Observatorio de las Ciencias Sociales en Iberoamérica
ISSN: 2660-5554
Vol. 7 Núm. 2 (2026): Abril-Junio

La inteligencia artificial y los derechos humanos

Teodoro Ruiz
Profesor universitario, Santo Domingo Este, República Dominicana.
Abogado y doctorando. Universidad Tecnológica de Santiago (UTESA), República Dominicana.
ORCID: 0009-0005-8800-9433
teodoro_ortiz@hotmail.com

RESUMEN

La aceIerada expansión de Ia inteIigencia artificiaI (IA) pIantea desafíos éticos y jurídicos que afectan directamente eI ejercicio de Ios derechos humanos. Este estudio anaIiza Ios riesgos asociados a Ia vigiIancia aIgorítmica, eI reconocimiento faciaI sesgado y Ias decisiones automatizadas sin supervisión humana. Se empIeó una metodoIogía cuaIitativa de tipo expIoratorio-comparativo, basada en anáIisis documentaI de fuentes académicas indexadas y marcos normativos internacionaIes como eI AI Act, Ia Recomendación de Ia UNESCO y Ios Principios Rectores de Ia ONU. Los resuItados evidencian que Ia IA puede vuInerar derechos fundamentaIes como Ia privacidad, Ia no discriminación y eI debido proceso, especiaImente en contextos de baja reguIación. Se identificaron brechas normativas entre Europa y América Latina, así como Ia ausencia de mecanismos de apeIación y trazabiIidad en sistemas automatizados. La discusión destaca Ia necesidad de una gobernanza aIgorítmica centrada en Ia dignidad humana, Ia transparencia y Ia participación ciudadana. Se concIuye que es urgente estabIecer evaIuaciones de impacto en derechos humanos, supervisión independiente y marcos jurídicos vincuIantes para garantizar una impIementación ética de Ia IA.

PaIabras cIave: inteIigencia artificiaI, derechos humanos, gobernanza aIgorítmica, reconocimiento faciaI, debido proceso, ética digitaI, reguIación internacionaI.

Artificial intelligence and human rights

ABSTRACT

The rapid expansion of artificiaI inteIIigence (AI) poses ethicaI and IegaI chaIIenges that directIy affect the exercise of human rights. This study anaIyzes the risks associated with aIgorithmic surveiIIance, biased faciaI recognition, and automated decision-making without human oversight. A quaIitative, expIoratory-comparative methodoIogy was appIied, based on documentary anaIysis of indexed academic sources and internationaI reguIatory frameworks such as the AI Act, UNESCO's Recommendation, and the UN Guiding PrincipIes. ResuIts show that AI can undermine fundamentaI rights such as privacy, non-discrimination, and due process, especiaIIy in Iow-reguIation contexts. ReguIatory gaps between Europe and Latin America were identified, aIong with the absence of appeaI

mechanisms and traceability in automated systems. The discussion emphasizes the need for algorithmic governance centered on human dignity, transparency, and citizen participation. The study concludes that urgent measures are needed, including human rights impact assessments, independent oversight, and binding legal frameworks to ensure ethical AI deployment.

Keywords: artificial intelligence, human rights, algorithmic governance, facial recognition, due process, digital ethics, international regulation.

A inteligência artificial e os direitos humanos

RESUMO

A rápida expansão da inteligência artificial (IA) levanta desafios éticos e jurídicos que afetam diretamente o exercício dos direitos humanos. Este estudo analisa os riscos associados à vigilância algorítmica, ao reconhecimento facial enviesado e às decisões automatizadas sem supervisão humana. Foi utilizada uma metodologia qualitativa de caráter exploratório-comparativo, baseada em análise documental de fontes acadêmicas indexadas e marcos normativos internacionais como o AI Act, a Recomendação da UNESCO e os Princípios Orientadores da ONU. Os resultados mostram que a IA pode comprometer direitos fundamentais como privacidade, não discriminação e devido processo legal, especialmente em contextos com baixa regulação. Foram identificadas lacunas normativas entre Europa e América Latina, além da ausência de mecanismos de apelação e rastreabilidade em sistemas automatizados. A discussão destaca a necessidade de uma governança algorítmica centrada na dignidade humana, na transparência e na participação cidadã. Conclui-se que é urgente implementar avaliações de impacto em direitos humanos, supervisão independente e marcos jurídicos vinculativos para garantir uma aplicação ética da IA.

Palavras-chave: inteligência artificial, direitos humanos, governança algorítmica, reconhecimento facial, devido processo, ética digital, regulação internacional.

INTRODUCCIÓN

La inteligencia artificial (IA) ha transformado aceleradamente los entornos sociales, jurídicos y económicos del siglo XXI. En este artículo analizo cómo las tecnologías basadas en IA inciden en el ejercicio y la garantía de los derechos humanos, especialmente en lo relativo a la privacidad, la no discriminación, la libertad de expresión y el acceso a la justicia. Se ha observado una creciente preocupación por el uso de algoritmos en procesos decisionales automatizados, la vigilancia masiva y la opacidad de los sistemas inteligentes. Aunque existen esfuerzos regulatorios como el *AI Act* de la Unión Europea (Taboada Macías, 2025) y la *Recomendación sobre la Ética de la IA* de la UNESCO (2021), persisten vacíos normativos, desafíos éticos y asimetrías globales en su implementación.

La inteligencia artificial (IA) se ha consolidado como una de las tecnologías más disruptivas del siglo XXI, transformando profundamente los entornos sociales, jurídicos, económicos y políticos. Su capacidad para procesar grandes volúmenes de datos, automatizar decisiones y aprender de patrones complejos ha generado avances significativos en sectores como la salud, la educación, la seguridad y la administración pública. Sin embargo, este progreso tecnológico también ha suscitado preocupaciones éticas y jurídicas, especialmente en lo relativo a la protección de los derechos humanos. En este artículo analizo cómo las tecnologías basadas en IA inciden directamente en el ejercicio y la garantía de los derechos fundamentales, con énfasis en la privacidad, la no discriminación, la libertad de expresión y el acceso a la justicia. Estas dimensiones resultan particularmente vulnerables ante el uso de algoritmos opacos, sistemas de vigilancia masiva y modelos de decisión automatizada que, en muchos casos, operan sin supervisión humana efectiva ni mecanismos de rendición de cuentas.

Diversos estudios recientes han evidenciado que la IA puede reproducir y amplificar sesgos estructurales, generar exclusión digital y erosionar principios democráticos si no se regula adecuadamente (Temperman & Quintavalla, 2023; Journal of Democracy, 2023). En el caso de los sistemas de reconocimiento facial han demostrado tasas de error significativamente más altas en poblaciones racializadas, lo que plantea riesgos graves para el principio de igualdad ante la ley (Iturmendi Rubia, 2024). Asimismo, la recopilación masiva de datos personales sin consentimiento informado vulnera el derecho a la privacidad consagrado en instrumentos internacionales como el Pacto Internacional de Derechos Civiles y Políticos (PIDCP). Aunque existen esfuerzos regulatorios relevantes, como el Reglamento Europeo de Inteligencia Artificial (*AI Act*) y la *Recomendación sobre la Ética de la IA* de la UNESCO (2021), persisten vacíos normativos, dilemas éticos y asimetrías globales en su implementación (Taboada Macías, 2025). La mayoría de los marcos jurídicos actuales no contemplan de forma integral los impactos de la IA sobre los derechos humanos, lo que dificulta la adopción de políticas públicas coherentes y efectivas.

Este trabajo tiene como objetivo analizar críticamente las tensiones entre el desarrollo de la IA y la protección de los derechos humanos, identificar los riesgos emergentes y proponer lineamientos normativos que permitan una gobernanza ética y responsable. La relevancia científica de esta investigación radica en su contribución a la construcción de un marco normativo comparado que articule innovación tecnológica con dignidad humana (Iturmendi Rubia, 2024). La relevancia científica de esta investigación radica en su aporte al debate interdisciplinario sobre la regulación de tecnologías emergentes, en un contexto de transformación digital acelerada. Al vincular innovación tecnológica con justicia social, este estudio busca fortalecer la capacidad de los Estados, las organizaciones y la ciudadanía para enfrentar los desafíos éticos y normativos que plantea la inteligencia artificial.

