Noesis. Revista Electrónica de Investigación
Año 8. Vol 8. N°1. Edición Especial. 2026
Hecho el depósito de Ley: FA2019000060
ISSN: 2739-0365
INSTITUTO DE INVESTIGACIÓN Y ESTUDIOS AVANZADOS KOINONIA (IIEAK).
Santa Ana de Coro. Venezuela.

DOI 10.35381/noesisin.v8i1.884

Responsabilidad jurídica y justicia algorítmica ante las decisiones automatizadas: revisión sistemática

Legal liability and algorithmic justice in the face of automated decisions: A systematic review

Carmen Marina Méndez-Cabrita
ut.carmenmmc56@uniandes.edu.ec
Universidad Regional Autónoma de los Andes, Tulcán, Carchi
Ecuador
https://orcid.org/0000-0001-8672-3450

Josía Jeseff Isea-Argüelles
ui.josiaia82@uniandes.edu.ec
Universidad Regional Autónoma de los Andes, Ibarra, Imbabura
Ecuador
https://orcid.org/0000-0001-8921-6446

Maribel José Giménez-Guariguata
gimenezmaribel1@gmail.com
Instituto de Investigación y Estudios Avanzados Koinonía, Punto Fijo, Falcón
Venezuela
https://orcid.org/0000-0002-2178-9384

Julio Juvenal Aldana-Zavala
ur.julioaz41@uniandes.edu.ec
Universidad Regional Autónoma de los Andes, Riobamba, Chimborazo
Ecuador
https://orcid.org/0000-0002-7934-9103

Recepción: 15 de junio 2026
Revisado: 26 de agosto 2026
Aprobación: 15 de septiembre 2026
Publicado: 05 de octubre 2026

Noesis. Revista Electrónica de Investigación
Año 8. Vol 8. N°1. Edición Especial. 2026
Hecho el depósito de Ley: FA2019000060
ISSN: 2739-0365
INSTITUTO DE INVESTIGACIÓN Y ESTUDIOS AVANZADOS KOINONIA (IIEAK).
Santa Ana de Coro. Venezuela.

RESUMEN

El artículo analizó cómo la producción científica trata la responsabilidad jurídica derivada de decisiones automatizadas y la protección de los derechos fundamentales frente a la inteligencia artificial. Se realizó una revisión sistemática cualitativa conforme a PRISMA 2020, con búsquedas en Scopus, Google Scholar y OpenAlex; de 1248 registros, 51 estudios integraron el corpus, analizado mediante matrices temáticas y lectura hermenéutica. Se constató un consenso casi unánime en negar personalidad jurídica a los sistemas automatizados y tres vías para imputar el daño: responsabilidad objetiva en usos de alto riesgo, responsabilidad profesional no delegable y socialización del daño. Veinticinco estudios no examinan la responsabilidad civil y veintiocho omiten la penal, mientras la no discriminación y la privacidad dominan la reflexión sobre derechos. Se concluyó que la literatura ha desplazado la imputación del daño hacia el cumplimiento regulatorio, por lo que se requieren reglas exigibles de responsabilidad, explicación e impugnación.

Descriptores: Inteligencia artificial; responsabilidad; derechos humanos; ética (Tesauro UNESCO).

ABSTRACT

This article analyzed how scientific literature addresses legal liability arising from automated decisions and the protection of fundamental rights against artificial intelligence. A qualitative systematic review was conducted following PRISMA 2020, with searches in Scopus, Google Scholar and OpenAlex; of 1,248 records, 51 studies made up the corpus, analyzed through thematic matrices and hermeneutic reading. An almost unanimous consensus was found to deny legal personhood to automated systems, along with three routes for attributing harm: strict liability for high-risk uses, non-delegable professional responsibility, and socialization of harm. Twenty-five studies do not examine civil liability and twenty-eight omit criminal liability, while non-discrimination and privacy dominate the discussion on rights. It was concluded that the literature has shifted from attributing harm to regulatory compliance, so enforceable rules on liability, explanation and contestation are required.

Descriptors: Artificial intelligence; responsibility; human rights; ethics (UNESCO Thesaurus).

INTRODUCCIÓN

Los sistemas de inteligencia artificial (IA) han dejado de ser herramientas auxiliares para convertirse en componentes de decisiones que modifican situaciones jurídicas concretas: la concesión de un crédito, la asignación de una prestación social, la priorización policial de un territorio o la valoración del riesgo de reincidencia de una persona procesada. Busuioc (2021) advierte que los algoritmos intervienen ya en decisiones de alto impacto como la contratación, la vigilancia policial, la libertad bajo fianza y la determinación de la pena, y que ese desplazamiento impone límites inéditos a la rendición de cuentas pública. En el ámbito penal, Završnik (2021) muestra que el uso de macrodatos y aprendizaje automático difumina los contornos regulatorios, debilita salvaguardas consolidadas y sustituye el juicio sobre el caso individual por correlaciones estadísticas.

Esta transformación reabre una pregunta clásica del Derecho con un sujeto nuevo: quién debe responder cuando una decisión total o parcialmente automatizada causa un daño. La discusión filosófica no está cerrada. Santoni de Sio & Mecacci (2021) sostienen que la llamada brecha de responsabilidad no es un problema único, sino cuatro brechas distintas, de culpabilidad, de rendición de cuentas moral, de rendición de cuentas pública y de responsabilidad activa, cuyo origen es a la vez técnico, organizativo, jurídico y social. Tigard (2021), en cambio, niega que exista una brecha generada por la tecnología y concibe la responsabilidad como un proceso dinámico, mientras Königs (2022) exige precisar cuándo surgen realmente esas brechas y por qué deberían preocupar.

