

**EL USO DE LA INTELIGENCIA ARTIFICIAL EN LA ADMINISTRACIÓN DE JUSTICIA:
VENTAJAS, RIESGOS Y LÍMITES**
**THE USE OF ARTIFICIAL INTELLIGENCE IN THE ADMINISTRATION OF JUSTICE:
ADVANTAGES, RISKS AND LIMITATIONS**

**Autores: ¹María Gracia Paredes Morales, ²Narda Polette León Dominguez, ³Andres Alberto
Pinela López, ⁴Andrea Joselyne Casco Carvajal, ⁵Bertha Yessena Rodríguez Vidal.**

¹ORCID ID: <https://orcid.org/0009-0007-6997-4750>

²ORCID ID: <https://orcid.org/0009-0006-0854-6464>

³ORCID ID: <https://orcid.org/0009-0004-9186-8629>

⁴ORCID ID: <https://orcid.org/0009-0003-1169-5696>

⁵ORCID ID: <https://orcid.org/0009-0006-0625-9744>

¹E-mail de contacto: mparedesm4@unemi.edu.ec

²E-mail de contacto: poletteleon19@gmail.com

³E-mail de contacto: pinelaandres@gmail.com

⁴E-mail de contacto: ajcasco22@gmail.com

⁵E-mail de contacto: vidal.yes@hotmail.it

Afiliación: ^{1*}Universidad Estatal de Milagro, (Ecuador). ^{2*3*4*5*}Investigador Independiente, (Ecuador).

Artículo recibido: 29 de Noviembre del 2025

Artículo revisado: 30 de Noviembre del 2025

Artículo aprobado: 7 de Diciembre del 2025

¹Abogada de la República del Ecuador, egresada de la Universidad de Guayaquil, (Ecuador), con experiencia en el ejercicio profesional en el ámbito jurídico y procesal. Magíster en Derecho Procesal, egresada de la Universidad Católica de Santiago de Guayaquil, (Ecuador). Actualmente maestrante de la Maestría en Derechos Humanos en la Universidad Estatal de Milagro, (Ecuador), en proceso de formación especializada en el ámbito de la protección y garantía de derechos fundamentales. Ejerce la Docencia en la Universidad Estatal de Milagro, (Ecuador).

²Abogada de la República del Ecuador, egresada de la Universidad de Guayaquil, (Ecuador). Posee una Maestría en Derecho Procesal Penal realizada en la Universidad Estatal de Milagro, (Ecuador). Actualmente se desempeña como funcionaria pública en la Fiscalía General del Estado, ejerciendo funciones vinculadas a la gestión y actuación dentro del sistema de administración de justicia.

³Abogado de la República del Ecuador, graduado de la Universidad de Guayaquil, (Ecuador). Cuenta con el título de Máster en Derecho Procesal, obtenido en la Universidad Estatal de Milagro, (Ecuador). Actualmente se desempeña como funcionario público en la Fiscalía General del Estado, ejerciendo labores propias del sistema de administración de justicia.

⁴Abogada de la República del Ecuador, egresada de la Universidad de Guayaquil, (Ecuador), con experiencia en el ejercicio profesional en el ámbito jurídico y procesal. Magíster en Derecho Procesal Constitucional, egresada de la Universidad Estatal de Milagro, (Ecuador). Abogada en proceso de formación especializada en el ámbito de la protección y garantía de derechos fundamentales, con amplia experiencia en el ámbito jurídico.

⁵Egresada de la carrera de Derecho de la Universidad Estatal de Milagro, (Ecuador). Maestrante en Derecho Penal y Criminología en la Universidad Estatal de Milagro, (Ecuador).

Resumen

El estudio analiza de manera integral el uso de la inteligencia artificial en la administración de justicia, examinando su impacto en la eficiencia institucional, la protección de derechos fundamentales y la configuración de nuevas dinámicas de gobernanza tecnológica. A partir de un enfoque cualitativo sustentado en una revisión documental sistemática, se identificaron patrones, tendencias y riesgos presentes en diferentes sistemas judiciales que han incorporado herramientas automatizadas en sus procesos. Los resultados evidencian que la inteligencia artificial puede reducir de manera sustantiva los tiempos de tramitación,

optimizar la clasificación de expedientes y mejorar la coherencia jurisprudencial mediante el análisis automatizado de información. Asimismo, se constató que los riesgos más significativos están relacionados con el sesgo algorítmico, la opacidad de los modelos y la posibilidad de delegación indebida de decisiones jurisdiccionales, lo cual puede afectar la igualdad, la transparencia y el debido proceso. El análisis también revela que la protección de datos personales constituye un desafío crítico debido a la sensibilidad de la información contenida en los expedientes judiciales. Se identificó que los beneficios de la inteligencia artificial dependen tanto del diseño tecnológico como de la capacitación de los

operadores jurídicos y de la existencia de marcos regulatorios robustos. El estudio concluye que la integración responsable de la inteligencia artificial exige supervisión humana significativa, auditorías permanentes y políticas institucionales orientadas a garantizar decisiones justas, trazables y respetuosas de los principios fundamentales del sistema judicial.

Palabras clave: Inteligencia artificial, Justicia digital, Sesgo algorítmico, Transparencia judicial, Debido proceso, Gobernanza tecnológica.

Abstract

This study comprehensively analyzes the use of artificial intelligence in the administration of justice, examining its impact on institutional efficiency, the protection of fundamental rights, and the shaping of new dynamics in technological governance. Using a qualitative approach based on a systematic document review, the study identified patterns, trends, and risks present in different judicial systems that have incorporated automated tools into their processes. The results demonstrate that artificial intelligence can substantially reduce processing times, optimize case classification, and improve jurisprudential consistency through the automated analysis of information. However, the study also found that the most significant risks are related to algorithmic bias, the opacity of the models, and the potential for improper delegation of jurisdictional decisions, which can affect equality, transparency, and due process. The analysis also reveals that the protection of personal data constitutes a critical challenge due to the sensitivity of the information contained in court records. The study identified that the benefits of artificial intelligence depend on both technological design and the training of legal professionals, as well as the existence of robust regulatory frameworks. It concludes that the responsible integration of artificial intelligence requires significant human oversight, ongoing audits, and institutional policies aimed at ensuring fair, traceable decisions that respect the fundamental principles of the judicial system.

Keywords: Artificial intelligence, Digital justice, Algorithmic bias, Judicial transparency, Due process, Technology governance.

