

Justicia social frente a la IA

Sesgos y vulneraciones de derechos humanos en México, Brasil y Colombia

Yolanda del Sol Ortega Cruz

Universidad Autónoma de Zacatecas, México
<https://orcid.org/0009-0007-5800-0715>
sol.ortega@uaz.edu.mx

Francisco Rosendo Olivares

Universidad Nacional Rosario Castellanos, México
<https://orcid.org/0000-0002-5580-2401>
franciscorosendo31@rcastellanos.cdmx.gob.mx

Jesús Tomás Portillo Juárez

Universidad Veracruzana, México
<https://orcid.org/0000-0003-2785-0390>
zs23000433@estudiantes.uv.mx

Fecha de recepción: 17/4/2026
Fecha de aceptación: 21/5/2026

Resumen

La implementación de la Inteligencia Artificial (IA) en los sistemas de justicia y administración pública de México, Brasil y Colombia se presenta como solución técnica hacia la eficiencia procesal. Empero, bajo la lógica del capitalismo digital, este proceso conlleva riesgos para la justicia social en regiones con desigualdades estructurales. Desde una perspectiva situada de derechos humanos se analiza cómo los sesgos algorítmicos de clase y raza se infiltran en los sistemas de decisión automatizada. Se concluye que la autoridad algorítmica, al carecer de transparencia y supervisión humana estricta, actúa como una barrera para la tutela judicial efectiva con derechos humanos.

Tramas
y Redes
Jun. 2026
N°10
ISSN
2796-9096

Palabras clave

1| inteligencia artificial 2| sesgos algorítmicos 3| derechos humanos 4| justicia social
5| América Latina

Cita sugerida

Ortega Cruz, Yolanda del Sol; Rosendo Olivares, Francisco y Portillo Juárez, Jesús Tomás (2026). Justicia social frente a la IA: sesgos y vulneraciones de derechos humanos en México, Brasil y Colombia. *Tramas y Redes*, (10), 163-180, 10ai. 10.54871/cl4c10ai



Esta obra está bajo licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional https://creativecommons.org/licenses/by-nc-sa/4.0/deed.es_AR

Justiça social versus IA: preconceitos e violações de direitos humanos no México, Brasil e Colômbia

Resumo

A implementação da Inteligência Artificial (IA) nos sistemas de justiça e administração pública do México, Brasil e Colômbia apresenta-se como solução técnica para a eficiência processual. Contudo, sob a lógica do capitalismo digital, esse processo traz riscos à justiça social em regiões com desigualdades estruturais. A partir de uma perspectiva situada de direitos humanos, analisa-se como vieses algorítmicos de classe e raça se infiltram nos sistemas de decisão automatizada. Conclui-se que a autoridade algorítmica, sem transparência e supervisão humana rigorosa, atua como barreira à tutela judicial efetiva e aos direitos humanos.

Palavras-chave

1| inteligência artificial 2| vieses algorítmicos 3| direitos humanos 4| justiça social
5| América Latina

Social justice vs. AI: biases and human rights violations in Mexico, Brazil, and Colombia

Abstract

The implementation of Artificial Intelligence (AI) in the justice and public administration systems of Mexico, Brazil, and Colombia is presented as a technical solution for procedural efficiency. However, under the logic of digital capitalism, this process entails risks for social justice in regions marked by structural inequalities. From a situated human rights perspective, this study analyzes how class- and race-based algorithmic biases infiltrate automated decision-making systems. It concludes that algorithmic authority, lacking transparency and strict human oversight, acts as a barrier to effective judicial protection and human rights.

Keywords

1| artificial intelligence 2| algorithmic bias 3| human rights 4| social justice 5| Latin America

Introducción

La administración de justicia en América Latina atraviesa una transformación paradigmática impulsada por la crisis de eficiencia procesal y el auge del capitalismo digital. En una región marcada por la congestión judicial, donde millones de procesos permanecen detenidos por años, la IA ha surgido como promesa de modernización y “solucionismo” estatal. Sin embargo, esta transición no ocurre en un vacío de poder, sino en un régimen global donde los datos se convierten en capital y la infraestructura tecnológica en un nuevo espacio de disputa por la soberanía y los derechos humanos.