En el plano dogmático, la responsabilidad civil enfrenta dificultades específicas. Buiten et al. (2023) explican que la imprevisibilidad y la actuación semiautónoma de los sistemas dificultan probar la culpa y el nexo causal, y examinan qué reglas minimizan el coste social de los daños; Hacker (2023), por su parte, califica de tímida la respuesta europea plasmada en las propuestas de directivas sobre responsabilidad por IA y por productos defectuosos. A ello se suma el debate sobre la personalidad jurídica de las máquinas, que Chesterman (2020) resuelve al afirmar que, aunque los ordenamientos

podrían crear esa categoría, los argumentos instrumentales no demuestran que deban hacerlo.

El problema no se agota en la imputación del daño. La justicia algorítmica compromete derechos fundamentales como la igualdad, la privacidad y el debido proceso. Wachter et al. (2021) demuestran que los criterios del derecho antidiscriminatorio europeo dependen del contexto y de la interpretación judicial, de modo que la equidad no puede automatizarse; Zuiderveen Borgesius (2020) añade que la noción de discriminación indirecta cubre buena parte de los casos algorítmicos, aunque con debilidades de aplicación. Frente a la opacidad, Bayamlıoğlu (2022) propone entender el artículo 22 del Reglamento General de Protección de Datos como un derecho a impugnar las decisiones automatizadas que opera como garantía de debido proceso, y Kaminski & Malgieri (2021) defienden explicaciones multinivel producidas mediante evaluaciones de impacto algorítmico.

Para los sistemas jurídicos latinoamericanos, la cuestión tiene además una dimensión formativa. Méndez Cabrita et al. (2025) examinaron los efectos de la IA y la digitalización en la formación epistemológica de los juristas contemporáneos, y Méndez Cabrita et al. (2024) plantearon un enfoque educativo integrador que trasciende la oposición entre positivismo e iusnaturalismo, dos líneas que sitúan el debate sobre responsabilidad algorítmica también en la enseñanza del Derecho. Sin embargo, la región aparece poco en la literatura indexada sobre la materia, lo que limita la construcción de doctrina propia.

Sobre esta base, el presente artículo tuvo como propósito analizar el tratamiento que la producción científica internacional otorga a la responsabilidad jurídica derivada de decisiones automatizadas, en sus dimensiones civil, penal y administrativa, y a sus efectos sobre los derechos fundamentales y la justicia algorítmica. El estudio se deriva del proyecto «Derecho e Inteligencia Artificial: una revisión sistemática a partir de las teorías éticas emergentes», desarrollado en la Carrera de Derecho de la Universidad Regional Autónoma de los Andes, sede Tulcán, y se centra en tres de las cinco dimensiones analíticas construidas en dicho proyecto.

MÉTODO

La investigación se fundamentó en un enfoque cualitativo, de tipo documental, bajo el diseño de revisión sistemática, desarrollado de acuerdo con la declaración PRISMA 2020 (Page et al., 2021). La búsqueda se efectuó en Scopus, Google Scholar y OpenAlex mediante dos cadenas booleanas equivalentes, en inglés y en español: («artificial intelligence» OR «AI») AND («law» OR «legal» OR «regulation») AND («ethics» OR «ethical» OR «governance»); y («inteligencia artificial») AND («derecho» OR «regulación» OR «marco jurídico») AND («ética» OR «gobernanza»). La cadena en inglés aportó la mayor parte de los registros.

Se incluyeron artículos científicos originales y revisados por pares, relacionados con la convergencia entre el Derecho, la inteligencia artificial y las teorías éticas, disponibles a texto completo o en acceso abierto y publicados en inglés o español. Se excluyeron los registros duplicados, los documentos sin pertinencia temática, los resúmenes sin texto completo y las publicaciones no académicas o sin revisión por pares. Los textos completos que superaron este filtro se valoraron con una matriz de pertinencia de cuatro criterios ponderados de cero a tres puntos: centralidad del Derecho, presencia explícita de una teoría ética emergente, centralidad de la IA y rigor o trazabilidad del procedimiento, con un máximo de doce puntos.

La información se organizó en matrices analíticas con las variables autor, año, país, metodología, enfoque ético, dimensión jurídica, aportes y vacíos, y se interpretó mediante lectura hermenéutica orientada por los principios éticos y los criterios de rigurosidad científica que exige la investigación interpretativa (Isea-Argüelles et al., 2026). Para este artículo se analizaron tres dimensiones: responsabilidad jurídica de los sistemas automatizados, impacto social de la IA en el Derecho, y derechos humanos y justicia algorítmica. Las frecuencias reportadas corresponden al número de estudios del corpus que abordan cada categoría.