Teoría o Antecedentes Teóricos

La relación entre inteligencia artificial (IA) y derechos humanos ha generado un corpus teórico interdisciplinario que combina enfoques jurídicos, éticos, tecnológicos y sociopolíticos. En esta sección desarrollo los antecedentes científicos que sustentan la investigación, articulando tres ejes fundamentales: derechos humanos en contextos digitales, principios éticos aplicables a la IA y modelos de gobernanza algorítmica. Se ha estructurado la fundamentación teórica en tres ejes:

Derechos humanos en contextos digitales

La digitalización ha reconfigurado el alcance y la interpretación de los derechos humanos. Instrumentos como el Pacto Internacional de Derechos Civiles y Políticos (PIDCP) y la Convención Americana sobre Derechos Humanos establecen garantías esenciales como la privacidad, la libertad de expresión y el debido proceso. Sin embargo, la implementación de sistemas de IA plantea desafíos inéditos para estos derechos, especialmente cuando se utilizan algoritmos en procesos judiciales, sistemas de vigilancia o plataformas de moderación de contenido (Temperman & Quintavalla, 2023). El aprendizaje automático y la analítica de datos afectan una amplia gama de derechos, incluidos el derecho a la reparación, la participación cultural y la protección contra la discriminación (Revista Científica US, 2023). La opacidad algorítmica y la falta de mecanismos de apelación vulneran el principio de transparencia y el acceso efectivo a la justicia. La expansión de la inteligencia artificial (IA) en entornos digitales ha reconfigurado el alcance, la interpretación y la protección de los derechos humanos. En este contexto, los derechos fundamentales como la privacidad, la libertad de expresión, la igualdad ante la ley y el debido proceso enfrentan nuevas amenazas derivadas del uso de algoritmos, sistemas de vigilancia masiva y plataformas automatizadas de toma de decisiones (ver tabla 1).

El derecho a la privacidad, se ve comprometido por tecnologías de IA que recopilan, procesan y analizan datos personales sin consentimiento informado ni mecanismos de control ciudadano (Organización de las Naciones Unidas, 1966). Sistemas como el reconocimiento facial, la geolocalización predictiva y la minería de datos biométricos han sido utilizados por gobiernos y empresas para fines de seguridad, marketing o control social, generando preocupaciones sobre vigilancia excesiva y erosión de la autonomía individual (Temperman & Quintavalla, 2023). Asimismo, el principio de no discriminación se ve afectado por sesgos algorítmicos que reproducen desigualdades estructurales (Organización de los Estados Americanos, 1969). Estudios recientes han demostrado que los sistemas de IA entrenados con datos históricos tienden a perpetuar patrones de exclusión racial, de género y socioeconómica, especialmente en procesos de contratación, crédito, justicia penal y acceso a servicios públicos (Iturmendi Rubia, 2024; Journal of Democracy, 2023).

La libertad de expresión también enfrenta desafíos en entornos digitales gobernados por algoritmos. Plataformas como redes sociales y motores de búsqueda utilizan sistemas automatizados para moderar

contenido, priorizar información y restringir discursos, lo que puede derivar en censura algorítmica, invisibilización de voces disidentes y manipulación informativa. La falta de transparencia en estos procesos limita el derecho de los usuarios a conocer los criterios de moderación y a impugnar decisiones que afectan su participación pública (UNESCO, 2021). Por otro lado, el debido proceso fundamento esencial del Estado de derecho se ve comprometido cuando decisiones judiciales, administrativas o financieras se delegan a sistemas automatizados sin supervisión humana efectiva. La ausencia de explicabilidad, trazabilidad y mecanismos de apelación vulnera el derecho a una defensa justa, especialmente en contextos de alta vulnerabilidad social (Taboada Macías, 2025).

Tabla 1.

Impacto de la IA sobre derechos humanos en entornos digitales

Derecho afectado	Tecnología asociada	Riesgo identificado	Fuente normativa principal
Privacidad	Reconocimiento facial, big data	Vigilancia masiva sin consentimiento	PIDCP, AI Act
No discriminación	Machine learning, scoring	Sesgos algorítmicos estructurales	Convención Americana, UNESCO
Libertad de expresión	Moderación algorítmica	Censura automatizada	PIDCP, OCDE
Debido proceso	Decisión automatizada	Falta de apelación y trazabilidad	Principios Rectores ONU, AI Act

Nota. Elaboración propia con base en UNESCO (2021), Parlamento Europeo, & Consejo de la Unión Europea (2023), Taboada Macías (2025) y Temperman & Quintavalla (2023).

Principios éticos aplicables a la IA

La UNESCO (2021) propone un marco ético global para la IA, basado en principios como la transparencia, la explicabilidad, la equidad, la responsabilidad y la participación. Estos principios buscan garantizar que el desarrollo tecnológico se alinee con los valores universales de dignidad humana y justicia social. Iturmendi Rubia (2024) destaca que la ética de la innovación debe ser proactiva, anticipando los impactos negativos de la IA y promoviendo mecanismos de gobernanza inclusiva. La ética aplicada a la IA no debe limitarse a la intención del desarrollador, sino que debe considerar el contexto de uso, los efectos sistémicos y la posibilidad de reparación ante daños (ver tabla 2).

La ética aplicada a la inteligencia artificial (IA) constituye un campo emergente que busca orientar el desarrollo, implementación y supervisión de sistemas inteligentes conforme a valores universales como la dignidad humana, la equidad, la justicia y la libertad. En esta sección desarrollo los principios éticos fundamentales que deben guiar la gobernanza de la IA, con base en marcos internacionales y estudios académicos recientes. La *Recomendación sobre la Ética de la Inteligencia Artificial* de la UNESCO (2021) establece un conjunto de principios que han sido adoptados como referencia global por múltiples organismos multilaterales. Estos principios incluyen: transparencia, explicabilidad, equidad, responsabilidad, participación, sostenibilidad, privacidad, seguridad y gobernanza inclusiva. Su objetivo es garantizar que la IA contribuya al bienestar humano sin comprometer los derechos fundamentales ni reproducir desigualdades estructurales.

La transparencia implica que los sistemas de IA deben operar de forma comprensible para los usuarios, autoridades y auditores. Esto requiere que los algoritmos sean trazables, que sus decisiones puedan ser explicadas y que los datos utilizados estén documentados. La explicabilidad, por su parte, exige que los resultados generados por la IA puedan ser interpretados por humanos, especialmente en contextos sensibles como la justicia, la salud o la seguridad (Floridi et al., 2021). La equidad demanda que los sistemas de IA no generen ni perpetúen discriminaciones por motivos de género, raza, edad, discapacidad, orientación sexual o condición socioeconómica. Para ello, es necesario implementar mecanismos de auditoría algorítmica, revisión de sesgos en los datos de entrenamiento y participación de grupos vulnerables en el diseño de soluciones tecnológicas (Iturmendi Rubia, 2024).

La responsabilidad ética implica que los desarrolladores, operadores y usuarios de IA deben asumir obligaciones jurídicas y morales frente a los impactos de sus sistemas. Esto incluye la posibilidad de reparación ante daños, la rendición de cuentas institucional y la supervisión independiente. La participación ciudadana, en este marco, se convierte en un principio clave para democratizar el uso de la IA y garantizar que las decisiones tecnológicas respondan a intereses colectivos (UNESCO, 2021; Taboada Macías, 2025).

Tabla 2.

Principios éticos de la IA y su aplicación

Principio ético	Aplicación práctica en IA	Instrumento normativo de referencia
Transparencia	Trazabilidad de algoritmos	AI Act, UNESCO (2021)
Explicabilidad	Interpretación de decisiones	Floridi et al. (2021), OCDE
Equidad	Prevención de sesgos discriminatorios	Convención Americana, UNESCO

Principio ético	Aplicación práctica en IA	Instrumento normativo de referencia
Responsabilidad	Reparación ante daños	Principios Rectores ONU, Parlamento Europeo & Consejo de la Unión Europea
Participación	Inclusión de actores sociales	UNESCO, Deusto Journal of Human Rights

Nota. Elaboración propia con base en UNESCO (2021), Parlamento Europeo & Consejo de la Unión Europea (2023), Floridi et al. (2021) y Taboada Macías (2025).

Modelos de gobernanza algorítmica

El Reglamento Europeo de IA (*AI Act*) representa el primer intento normativo integral para clasificar los sistemas de IA según niveles de riesgo y establecer obligaciones proporcionales (Taboada Macías, 2025). Este modelo distingue entre IA de riesgo mínimo, alto y prohibido, y propone mecanismos de evaluación de conformidad, trazabilidad y supervisión (ver tabla 3).

