veracidad de esa información (Polo Herazo & Padilla Maussa, 2025). La falta de una motivación humana coherente hace que la decisión sea ilegítima y susceptible de nulidad (Serna, 2025).

El derecho de defensa se fortalece cuando se permite a los abogados litigantes cuestionar la "evidencia algorítmica" en igualdad de armas. La defensa debe contar con los

recursos técnicos para impugnar una identificación facial o un reporte predictivo, lo que exige que el Estado proporcione peritos expertos en informática jurídica (Yáñez García-Bernalt, 2024). Una implementación garantista proscribiera cualquier escenario donde el resultado del algoritmo sea considerado una verdad absoluta e incuestionable en el juicio oral (Serventich, 2022).

La formación y alfabetización digital de los jueces penales es un criterio preventivo contra el uso acrítico de la tecnología. Los operadores deben poseer competencias cognitivas e instrumentales que les permitan comprender el funcionamiento de los sistemas que utilizan, identificando sesgos y riesgos epistémicos antes de que estos influyan en el fallo (Guerrero Gualdrón, 2024). Sin una capacitación técnica y ética transversal, el juez corre el riesgo de convertirse en un espectador pasivo del desarrollo tecnológico, cediendo su autonomía a los ingenieros de datos (Dictamen CIEJ, 2023).

El diálogo interdisciplinario es indispensable para diseñar arquitecturas de IA que respeten el derecho penal. Los algoritmos no deben ser creados únicamente por científicos de datos, sino por equipos que incluyan a juristas y expertos en ética para asegurar que los modelos matemáticos se ajusten a principios como la mínima intervención y la presunción de inocencia (Miró Llinars, 2018). Este enfoque de "ética por diseño" garantiza que los límites legales se integren en el software desde su fase de entrenamiento (Guerrero Gualdrón, 2024).

El uso de la IA en la justicia penal debe limitarse a tareas que no comprometan la libertad de forma directa sin revisión humana inmediata. Mientras que la automatización de notificaciones o la búsqueda de jurisprudencia agiliza el proceso sin mayores riesgos, la determinación de medidas cautelares o la cuantificación de penas requiere un escrutinio judicial riguroso (Dictamen CIEJ, 2023). La tecnología debe agilizar el acceso, pero el acto

de privar de la libertad a una persona siempre debe pasar por el tamiz de la sensibilidad y racionalidad humana (Serventich, 2022).

La transparencia algorítmica implica también la publicación periódica de evaluaciones de impacto sobre el uso de estas herramientas. Los sistemas judiciales deben informar a la sociedad sobre la efectividad y los errores de los algoritmos implementados, permitiendo una fiscalización ciudadana sobre cómo la tecnología está afectando la equidad del sistema (Yáñez García-Bernalt, 2024). La confianza en la justicia digital se construye sobre la base de la honestidad institucional respecto a los límites y fallos de la innovación (Serna, 2025).

La garantía del juez imparcial prohíbe que la IA sea alimentada con datos que reflejen intereses comerciales o políticos de empresas privadas proveedoras. La dependencia de proveedores externos puede generar conflictos de interés y comprometer la independencia judicial si el algoritmo está diseñado para favorecer ciertos resultados (Dictamen CIEJ, 2023). Por ello, es recomendable que los sistemas de IA judicial sean de desarrollo estatal o de código abierto, asegurando que su única finalidad sea el cumplimiento de la ley (Serventich, 2022).

El debido proceso probatorio exige que el juez verifique la licitud de las pruebas obtenidas mediante IA, descartando aquellas que provengan de vigilancias masivas no autorizadas o de bases de datos obtenidas ilegalmente (Cantos Pardo, 2024). La incorporación de "evidencia digital" debe seguir protocolos estrictos de custodia y autenticidad, asegurando que los datos no hayan sido manipulados mediante técnicas de deepfake o alteraciones algorítmicas (Yáñez García-Bernalt, 2024).

La implementación garantista también requiere mecanismos de revisión y recurso efectivo contra las decisiones basadas en IA. El justiciable tiene derecho a apelar no solo el

fondo de la decisión, sino también el uso de la herramienta tecnológica si considera que esta introdujo sesgos o errores en el proceso (Polo Herazo & Padilla Maussa, 2025). La doble instancia debe permitir una revisión humana integral que evalúe tanto los hechos como la fiabilidad de la tecnología aplicada (Yáñez García-Bernalt, 2024).