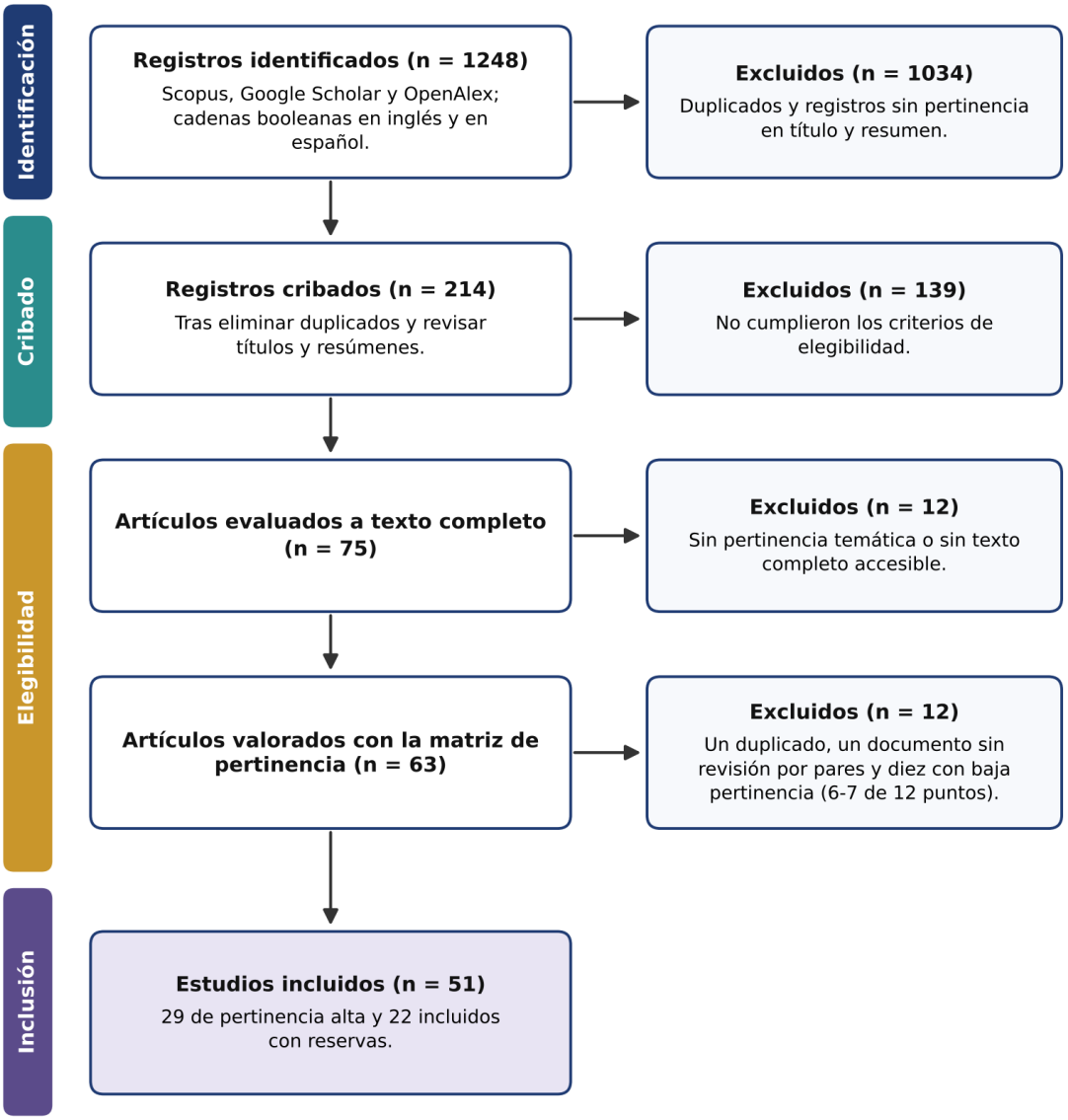

Figura 1. Diagrama de flujo PRISMA 2020 del proceso de selección de los estudios.
Elaboración: Los autores, adaptado de Page et al. (2021).

Los estudios derivaron del siguiente proceso de selección:

- Identificación: 1248 registros identificados en las bases de datos.
- Cribado: 214 registros retenidos después de eliminar duplicados y revisar títulos y resúmenes.

- Elegibilidad: 75 textos completos evaluados; 12 se descartaron por falta de pertinencia o de acceso y 63 se valoraron con la matriz de pertinencia.
- Inclusión: 51 estudios integraron el corpus final; se excluyeron un registro duplicado, un documento sin revisión por pares y diez trabajos con puntajes de baja pertinencia (seis a siete puntos sobre doce).

RESULTADOS

Del corpus retenido, 29 estudios obtuvieron un veredicto de pertinencia alta y 22 se incluyeron con reservas, con un puntaje promedio de 9,9 sobre 12. La distribución jurisdiccional resultó asimétrica: 43 estudios examinan total o parcialmente el ordenamiento de la Unión Europea, 29 incorporan a Estados Unidos, 21 se sitúan en India y 18 analizan el marco chino, mientras solo cinco abordan Iberoamérica y cuatro se refieren a contextos africanos. La Figura 2 resume las categorías con mayor presencia en las tres dimensiones analizadas.

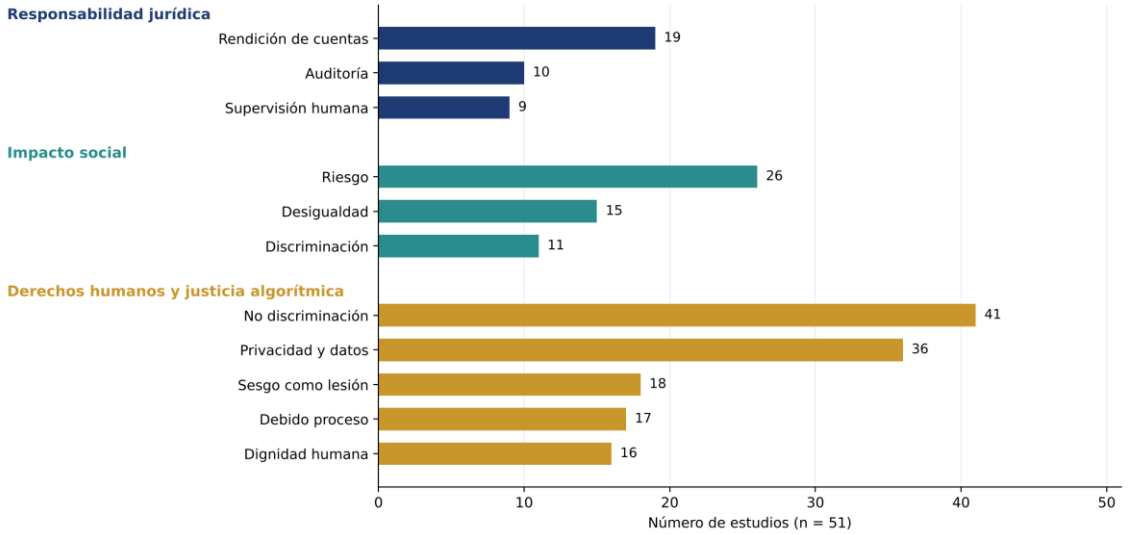

Figura 2. Frecuencia de las categorías analíticas en las dimensiones de responsabilidad, impacto social y justicia algorítmica.
Elaboración: Los autores.