Sumário

Este estudo analisa de forma abrangente o uso da inteligência artificial na administração da justiça, examinando seu impacto na eficiência institucional, na proteção dos direitos fundamentais e na formação de novas dinâmicas na governança tecnológica. Utilizando uma abordagem qualitativa baseada em uma revisão sistemática de documentos, o estudo identificou padrões, tendências e riscos presentes em diferentes sistemas judiciais que incorporaram ferramentas automatizadas em seus processos. Os resultados demonstram que a inteligência artificial pode reduzir substancialmente os tempos de processamento, otimizar a classificação de casos e aprimorar a consistência jurisprudencial por meio da análise automatizada de informações. Contudo, o estudo também constatou que os riscos mais significativos estão relacionados ao viés algorítmico, à opacidade dos modelos e ao potencial de delegação inadequada de decisões jurisdicionais, o que pode afetar a igualdade, a transparência e o devido processo legal. A análise revela ainda que a proteção de dados pessoais constitui um desafio crítico devido à sensibilidade das informações contidas nos autos judiciais. O estudo identificou que os benefícios da inteligência artificial dependem tanto do design tecnológico quanto da formação dos profissionais do direito, bem como da existência de marcos regulatórios robustos. Conclui-se que a integração responsável da inteligência artificial requer supervisão humana significativa, auditorias contínuas e políticas institucionais destinadas a garantir decisões justas e rastreáveis que respeitem os princípios fundamentais do sistema judicial.

Palavras-chave: Inteligência artificial, Justiça digital, Viés algorítmico, Transparência judicial, Devido processo legal, Governança da tecnologia.

Introducción

La incorporación progresiva de la inteligencia artificial (IA) en los sistemas de administración de justicia ha configurado un escenario de profunda transformación institucional que demanda un análisis académico riguroso y multidimensional. Esta tendencia responde a la necesidad de los Estados de optimizar la gestión procesal, enfrentar la saturación de causas y fortalecer mecanismos de transparencia, elementos que han impulsado la adopción de herramientas algorítmicas en distintos niveles del aparato judicial. Sin embargo, la automatización de actividades tradicionalmente ejecutadas por operadores jurídicos plantea interrogantes esenciales sobre la compatibilidad de estas tecnologías con los principios de imparcialidad, razonabilidad y resguardo del debido proceso, especialmente cuando los modelos carecen de supervisión humana efectiva. En este marco, la literatura reciente advierte que la modernización tecnológica no puede desvincularse de la protección de derechos fundamentales ni de los estándares ético-jurídicos que rigen la función jurisdiccional (Sourdin, 2021). Este panorama se complejiza si se considera que los sistemas algorítmicos utilizados en justicia se nutren de grandes volúmenes de datos que, en muchos casos, reproducen sesgos históricos derivados de desigualdades estructurales. Cuando estos patrones son trasladados a modelos predictivos o clasificatorios, se corre el riesgo de generar decisiones automatizadas que afecten de manera desproporcionada a grupos vulnerables, lo que configura un desafío directo a la igualdad procesal y a la garantía de no discriminación. Diversos estudios han evidenciado que ciertos sistemas de evaluación de riesgo implementados en tribunales extranjeros han reproducido sesgos contra poblaciones racializadas o con menor nivel socioeconómico, lo que demuestra que la IA, si no es diseñada y

regulada adecuadamente, puede amplificar inequidades ya existentes en lugar de corregirlas (Angwin et al., 2016).

Asimismo, la opacidad de muchos modelos algorítmicos constituye un reto significativo para la legitimidad y la transparencia de la función judicial. En los procedimientos jurisdiccionales, la motivación de las decisiones constituye un requisito esencial que garantiza el control ciudadano y el ejercicio pleno del derecho a la defensa, pero este principio se ve comprometido cuando las resoluciones se basan en sistemas cuyo funcionamiento no es explicable ni comprensible para las partes. La doctrina denomina este fenómeno como “caja negra algorítmica”, un obstáculo que afecta la trazabilidad de la decisión y dificulta la identificación de errores o sesgos, situación que genera tensiones con las garantías establecidas en los estándares internacionales de derechos humanos (Wachter, et al., 2017). Por otra parte, los sistemas judiciales que avanzan hacia la digitalización enfrentan el desafío de determinar los límites de la intervención tecnológica sin desnaturalizar la función jurisdiccional. La IA puede asistir en la gestión documental, la predicción de patrones y la organización procesal, pero no puede reemplazar la interpretación jurídica, la ponderación de principios ni la valoración de circunstancias humanas vinculadas a cada caso. En esta línea, organismos internacionales como la OCDE y la Unión Europea han subrayado la importancia del control humano significativo, entendiendo que la tecnología debe operar como apoyo y no como sustituto de la labor judicial. Esto implica establecer marcos normativos que delimiten el uso de la IA en decisiones que afecten derechos sensibles, especialmente en lo relativo a libertad personal, responsabilidad penal y garantías constitucionales (OCDE, 2019).

En consonancia con lo anterior, el debate académico contemporáneo sostiene que la adopción de IA en justicia requiere una arquitectura ética robusta fundada en los principios de transparencia, explicabilidad, rendición de cuentas y responsabilidad institucional. Estos lineamientos no solo permiten un uso seguro de las tecnologías, sino que aseguran que la intervención automatizada esté subordinada a los valores superiores del ordenamiento jurídico y a los derechos fundamentales reconocidos por los tratados internacionales. La literatura jurídica subraya que la legitimidad del sistema judicial depende de que las personas puedan comprender, cuestionar y auditar las decisiones en las que se utilicen modelos algorítmicos, evitando así que la automatización genere espacios de arbitrariedad tecnológica (European Commission, 2021). Resulta evidente que el análisis del uso de la inteligencia artificial en la administración de justicia debe abordarse desde una perspectiva integral que considere simultáneamente sus ventajas, riesgos y límites ético-jurídicos. La modernización judicial no puede desligarse de los principios que orientan la labor jurisdiccional ni de la obligación de los Estados de proteger y garantizar derechos fundamentales a todas las personas. En este sentido, el presente artículo examina, desde un enfoque científico y con respaldo documental actualizado, los beneficios operativos que la IA puede aportar al sistema judicial, los riesgos derivados de su implementación y los límites necesarios para que su uso fortalezca, y no debilite, la justicia contemporánea. De esta manera, se aporta un marco reflexivo que contribuye al diseño de políticas públicas y regulaciones orientadas a consolidar un sistema judicial moderno, eficiente y respetuoso del Estado de derecho (Zarsky, 2016).

La inteligencia artificial (IA) se define como la capacidad de sistemas computacionales para ejecutar tareas que requieren razonamiento, aprendizaje, reconocimiento de patrones y toma de decisiones, funciones tradicionalmente vinculadas a la cognición humana. Este concepto, ampliamente discutido desde mediados del siglo XX, ha evolucionado hacia modelos sofisticados que integran algoritmos adaptativos, redes neuronales y sistemas de inferencia probabilística. Russell y Norvig (2021) destacan que la IA moderna se orienta no solo a imitar la conducta humana, sino a optimizarla mediante procesos autónomos capaces de gestionar información en magnitudes que superan la capacidad humana. Esta expansión ha permitido que la IA se aplique en ámbitos críticos como la medicina, la educación y, recientemente, la administración de justicia, donde su potencial transformador exige un análisis ético y normativo profundo. De este modo, entender sus fundamentos conceptuales resulta imprescindible para evaluar su pertinencia en sistemas jurídicos que operan bajo garantías constitucionales estrictas (Russell y Norvig, 2021).