Los límites éticos de la IA deben armonizarse con los tratados internacionales de derechos humanos, que prohíben cualquier forma de tortura o trato cruel mediado por la tecnología. El uso de la IA en la ejecución de la pena, como en la vigilancia electrónica o el tratamiento penitenciario, debe estar orientado a la resocialización y nunca a la deshumanización o el control excesivo del individuo (Miró Llinares, 2018; Guerrero Gualdrón, 2024). La tecnología debe ser una aliada de la dignidad y no una herramienta de opresión (Serna, 2025).

El futuro de la resolución de procesos penales bajo IA debe ser una "Justicia Inteligente" que combine lo mejor de la máquina y lo mejor del humano. Este binomio permite que la capacidad de procesamiento del algoritmo libere al juez de cargas administrativas, potenciando su rol como guardián de las garantías y líder ético de la sociedad (Miró Llinares, 2018; Guerrero Gualdrón, 2024). El límite final es, por tanto, el mantenimiento de una justicia centrada en la persona humana (Yáñez García-Bernalt, 2024).

En conclusión, los límites del uso de la IA en los procesos penales se definen por un marco garantista que prioriza la racionalidad humana, la transparencia absoluta y la protección de los derechos fundamentales sobre la mera eficiencia técnica. El juez penal debe erigirse como el filtro crítico que asegure que la inteligencia artificial potencie la justicia sin comprometer su esencia humana, garantizando que cada decisión sea el resultado de un razonamiento ético, justo y constitucionalmente legítimo (Serna, 2025;

Guerrero Gualdrón, 2024; Polo Herazo & Padilla Maussa, 2025; Yáñez García-Bernalt, 2024).

Conclusiones

La integración de la inteligencia artificial en la justicia penal representa una transición estructural que pone a prueba la vigencia del debido proceso frente a la búsqueda de celeridad administrativa. Si bien herramientas como Prometea o PretorIA han demostrado una capacidad notable para optimizar tiempos y reducir la congestión mediante la automatización de tareas repetitivas, esta eficiencia no puede ser el único criterio de éxito en un sistema que debe garantizar la justicia material. El valor jurídico de estas tecnologías reside únicamente en su carácter auxiliar, pues el razonamiento lógico y la interpretación del espíritu de la norma son funciones humanas indelegables que aseguran la legitimidad de las decisiones judiciales.

La pretendida neutralidad de los algoritmos es, en la práctica, una noción cuestionable que puede ocultar la reproducción de sesgos históricos y discriminaciones estructurales. Sistemas predictivos como COMPAS o Prisma, aunque buscan objetivar la valoración del riesgo, pueden terminar basando sus resultados en variables que estigmatizan a colectivos vulnerables por factores socioeconómicos o raciales. La opacidad de estos modelos de "caja negra" impide que el procesado comprenda el origen de su clasificación de riesgo, lo que vulnera el derecho a la defensa y la garantía de contradicción, convirtiendo la prueba algorítmica en una verdad inescrutable e inimpugnable.

Ante la expansión de la automatización, la legitimidad social de la justicia penal depende de que el juez asuma un rol activo y crítico, poseyendo un discernimiento tecnológico que le permita supervisar eficazmente los sistemas que emplea. La

jurisprudencia colombiana ha sido clara al establecer que el uso de la inteligencia artificial es aceptable solo si no sustituye la racionalidad del juzgador ni compromete la independencia judicial. Un sistema que delega la motivación de sus fallos a un software sin verificación humana rigurosa no solo incurre en riesgos de "alucinaciones" o errores técnicos, sino que se despoja de la sensibilidad ética necesaria para juzgar la complejidad del comportamiento humano.

Finalmente, una implementación garantista exige un marco regulatorio robusto que establezca límites claros y priorice la dignidad humana sobre la técnica. Normas como el Reglamento Europeo de Inteligencia Artificial marcan un camino necesario al prohibir prácticas invasivas como la vigilancia biométrica masiva y clasificar los sistemas judiciales como de alto riesgo, exigiendo transparencia y supervisión humana obligatoria. En última instancia, la inteligencia artificial debe ser un socio estratégico que libere al operador jurídico de lo mecánico, permitiéndole recuperar su función deliberativa y creativa, asegurando que la justicia siga siendo una virtud humana al servicio de las personas.