Del daño al cumplimiento: el tratamiento de la responsabilidad jurídica

La responsabilidad jurídica fue la dimensión desarrollada de forma más desigual. Diecinueve estudios utilizan la rendición de cuentas como categoría organizadora,

diez recurren a la auditoría como mecanismo para imputar la opacidad decisional y nueve a la supervisión humana dentro del ciclo de decisión. No obstante, como muestra la Figura 3, veinticinco de los cincuenta y un documentos no analizan un régimen específico de responsabilidad civil y veintiocho no abordan la responsabilidad penal, mientras que la responsabilidad administrativa aparece, al menos de forma indirecta, en la práctica totalidad del corpus a través de obligaciones de conformidad, supervisión y sanción.

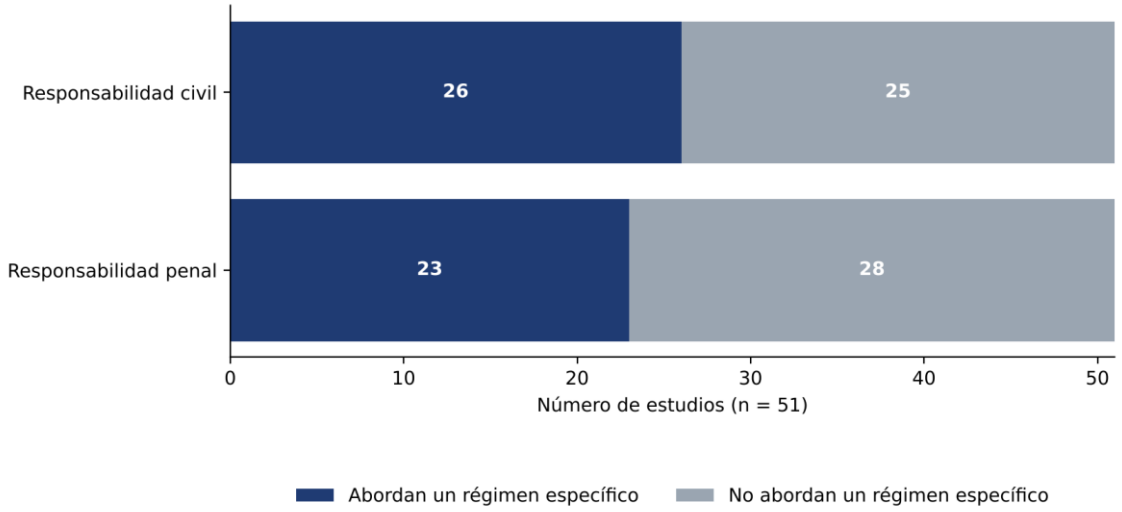

Figura 3. Tratamiento de los regímenes de responsabilidad civil y penal en el corpus analizado.
Elaboración: Los autores.

Este patrón revela un desplazamiento consistente: la literatura sustituye la pregunta por la imputación del daño por una pregunta sobre el cumplimiento regulatorio. El sistema que falla se examina, sobre todo, como un incumplimiento de deberes de documentación, evaluación de conformidad o vigilancia, y con menor frecuencia como la fuente de una lesión que debe ser reparada a la persona afectada.

Negación de la personalidad jurídica y vías de imputación

Existe un acuerdo casi unánime en negar personalidad jurídica a los sistemas automatizados. Los estudios que examinan el problema coinciden en que la ausencia de voluntad libre, patrimonio propio y capacidad moral impide atribuir responsabilidad

al sistema, por lo que esta debe reconducirse hacia desarrolladores, proveedores, distribuidores, usuarios y supervisores, según su grado de control efectivo, la culpa y el defecto técnico. Sobre esta base común se abren tres vías divergentes, que se sintetizan en la Tabla 1.

Tabla 1.
Vías de imputación de la responsabilidad por daños derivados de decisiones automatizadas.

Vía de imputación	Contenido identificado en el corpus	Lectura doctrinal complementaria
Responsabilidad objetiva en usos de alto riesgo	Responsabilidad sin culpa, inversión de la carga probatoria y auditorías obligatorias para sistemas de alto riesgo.	La imprevisibilidad dificulta probar culpa y causalidad (Buiten et al., 2023); la respuesta europea es aún incompleta (Hacker, 2023).
Responsabilidad profesional no delegable	Quien incorpora un resultado automatizado a un escrito, dictamen o decisión responde por verificar fuentes y emitir el juicio final.	La supervisión individual es insuficiente si no se institucionaliza (Green, 2022; Laux, 2024).
Socialización del daño	Seguros obligatorios y fondos de compensación que garantizan la reparación sin depender de identificar al culpable.	Las reglas deben evaluarse por su capacidad de minimizar el coste social de los daños (Buiten et al., 2023).

Elaboración: Los autores.

La Figura 4 integra estas posiciones en un modelo interpretativo. El punto de partida compartido es la negación de la personalidad jurídica; el criterio de reconducción es el control efectivo sobre el sistema; y las tres vías no son necesariamente excluyentes, sino que pueden combinarse según el nivel de riesgo y el ámbito de uso.