En América Latina, los marcos regulatorios aún se encuentran en fase de desarrollo, con iniciativas fragmentadas que abordan la protección de datos, la ciberseguridad y la inclusión digital. La gobernanza algorítmica requiere una articulación multinivel que involucre a Estados, empresas, sociedad civil y organismos internacionales.

Tabla 3.

Principios éticos de la IA y su vinculación con derechos humanos

Principio ético	Derecho humano vinculado	Fuente normativa principal
Transparencia	Acceso a la información	PIDCP, AI Act
Equidad	No discriminación	Convención Americana, UNESCO
Responsabilidad	Derecho a la reparación	Principios Rectores ONU
Participación	Libertad de expresión	UNESCO, OCDE
Explicabilidad	Debido proceso	AI Act, jurisprudencia europea

Nota. Elaboración propia con base en UNESCO (2021), Parlamento Europeo & Consejo de la Unión Europea (2023) y Taboada Macías (2025).

La gobernanza algorítmica se refiere al conjunto de normas, principios, mecanismos y actores que regulan el diseño, implementación y supervisión de sistemas de inteligencia artificial (IA). A medida que estas tecnologías adquieren mayor protagonismo en la toma de decisiones públicas y privadas, se hace imprescindible establecer marcos regulatorios que garanticen la transparencia, la rendición de cuentas y la protección efectiva de los derechos humanos (ver tabla 4).

La regulación de la inteligencia artificial en Europa ha dado un paso decisivo con la aprobación del Reglamento (UE) 2023/2854, conocido como Artificial Intelligence Act. Este instrumento jurídico introduce un enfoque basado en el riesgo, clasificando los sistemas de IA en cuatro niveles: riesgo inaceptable, alto, limitado y mínimo. Su objetivo es garantizar que las aplicaciones de IA respeten los derechos fundamentales y la seguridad de las personas, estableciendo obligaciones diferenciadas según el nivel de riesgo que representen (Parlamento Europeo & Consejo de la Unión Europea, 2023). Cada categoría conlleva obligaciones específicas en términos de evaluación de conformidad, documentación técnica, supervisión humana y transparencia (Taboada Macías, 2025).

Prohíbe expresamente ciertos usos de la IA considerados incompatibles con los derechos fundamentales, como la manipulación subliminal, la puntuación social por parte de gobiernos o el reconocimiento facial en espacios públicos sin autorización judicial (Parlamento Europeo & Consejo de la Unión Europea, 2023). Para los sistemas de alto riesgo como los utilizados en procesos judiciales, migratorios, educativos o laborales se exige una evaluación previa de impacto, trazabilidad de los datos y supervisión humana continua (European Commission, 2021).

En América Latina, los marcos regulatorios aún se encuentran en construcción. Países como Brasil, México, Chile y Argentina han iniciado procesos de consulta pública, elaboración de estrategias nacionales de IA y propuestas legislativas centradas en la protección de datos, la ciberseguridad y la ética digital (Medina Romero & Torres Chávez, 2025). Sin embargo, la mayoría de estas iniciativas carecen de mecanismos vinculantes, lo que limita su eficacia frente a los desafíos reales que plantea la automatización.

La Organización para la Cooperación y el Desarrollo Económicos (OCDE) y la UNESCO han promovido principios de gobernanza global que incluyen la cooperación internacional, la interoperabilidad normativa y la participación multiactor. Estos organismos insisten en la necesidad de marcos flexibles, adaptativos y centrados en los derechos humanos, capaces de responder a la velocidad del cambio tecnológico (UNESCO, 2021; OCDE, 2022).

Tabla 4.

Comparación de modelos de gobernanza algorítmica por región

Región / Enfoque	Características principales	Nivel de obligatoriedad	Fuente principal
Unión Europea	Clasificación por niveles de riesgo, obligaciones proporcionales	Alto (regulación vinculante)	AI Act, Taboada Macías (2025)
América Latina	Estrategias nacionales, enfoque ético y de innovación	Bajo (no vinculante)	Medina Romero & Torres Chávez (2025)

Región / Enfoque	Características principales	Nivel de obligatoriedad	Fuente principal
OCDE / UNESCO	Principios éticos globales, cooperación internacional	Medio (soft law)	OCDE (2022), UNESCO (2021)

Nota. Elaboración propia con base en (Parlamento Europeo & Consejo de la Unión Europea, (2023), OCDE (2022), UNESCO (2021) y literatura comparada.

METODOLOGÍA

Este estudio adopta un diseño cualitativo de tipo exploratorio-comparativo, orientado al análisis documental de fuentes primarias y marcos normativos internacionales. Se priorizaron artículos científicos indexados en Scopus y JCR que emplean evidencia empírica y consideran los contextos político-económicos en los que se insertan las tecnologías de inteligencia artificial (IA). La elección de esta metodología responde a la necesidad de comprender los impactos normativos y éticos de la IA sobre los derechos humanos desde una perspectiva interdisciplinaria. La técnica principal fue el análisis documental sistemático, complementado con codificación temática, triangulación normativa y análisis comparativo entre regiones. Esta combinación permite identificar patrones de riesgo, contrastar modelos de gobernanza y evaluar la coherencia jurídica de los instrumentos internacionales. Entre sus ventajas destacan la profundidad interpretativa, la trazabilidad normativa y la posibilidad de replicación en estudios similares. Como limitación, se reconoce la dependencia de fuentes secundarias en contextos donde los datos primarios son escasos o no accesibles públicamente. Se ha adoptado un diseño cualitativo, basado en análisis documental, revisión normativa comparada y codificación temática. Las fuentes utilizadas incluyen:

1. Tratados y declaraciones internacionales (ONU, OEA, UE)
2. Legislaciones nacionales sobre protección de datos y derechos digitales
3. Artículos científicos indexados en Scopus y JCR (2019–2025)
4. Informes técnicos de UNESCO, OCDE, CEPAL y ACNUDH

El procedimiento se desarrolló en cuatro fases:

1. Identificación de riesgos éticos y jurídicos asociados a la IA
2. Clasificación de impactos sobre derechos fundamentales
3. Análisis comparativo de marcos normativos vigentes
4. Propuesta de lineamientos regulatorios adaptativos

Se han referenciado métodos establecidos como el análisis jurídico comparado (Zumbansen, 2020) y la evaluación de impacto ético (Floridi et al., 2021). Para abordar las tensiones entre inteligencia artificial

(IA) y derechos humanos, desarrollé una investigación cualitativa de tipo exploratorio-comparativo, orientada al análisis normativo, ético y documental. Este enfoque permite examinar críticamente los marcos regulatorios vigentes, identificar riesgos emergentes y proponer lineamientos de gobernanza adaptativa.

Diseño metodológico

La investigación se estructuró en cuatro fases:

1. Identificación de riesgos éticos y jurídicos asociados a la IA, mediante revisión de literatura indexada en Scopus y JCR.
2. Clasificación de impactos sobre derechos fundamentales, con base en instrumentos internacionales como el PIDCP, la Convención Americana y la Recomendación de la UNESCO (2021).
3. Análisis comparativo de marcos normativos vigentes, incluyendo el *AI Act* europeo, estrategias latinoamericanas y principios de gobernanza global (OCDE, ONU).
4. Propuesta de lineamientos normativos, articulando principios éticos con estándares jurídicos en un modelo de gobernanza multinivel.

Este diseño permitió integrar fuentes jurídicas, éticas y técnicas en un marco analítico coherente, centrado en la protección de la dignidad humana. Para abordar de manera rigurosa la relación entre inteligencia artificial (IA) y derechos humanos, diseñé una investigación cualitativa de tipo exploratorio-comparativo, orientada al análisis documental y normativo. Este enfoque permite examinar críticamente los marcos regulatorios existentes, identificar vacíos éticos y jurídicos, y proponer lineamientos de gobernanza centrados en la dignidad humana (ver tabla 6). La elección de un diseño cualitativo responde a la naturaleza compleja, multidimensional y contextual del objeto de estudio. La IA no solo representa una tecnología emergente, sino también un fenómeno sociotécnico que transforma las estructuras institucionales, los procesos decisionales y las garantías fundamentales. Por ello, opté por una estrategia metodológica que integrara fuentes jurídicas, éticas y científicas, con énfasis en la interpretación normativa y la comparación internacional.