El capitalismo digital, como fase informacional que intensifica las lógicas del neoliberalismo, impulsa la mercantilización de la vida y la automatización de funciones estatales. En los Poderes Judiciales de México, Brasil y Colombia ello se expresa en Sistemas de Decisión Automatizada (SDA) destinados a clasificar, priorizar y, en algunos casos, proyectar resoluciones jurídicas. Aunque la narrativa oficial destaca celeridad y objetividad, la delegación de funciones cognitivas complejas a algoritmos implica riesgos de reproducir desigualdades estructurales. Estos riesgos pueden observarse en dos vertientes: una derivada de la naturaleza misma de la IA en el capitalismo digital y otra vinculada con sus capacidades operacionales.

En un primer sentido, autores como Karl Polanyi, Nancy Fraser y Jürgen Habermas han advertido tendencias contrarias a la tutela de derechos: pérdida de autonomía frente a la dependencia tecnológica (Illich, 1974), expansión del poder privado sobre el público (Gluj, 2024) y colonización del mundo de la vida mediante la tecnificación en detrimento de la deliberación democrática (Habermas, 1984). Desde esta perspectiva, el uso de IA para la tutela de derechos contiene una tensión estructural entre automatización y garantía democrática.

En el segundo eje, los algoritmos no son neutrales, sino constructos humanos que heredan y amplifican sesgos de clase, raza y género presentes en los datos de entrenamiento. En América Latina, donde la justicia ha sido históricamente esquiva para poblaciones marginadas, la “caja negra” algorítmica puede convertirse en barrera para la tutela judicial efectiva y el debido proceso. Además, la interacción entre IA y operadores humanos transforma a la primera en un actante con papel activo dentro del sistema jurídico (Greimas, 1966), capaz de reproducir sesgos sin atravesar procesos deliberativos humanos. Existe amplia literatura sobre sesgos de género reproducidos por la IA (Ledesma, 2022; Pérez-Ugena, 2024; Carrera y Sánchez, 2024; Gordo López y Rubio Martín, 2024; Vásquez Pérez, 2024).

Este estudio desarrolla un análisis comparado y situado de tres experiencias emblemáticas: el impacto de la opacidad tecnológica en el juicio de amparo mexicano, el uso del sistema VICTOR frente al racismo sistémico en el Supremo Tribunal Federal de Brasil y la automatización del triaje de derechos fundamentales mediante PretorIA en la Corte Constitucional de Colombia. Desde una perspectiva de derechos humanos, se examina cómo la autoridad algorítmica desafía el control de convencionalidad y exige una supervisión humana significativa, capaz de subordinar la eficiencia técnica a la dignidad humana.

Marco teórico

La expansión de la IA en el ámbito jurídico no puede comprenderse únicamente como una innovación tecnológica orientada a optimizar tareas, sino como parte de una mutación más amplia del capitalismo contemporáneo, en la que la información se convierte en el principal insumo de acumulación, control y predicción. En ese contexto, la datificación de lo jurídico supone la transformación de conflictos, pruebas, decisiones y trayectorias procesales en conjuntos de datos susceptibles de ser clasificados, correlacionados y procesados por sistemas automatizados. Esta dinámica no es neutral, porque desplaza progresivamente el centro de gravedad del razonamiento jurídico desde la interpretación argumentativa hacia la optimización computacional, con el consiguiente riesgo de reducir la complejidad del derecho a indicadores de eficiencia o probabilidad estadística (Ciccolella, 2024).