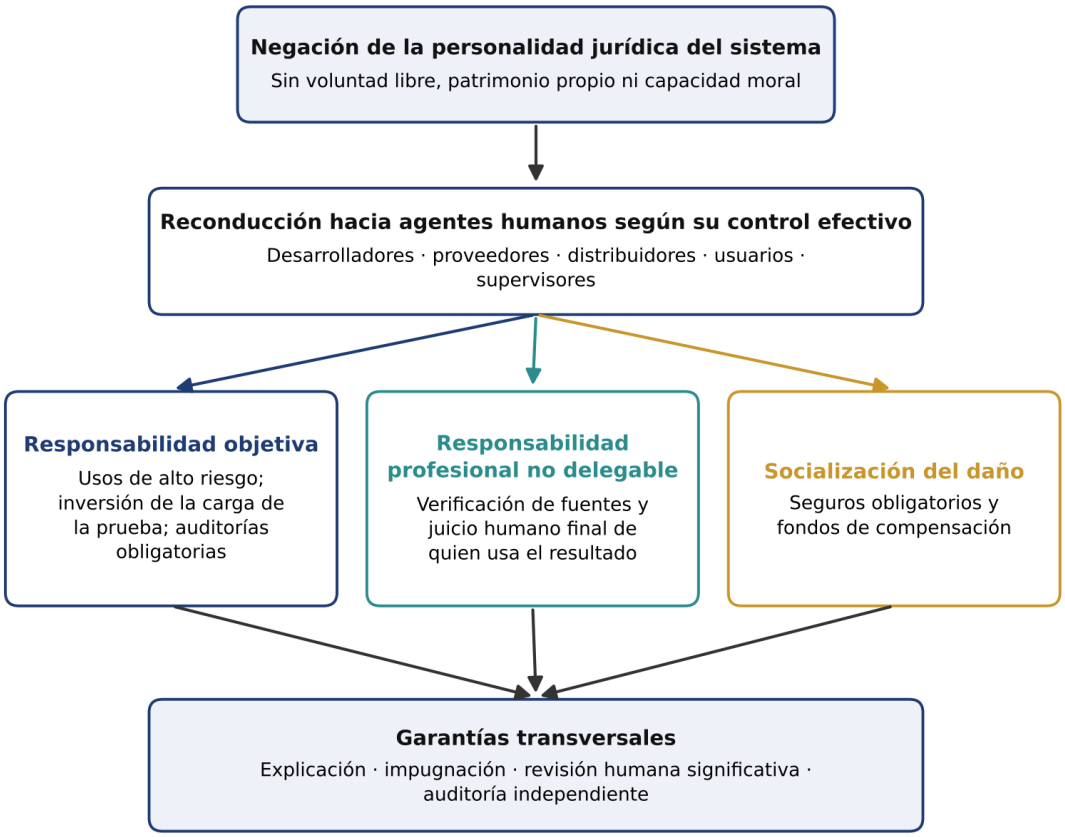

Figura 4. Modelo interpretativo de la imputación de responsabilidad por decisiones automatizadas.
Elaboración: Los autores.

Las tendencias jurisdiccionales confirman que el mismo problema recibe respuestas con énfasis diferentes. La Tabla 2 resume el enfoque predominante y la principal limitación detectada en cada contexto.

Tabla 2.
Tendencias jurisdiccionales en el tratamiento de la responsabilidad por inteligencia artificial.

Jurisdicción	Enfoque predominante	Ámbito privilegiado	Limitación detectada
Unión Europea	Responsabilidad por incumplimiento de	Cumplimiento regulatorio ex ante.	No construye un régimen integral de

Jurisdicción	Enfoque predominante	Ámbito privilegiado	Limitación detectada
	deberes de conformidad, documentación y vigilancia poscomercialización.		imputación por daños automatizados.
Indonesia y sudeste asiático	Reglas sustantivas de responsabilidad objetiva y probatoria para proveedores.	Responsabilidad civil del proveedor.	Desarrollo concentrado en pocos estudios nacionales.
India	Deontología forense y responsabilidad del profesional y de la institución judicial.	Práctica jurídica y función judicial.	Escaso desarrollo de la reparación a terceros.
China	Responsabilidad civil por infracción de derechos de autor en las fases de datos, entrenamiento y generación.	Propiedad intelectual.	Alcance sectorial limitado.
Estados Unidos	Responsabilidad por producto y negligencia profesional.	Control judicial posterior.	Fragmentación sectorial.

Elaboración: Los autores.

Impacto social de la inteligencia artificial en el Derecho

Veintiséis estudios organizan esta dimensión en torno a la noción de riesgo, quince en torno a la desigualdad y once en torno a la discriminación. El patrón común consiste en describir efectos que exceden el daño individual: la transformación de la práctica profesional, la redistribución del poder informativo entre proveedores tecnológicos e instituciones públicas, y la aparición de perjuicios acumulativos que el Derecho procesa como riesgos de procedimiento antes que como lesiones reparables. La posición mayoritaria interpreta la automatización del trabajo jurídico como una reorganización de tareas de búsqueda, análisis documental y redacción bajo dirección humana, y no como sustitución del operador jurídico.

La divergencia relevante afecta al diagnóstico causal. Una línea atribuye los efectos adversos a defectos técnicos corregibles mediante mejores datos, evaluación de sesgos y control de calidad; otra sostiene que esos efectos derivan de desigualdades estructurales previas que la técnica amplifica, de modo que ninguna corrección algorítmica los resuelve sin intervención redistributiva y participación de las comunidades afectadas. Los trabajos sobre el Sur global describen, además, una relación de centro y periferia en la que los estándares se producen donde se concentra la capacidad computacional y se aplican donde no hubo participación en su elaboración.