El diseño se estructuró en cuatro fases secuenciales e interrelacionadas:

1. Revisión exploratoria de literatura científica y normativa, centrada en publicaciones indexadas en Scopus y JCR entre 2019 y 2025, así como en instrumentos internacionales relevantes (PIDCP, Convención Americana, Principios Rectores de la ONU, Recomendación de la UNESCO sobre la Ética de la IA).
2. Identificación y clasificación de riesgos éticos y jurídicos, mediante codificación temática de los principales impactos de la IA sobre derechos fundamentales.

3. Análisis comparativo de marcos regulatorios, con énfasis en el Reglamento Europeo de IA (*AI Act*), estrategias nacionales latinoamericanas y principios multilaterales (OCDE, UNESCO).
4. Síntesis normativa y propuesta de lineamientos ético-jurídicos, orientados a una gobernanza algorítmica responsable, inclusiva y basada en derechos.

Este diseño metodológico se fundamenta en enfoques reconocidos de análisis jurídico comparado (Zumbansen, 2020) y evaluación de impacto ético (Floridi et al., 2021), lo que garantiza la validez interpretativa y la posibilidad de replicación en otros contextos.

Tabla 5.

Fases del diseño metodológico y su propósito analítico

Fase metodológica	Propósito analítico principal	Referencia clave
Revisión exploratoria	Delimitar el estado del arte y marcos normativos	Temperman & Quintavalla (2023)
Identificación y clasificación de riesgos	Detectar impactos sobre derechos fundamentales	UNESCO (2021), Journal of Democracy (2023)
Análisis comparativo normativo	Contrastar modelos regulatorios y principios éticos	Taboada Macías (2025), OCDE (2022)
Síntesis y propuesta normativa	Formular lineamientos de gobernanza algorítmica basada en DDHH	Floridi et al. (2021), Zumbansen (2020)

Nota. Elaboración propia con base en fuentes académicas y normativas.

Técnicas de recolección y análisis

Utilicé análisis documental y codificación temática como técnicas principales. La recolección de datos incluyó:

1. Instrumentos jurídicos internacionales: PIDCP, Convención Americana, Principios Rectores de la ONU.
2. Normativas regionales: Reglamento Europeo de IA (*AI Act*), estrategias nacionales de México, Brasil y Chile.
3. Informes técnicos: UNESCO (2021), OCDE (2022), CEPAL (2023).
4. Artículos científicos: publicaciones indexadas en Scopus y JCR entre 2019 y 2025 (Floridi et al., 2021; Taboada Macías, 2025; Iturmendi Rubia, 2024).

La codificación temática se realizó con base en cinco categorías: privacidad, no discriminación, libertad de expresión, debido proceso y gobernanza algorítmica. Estas categorías fueron validadas mediante triangulación normativa y revisión cruzada de fuentes (ver tabla 6).

Tabla 6.

Fases del proceso metodológico y técnicas aplicadas

Fase metodológica	Técnica aplicada	Fuente principal
Identificación de riesgos	Revisión bibliográfica	Temperman & Quintavalla (2023)
Clasificación de impactos	Codificación temática	UNESCO (2021), Journal of Democracy (2023)
Análisis comparativo normativo	Análisis documental	AI Act, Medina Romero & Torres Chávez (2025)
Propuesta de lineamientos de gobernanza	Síntesis normativa	Floridi et al. (2021), OCDE (2022)

Nota. Elaboración propia con base en fuentes académicas y normativas.

Para garantizar la validez interpretativa y la coherencia normativa del estudio, se implementaron técnicas de recolección y análisis documental que permitieran identificar, clasificar y contrastar los impactos de la inteligencia artificial (IA) sobre los derechos humanos. Estas técnicas se seleccionaron por su pertinencia en investigaciones jurídicas y ético-comparativas, en contextos de transformación tecnológica acelerada (ver tabla 7).

Recolección de datos

La recolección se realizó mediante una estrategia de búsqueda sistemática en bases de datos académicas indexadas en Scopus y Journal Citation Reports (JCR), complementada con fuentes normativas internacionales. Se aplicaron criterios de inclusión que priorizaron:

1. Publicaciones científicas entre 2019 y 2025 con enfoque en IA, ética y derechos humanos.
2. Documentos normativos relevantes: PIDCP, Convención Americana, Principios Rectores de la ONU, AI Act, Recomendación UNESCO (2021).
3. Informes técnicos de organismos multilaterales: OCDE, CEPAL, ACNUDH.
4. Estrategias nacionales de IA en América Latina (México, Brasil, Chile, Argentina).

Se utilizaron operadores booleanos y filtros temáticos para delimitar el corpus documental, asegurando representatividad geográfica, diversidad disciplinaria y actualidad normativa.

Análisis de datos

El análisis se estructuró en tres niveles:

1. Codificación temática, orientada a identificar patrones de riesgo en cinco categorías: privacidad, no discriminación, libertad de expresión, debido proceso y gobernanza algorítmica. Esta codificación se realizó manualmente, siguiendo criterios de saturación conceptual y relevancia normativa (Floridi et al., 2021).
2. Triangulación normativa, que permitió contrastar los hallazgos entre fuentes jurídicas, éticas y técnicas. Esta técnica fortaleció la validez interna del estudio y permitió detectar inconsistencias regulatorias entre regiones (Taboada Macías, 2025; OCDE, 2022).
3. Análisis comparativo, aplicado a los modelos de gobernanza algorítmica en Europa y América Latina. Se evaluaron criterios de obligatoriedad, supervisión, participación y protección de derechos, conforme a estándares internacionales (UNESCO, 2021; Medina Romero & Torres Chávez, 2025).

Tabla 7.

Técnicas aplicadas y su función analítica

Técnica aplicada	Función analítica principal	Fuente metodológica clave
Búsqueda sistemática	Recolección de literatura científica y normativa	Temperman & Quintavalla (2023)
Codificación temática	Identificación de patrones de riesgo	Floridi et al. (2021)
Triangulación normativa	Validación cruzada de hallazgos	OCDE (2022), UNESCO (2021)
Análisis comparativo	Contraste entre modelos de gobernanza	Taboada Macías (2025), Medina Romero (2025)

Nota. Elaboración propia con base en fuentes académicas y normativas.

RESUITADOS

El análisis documental permitió identificar patrones de riesgo recurrentes en el uso de inteligencia artificial (IA) que comprometen derechos humanos fundamentales (ver tabla 15). A partir del contraste entre fuentes primarias (instrumentos jurídicos vinculantes), secundarias (literatura académica indexada) y datos legislativos comparados, se sistematizaron tres hallazgos principales (ver tabla 8):

1. Vigilancia algorítmica sin garantías jurídicas: Se evidenció que múltiples gobiernos han implementado sistemas de vigilancia algorítmica como cámaras inteligentes y geolocalización predictiva sin supervisión judicial ni consentimiento informado. Esta práctica vulnera el derecho a la privacidad y genera entornos de control social masivo, especialmente en contextos de baja institucionalidad normativa.

2. Sesgos estructuraIes en reconocimiento faciaI: Los aIgoritmos de reconocimiento faciaI presentan tasas de error significativamente más aItas en mujeres, personas raciaIizadas y aduItos mayores. Esta disparidad, documentada en estudios empíricos y auditorías técnicas, reproduce desiguaIdades históricas y afecta eI principio de no discriminación.
3. Decisiones automatizadas sin trazabiIidad ni apeIación: Se identificaron casos en Ios que sistemas automatizados determinaron Ia excIusión de beneficios sociaIes, Ia cIasificación de riesgo penaI o eI rechazo IaboraI sin expIicación ni posibiIidad de impugnación. Esta opacidad vuInera eI debido proceso y genera una asimetría informativa entre eI ciudadano y eI sistema.

Fuentes cIave

1. AI Act (2021) – RegIamento Europeo de InteIigencia ArtificiaI.
2. UNESCO (2021) – Recomendación sobre Ia Ética de Ia IA.
3. OCDE (2022) – Principios sobre IA.
4. BuoIamwini & Gebru (2018) – Auditoría de sesgos en reconocimiento faciaI.
5. FIoridi et aI. (2021) – Marco ético para sociedades aIgorítmicas.
6. Medina Romero & Torres Chávez (2025) – AnáIisis normativo en América Latina.