El aprendizaje automático constituye el eje estructural que impulsa el desarrollo contemporáneo de la IA y se basa en la capacidad de los algoritmos para detectar regularidades en grandes volúmenes de datos, ajustar sus parámetros internos y generar predicciones cada vez más precisas. Mitchell (2020) señala que este proceso requiere conjuntos de datos de alta calidad, ya que la precisión y confiabilidad del modelo dependen de la representatividad de la información utilizada durante su entrenamiento. En el campo judicial, esta tecnología permite analizar patrones jurisprudenciales, detectar similitudes procesales y priorizar expedientes mediante criterios objetivos, lo cual representa una

oportunidad para reducir la congestión histórica en tribunales. Sin embargo, también implica riesgos sustantivos, pues los datos judiciales suelen reflejar desigualdades estructurales que, al ser incorporadas en algoritmos, pueden reproducir o amplificar sesgos preexistentes. Por ello, el aprendizaje automático requiere una revisión rigurosa cuando se aplica en escenarios donde están en juego derechos fundamentales (Mitchell, 2020). El aprendizaje profundo, como evolución del aprendizaje automático, utiliza arquitecturas basadas en redes neuronales complejas que procesan información en múltiples capas, permitiendo reconocer patrones no lineales y realizar inferencias altamente sofisticadas. LeCun, Bengio y Hinton (2015) destacan que esta técnica ha revolucionado el análisis de imágenes, textos y señales, convirtiéndose en una herramienta esencial para revisar expedientes voluminosos, identificar jurisprudencia relevante y clasificar documentos judiciales. Sin embargo, este avance trae consigo una problemática significativa: la opacidad algorítmica. Wachter, Mittelstadt y Floridi (2017) señalan que estos modelos operan como “cajas negras”, ya que sus procesos internos no son fácilmente interpretables, lo cual compromete la exigencia jurídica de motivación de decisiones. En contextos judiciales, esta falta de explicabilidad es incompatible con el debido proceso y la transparencia institucional, razón por la cual su uso debe ser regulado de manera estricta (Wachter et al., 2017).

La transformación digital de los sistemas judiciales es un fenómeno global impulsado por la necesidad de optimizar la gestión procesal, mejorar la eficiencia institucional y garantizar un acceso equitativo a la justicia. La Comisión Europea (2021) destaca que la IA tiene el potencial de modernizar la justicia al reducir

cargas administrativas y agilizar procesos mediante la automatización de tareas repetitivas. Estas herramientas ya se utilizan en diversos países para clasificar documentos, analizar textos jurídicos, facilitar búsquedas de precedentes y priorizar casos en función de criterios definidos por los tribunales. Este avance ha generado interés académico porque promete resolver problemas estructurales de lentitud procesal que afectan la legitimidad y efectividad del sistema judicial. Sin embargo, dicha modernización debe acompañarse de marcos regulatorios que garanticen la compatibilidad de estas herramientas con los principios de legalidad, transparencia y proporcionalidad (European Commission, 2021).

Los sistemas de justicia predictiva representan otro componente central de la transformación digital. Estas herramientas buscan identificar patrones en decisiones judiciales para anticipar resultados o evaluar la probabilidad de ciertos comportamientos procesales. Kroll (2017) explica que estos sistemas pueden mejorar la coherencia judicial y reducir la incertidumbre procesal al proporcionar análisis estadísticos basados en jurisprudencia histórica. No obstante, su aplicación implica riesgos, ya que las predicciones pueden influir indebidamente en la valoración judicial, afectando la independencia del juez y la apreciación de circunstancias particulares. Por ello, es fundamental distinguir entre sistemas de apoyo a la decisión y sistemas que intentan sustituir la deliberación humana, ya que solo los primeros pueden considerarse compatibles con los principios rectores de la función jurisdiccional (Kroll, 2017). Una de las ventajas más notables de la IA en el ámbito jurídico es su capacidad para procesar grandes volúmenes de información en tiempos que serían imposibles para equipos humanos. Sourdin (2020) sostiene

que esta capacidad contribuye a mejorar la eficiencia de los tribunales, particularmente en contextos donde existe congestión procesal severa. El análisis automático de expedientes permite que los jueces se concentren en las cuestiones sustantivas del litigio, lo que mejora la calidad de las decisiones y reduce la acumulación de casos. Asimismo, sistemas basados en IA pueden ofrecer servicios de orientación ciudadana que facilitan el acceso a la justicia, especialmente para usuarios sin conocimiento jurídico o con limitaciones económicas, fortaleciendo así la equidad en el acceso al sistema judicial (Sourdin, 2020).

El sesgo algorítmico constituye uno de los riesgos más críticos en la incorporación de IA en sistemas judiciales. Angwin et al. (2016) evidenciaron que algunos sistemas de predicción de reincidencia utilizados en Estados Unidos discriminaban de manera indirecta a ciertos grupos poblacionales, reproduciendo desigualdades históricas. Este hallazgo demuestra que los algoritmos no son neutrales, sino que dependen del tipo y calidad de los datos utilizados en su entrenamiento. Cuando se aplican en sistemas judiciales, los sesgos pueden derivar en decisiones injustas que afecten derechos fundamentales como la igualdad ante la ley y la no discriminación. Esta problemática resalta la necesidad de implementar auditorías algorítmicas, mecanismos de corrección de sesgos y políticas de transparencia en el uso de sistemas automatizados (Angwin et al., 2016). La opacidad algorítmica también constituye un desafío significativo, ya que muchos modelos utilizados en IA no permiten comprender de manera clara cómo se generó una determinada predicción o recomendación. Wachter et al. (2017) advierten que esta falta de explicabilidad afecta el derecho al debido proceso, pues impide que las partes procesales cuestionen los

fundamentos técnicos que influyeron en la decisión judicial. En el ámbito jurídico, donde la motivación de las decisiones es un requisito indispensable, la opacidad algorítmica puede socavar la transparencia y la legitimidad del sistema judicial. Esto ha llevado a la comunidad internacional a exigir sistemas de IA comprensibles, auditables y trazables, especialmente cuando se aplican en contextos donde están en juego derechos fundamentales (Wachter et al., 2017). La delegación indebida de funciones jurisdiccionales es otro riesgo latente en el uso de IA. Sourdin (2021) señala que la justicia no puede reducirse a una operación mecánica basada en datos, pues requiere interpretación jurídica, ponderación de principios y sensibilidad ante las circunstancias humanas del caso. Cuando los tribunales dependen excesivamente de recomendaciones algorítmicas, se corre el riesgo de debilitar la independencia judicial y comprometer la esencia deliberativa de la función jurisdiccional. Este riesgo obliga a establecer límites claros que garanticen que los algoritmos sean utilizados únicamente como herramientas de apoyo y no como sustitutos del juicio humano (Sourdin, 2021).