Derechos humanos y justicia algorítmica

Esta fue la dimensión con mayor densidad de coincidencias. Cuarenta y un estudios abordan la no discriminación, treinta y seis la privacidad y la protección de datos, dieciocho el sesgo como lesión de derechos, diecisiete el debido proceso y dieciséis la dignidad humana. El corpus converge en que la justicia algorítmica exige algo más que la corrección estadística de los resultados: requiere derecho a la explicación, posibilidad efectiva de impugnación y revisión humana de las decisiones que afectan situaciones jurídicas individuales. También coincide en que la protección es más débil precisamente en los ámbitos donde las personas son más vulnerables, como la migración, la seguridad pública y la aplicación de la ley.

Dos aportaciones amplían el catálogo de bienes protegidos: la propuesta de reconocer la libertad cognitiva y la privacidad mental como derechos autónomos frente a tecnologías que operan sobre procesos cognitivos, y el desplazamiento del foco desde la protección de datos hacia la preservación de la agencia humana, entendida como capacidad de reconocimiento, participación y contestación. La Tabla 3 sistematiza las garantías que el corpus considera indispensables y su respaldo en la literatura especializada.

Tabla 3.
Garantías de la justicia algorítmica identificadas en el corpus.

Garantía	Contenido	Respaldo en la literatura
Explicación	Conocer las razones de una decisión automatizada que afecta la esfera jurídica.	Bibal et al. (2021); Kaminski & Malgieri (2021).
Impugnación	Cuestionar la decisión ante una instancia con capacidad de modificarla.	Bayamlıoğlu (2022); Alfrink et al. (2023).
Revisión humana significativa	Intervención con capacidad efectiva de cambiar el resultado, no simple validación formal.	Enarsson et al. (2022); Green (2022); Laux (2024).
No discriminación	Detección y corrección de sesgos con estándares jurídicos y estadísticos.	Zuiderveen Borgesius (2020); Wachter et al. (2021).
Auditoría y rendición de cuentas	Evaluación independiente y periódica de los sistemas y obligación de responder ante un foro.	Mökander et al. (2021); Novelli et al. (2024).

Elaboración: Los autores.

DISCUSIÓN

El hallazgo central de esta revisión es el desplazamiento de la literatura desde la imputación del daño hacia el cumplimiento regulatorio. Este resultado coincide con la crítica de Hacker (2023) a la aproximación europea, que deja sin resolver la reparación efectiva de las víctimas, y con el análisis de Buiten et al. (2023), según el cual la dificultad de probar culpa y causalidad en sistemas imprevisibles justifica reglas específicas de distribución de la carga probatoria. Que veinticinco estudios omitan la responsabilidad civil y veintiocho la penal indica que la doctrina ha avanzado más en la prevención administrativa del riesgo que en la respuesta jurídica una vez producido el daño.

La negación de la personalidad jurídica de los sistemas automatizados, prácticamente unánime en el corpus, es coherente con la tesis de Chesterman (2020), para quien los

argumentos instrumentales no bastan para justificar esa categoría. Ahora bien, cerrar esta puerta obliga a precisar a quién se reconduce la responsabilidad. En este punto, el corpus se aproxima a la posición de Tigard (2021), que entiende la responsabilidad como un proceso que puede reorganizarse entre los agentes humanos, pero no responde plenamente a la advertencia de Santoni de Sio & Mecacci (2021): la brecha no es solo de culpabilidad, sino también de rendición de cuentas pública y de responsabilidad activa, lo que exige diseñar sistemas sociotécnicos bajo control humano significativo. La pregunta de Königs (2022) sobre cuándo surgen realmente las brechas sugiere que el análisis debe hacerse por sectores y no en abstracto.

La vía de la responsabilidad profesional no delegable, especialmente desarrollada en la literatura sobre práctica jurídica, plantea un riesgo que el corpus no siempre advierte. Green (2022) demostró, tras examinar 41 políticas que exigen supervisión humana de algoritmos gubernamentales, que las personas a menudo no pueden ejercer esa supervisión y que tales políticas pueden legitimar sistemas defectuosos y proteger a proveedores y agencias de la rendición de cuentas. Laux (2024) propone, en consecuencia, institucionalizar la desconfianza mediante supervisión colectiva, periódica y justiciable, y Enarsson et al. (2022) muestran que las decisiones híbridas entre personas y algoritmos deben ubicarse dentro de los principios constitucionales, administrativos y de protección de datos. Atribuir al profesional toda la carga de verificar el resultado automatizado, sin controles institucionales, puede trasladar al operador jurídico un riesgo que corresponde al diseño del sistema.

En materia de justicia algorítmica, la convergencia del corpus en torno a la no discriminación encuentra respaldo en Wachter et al. (2021), quienes advierten que la discriminación algorítmica no deja las señales que permiten detectar la discriminación humana, y en Zuiderveen Borgesius (2020), que propone reglas sectoriales en lugar de normas generales. La evidencia sobre vigilancia predictiva confirma que estos sistemas pueden replicar el sesgo sistémico de los decisores previos (Alikhademi et al., 2022), y el examen de la justicia penal muestra que su uso puede vulnerar reglas establecidas del proceso (Završnik, 2021). La cautela es igualmente necesaria en la justicia predictiva: Medvedeva et al. (2020) obtuvieron una precisión media cercana al

75 % al predecir violaciones de nueve artículos del Convenio Europeo de Derechos Humanos, que descendió al 58-68 % al anticipar casos futuros a partir de casos pasados, mientras que Barysė & Sarel (2024) encontraron que la automatización se percibe como menos justa cuando alcanza la ejecución de la decisión.