Referencias

- Acosta Morales, M. G. (2025). La inteligencia artificial y su impacto en el derecho penal ecuatoriano: análisis desde una revisión sistemática. *Serie Científica de la Universidad de las Ciencias Informáticas*, 18(2), 125-139.
http://scielo.sld.cu/scielo.php?pid=S2306-24952025000200125&script=sci_abstract&tlng=en
- Beltrán, S. C., & Quesada, D. L. J. (2025). *La protección de datos en el contexto de la cooperación transfronteriza de las autoridades policiales: nuevos desafíos desde la perspectiva del constitucionalismo europeo multinivel*. (Disertación Doctoral

Universidad de Valencia). <https://webges.uv.es/public/uvEntreuWeb/tesis/tesis-3695330-S27IJ338FE0UPOP8.pdf>

Blázquez, R. B. (2020). El sesgo de la máquina en la toma de decisiones en el proceso penal. *Ius et scientia*, 6(2), 54-71.

<https://revistascientificas.us.es/index.php/ies/article/view/14328>

Cantos Pardo, M. (2024). La algoritmización del dictamen pericial: ¿puerta de entrada para la aparición del “perito-robot”? *Actualidad Jurídica Iberoamericana*, (21), 434-467.

Casabona, C. M. R., & Martín, M. Á. R. (2023). *Derecho penal, ciberseguridad, ciberdelitos e inteligencia artificial*. Editorial Comares.

Comisión Iberoamericana de Ética Judicial [CIEJ]. (2023). Vigésimocuarto dictamen sobre el uso ético de la inteligencia artificial en la labor jurisdiccional. Ponente: Luis Porfirio Sánchez Rodríguez.

Corte Constitucional de Colombia. (2024). Sentencia T-323 de 2024. M. P. Juan Carlos Cortés González. <https://www.corteconstitucional.gov.co/relatoria/2024/T-323-24.htm>

Cruz Prieto, B. V., & León Daza, N. V. (2023). Implementación de la inteligencia artificial en las decisiones judiciales. ¿Una herramienta para el Juez o su remplazo absoluto en el Proceso Penal Colombiano?.

<http://repository.unilibre.edu.co/handle/10901/28159>

Díaz, C. V. M. (2024). Transformación del derecho penal: el papel de la inteligencia artificial en la legislación colombiana. *Revista Ciencias Humanas*, 17, 75-82.

<https://revistas.usb.edu.co/index.php/CienciasHumanas/article/view/7315>

- Espinoza, J. V., Durán, J. C. B., & García, M. R. (2024). Inteligencia artificial y nuevas formas de derecho en el siglo XXI. *Revista de Ciencias sociales*, 30(9), 447-458.
<https://dialnet.unirioja.es/servlet/articulo?codigo=9645081>
- Gallegos, S. B. G. (2022). La imputación objetiva en el delito de tránsito. *Universidad y Sociedad*, 14(S2), 194-200. <https://rus.ucf.edu.cu/index.php/rus/article/view/2754>
- García Falconí, R. J., & Barona Pazmiño, K. F. (2024). Inteligencia artificial y proceso penal. *Revista San Gregorio*, 1(58), 87-100.
http://scielo.senescyt.gob.ec/scielo.php?pid=S2528-79072024000200087&script=sci_arttext
- Gélvez-Ferreira, J. D., Helfer, P. M., Rodríguez, M. P. N., & Ruiz, C. A. R. (2021). *Predicción del delito en Colombia: experiencia en ciudades intermedias*. Dirección de Estudios Económicos.
- Guerrero Gualdrón, A. (2024). Competencias de los jueces penales ante la integración de la inteligencia artificial en el proceso penal [Trabajo de grado de Maestría, Universidad Libre].
- Gutiérrez, J. D. (2020). Retos éticos de la inteligencia artificial en el proceso judicial. *Instituto Colombiano de Derecho Procesal*, 499-516.
- Jiménez, L. C. P. (2024). Polic (ia): Un desafío para la política criminal de Colombia. *Informática y Derecho: Revista Iberoamericana de Derecho Informático (segunda época)*, (15), 171-181. <https://dialnet.unirioja.es/servlet/articulo?codigo=9887894>
- Lara Noboa, C. S. (2024). El uso de la inteligencia artificial en la toma de decisiones legales en los procesos civiles: Implicaciones éticas y jurídicas (Bachelor's thesis, Riobamba). <http://dspace.unach.edu.ec/handle/51000/14010>