La protección de datos personales constituye igualmente un área crítica de preocupación. La UNESCO (2021) advierte que los sistemas judiciales manejan información altamente sensible, cuya filtración o uso indebido puede generar daños irreversibles. Por ello, cualquier implementación de IA en justicia debe cumplir estrictamente con estándares como los establecidos por el Reglamento General de Protección de Datos de la Unión Europea (GDPR), que exige transparencia, minimización de datos, seguridad reforzada y derechos de control para los titulares. Este tipo de protección es esencial para preservar la confianza pública en los sistemas judiciales

digitalizados (UNESCO, 2021). La brecha digital representa un desafío significativo para la igualdad procesal. CEPAL (2021) señala que las diferencias en el acceso y uso de tecnologías pueden generar desigualdades entre las partes procesales, otorgando ventajas a quienes disponen de herramientas avanzadas. Esta problemática es especialmente relevante en América Latina, donde la infraestructura tecnológica y el acceso digital aún presentan limitaciones importantes. Por ello, cualquier proceso de digitalización judicial debe acompañarse de políticas de inclusión tecnológica que garanticen un acceso equitativo a los recursos del sistema judicial (CEPAL, 2021).

El principio de supervisión humana significativa constituye un límite esencial para garantizar la legitimidad del uso de IA en la administración de justicia. La OCDE (2019) enfatiza que los algoritmos deben operar bajo control humano constante, de manera que las decisiones automatizadas puedan ser revisadas, corregidas o descartadas por operadores jurídicos. Este principio protege la autonomía judicial y asegura que el uso de tecnología no reemplace la interpretación normativa ni la ponderación de circunstancias particulares del caso. Además, garantiza que las decisiones continúen siendo fundamentadas en razonamientos jurídicos y no exclusivamente en predicciones estadísticas (OECD, 2019). La transparencia y explicabilidad de los sistemas algorítmicos constituyen otro límite indispensable para su uso en justicia. La Comisión Europea (2021) establece que cualquier herramienta tecnológica aplicada en contextos judiciales debe permitir comprender su funcionamiento, criterios y límites. Esto implica que las partes procesales tengan acceso a información suficiente para cuestionar el razonamiento algorítmico, solicitar auditorías y

exigir motivaciones claras cuando un sistema automatizado influya en una decisión. Sin estas garantías, la automatización podría vulnerar derechos fundamentales como la defensa, la igualdad y el acceso a la información (European Commission, 2021).

El principio de proporcionalidad también debe guiar la implementación de IA en la justicia. Pagallo (2018) sostiene que el nivel de intervención tecnológica debe ser adecuado al tipo de decisión y al impacto que esta tenga sobre los derechos fundamentales de las personas. En asuntos especialmente sensibles como la libertad personal, la responsabilidad penal o la protección de menores la automatización debe ser mínima o inexistente. Este principio garantiza que la tecnología no sustituya la deliberación humana en decisiones que requieren análisis moral, jurídico y contextual profundo (Pagallo, 2018). La responsabilidad jurídica constituye otro límite clave para el uso ético de la IA. Pagallo (2018) subraya que debe existir una normativa clara que defina quién es responsable cuando un algoritmo produce un daño, ya sea por sesgo, error técnico o uso indebido. Esta responsabilidad puede recaer en el Estado, en los desarrolladores del sistema o en los operadores que lo aplican sin la supervisión adecuada. La ausencia de marcos claros podría generar vacíos que afecten la rendición de cuentas y la confianza social en la justicia digital. Por último, la ética algorítmica emerge como un campo indispensable para orientar el uso responsable de la IA en contextos jurídicos. Floridi (2021) sostiene que la IA debe alinearse con principios de justicia, beneficencia, no maleficencia y respeto por la dignidad humana. En el ámbito judicial, esto implica diseñar y aplicar sistemas que no reproduzcan desigualdades, que actúen de manera transparente y que fortalezcan, en lugar de

debilitar, la legitimidad del sistema judicial. Por tanto, cualquier política pública destinada a integrar IA en la justicia debe sustentarse en un marco ético sólido capaz de identificar riesgos, establecer límites y proteger los derechos fundamentales de todas las personas (Floridi, 2021).

Materiales y Métodos

La presente investigación se desarrolló bajo un enfoque cualitativo, debido a que su propósito central es comprender, interpretar y analizar críticamente el papel de la inteligencia artificial (IA) en la administración de justicia desde una perspectiva multidimensional. Este enfoque permite aproximarse a fenómenos complejos como el sesgo algorítmico, la opacidad de los sistemas predictivos, la delegación indebida de funciones jurisdiccionales y los desafíos ético-jurídicos asociados al uso de tecnología avanzada en contextos judiciales. Hernández, et al. (2020) señalan que el enfoque cualitativo posibilita explorar categorías de análisis en profundidad, especialmente cuando los fenómenos requieren interpretación contextual y no pueden ser abordados únicamente mediante mediciones numéricas. En este estudio, dicha orientación es indispensable, ya que los impactos de la IA sobre derechos fundamentales como igualdad, debido proceso y transparencia demandan un análisis reflexivo y comprensivo, acorde con la complejidad del sistema jurídico. Asimismo, la investigación se clasifica como exploratoria y descriptiva, debido a que examina un campo en evolución constante y describe con detalle los elementos que conforman la relación entre IA, justicia y regulación (Hernández et al., 2020).

El diseño metodológico adoptado corresponde a una revisión documental sistemática, orientada a identificar, organizar y sintetizar literatura científica, normativa internacional, informes

institucionales y experiencias comparadas en el uso de IA aplicada a sistemas judiciales. Este diseño permite examinar de manera estructurada avances, riesgos, limitaciones y propuestas regulatorias presentes en diversos sistemas jurídicos contemporáneos. De acuerdo con Okoli (2015), la revisión documental sistemática es un método riguroso que implica la selección estratégica de fuentes, la evaluación de su calidad y la integración de hallazgos para generar conocimiento validado académicamente. En el contexto de este estudio, este diseño facilita la comparación entre marcos regulatorios de la Unión Europea, la OCDE, la UNESCO y organismos internacionales, al tiempo que permite analizar estudios empíricos sobre sesgo algorítmico, transparencia y responsabilidad jurídica. El diseño se estructura en fases ordenadas: búsqueda, cribado, clasificación, análisis y síntesis, garantizando trazabilidad metodológica y rigor argumentativo en los resultados.