En los poderes judiciales, la incorporación de sistemas de IA ha sido presentada como una vía para mejorar la gestión administrativa, reducir tiempos de tramitación, ordenar grandes volúmenes documentales y facilitar la relación con las personas usuarias. El informe de CEJA distingue con claridad entre usos administrativos, que en principio pueden considerarse de menor riesgo, y usos vinculados a la función jurisdiccional, en los que la tecnología puede incidir directamente en la interpretación normativa, la valoración probatoria, la construcción del convencimiento judicial y, en ciertos casos, incluso en la redacción de resoluciones o en la predicción de resultados procesales (CEJA, 2025). Esa distinción es central para el presente estudio, porque permite advertir que no toda aplicación de IA en justicia tiene el mismo impacto sobre los derechos fundamentales, ni genera el mismo grado de afectación a la independencia judicial, al debido proceso o a la motivación de las decisiones (CEJA, 2025).

La literatura revisada coincide en que los algoritmos no son instrumentos neutrales. Por el contrario, su funcionamiento depende de elecciones humanas previas relativas a la selección de datos, la definición

de variables relevantes, el diseño del modelo y el criterio de optimización adoptado. El Consejo de Estado italiano, al analizar la legitimidad del uso de algoritmos en la administración pública, sostuvo que la aparente objetividad de estas herramientas encubre decisiones valorativas previas, pues el modo en que los datos se recogen, sistematizan e interpretan responde a elecciones humanas conscientes o inconscientes, de manera que el algoritmo no elimina la discrecionalidad, sino que la reubica en fases menos visibles del proceso decisorio (Consiglio di Stato, Sent. n. 8472/2019). En el mismo sentido, el trabajo del Consejo de Europa sobre IA y sistemas judiciales subraya que los métodos de aprendizaje automático pueden reproducir patrones del pasado sin comprender las causalidades jurídicas subyacentes, lo que vuelve imprescindible una supervisión humana real y no meramente formal (CEPEJ, 2018/2019).

Desde esta perspectiva, los sesgos algorítmicos no constituyen anomalías excepcionales, sino riesgos estructurales inherentes al uso de datos históricos en contextos judiciales marcados por desigualdad. Si los sistemas se entrenan con decisiones previamente tomadas en entornos donde han operado discriminaciones de clase, raza, territorio o género, el resultado puede ser la reproducción y legitimación de esas mismas asimetrías bajo una apariencia técnica de objetividad. El informe de CEJA advierte que los algoritmos pueden amplificar patrones discriminatorios presentes en los datos de entrenamiento y, en ciertos supuestos, reforzar desigualdades ya arraigadas en el sistema de justicia (CEJA, 2025). De manera convergente, el documento del Consejo de Europa destaca que los sistemas predictivos pueden generar efectos particularmente problemáticos en materia penal, donde el uso de variables indirectas o socioeconómicas puede derivar en estigmatización de poblaciones vulnerables y en la consolidación de perfiles de riesgo que afectan el principio de igualdad ante la ley (CEPEJ, 2018/2019).

El debate adquiere mayor intensidad cuando la IA se utiliza en decisiones que inciden sobre derechos, cargas procesales o consecuencias penales. En esos supuestos, la opacidad técnica de los modelos complejos plantea una tensión directa con el debido proceso, pues toda decisión que afecte derechos debe ser motivada, comprensible e impugnable. La llamada “caja negra” algorítmica dificulta o incluso impide conocer el camino lógico que condujo a un determinado resultado, lo que obstaculiza el ejercicio del derecho de defensa y limita el control jurisdiccional de la decisión. El Consejo de Europa insiste en que la justicia asistida por IA debe respetar principios como transparencia, no discriminación, calidad y seguridad de los datos, y control humano significativo, evitando que la herramienta sustituya o vacíe de contenido la intervención judicial (CEPEJ, 2018/2019). En esa misma línea, CEJA sostiene que las herramientas

que apoyan funciones jurisdiccionales críticas deben someterse a controles estrictos y que su utilización debe ser compatible con la prohibición de delegar funciones propias del juez en sistemas automatizados (CEJA, 2025).