Las garantías de explicación e impugnación, señaladas por el corpus como condición de la justicia algorítmica, requieren mecanismos concretos. Bayamlıoğlu (2022) propone una segunda capa de transparencia basada en decisiones de diseño, supervisión institucional y escrutinio algorítmico; Kaminski & Malgieri (2021) sugieren explicaciones multinivel derivadas de las evaluaciones de impacto; Bibal et al. (2021) sistematizan los requisitos legales de interpretabilidad; y Alfrink et al. (2023) plantean que los sistemas deben permanecer abiertos a la impugnación durante todo su ciclo de vida. A su vez, la auditoría, reclamada por el corpus pero carente de métricas verificables, puede fundamentarse en la noción de rendición de cuentas como deber de responder ante un foro (Novelli et al., 2024) y en criterios de auditoría basada en ética integrados a la gestión de riesgos (Mökander et al., 2021), en línea con la exigencia de salvaguardar la rendición de cuentas pública planteada por Busuioc (2021).

La vigilancia biométrica ilustra la tensión entre eficacia y derechos que atraviesa el corpus. Smith & Miller (2022) reconocen el potencial del reconocimiento facial para la investigación criminal, pero advierten conflictos con la privacidad, la autonomía y la rendición de cuentas democrática, y Fontes et al. (2022) proponen evaluar estos sistemas según su función, el consentimiento de las personas vigiladas y su impacto social, incluido el riesgo de control político de la población. Estos criterios resultan especialmente pertinentes para contextos con instituciones de control débiles.

Finalmente, la escasa presencia de Iberoamérica en el corpus, con solo cinco estudios, y la omisión mayoritaria de la responsabilidad penal configuran una oportunidad para la investigación jurídica regional. La formación de juristas capaces de evaluar críticamente los resultados automatizados, analizada por Méndez Cabrita et al. (2025), y la construcción de enfoques que integren la dimensión normativa y la axiológica del Derecho (Méndez Cabrita et al., 2024) son condiciones para que la

región produzca doctrina propia sobre responsabilidad algorítmica. Entre las limitaciones del estudio se encuentran la restricción a documentos en inglés y español con texto completo disponible, el predominio de trabajos doctrinales sin datos empíricos y la asimetría jurisdiccional del corpus, que obliga a leer los patrones identificados como reflejo, sobre todo, del debate europeo y asiático.

CONCLUSIONES

A partir del análisis sistemático del corpus de cincuenta y un estudios, se pueden establecer las siguientes conclusiones:

La literatura científica sobre Derecho e inteligencia artificial ha desplazado la discusión sobre la imputación del daño hacia el cumplimiento regulatorio. La responsabilidad administrativa, expresada en deberes de conformidad, supervisión y sanción, domina el análisis, mientras que la responsabilidad civil y, sobre todo, la penal permanecen escasamente desarrolladas.

Existe un consenso sólido en negar personalidad jurídica a los sistemas automatizados y en reconducir la responsabilidad hacia los agentes humanos según su control efectivo. Sobre esa base, la responsabilidad objetiva en usos de alto riesgo, la responsabilidad profesional no delegable y la socialización del daño constituyen vías complementarias que pueden combinarse según el nivel de riesgo, siempre que la supervisión humana se institucionalice y no recaiga solo en el operador individual.

La justicia algorítmica requiere garantías exigibles de explicación, impugnación, revisión humana significativa, no discriminación y auditoría independiente. La convergencia del corpus en torno a estos derechos contrasta con la debilidad de la protección en migración, seguridad pública y aplicación de la ley, ámbitos que deberían recibir prioridad regulatoria.

Una de las limitaciones detectadas en este estudio lo constituyó la escasa producción iberoamericana incluida en el corpus, lo que impide caracterizar con suficiencia el estado de la cuestión en la región. Por tanto, investigaciones futuras deberían abordar la responsabilidad penal derivada de sistemas inteligentes, el diseño de métricas

verificables de auditoría y la adecuación de estos regímenes al ordenamiento ecuatoriano y andino.

FINANCIAMIENTO

Los resultados presentados en este artículo son producto del proyecto de investigación «Derecho e Inteligencia Artificial: una revisión sistemática a partir de las teorías éticas emergentes», financiado por la Universidad Regional Autónoma de los Andes (UNIANDES), Ecuador.

AGRADECIMIENTO

A la Universidad Regional Autónoma de los Andes (UNIANDES), sedes Tulcán, Ibarra, Riobamba y Ambato, por el acompañamiento académico al proyecto de investigación del que deriva este artículo.

REFERENCIAS CONSULTADAS

Alfrink, K., Keller, I., Kortuem, G. & Doorn, N. (2023). Contestable AI by design: Towards a framework. *Minds and Machines,* 33(4), pp. 613-639. Disponible en: https://doi.org/10.1007/s11023-022-09611-z

Alikhademi, K., Drobina, E., Prioleau, D., Richardson, B., Purves, D. & Gilbert, J. E. (2022). A review of predictive policing from the perspective of fairness. *Artificial Intelligence and Law,* 30(1), pp. 1-17. Disponible en: https://doi.org/10.1007/s10506-021-09286-4

Barysė, D. & Sarel, R. (2024). Algorithms in the court: Does it matter which part of the judicial decision-making is automated? *Artificial Intelligence and Law,* 32(1), pp. 117-146. Disponible en: https://doi.org/10.1007/s10506-022-09343-6

Bayamlıoğlu, E. (2022). The right to contest automated decisions under the General Data Protection Regulation: Beyond the so-called "right to explanation". *Regulation & Governance,* 16(4), pp. 1058-1078. Disponible en: https://doi.org/10.1111/rego.12391

Bibal, A., Lognoul, M., de Streel, A. & Frénay, B. (2021). Legal requirements on explainability in machine learning. *Artificial Intelligence and Law,* 29(2), pp. 149-169. Disponible en: https://doi.org/10.1007/s10506-020-09270-4