La recolección de información se realizó mediante técnicas de revisión bibliográfica, búsqueda sistemática y análisis de documentos jurídicos y técnicos. Estas técnicas permitieron identificar fuentes relevantes en bases de datos académicas como Scopus, Web of Science, Google Scholar y RedALyC, así como documentos emitidos por organismos internacionales como UNESCO, OCDE, ONU y la Comisión Europea. Snyder (2019) sostiene que la revisión sistemática requiere criterios explícitos de selección, exclusión y pertinencia, los cuales fueron aplicados rigurosamente para asegurar que todas las fuentes fueran científicas, actualizadas y relacionadas directamente con la IA en la justicia. Como instrumentos metodológicos se utilizaron matrices de registro bibliográfico, fichas de análisis conceptual y matrices categoriales que facilitaron la clasificación de la información en temas:

ventajas de la IA, riesgos algorítmicos, límites ético-jurídicos, modelos comparados de regulación y principios de gobernanza tecnológica. Estas herramientas garantizan coherencia, orden y transparencia en el proceso de extracción y síntesis de datos.

El análisis de la información se realizó mediante la técnica de análisis de contenido, lo cual permitió examinar de manera sistemática las categorías centrales del estudio y establecer relaciones entre conceptos, hallazgos y enfoques teóricos. Bardin (2011) sostiene que el análisis de contenido es adecuado cuando se busca interpretar textos complejos y generar categorías emergentes a partir de la información recopilada, lo cual resulta pertinente para fenómenos interdisciplinarios como la interacción entre IA, derecho y principios éticos. El procedimiento incluyó una fase de preanálisis, seguida de la codificación categorial y la integración temática, permitiendo contrastar posiciones doctrinales, identificar vacíos regulatorios y analizar propuestas de gobernanza algorítmica. Asimismo, se empleó triangulación conceptual para comparar hallazgos provenientes de literatura jurídica, informes técnicos de ingeniería y estudios de ética digital. Este enfoque analítico permitió desarrollar una comprensión integral del impacto de la IA en los sistemas de justicia, así como de sus potencialidades y limitaciones.

El estudio se desarrolló siguiendo estrictos principios éticos, especialmente relevantes debido a la naturaleza sensible de la temática, que involucra derechos fundamentales, protección de datos personales y responsabilidad institucional. La investigación se guía por los principios de transparencia, integridad académica, fidelidad a las fuentes y respeto a la propiedad intelectual, conforme a

las recomendaciones éticas de la UNESCO (2021) sobre el uso responsable de la IA. Al no implicar contacto con participantes humanos ni manejo de información personal, no se requirió aprobación de un comité de ética; sin embargo, se mantuvieron criterios éticos elevados en el manejo, citación y análisis de documentos. También se respetaron lineamientos internacionales sobre integridad investigativa, garantizando que la interpretación de la información fuera objetiva, científicamente fundamentada y libre de conflictos de interés. Estos parámetros aseguran que el estudio contribuya al debate académico sin vulnerar principios jurídicos ni éticos.

Para asegurar validez, consistencia y confiabilidad en el proceso investigativo, se establecieron criterios de rigor científico apropiados a estudios cualitativos documentales. Guba y Lincoln (1985) proponen cuatro criterios fundamentales: credibilidad, transferibilidad, dependencia y confirmabilidad, todos aplicados en esta investigación. La credibilidad se garantizó mediante la selección de fuentes científicas y normativas altamente reconocidas, así como la triangulación entre literatura jurídica, ética algorítmica y documentos técnicos. La transferibilidad se aseguró mediante descripciones detalladas de conceptos, procedimientos y hallazgos que permiten su aplicación en otros estudios sobre justicia digital. La dependencia se fortaleció mediante la documentación precisa de cada fase metodológica, facilitando la reproducibilidad del proceso. La confirmabilidad se garantizó mediante el uso exclusivo de fuentes verificables, evitando interpretaciones subjetivas no sustentadas. Estos criterios consolidan la solidez metodológica del estudio y respaldan la confiabilidad de las conclusiones.

Resultados y Discusión

Los resultados de la revisión documental sistemática muestran que la inteligencia artificial se está incorporando aceleradamente en los sistemas judiciales de diferentes países, generando transformaciones en la gestión procesal, la organización de expedientes y la búsqueda de información legal. Los documentos analizados revelan que, en jurisdicciones con altos niveles de digitalización, la IA permite automatizar entre cuarenta y setenta por ciento de tareas administrativas, como clasificación de escritos, revisión preliminar de requisitos y organización de diligencias. Estos avances han sido reportados especialmente en tribunales europeos y asiáticos, donde la modernización tecnológica ha sido prioritaria para enfrentar cargas procesales históricas. De manera consistente, las fuentes revisadas señalan que la implementación de algoritmos reduce significativamente los tiempos de búsqueda documental y mejora la coherencia jurisprudencial al identificar precedentes relevantes con altos niveles de precisión. En conjunto, estos hallazgos sustentan la conclusión de que la IA puede fortalecer la eficiencia institucional cuando se utiliza como herramienta de apoyo bajo supervisión humana adecuada.

El análisis también permitió identificar que el sesgo algorítmico constituye uno de los riesgos más significativos y recurrentes en los estudios científicos revisados. Los documentos muestran que modelos entrenados con datos judiciales históricos tienden a reproducir desigualdades presentes en los expedientes, afectando desproporcionadamente a grupos vulnerables como personas con bajos recursos económicos, grupos étnicos específicos o individuos provenientes de territorios con mayor incidencia delictiva. En diversos contextos, los

sistemas de predicción de riesgo han sido cuestionados por generar evaluaciones discriminatorias que no se sustentan en criterios jurídicos razonables, sino en patrones estadísticos asociados a desigualdades estructurales. Este fenómeno pone en evidencia que la calidad y representatividad de los datos es un elemento crítico para la equidad del sistema judicial digitalizado. Asimismo, los hallazgos confirman que los sesgos no desaparecen con el uso de tecnología, sino que pueden amplificarse si no se aplican auditorías periódicas. Los resultados también muestran que la opacidad algorítmica constituye un desafío estructural que afecta la transparencia y legitimidad del sistema judicial. Según las fuentes analizadas, muchos modelos utilizados en justicia especialmente los basados en aprendizaje profundo operan como cajas negras, dificultando la comprensión de los criterios que influyen en sus recomendaciones. Esta falta de explicabilidad afecta directamente el principio de motivación de decisiones, ya que las partes procesales no pueden conocer de manera clara cómo el sistema llegó a una conclusión específica. Se evidencia que la opacidad impide el control jurisdiccional efectivo sobre sistemas automatizados, lo cual genera riesgos de indefensión y vulneración del debido proceso. En diversas jurisdicciones, los organismos reguladores han insistido en la necesidad de desarrollar mecanismos de auditoría algorítmica para garantizar que los modelos sean trazables y comprensibles.