- Miró Llinares, F. (2018). Inteligencia artificial y justicia penal: Más allá de los resultados lesivos causados por robots. *Revista de Derecho Penal y Criminología*, (20), 87-130.
<https://doi.org/10.5944/rdpc.20.2018.26446>
- Morales Moreno, Á. M. (2021). *Inteligencia Artificial Y Derecho Penal: Primeras Aproximaciones*. *Revista jurídica de Castilla y León*, (53).
- Morales, M. G. P., Dominguez, N. P. L., López, A. A. P., Carvajal, A. J. C., & Vidal, B. Y. R. (2025). El uso de la inteligencia artificial en la administración de justicia: Ventajas, riesgos y límites. *Ciencia y Educación*, 6(12.1), 270-286.
<https://www.cienciayeducacion.com/index.php/journal/article/view/2082>
- Moreno, J. D. L., Soto, F. F. S., & Escaleras, Á. M. H. (2024). Revolucionando la justicia: el impacto de la inteligencia artificial en el Derecho Penal. *Estudios y Perspectivas Revista Científica y Académica*, 4(3), 2155-2174.
<https://estudiosyperspectivas.org/index.php/EstudiosyPerspectivas/article/view/537>
- Nounou, L. A. (2024). ¿Construyendo puentes o barreras? El papel de la inteligencia artificial como herramienta en el acceso a la justicia [Seminario de Grado, Universidad Nacional de La Plata].
- Ojeda Yanez, R. J. (2026). Los delitos a partir del uso de IA y su tipificación en la legislación penal ecuatoriana (Bachelor's thesis, Riobamba).
- Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 por el que se establecen normas armonizadas en materia de inteligencia artificial.
- Polo Herazo, G., & Padilla Maussa, K. S. (2025). Alcance y efectos jurídicos del uso de la inteligencia artificial en la administración de justicia en Colombia, en el marco jurídico de la sentencia T 323 de 2024 [Monografía de grado, Universidad Cooperativa de Colombia].

- Puig, S. M. (2003). Significado y alcance de la imputación objetiva en derecho penal. *Revista electrónica deficiencia penal y criminología*, 5, 1-19.
<http://criminnet.ugr.es/recpc/05/recpc05-05>
- Román, M. C., & Domínguez, S. P. (2025). Explorando las percepciones de jueces y profesionales penitenciarios sobre el empleo de inteligencia artificial para el ejercicio de sus funciones. *Revista de Estudios Jurídicos y Criminológicos*, (12), 283-309. <https://dialnet.unirioja.es/servlet/articulo?codigo=10666243>
- Serna, C. (2025). El impacto de la inteligencia artificial en el derecho penal. *Revista Principia Iuris*, 22(46), 13-40.
- Serventich, C. (2022). Inteligencia artificial en el proceso penal. ¿Más vale humano conocido o algoritmo por conocer?. *Revista Jurídica Austral*, 3(2), 869-880.
<https://doi.org/10.26422/RJA.2022.0302.ser>
- Valeria, J. S. Z. Y., Abril, T., Vásquez, S., & Simijaca, G. (2022). Herramientas De La Inteligencia Artificial Dentro Del Sistema Judicial Colombiano; Estudio Del Caso Pretoria Y Prisma. *Principia Iuris*, 19(40), 47-66.
<http://revistas.santototunja.edu.co/index.php/piuris/article/view/2720>
- Yáñez García-Bernalt, I. (2024). La inteligencia artificial en el proceso penal: eficiencia versus garantías. *IUS ET SCIENTIA: Revista electrónica de Derecho y Ciencia*, 10(2), 80-100. <https://doi.org/10.12795/IESTSCIENTIA.2024.i02.04>