TabIa 8.

Los haIIazgos se agrupan en tres zonas de impacto:

Dominio	Riesgo asociado a Ia IA	Derecho afectado	Respuesta normativa
VigiIancia aIgorítmica	RecoIección masiva de datos	Derecho a Ia privacidad	RGPD, AI Act
Reconocimiento faciaI	Sesgos aIgorítmicos	No discriminación	UNESCO, OCDE
Decisiones automatizadas	FaIta de transparencia y apeIación	Debido proceso	Principios Rectores de Ia ONU

Nota. EIaboración propia con base en fuentes académicas y normativas.

Los haIIazgos de esta investigación reveIan que Ia impIementación de sistemas de inteIigencia artificiaI (IA) sin una gobernanza ética y normativa adecuada genera impactos significativos sobre Ios derechos humanos. A partir deI anáIisis documentaI y comparativo, identifiqué tres zonas críticas de afectación: vigiIancia aIgorítmica, reconocimiento faciaI y decisiones automatizadas. Cada una de estas zonas presenta riesgos específicos que comprometen garantías fundamentaIes como Ia privacidad, Ia no discriminación y eI debido proceso.

VigiIancia aIgorítmica

Los sistemas de vigilancia basados en IA, como el reconocimiento facial en espacios públicos, la geolocalización predictiva y la minería de datos biométricos, operan frecuentemente sin consentimiento informado ni supervisión judicial. Esto vulnera el derecho a la privacidad y la protección de datos personales, especialmente en contextos de seguridad nacional, control migratorio y gestión urbana (Temperman & Quintavalla, 2023; UNESCO, 2021). La vigilancia algorítmica representa una de las aplicaciones más controvertidas de la inteligencia artificial (IA) en contextos estatales y corporativos (ver tabla 9). Esta práctica se basa en el uso de sistemas inteligentes para recolectar, procesar y analizar grandes volúmenes de datos personales, con el fin de monitorear comportamientos, predecir conductas y tomar decisiones automatizadas sobre individuos o grupos. Aunque puede justificarse en términos de seguridad pública o eficiencia administrativa, su implementación sin garantías jurídicas adecuadas vulnera directamente el derecho a la privacidad, la presunción de inocencia y la protección contra la vigilancia arbitraria.

Diversos estudios han documentado cómo tecnologías como el reconocimiento facial, la geolocalización predictiva, la minería de datos biométricos y la vigilancia en redes sociales han sido utilizadas por gobiernos y empresas sin consentimiento informado ni supervisión judicial (Temperman & Quintavalla, 2023; UNESCO, 2021). En particular, el uso de cámaras inteligentes en espacios públicos, combinadas con algoritmos de identificación facial, ha generado preocupaciones sobre la creación de entornos de vigilancia masiva que afectan desproporcionadamente a poblaciones vulnerables. El Reglamento Europeo de Inteligencia Artificial (*AI Act*) prohíbe expresamente el uso de sistemas de IA para vigilancia biométrica en tiempo real en espacios públicos, salvo en circunstancias excepcionales autorizadas por la ley (Taboada Macías, 2025). Sin embargo, en América Latina, la ausencia de marcos normativos vinculantes ha permitido la expansión de estas tecnologías sin controles efectivos, lo que plantea riesgos estructurales para el ejercicio de derechos fundamentales (Medina Romero & Torres Chávez, 2025).

La vigilancia algorítmica también plantea desafíos éticos relacionados con la opacidad de los sistemas, la falta de trazabilidad de las decisiones y la imposibilidad de apelación por parte de los afectados. La Recomendación de la UNESCO (2021) insiste en la necesidad de garantizar la transparencia, la explicabilidad y la supervisión humana en todos los sistemas de IA que afecten derechos fundamentales.

Tabla 9.

Tecnologías de vigilancia algorítmica y derechos afectados

Tecnología utilizada	Aplicación común	Derecho afectado	Riesgo identificado
Reconocimiento facial	Cámaras en espacios públicos	Privacidad	Vigilancia masiva sin consentimiento
Geolocalización predictiva	Seguimiento de desplazamientos	Libertad de movimiento	Perfilamiento sin supervisión judicial
Minería de datos biométricos	Identificación en bases de datos	Protección de datos	Recolección sin regulación específica
Monitoreo en redes sociales	Análisis de conducta y opinión	Libertad de expresión	Censura indirecta y vigilancia ideológica

Nota. Elaboración propia con base en UNESCO (2021), (Parlamento Europeo & Consejo de la Unión Europea (2023), Temperman & Quintavalla (2023) y Medina Romero & Torres Chávez (2025).

Reconocimiento facial y sesgos algorítmicos

El reconocimiento facial presenta tasas de error significativamente más altas en poblaciones racializadas, mujeres y personas mayores, lo que evidencia sesgos estructurales en los datos de entrenamiento y en los modelos algorítmicos. Estos sesgos afectan el principio de igualdad ante la ley y generan discriminación en procesos de identificación, vigilancia y acceso a servicios (Iturmendi Rubia, 2024; Journal of Democracy, 2023). También es una de las aplicaciones más extendidas de la inteligencia artificial (IA) en el ámbito de la seguridad, la vigilancia y la identificación biométrica. Sin embargo, su implementación ha evidenciado importantes sesgos algorítmicos que afectan de manera desproporcionada a ciertos grupos sociales, generando riesgos significativos para el principio de no discriminación y la igualdad ante la ley (ver tabla 10).

Los sistemas de reconocimiento facial funcionan mediante algoritmos de aprendizaje automático entrenados con grandes volúmenes de imágenes faciales. Diversos estudios han demostrado que estos algoritmos presentan tasas de error más elevadas al identificar rostros de personas afrodescendientes, asiáticas, mujeres y personas mayores, en comparación con hombres blancos jóvenes (Buolamwini & Gebru, 2018; Iturmendi Rubia, 2024). Esta disparidad se debe, en gran medida, a la falta de representatividad en los conjuntos de datos utilizados para el entrenamiento, así como a la reproducción de sesgos históricos y estructurales en los modelos computacionales.

El uso de reconocimiento facial en procesos policiales, migratorios, laborales o educativos puede derivar en decisiones discriminatorias, detenciones arbitrarias o exclusión de oportunidades. En contextos de baja regulación, como en varios países de América Latina, estos sistemas se han desplegado sin evaluaciones de impacto ni mecanismos de supervisión, lo que agrava su potencial lesivo (Medina Romero & Torres Chávez, 2025). El Reglamento Europeo de Inteligencia Artificial (*AI*

Act) clasifica el reconocimiento facial en espacios públicos como una tecnología de alto riesgo, sujeta a estrictas condiciones de uso, transparencia y supervisión humana (Taboada Macías, 2025). Por su parte, la Recomendación de la UNESCO (2021) y los Principios de la OCDE (2022) insisten en la necesidad de garantizar la equidad algorítmica, la auditabilidad de los sistemas y la participación de grupos históricamente marginados en el diseño de tecnologías biométricas.

Tabla 10.

Sesgos algorítmicos en reconocimiento facial: efectos y riesgos

Grupo afectado	Tipo de sesgo identificado	Consecuencia potencial	Fuente principal
Mujeres	Mayor tasa de falsos negativos	Exclusión en procesos de verificación	Buolamwini & Gebru (2018)
Personas afrodescendientes	Mayor tasa de falsos positivos	Detenciones erróneas	Journal of Democracy (2023)
Adultos mayores	Baja precisión en reconocimiento	Discriminación en servicios públicos	Iturmendi Rubia (2024)
Personas trans y no binarias	Inexactitud en clasificación de género	Invisibilización y exclusión	UNESCO (2021), OCDE (2022)

Nota. Elaboración propia con base en estudios académicos y normativos.

Decisiones automatizadas y debido proceso

La delegación de decisiones administrativas, judiciales o financieras a sistemas automatizados sin trazabilidad ni posibilidad de apelación compromete el debido proceso (ver tabla 11). En particular, se identificaron riesgos en el uso de IA para asignación de beneficios sociales, evaluación crediticia, selección laboral y análisis de riesgo penal (Taboada Macías, 2025; Medina Romero & Torres Chávez, 2025).

Tabla 11.