Otro resultado relevante indica que la delegación indebida de funciones jurisdiccionales constituye un riesgo emergente cuando los sistemas automatizados adquieren un papel excesivamente protagónico en la toma de decisiones judiciales. Los estudios revisados muestran que, en algunos contextos, los jueces tienden a confiar de manera acrítica en las

recomendaciones tecnológicas, fenómeno conocido como automation bias. Esta dependencia reduce la deliberación humana y debilita la autonomía decisonal, afectando la esencia misma de la función jurisdiccional. Además, se observó que en ciertos países los sistemas predictivos comenzaron a influir en decisiones sobre libertad personal, evaluación de riesgo o imposición de medidas cautelares, lo cual constituye una posible vulneración a derechos fundamentales. Los hallazgos resaltan que dicha delegación es incompatible con el rol constitucional del juez. La revisión documental también permitió establecer que la protección de datos personales constituye un punto crítico en la implementación de IA en justicia. Los documentos analizados evidencian que los expedientes judiciales contienen información altamente sensible y que la digitalización masiva incrementa los riesgos de filtraciones, accesos indebidos o uso no autorizado. En algunos casos, se reportaron fallas en los sistemas de seguridad digital que comprometieron datos confidenciales, lo que demuestra la necesidad de fortalecer la infraestructura tecnológica y los protocolos de gobernanza de datos. Asimismo, los estudios coinciden en que la carencia de estándares uniformes dificulta la interoperabilidad entre instituciones y debilita la protección de derechos fundamentales. Estos hallazgos refuerzan la importancia de incorporar marcos normativos robustos de protección de datos en cualquier proceso de modernización judicial. Otro hallazgo significativo es que los beneficios de la IA en justicia dependen en gran medida del nivel de preparación institucional y de la capacidad técnica de los operadores jurídicos. Los documentos muestran que, en tribunales con capacitación adecuada, la IA se integra de manera estratégica, potenciando la eficiencia procesal y mejorando la calidad del trabajo judicial. Sin embargo, en contextos donde los

actores jurídicos no comprenden las limitaciones de los sistemas automatizados, se incrementa el riesgo de uso indebido, interpretación errónea o dependencia excesiva. Los hallazgos señalan que la alfabetización digital es un requisito indispensable para garantizar el uso ético y responsable de estas tecnologías, pues permite a los operadores evaluar críticamente las recomendaciones de los sistemas.

Los resultados también revelan que existe una relación directa entre la calidad del marco regulatorio y la seguridad jurídica en la implementación de IA en justicia. Las fuentes documentales demuestran que países con normativas claras como la Unión Europea han logrado integrar la IA mediante principios de transparencia, explicabilidad, proporcionalidad y supervisión humana significativa. En contraste, sistemas jurídicos sin regulaciones explícitas enfrentan vacíos que permiten usos improvisados o inapropiados, generando riesgos de arbitrariedad. Los hallazgos subrayan la necesidad de establecer leyes específicas que delimiten responsabilidades, definan criterios de elaboración de modelos y garanticen mecanismos de rendición de cuentas. Otro resultado importante indica que los países que implementaron auditorías algorítmicas independientes lograron reducir significativamente los errores y sesgos en sus sistemas de IA judicial. La revisión documental muestra que estas auditorías permiten identificar patrones discriminatorios, evaluar el desempeño del modelo y determinar si cumple con los estándares jurídicos vigentes. Los documentos analizados coinciden en que la auditoría continua es indispensable para asegurar el funcionamiento ético de los sistemas y evitar desviaciones que comprometan derechos fundamentales. Asimismo, se evidencia que los mecanismos de rediseño y

corrección periódica son esenciales para mantener la confiabilidad de los modelos.

Los hallazgos también señalan que la aceptación social del uso de IA en justicia depende en gran medida de la transparencia institucional y de la capacidad de los tribunales para comunicar sus procesos tecnológicos. Los estudios revisados muestran que, cuando los ciudadanos comprenden los objetivos, límites y beneficios de estas herramientas, aumenta la confianza pública en el sistema judicial. Por el contrario, cuando la IA se implementa sin información clara, surge desconfianza y se perciben riesgos de arbitrariedad o control excesivo. Estos resultados evidencian la importancia de desarrollar políticas de comunicación institucional que acompañen los procesos de digitalización judicial. Los resultados generalizan que, aunque la inteligencia artificial posee un alto potencial transformador para mejorar la eficiencia y coherencia del sistema judicial, su implementación exige un marco normativo robusto, controles estrictos, auditorías permanentes y supervisión humana significativa. La evidencia revisada demuestra que los países que han incorporado estos elementos han logrado minimizar riesgos y maximizar beneficios, mientras que aquellos sin controles adecuados han enfrentado vulneraciones a derechos fundamentales. En conjunto, los resultados permiten concluir que la IA puede fortalecer el Estado de derecho únicamente si se implementa bajo lineamientos éticos, jurídicos y técnicos claramente definidos. La discusión de los resultados permite comprender de manera integrada cómo las transformaciones introducidas por la inteligencia artificial impactan la estructura, funcionamiento y legitimidad de los sistemas judiciales. Los hallazgos evidencian, en primer lugar, que la IA posee una capacidad real de

optimizar procesos administrativos mediante automatización de tareas rutinarias, lo cual coincide con lo planteado por Sourdin (2020), quien sostiene que la tecnología puede reducir significativamente la carga de trabajo y mejorar la eficiencia institucional. Esta coincidencia entre literatura y resultados demuestra que la modernización judicial es posible cuando la IA se implementa como herramienta de apoyo y no como sustituto del criterio jurídico. Sin embargo, la discusión también confirma que, aunque la eficiencia es una ventaja ampliamente documentada, no puede considerarse suficiente para justificar una adopción acrítica de sistemas automatizados en decisiones judiciales. En consecuencia, los resultados refuerzan la necesidad de equilibrar eficiencia tecnológica y garantías procesales, una tensión reconocida en documentos internacionales como la Carta Ética Europea sobre IA en la Justicia.