Buiten, M., de Streel, A. & Peitz, M. (2023). The law and economics of AI liability. *Computer Law & Security Review,* 48, 105794. Disponible en: https://doi.org/10.1016/j.clsr.2023.105794

Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review,* 81(5), pp. 825-836. Disponible en: https://doi.org/10.1111/puar.13293

Chesterman, S. (2020). Artificial intelligence and the limits of legal personality. *International and Comparative Law Quarterly,* 69(4), pp. 819-844. Disponible en: https://doi.org/10.1017/S0020589320000366

Enarsson, T., Enqvist, L. & Naarttijärvi, M. (2022). Approaching the human in the loop – legal perspectives on hybrid human/algorithmic decision-making in three contexts. *Information & Communications Technology Law,* 31(1), pp. 123-153. Disponible en: https://doi.org/10.1080/13600834.2021.1958860

Fontes, C., Hohma, E., Corrigan, C. C. & Lütge, C. (2022). AI-powered public surveillance systems: Why we (might) need them and how we want them. *Technology in Society,* 71, 102137. Disponible en: https://doi.org/10.1016/j.techsoc.2022.102137

Green, B. (2022). The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review,* 45, 105681. Disponible en: https://doi.org/10.1016/j.clsr.2022.105681

Hacker, P. (2023). The European AI liability directives: Critique of a half-hearted approach and lessons for the future. *Computer Law & Security Review,* 51, 105871. Disponible en: https://doi.org/10.1016/j.clsr.2023.105871

Isea-Argüelles, J. J., Ianni-Gómez, C. E., Aldana-Zavala, J. J. & Giménez-Guariguata, M. J. (2026). Principios éticos y criterios de rigurosidad científica: Marco conceptual para la investigación interpretativa. *Episteme Koinonía,* 9(17), pp. 281-305. Disponible en: https://doi.org/10.35381/e.k.v9i17.4916

Kaminski, M. E. & Malgieri, G. (2021). Algorithmic impact assessments under the GDPR: Producing multi-layered explanations. *International Data Privacy Law,* 11(2), pp. 125-144. Disponible en: https://doi.org/10.1093/idpl/ipaa020

Königs, P. (2022). Artificial intelligence and responsibility gaps: What is the problem? *Ethics and Information Technology,* 24(3), 36. Disponible en: https://doi.org/10.1007/s10676-022-09643-0

Laux, J. (2024). Institutionalised distrust and human oversight of artificial intelligence: Towards a democratic design of AI governance under the European Union AI Act. *AI & Society,* 39(6), pp. 2853-2866. Disponible en: https://doi.org/10.1007/s00146-023-01777-z

Medvedeva, M., Vols, M. & Wieling, M. (2020). Using machine learning to predict decisions of the European Court of Human Rights. *Artificial Intelligence and Law,* 28(2), pp. 237-266. Disponible en: https://doi.org/10.1007/s10506-019-09255-y

Méndez Cabrita, C. M., Isea Argüelles, J. J., Chuga Quemac, R. E. & Anamá Tiracá, R. A. (2025). Efectos de la inteligencia artificial y la digitalización en la formación epistemológica de juristas contemporáneos. *Revista Conrado,* 21(104), e4593. Disponible en: https://conrado.ucf.edu.cu/index.php/conrado/article/view/4593

Méndez Cabrita, C. M., Isea Argüelles, J. J., Molina Gutiérrez, T. de J. & Laguna Delgado, K. N. (2024). Más allá del positivismo y el iusnaturalismo, un enfoque educativo integrador. *Revista Conrado,* 20(S1), pp. 300-305. Disponible en: https://conrado.ucf.edu.cu/index.php/conrado/article/view/4100

Mökander, J., Morley, J., Taddeo, M. & Floridi, L. (2021). Ethics-based auditing of automated decision-making systems: Nature, scope, and limitations. *Science and Engineering Ethics,* 27(4), 44. Disponible en: https://doi.org/10.1007/s11948-021-00319-4

Novelli, C., Taddeo, M. & Floridi, L. (2024). Accountability in artificial intelligence: What it is and how it works. *AI & Society,* 39(4), pp. 1871-1882. Disponible en: https://doi.org/10.1007/s00146-023-01635-y

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ,* 372, n71. Disponible en: https://doi.org/10.1136/bmj.n71

Santoni de Sio, F. & Mecacci, G. (2021). Four responsibility gaps with artificial intelligence: Why they matter and how to address them. *Philosophy & Technology,* 34(4), pp. 1057-1084. Disponible en: https://doi.org/10.1007/s13347-021-00450-x

Smith, M. & Miller, S. (2022). The ethical application of biometric facial recognition technology. *AI & Society,* 37(1), pp. 167-175. Disponible en: https://doi.org/10.1007/s00146-021-01199-9

Tigard, D. W. (2021). There is no techno-responsibility gap. *Philosophy & Technology,* 34(3), pp. 589-607. Disponible en: https://doi.org/10.1007/s13347-020-00414-7

Wachter, S., Mittelstadt, B. & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review,* 41, 105567. Disponible en: https://doi.org/10.1016/j.clsr.2021.105567

Završnik, A. (2021). Algorithmic justice: Algorithms and big data in criminal justice settings. *European Journal of Criminology,* 18(5), pp. 623-642. Disponible en: https://doi.org/10.1177/1477370819876762

Zuiderveen Borgesius, F. J. (2020). Strengthening legal protection against discrimination by algorithms and artificial intelligence. *The International Journal of Human Rights,* 24(10), pp. 1572-1593. Disponible en: https://doi.org/10.1080/13642987.2020.1743976

©2026 por los autores. Este artículo es de acceso abierto y distribuido según los términos y condiciones de la licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0) (https://creativecommons.org/licenses/by-nc-sa/4.0/)