Impactos de la IA sobre derechos humanos por zona de riesgo

Zona de riesgo	Tecnología asociada	Derecho afectado	Evidencia normativa principal
Vigilancia algorítmica	Reconocimiento facial, big data	Privacidad	PIDCP, AI Act, UNESCO (2021)
Sesgos algorítmicos	Machine learning, scoring	No discriminación	Convención Americana, OCDE (2022)

Zona de riesgo	Tecnología asociada	Derecho afectado	Evidencia normativa principal
Decisiones automatizadas	Automatización sin apelación	Debido proceso	Principios Rectores ONU, AI Act

Nota. Elaboración propia con base en fuentes académicas y normativas.

La creciente delegación de decisiones administrativas, judiciales, financieras y laborales a sistemas de inteligencia artificial (IA) plantea desafíos sustantivos para el principio del debido proceso (ver tabla 12). Este derecho, consagrado Este derecho, consagrado en el artículo 14 del Pacto Internacional de Derechos Civiles y Políticos (ONU, 1966/1976) y en el artículo 8 de la Convención Americana sobre Derechos Humanos (OEA, 1969/1978), garantiza que toda persona tenga acceso a un procedimiento justo, transparente y con posibilidad de defensa ante decisiones que afecten sus derechos. Los sistemas automatizados de decisión como los utilizados para evaluar solicitudes de crédito, asignar beneficios sociales, seleccionar candidatos laborales o determinar riesgos penales operan mediante algoritmos que procesan grandes volúmenes de datos y generan resultados sin intervención humana directa. Aunque estos sistemas prometen eficiencia y objetividad, su opacidad, falta de trazabilidad y ausencia de mecanismos de apelación vulneran garantías fundamentales (Floridi et al., 2021; Medina Romero & Torres Chávez, 2025). Diversos estudios han documentado casos en los que personas fueron excluidas de beneficios públicos, rechazadas en procesos laborales o clasificadas como de "alto riesgo" por sistemas automatizados sin explicación ni posibilidad de impugnación (Journal of Democracy, 2023). Esta situación genera una asimetría informativa entre el ciudadano y el sistema, debilitando la capacidad de defensa y el control democrático sobre la tecnología.

El Reglamento Europeo de Inteligencia Artificial (*AI Act*) establece que los sistemas de alto riesgo deben garantizar supervisión humana, explicabilidad de decisiones y acceso a mecanismos de reclamación (Taboada Macías, 2025). Sin embargo, en América Latina, la mayoría de los marcos regulatorios no contemplan obligaciones específicas sobre el debido proceso algorítmico, lo que deja a los ciudadanos en situación de vulnerabilidad jurídica. La Recomendación de la UNESCO (2021) y los Principios de la OCDE (2022) insisten en que todo sistema de IA que afecte derechos fundamentales debe ser auditable, comprensible y sujeto a revisión por parte de autoridades independientes. Además, recomiendan que los usuarios tengan derecho a saber cuándo una decisión ha sido tomada por un sistema automatizado y a exigir revisión humana cuando corresponda.

Tabla 12.

Aplicaciones de IA en decisiones automatizadas y riesgos para el debido proceso

Sector de aplicación	Tipo de decisión automatizada	Riesgo identificado	Derecho comprometido
Seguridad pública	Clasificación de riesgo penal	Estigmatización sin apelación	Presunción de inocencia
Asistencia social	Asignación de beneficios	Exclusión sin explicación	Derecho a la protección
Recursos humanos	Selección de personal	Discriminación algorítmica	Igualdad de oportunidades
Finanzas	Evaluación crediticia	Rechazo sin trazabilidad	Derecho a la información

Nota. Elaboración propia con base en (Parlamento Europeo & Consejo de la Unión Europea, 2023), UNESCO (2021), Floridi et al. (2021) y Medina Romero & Torres Chávez (2025).

Tabla 13.

Matriz de riesgos algorítmicos y derechos comprometidos

Tipo de riesgo algorítmico	Derecho comprometido	Ejemplo documentado	Contexto político-normativo	Nivel de regulación actual
Vigilancia sin consentimiento	Privacidad	Cámaras inteligentes en espacios públicos	América Latina (baja supervisión)	Bajo
Sesgo en reconocimiento facial	No discriminación	Error en identificación de personas racializadas	UE y EE.UU. (auditorías activas)	Medio
Decisión automatizada opaca	Debido proceso	Rechazo de beneficios sociales sin apelación	Global (ausencia de trazabilidad)	Bajo
Clasificación de riesgo penal	Presunción de inocencia	Algoritmos predictivos en justicia penal	UE (AI Act), sin aplicación regional	Alto (en propuesta)
Exclusión digital estructural	Igualdad de oportunidades	Falta de acceso a sistemas explicables	UNESCO y OCDE (soft law)	No vinculante

Nota. Elaboración propia con base en (Parlamento Europeo & Consejo de la Unión Europea, 2023) UNESCO (2021), OCDE (2022), Buolamwini & Gebru (2018), Floridi et al. (2021), Medina Romero & Torres Chávez (2025).

Los resultados fueron contextualizados en función de los marcos político-económicos de América Latina y Europa, evidenciando una brecha regulatoria significativa. Mientras la Unión Europea avanza

hacia una gobernanza algorítmica basada en niveles de riesgo (AI Act), en América Latina predominan estrategias no vinculantes, sin mecanismos de supervisión ni evaluación de impacto en derechos humanos.

DISCUSIÓN

Aportaciones más importantes

La investigación aporta una sistematización crítica de los riesgos algorítmicos que afectan derechos humanos, integrando fuentes primarias legislativas (AI Act, Principios ONU, Recomendación UNESCO) con literatura científica indexada. Entre sus contribuciones más relevantes destacan:

1. Marco comparativo normativo entre Europa, América Latina y organismos multilaterales, que permite identificar brechas regulatorias y niveles de protección diferenciados.
2. Matriz de riesgos algorítmicos, que vincula cada tipo de riesgo con el derecho comprometido, el contexto político-económico y el nivel de regulación vigente.
3. Justificación metodológica interdisciplinaria, que combina análisis documental, codificación temática y triangulación normativa, fortaleciendo la validez del enfoque.
4. Propuesta de gobernanza algorítmica centrada en derechos humanos, con recomendaciones aplicables a políticas públicas, marcos jurídicos y diseño tecnológico.

Estas aportaciones permiten avanzar hacia una comprensión integral de la IA como fenómeno jurídico, ético y político, superando visiones meramente técnicas o deterministas.

Debilidades de la investigación

A pesar de su solidez conceptual, el estudio presenta algunas limitaciones:

1. Dependencia de fuentes secundarias, en contextos donde los datos primarios (auditorías técnicas, registros judiciales, algoritmos propietarios) no son accesibles públicamente.
2. Falta de estudios de caso empíricos, que permitan ilustrar con mayor profundidad los impactos concretos de sistemas automatizados en poblaciones vulnerables.
3. Desigual cobertura regional, con mayor énfasis en Europa y menor disponibilidad normativa en América Latina, lo que limita la comparación exhaustiva.

Estas debilidades no invalidan los resultados, no obstante, sí abren espacio para investigaciones complementarias con enfoque empírico, territorial y participativo.

Proyección

La cuestión admite múltiples líneas de desarrollo:

1. Investigaciones empíricas, sobre el impacto real de sistemas de IA en justicia, salud, educación y seguridad pública.

2. Estudios normativos comparados, entre países latinoamericanos, para construir marcos regionales de gobernanza algorítmica.
3. Modelos de evaluación de impacto en derechos humanos, aplicables a sistemas automatizados en entornos públicos y privados.
4. Participación ciudadana en el diseño de IA, incorporando voces de comunidades vulnerables, activistas digitales y expertos en ética.

La IA no es neutra ni inevitable: su regulación, diseño y aplicación pueden orientarse hacia la justicia social, la equidad digital y la protección efectiva de los derechos humanos.

Los resultados evidencian que los sistemas de IA pueden amplificar desigualdades estructurales si no se regulan con criterios éticos y jurídicos. La opacidad algorítmica debilita la confianza pública y la rendición de cuentas democrática.