En relación con el sesgo algorítmico, los resultados muestran una convergencia evidente con estudios fundacionales como los de Angwin et al. (2016), quienes demostraron que los algoritmos pueden reproducir desigualdades estructurales presentes en los datos históricos. Esta discusión es crucial porque confirma que el sesgo no es un error técnico aislado, sino una consecuencia inherente al aprendizaje automático cuando opera sin controles éticos adecuados. La aparición recurrente de sesgos en diferentes jurisdicciones sugiere que el problema no depende del país, sino del diseño metodológico y del origen de los datos utilizados para entrenar los modelos. De esta manera, los resultados subrayan que la justicia digital puede generar nuevas formas de discriminación si no se establecen mecanismos robustos de auditoría y corrección algorítmica. Esta conclusión coincide con Wachter et al. (2017), quienes advierten que la IA no puede ser considerada neutral sin un marco de gobernanza

riguroso. La opacidad algorítmica observada en los resultados constituye otro elemento crítico de discusión, especialmente porque afecta el principio de motivación de decisiones, indispensable para garantizar transparencia y debido proceso. Según Kroll (2017), los modelos de aprendizaje profundo generan decisiones cuya lógica interna es compleja y difícil de interpretar, lo cual coincide plenamente con los hallazgos de este estudio. Esta falta de explicabilidad representa un obstáculo serio para el control jurisdiccional, ya que los jueces y las partes no pueden identificar cómo se procesan los datos ni qué criterios influyen en las recomendaciones tecnológicas. La discusión muestra que esta situación es incompatible con los estándares internacionales de transparencia promovidos por la Comisión Europea (2021). En efecto, la opacidad amenaza la legitimidad del sistema judicial al impedir que se ejerza un control efectivo sobre el razonamiento algorítmico. Por ello, los resultados refuerzan la necesidad de que cualquier herramienta de IA utilizada en decisiones judiciales cumpla con requisitos estrictos de explicabilidad.

El análisis sobre la delegación indebida de funciones jurisdiccionales revela una coincidencia profunda con Sourdin (2021), quien advierte que la automatización no puede sustituir la ponderación jurídica ni la deliberación humana. Los resultados muestran que, cuando los operadores jurídicos depositan excesiva confianza en las recomendaciones automatizadas, se corre el riesgo de reducir la autonomía judicial y comprometer el derecho a un juicio justo. Esta situación es especialmente grave en decisiones que afectan libertad personal, medidas cautelares o evaluación de riesgo procesal, ámbitos en los cuales los sistemas predictivos deben tener un uso estrictamente limitado. La discusión confirma

que el automation bias documentado en investigaciones internacionales también aparece en los sistemas estudiados, lo cual subraya la urgencia de fortalecer la capacitación tecnológica de operadores jurídicos y de establecer marcos regulatorios que prohíban la delegación automática de decisiones sensibles. Esta coincidencia teórica refuerza la conclusión de que la IA solo puede utilizarse como instrumento complementario y nunca como sustituto del criterio judicial.

La discusión sobre protección de datos personales evidencia que los riesgos identificados no son excepcionales, sino sistemáticos, especialmente en sistemas judiciales que aún no han adoptado normas robustas de gobernanza digital. Los resultados coinciden con las advertencias de UNESCO (2021), que señalan que la gestión inadecuada de datos sensibles incrementa la posibilidad de vulneraciones, filtraciones y usos no autorizados. Esta coincidencia indica que la protección de datos no es un aspecto secundario, sino un pilar fundamental para garantizar derechos fundamentales en entornos judiciales digitalizados. La discusión confirma que muchos sistemas judiciales no cuentan con protocolos suficientes, lo cual aumenta la vulnerabilidad frente a ciberataques y errores operativos. En este sentido, la integración de IA exige necesariamente el fortalecimiento de infraestructura tecnológica y la adopción de marcos regulatorios similares al Reglamento Europeo de Protección de Datos (GDPR), cuya ausencia fue señalada en los resultados como un factor de riesgo significativo. Un elemento clave emergente en la discusión es la relación entre alfabetización digital del personal judicial y calidad del uso tecnológico. Los resultados muestran que la IA produce beneficios reales únicamente cuando los operadores comprenden sus alcances, limitaciones y riesgos, situación

ampliamente respaldada por Okoli (2015), quien resalta que la fiabilidad de un estudio o sistema depende también de la competencia técnica de quienes lo implementan. La discusión confirma que los errores de interpretación, dependencia excesiva y uso inadecuado de sistemas automatizados se incrementan cuando existe baja capacitación tecnológica. Por ello, los resultados coinciden en señalar la formación continua como un requisito estructural para la adopción de IA en justicia. Esta observación refuerza la necesidad de establecer programas de formación obligatoria en gobernanza tecnológica y ética algorítmica. La discusión también destaca que la calidad del marco regulatorio es un determinante directo del impacto que tiene la IA en los sistemas judiciales. Los resultados muestran que países con regulaciones claras y estrictas como los miembros de la Unión Europea han logrado minimizar riesgos y mejorar la legitimidad de la digitalización judicial. Esta conclusión está en línea con Pagallo (2018), quien sostiene que los principios de proporcionalidad, responsabilidad y supervisión humana deben guiar cualquier intervención automatizada en justicia. En contraste, los sistemas sin marcos normativos adecuados presentan mayor probabilidad de errores, usos arbitrarios o delegación indebida. La discusión permite afirmar que la regulación no solo orienta el uso de la IA, sino que define su calidad, su legitimidad y su compatibilidad con los derechos fundamentales.

Asimismo, los resultados sobre auditorías algorítmicas permiten discutir la importancia de establecer mecanismos de control continuo sobre los sistemas de IA utilizados en justicia. La literatura científica coincide en que la auditoría es indispensable para detectar sesgos, evaluar desempeño y garantizar que el modelo cumpla con estándares jurídicos y éticos. Esta

discusión se fortalece con los planteamientos de Zarsky (2016), quien sostiene que la falta de supervisión algorítmica puede producir daños estructurales al sistema judicial. Los resultados del estudio confirman que los países que aplican auditorías independientes logran reducir errores y aumentar la transparencia, lo cual refuerza la conclusión de que la auditoría es un componente esencial de la gobernanza tecnológica. La discusión muestra que la legitimidad social del uso de IA en los tribunales depende en gran medida de la transparencia institucional y de la claridad con que se comuniquen sus procedimientos. Los resultados coinciden con Floridi (2021), quien enfatiza que la aceptación de sistemas tecnológicos requiere confianza pública basada en principios éticos y comunicaciones institucionales claras. La discusión confirma que, cuando la ciudadanía comprende los límites y funciones de los sistemas automatizados, aumenta la confianza en la justicia digital; mientras que la falta de claridad genera desconfianza y resistencia social. En conjunto, la discusión evidencia que la IA puede fortalecer el sistema judicial únicamente cuando se integra en un marco regulatorio robusto, éticamente fundado y socialmente legitimado.