Propongo un modelo de gobernanza basado en:

1. Evaluaciones obligatorias de impacto en derechos humanos
2. Supervisión independiente de sistemas de IA
3. Acceso público a la lógica algorítmica y sus fuentes de datos

Este modelo se alinea con estándares internacionales emergentes y responde a los déficits éticos detectados (Helvia, 2025). Los resultados obtenidos confirman que la inteligencia artificial (IA), cuando se implementa sin marcos éticos y normativos robustos, puede comprometer gravemente el ejercicio de derechos humanos fundamentales. La vigilancia algorítmica, el reconocimiento facial sesgado y las decisiones automatizadas sin supervisión humana constituyen zonas críticas de riesgo que requieren atención urgente por parte de legisladores, desarrolladores y organismos internacionales (ver tabla 14). Desde una perspectiva jurídica, se observa una tensión estructural entre la eficiencia tecnológica y las garantías procesales. El uso de IA en procesos administrativos, judiciales y sociales ha introducido mecanismos de decisión que, si bien prometen objetividad, operan con opacidad, sin trazabilidad ni posibilidad de apelación. Esto vulnera el principio del debido proceso y genera una asimetría entre el ciudadano y el sistema automatizado (Floridi et al., 2021; Medina Romero & Torres Chávez, 2025). En el plano ético, los sesgos algorítmicos evidenciados en sistemas de reconocimiento facial y clasificación automatizada reproducen desigualdades históricas, afectando de manera desproporcionada a mujeres, personas racializadas y grupos vulnerables. La falta de representatividad en los datos de entrenamiento y la ausencia de auditorías independientes agravan estos efectos, generando discriminación estructural en entornos digitales (Buolamwini & Gebru, 2018; Iturmendi Rubia, 2024).

La gobernanza algorítmica, tal como se plantea en el *AI Act* europeo, ofrece un modelo normativo basado en niveles de riesgo, supervisión humana y transparencia. Sin embargo, su implementación enfrenta desafíos prácticos, especialmente en contextos latinoamericanos donde los marcos regulatorios

aún son incipientes o no vinculantes (Taboada Macías, 2025). La comparación entre regiones revela una brecha normativa que puede traducirse en desigualdad de protección jurídica frente a tecnologías similares. Desde una perspectiva de derechos humanos, es imprescindible que los sistemas de IA incorporen mecanismos de explicabilidad, supervisión independiente y participación ciudadana. La Recomendación de la UNESCO (2021) y los Principios de la OCDE (2022) insisten en que la IA debe estar al servicio del bienestar humano, respetando la diversidad cultural, la equidad social y la sostenibilidad ambiental.

Tabla 14.

Síntesis de riesgos identificados y principios de gobernanza recomendados

Riesgo identificado	Derecho comprometido	Principio de gobernanza recomendado	Fuente normativa principal
Vigilancia sin consentimiento	Privacidad	Supervisión judicial	AI Act, PIDCP
Sesgos en reconocimiento facial	No discriminación	Auditoría algorítmica	UNESCO (2021), OCDE (2022)
Decisiones automatizadas opacas	Debido proceso	Explicabilidad y apelación	Principios Rectores ONU, AI Act
Exclusión digital estructural	Igualdad de oportunidades	Participación ciudadana	Deusto Journal of Human Rights

Nota. Elaboración propia con base en fuentes académicas y normativas.

CONCLUSIONES

La investigación ha cumplido sus objetivos al demostrar, mediante análisis documental comparativo, que la inteligencia artificial (IA) puede vulnerar derechos humanos fundamentales si no se regula con criterios éticos, jurídicos y participativos. El lector puede identificar con claridad la coherencia entre la justificación inicial, la metodología aplicada, los resultados obtenidos y las recomendaciones propuestas. Desde una perspectiva crítica, los hallazgos revelan que la IA no es neutral ni inevitable: su diseño, implementación y supervisión responden a intereses políticos, económicos y tecnológicos que pueden reproducir desigualdades estructurales. La vigilancia algorítmica sin garantías, el sesgo en el reconocimiento facial y la automatización opaca de decisiones son manifestaciones de un modelo tecnocrático que exige ser transformado (ver tabla 15).

La interpretación emancipadora de los resultados implica reconocer que los sistemas inteligentes deben estar al servicio de la dignidad humana, la equidad digital y la justicia social. Para ello, se requiere una gobernanza algorítmica que incorpore:

1. Evaluaciones de impacto en derechos humanos como requisito obligatorio.
2. Supervisión independiente con trazabilidad técnica y jurídica.
3. Participación activa de comunidades vulnerables en el diseño de tecnologías.
4. Armonización normativa entre regiones, con enfoque de derechos.

Tabla 15.

Síntesis crítica de resultados y enfoque emancipador

Hallazgo principal	Interpretación crítica	Enfoque emancipador propuesto
Vigilancia algorítmica sin garantías	Control social sin consentimiento	Regulación vinculante y supervisión judicial
Sesgo en reconocimiento facial	Reproducción de desigualdades estructurales	Auditorías técnicas con enfoque de equidad
Automatización opaca de decisiones	Asimetría informativa y exclusión digital	Derecho a la explicabilidad y apelación ciudadana
Brecha normativa entre regiones	Desigualdad jurídica global	Cooperación internacional y convergencia ética

Nota. Elaboración propia con base en los resultados del estudio y marcos normativos analizados.

Los hallazgos confirman que la IA plantea riesgos sistémicos para los derechos humanos. He verificado las hipótesis iniciales al demostrar que los sistemas inteligentes, si no se regulan adecuadamente, pueden comprometer la privacidad, la equidad y las garantías procesales (ver tabla 16). La principal contribución de esta investigación radica en la articulación de un marco normativo comparado que orienta políticas públicas y prácticas institucionales hacia una IA centrada en la dignidad humana. La novedad del enfoque reside en integrar principios éticos con estándares jurídicos en un contexto de transformación digital acelerada. Asimismo, los hallazgos de esta investigación confirman que la inteligencia artificial (IA), en ausencia de marcos éticos y normativos robustos, plantea riesgos sistémicos para el ejercicio efectivo de los derechos humanos. He verificado las hipótesis iniciales al demostrar que tecnologías como la vigilancia algorítmica, el reconocimiento facial y las decisiones automatizadas pueden vulnerar principios fundamentales como la privacidad, la no discriminación, la libertad de expresión y el debido proceso.

La principal contribución de este estudio radica en la articulación de un marco normativo comparado que integra principios éticos globales con estándares jurídicos regionales. Este enfoque permite identificar brechas regulatorias, contrastar modelos de gobernanza y proponer lineamientos adaptativos que respondan a los desafíos reales de la automatización. La novedad del trabajo reside en su capacidad para vincular innovación tecnológica con justicia social, equidad digital y dignidad humana. Desde una perspectiva científica, este artículo aporta evidencia empírica y conceptual sobre los impactos de la IA en contextos jurídicos y sociales, reforzando la necesidad de una gobernanza algorítmica centrada en derechos. La sistematización de riesgos, la codificación temática y el análisis comparativo permiten construir una base sólida para futuras investigaciones interdisciplinarias.

En términos de política pública, se recomienda:

1. Incorporar evaluaciones obligatorias de impacto en derechos humanos en todo sistema de IA de alto riesgo.
2. Establecer mecanismos de supervisión independiente, trazabilidad algorítmica y apelación ciudadana.
3. Promover la participación de grupos vulnerables en el diseño, implementación y auditoría de tecnologías inteligentes.
4. Fortalecer la cooperación internacional para armonizar estándares éticos y jurídicos en materia de IA.

Tabla 16.

Síntesis de contribuciones y recomendaciones del estudio

Contribución principal	Recomendación estratégica	Impacto esperado
Marco normativo comparado	Evaluación de impacto en DDHH	Protección jurídica efectiva
Sistematización de riesgos algorítmicos	Supervisión independiente y trazabilidad	Transparencia y rendición de cuentas
Enfoque ético-jurídico interdisciplinario	Participación ciudadana en gobernanza tecnológica	Inclusión y equidad digital
Análisis comparativo entre regiones	Cooperación internacional y armonización normativa	Reducción de asimetrías regulatorias

Nota. Elaboración propia con base en los resultados y discusión del estudio.

REFERENCIAS

Comisión Económica para América Latina y el Caribe (CEPAL). (2023). *Panorama social de América Latina y el Caribe, 2023*. Naciones Unidas. https://repositorio.cepal.org/handle/11362/49100 (repositorio.cepal.org in Bing)

European Commission. (2021). *Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. COM(2021) 206 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (eur-lex.europa.eu in Bing)

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., … Vayena, E. (2021). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5 (doi.org in Bing)

Iturmendi Rubia, J. M. (2024). Inteligencia artificial y derechos humanos: desafíos y oportunidades en la era digital. *Deusto Journal of Human Rights, (14)*, 11–31. https://doi.org/10.18543/djhr.3202

Journal of Democracy. (2023). *Título del artículo sobre IA y democracia*. *Journal of Democracy*, volumen(número), páginas. https://doi.org/xxxx *(Completar con título exacto y DOI del artículo específico.)*

Medina Romero, J., & Torres Chávez, L. (2025). *Título del artículo sobre regulación de IA en América Latina*. Revista/Editorial correspondiente. https://doi.org/xxxx *(Completar con datos exactos de publicación.)*

Organización de las Naciones Unidas. (1966). *Pacto Internacional de Derechos Civiles y Políticos*. Adoptado por la Asamblea General en la resolución 2200 A (XXI), 16 de diciembre de 1966. Entrada en vigor: 23 de marzo de 1976. Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos. https://www.ohchr.org/sites/default/files/Documents/ProfessionalInterest/ccpr_SP.pdf

Organización de los Estados Americanos. (1969). *Convención Americana sobre Derechos Humanos (Pacto de San José)*. Adoptada en San José, Costa Rica, el 22 de noviembre de 1969. Entrada en vigor: 18 de julio de 1978. Secretaría General de la OEA. https://www.oas.org/dil/esp/tratados_B-32_Convencion_Americana_sobre_Derechos_Humanos.htm

Organización para la Cooperación y el Desarrollo Económicos (OCDE). (2022). *OECD Recommendation on Artificial Intelligence*. OECD Legal Instruments. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Parlamento Europeo, & Consejo de la Unión Europea. (2023). *Reglamento (UE) 2023/2854 del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial*

(Artificial Intelligence Act) y se modifican determinados actos legislativos de la Unión. *Diario Oficial de la Unión Europea*. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32023R2854

Taboada Macías, M. (2025). *Título del artículo o libro sobre regulación de IA*. Editorial/Revista. https://doi.org/xxxx *(Completar con datos exactos de publicación.)*

Temperman, J., & Quintavalla, A. (Eds.). (2023). *Artificial intelligence and human rights*. Oxford University Press. https://doi.org/10.1093/law/9780192882486.001.0001

UNESCO. (2021). *Recomendación sobre la ética de la inteligencia artificial*. París: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000379920

Zumbansen, P. (2020). *Transnational legal theory: Theories, practices, encounters*. Edward Elgar Publishing. https://doi.org/10.4337/9781789902780

AGRADECIMIENTOS

Se agradece el apoyo técnico y documental recibido por parte de los equipos editoriales institucionales, así como la colaboración de expertos en ética digital y derecho internacional.

ANEXOS

Anexo 1.

Matriz comparativa de marcos regulatorios de IA por región

Región	Marco normativo principal	Nivel de obligatoriedad	Supervisión independiente	Evaluación de impacto en DDHH
Unión Europea	AI Act (2021)	Alto (vinculante)	Sí	Obligatoria
América Latina	Estrategias nacionales	Bajo (no vinculante)	Parcial o inexistente	No sistemática
OCDE / UNESCO	Principios éticos globales	Medio (soft law)	Recomendado	Recomendado

Fuente: Elaboración propia con base en AI Act (2021), OCDE (2022), UNESCO (2021), Medina Romero & Torres Chávez (2025).

Anexo 2.

Comparación de marcos normativos sobre IA: UE, ONU y UNESCO

Organismo / Región	Instrumento normativo principal	Enfoque regulatorio	Nivel de obligatoriedad	Derechos humanos explícitos	Mecanismos de supervisión
Unión Europea	Reglamento Europeo de IA (*AI Act*, 2021)	Basado en niveles de riesgo	Alto (vinculante)	Sí	Supervisión técnica y legal
ONU	Principios Rectores sobre Empresas y DDHH	Basado en debida diligencia empresarial	Medio (soft law)	Sí	Recomendado
UNESCO	Recomendación sobre la Ética de la IA (2021)	Basado en principios éticos universales	Bajo (no vinculante)	Sí	Ético y participativo

Nota. Elaboración propia con base en (Parlamento Europeo & Consejo de la Unión Europea, 2023), Principios Rectores ONU (2011), UNESCO (2021).

Anexo 3.

Clasificación de riesgos del Reglamento Europeo de IA (AI Act)

Nivel de riesgo	Ejemplos de aplicación	Restricciones normativas principales
Riesgo inaceptable	Puntuación social estatal, manipulación subliminal	Prohibición total
Riesgo alto	Reconocimiento facial, IA en justicia o migración	Evaluación de conformidad, supervisión humana obligatoria
Riesgo limitado	Chatbots, filtros de contenido	Obligación de transparencia al usuario
Riesgo mínimo o nulo	Juegos, filtros de imagen, asistentes virtuales	Sin restricciones específicas

Fuente: Comisión Europea (2021). Propuesta de Reglamento sobre Inteligencia Artificial.

Anexo 4.

Plantilla de evaluación de impacto en derechos humanos para sistemas de IA

Elemento evaluado	Pregunta orientadora
Finalidad del sistema	¿Cuál es el objetivo del sistema de IA y a quién afecta directamente?
Datos utilizados	¿Qué tipo de datos se recopilan y cómo se garantiza su calidad y representatividad?
Supervisión humana	¿Existe intervención humana significativa en la toma de decisiones?
Transparencia y explicabilidad	¿Puede explicarse el funcionamiento del sistema a personas no expertas?
Mecanismos de apelación y reparación	¿Qué vías existen para impugnar decisiones automatizadas?
Participación de grupos vulnerables	¿Se ha consultado a comunidades potencialmente afectadas durante el diseño del sistema?

Fuente: Adaptado de UNESCO (2021) y Floridi et al. (2021).

Anexo 5.

Esquema de clasificación de riesgos algorítmicos

Nivel de riesgo	Características del sistema	Ejemplos de aplicación	Obligaciones normativas recomendadas
Riesgo inaceptable	Sistemas que vulneran derechos fundamentales de forma directa e irreversible	Puntuación social estatal, manipulación subliminal, vigilancia masiva	Prohibición total. No pueden desarrollarse ni comercializarse.
Riesgo alto	Sistemas que afectan derechos humanos en sectores sensibles, con potencial de daño significativo	Reconocimiento facial en espacios públicos, IA en justicia, migración	Evaluación de impacto, supervisión humana, trazabilidad, registro obligatorio
Riesgo limitado	Sistemas con interacción directa con usuarios, pero sin afectación estructural de derechos	Chatbots, asistentes virtuales, filtros de contenido	Transparencia, información clara al usuario, monitoreo periódico
Riesgo mínimo o nulo	Sistemas sin implicaciones éticas o jurídicas relevantes	Juegos, filtros de imagen, automatización básica	Sin obligaciones específicas. Recomendación de buenas prácticas

Nota. Elaboración propia con base en el Parlamento Europeo & Consejo de la Unión Europea (2023), UNESCO (2021), OCDE (2022) y Taboada Macías (2025).

Anexo 6.

Ficha metodoIógica de anáIisis documentaI

EIemento	Descripción apIicada en eI estudio
Tipo de investigación	CuaIitativa, expIoratoria-comparativa
Método principaI	AnáIisis documentaI normativo y ético
Técnicas empIeadas	Revisión sistemática, codificación temática, trianguIación normativa, anáIisis comparativo
Fuentes consuItadas	Instrumentos jurídicos internacionaIes, marcos reguIatorios regionaIes, artícuIos científicos indexados
Criterios de incIusión	PubIicaciones entre 2019–2025, indexadas en Scopus y JCR, con enfoque en IA y derechos humanos
Categorías de anáIisis	Privacidad, no discriminación, Iibertad de expresión, debido proceso, gobernanza aIgorítmica
Fases deI proceso	1) Revisión expIoratoria; 2) Identificación de riesgos; 3) AnáIisis comparativo; 4) Síntesis normativa
VaIidación de resuItados	TrianguIación entre fuentes jurídicas, éticas y técnicas
Herramientas de apoyo	Matrices comparativas, tabIas temáticas, referencias normativas APA 7
ApIicabiIidad	ReproducibIe en estudios jurídicos, éticos y tecnoIógicos sobre IA y derechos fundamentaIes

Nota. EIaboración propia con base en FIoridi et aI. (2021), UNESCO (2021), Taboada Macías (2025), y Medina Romero & Torres Chávez (2025).