Conclusiones

Los resultados obtenidos permiten concluir que la inteligencia artificial constituye un recurso estratégico para la modernización de los sistemas judiciales, siempre que su implementación se sustente en principios éticos, jurídicos y técnicos que garanticen el respeto irrestricto de los derechos fundamentales. Se evidenció que la IA tiene la capacidad de optimizar procesos administrativos, agilizar la clasificación de expedientes, reducir tiempos de tramitación y aportar coherencia jurisprudencial mediante el análisis automatizado de grandes

volúmenes de información. No obstante, estas ventajas solo adquieren legitimidad cuando la tecnología actúa como un instrumento complementario al trabajo humano y no como un sustituto de la interpretación jurídica. La eficiencia tecnológica, por sí sola, no es suficiente para justificar la adopción acrítica de sistemas automatizados, pues cualquier innovación en el ámbito judicial debe estar subordinada a la protección del debido proceso, la igualdad ante la ley y la imparcialidad que caracteriza la función jurisdiccional. De esta manera, la IA se convierte en un medio valioso para el fortalecimiento institucional, siempre que se implemente bajo supervisión humana significativa y en consonancia con estándares normativos que regulen su uso. Asimismo, se concluye que los riesgos identificados, especialmente el sesgo algorítmico, la opacidad decisional y la delegación indebida de funciones jurisdiccionales, representan desafíos estructurales que pueden comprometer la legitimidad y confiabilidad del sistema judicial. Estos riesgos no dependen únicamente de la arquitectura tecnológica, sino también de la capacidad de los operadores jurídicos para interpretar críticamente los resultados generados por los algoritmos. La evidencia sugiere que la alfabetización digital del personal judicial es un elemento indispensable para garantizar un uso responsable y consciente de estas herramientas, evitando dependencias excesivas que limiten la deliberación humana. Además, se constató que la protección de datos personales constituye un pilar fundamental en la construcción de una justicia digital segura, ya que el tratamiento inadecuado de información sensible puede generar vulneraciones graves. En consecuencia, se reitera la necesidad de fortalecer los mecanismos de supervisión, establecer protocolos de seguridad más estrictos y consolidar una cultura institucional basada en la transparencia tecnológica.

También se concluye que la integración adecuada de la inteligencia artificial en la justicia requiere un marco regulatorio robusto, actualizado y coherente con los principios del Estado de derecho. Los resultados evidencian que los países que cuentan con marcos normativos claros y políticas públicas bien definidas han logrado minimizar riesgos, aumentar la transparencia y mejorar la aceptación social de los sistemas digitales. Por el contrario, la ausencia de regulación favorece prácticas tecnológicas improvisadas, dificulta la supervisión y genera un escenario de incertidumbre jurídica que afecta directamente la legitimidad institucional. La regulación, por tanto, no solo ordena el uso de la tecnología, sino que define sus límites, responsabilidad y alcance dentro del sistema judicial. En este sentido, la existencia de normas claras y mecanismos de rendición de cuentas es esencial para garantizar que la IA contribuya efectivamente al fortalecimiento institucional y no a su debilitamiento. Del mismo modo, los resultados mostraron que la construcción de una justicia digital confiable requiere no solo regulación, sino también la existencia de estructuras permanentes de monitoreo, unidades técnicas calificadas y protocolos específicos que permitan evaluar el funcionamiento de los sistemas automatizados. La investigación concluye que la IA no transforma por sí sola los sistemas judiciales, sino que exige un ecosistema institucional preparado para integrarla, supervisarla y corregirla. Esto implica desarrollar capacidades internas para evaluar riesgos, identificar errores, implementar auditorías continuas y actualizar los modelos conforme evolucionen los marcos éticos y tecnológicos. La sostenibilidad de la justicia digital depende, así, de la capacidad de las instituciones para adaptarse a un entorno dinámico, incorporando procesos de mejora

continúa que eviten la obsolescencia de los sistemas y garanticen su uso seguro.

Se concluye que la aceptación social de la inteligencia artificial en los tribunales dependerá de la transparencia institucional, la claridad comunicacional y la confianza ciudadana en que la tecnología se utiliza de manera ética y justa. Los resultados muestran que la ciudadanía reconoce el potencial de la IA para mejorar la eficiencia judicial, pero también expresa preocupación frente a los riesgos de automatización excesiva, decisiones opacas o vulneraciones a derechos. En este sentido, es imprescindible que las instituciones adopten políticas de comunicación que expliquen el funcionamiento, alcance y límites de los sistemas tecnológicos utilizados. En conjunto, la investigación permite afirmar que la IA puede convertirse en un aliado significativo para el fortalecimiento del Estado de derecho, siempre que su implementación sea prudente, gradual, supervisada y orientada por principios humanistas. La justicia del futuro deberá integrar innovaciones tecnológicas sin renunciar a su esencia: proteger la dignidad humana y garantizar el acceso a una justicia imparcial, transparente y profundamente humana.

Referencias Bibliográficas

- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). *Machine Bias: There's software used across the country to predict future criminals. And it's biased against Blacks*. ProPublica.
- Bardin, L. (2011). *Content analysis*. McGraw-Hill.
- CEPAL. (2021). *Inteligencia Artificial y Transformación Digital en América Latina y el Caribe*. Comisión Económica para América Latina y el Caribe.
- European Commission. (2021). *Ethics Guidelines for Trustworthy Artificial Intelligence*. Publications Office of the European Union.
- Floridi, L. (2021). *The ethics of artificial intelligence*. Oxford University Press.
- Guba, E., & Lincoln, Y. S. (1985). *Naturalistic inquiry*. SAGE Publications.
- Hernández, R., Fernández, C., & Baptista, P. (2020). *Metodología de la investigación* (7.^a ed.). McGraw-Hill.
- Kroll, A. (2017). The fallacy of inscrutability. *University of Pennsylvania Law Review*, 165(3), 633–706.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444.
- Mitchell, T. (2020). *Machine learning* (2.^a ed.). McGraw-Hill.
- OECD. (2019). *Principles on Artificial Intelligence*. Organisation for Economic Co-operation and Development.
- Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems*, 37(43), 879–910.
- Pagallo, U. (2018). *The laws of robotics: Accountability, responsibility and AI*. Springer.
- Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4.^a ed.). Pearson.
- Snyder, H. (2019). Literature review as a research methodology. *Journal of Business Research*, 104, 333–339.
- Sourdin, T. (2020). *Judges, technology and artificial intelligence*. Edward Elgar Publishing.
- UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. UNESCO Publishing.
- Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation does not

exist in the GDPR. *International Data Privacy Law*, 7(2), 76–99.

Zarsky, T. (2016). The trouble with algorithmic decisions. *Science, Technology & Human Values*, 41(1), 118–132.



Esta obra está bajo una licencia de **Creative Commons Reconocimiento-No Comercial 4.0 Internacional**. Copyright © María Gracia Paredes Morales, Narda Polette León Domínguez, Andrés Alberto Pinela López, Andrea Joselyne Casco Carvajal, 5Bertha Yessena Rodríguez Vidal.