La experiencia comparada también muestra que la mera presencia de un ser humano en el proceso no basta para satisfacer las exigencias del debido proceso. Lo jurídicamente relevante es que la intervención humana sea sustantiva, es decir, capaz de revisar, validar, corregir o rechazar el resultado producido por el sistema. Esta exigencia aparece claramente en la Carta Ética Europea sobre el uso de la IA en los sistemas judiciales, donde se establece que las herramientas deben diseñarse de forma tal que respeten los derechos fundamentales, permitan la comprensión de sus resultados y mantengan bajo control humano las decisiones relevantes (CEPEJ, 2018/2019). De igual manera, CEJA propone que la adopción de IA en los poderes judiciales se acompañe de gobernanza institucional, protocolos de uso, auditorías, capacitación y evaluación de impacto algorítmico, precisamente para evitar que la eficiencia técnica debilite las garantías procesales y la legitimidad judicial (CEJA, 2025).

La desigualdad digital constituye otro eje decisivo para este marco teórico. El acceso a la justicia no depende solo de la existencia formal de tribunales, sino también de la posibilidad real de comprender, usar y controvertir las herramientas tecnológicas que median la relación entre personas y sistema de justicia. El estudio de Vivian Neptune Rivera sobre desigualdad digital y acceso a la justicia advierte que las brechas tecnológicas no se limitan a la conectividad, sino que incluyen la capacidad desigual para acceder a información, manejar plataformas y enfrentar con solvencia técnica la producción y valoración de evidencia digital (Neptune Rivera, 2015). En consecuencia, la expansión de la IA judicial puede profundizar barreras ya existentes si no se acompaña de políticas de alfabetización digital, accesibilidad, asesoría y protección reforzada de personas en situación de vulnerabilidad (Neptune Rivera, 2015; CEJA, 2025).

En términos regulatorios, el debate internacional refleja una orientación creciente hacia modelos basados en riesgos, transparencia y protección reforzada de derechos fundamentales. El Consejo de Europa, a través de su Carta Ética, propone cinco principios rectores: transparencia, no discriminación, respeto de los derechos fundamentales, calidad y seguridad de los datos, y control del usuario (CEPEJ, 2018/2019). Por su parte, CEJA sistematiza una serie de criterios para los poderes judiciales de América Latina y el Caribe, entre ellos la necesidad de separar con claridad la gestión administrativa de la función jurisdiccional, estratificar los riesgos según el tipo de herramienta, y promover usos que fortalezcan el acceso a la justicia sin afectar la independencia judicial ni la motivación

de las decisiones (CEJA, 2025). Estos desarrollos muestran que el problema no es la IA en sí misma, sino el modo en que se diseña, incorpora y controla institucionalmente.

Por lo que la IA plantea una tensión fundamental entre racionalización tecnológica y preservación de la racionalidad jurídica. Si bien puede aportar eficiencia, trazabilidad documental y mejor organización del servicio judicial, también puede desplazar la deliberación humana, reforzar sesgos históricos y convertir el juicio jurídico en una mera validación de resultados probabilísticos. Por lo que se afirma que el uso legítimo de IA en justicia exige condiciones estrictas de legalidad, explicabilidad, supervisión humana efectiva, no discriminación y control institucional. En ausencia de esas garantías, la automatización no moderniza la justicia, sino que corre el riesgo de tecnificar la desigualdad y de debilitar el contenido sustantivo del debido proceso (CEPEJ, 2018/2019; CEJA, 2025; Algoritmos y derechos humanos, s. f.).

Datos y métodos

El presente artículo adopta un enfoque cualitativo, documental y comparado para analizar la incorporación de sistemas de inteligencia artificial en la justicia desde una perspectiva jurídico-constitucional y de derechos humanos. Más que describir experiencias de automatización judicial, examina sus efectos en el debido proceso, la defensa técnica, la motivación de las decisiones y la protección de grupos vulnerables, mediante una lectura crítica de normas, directrices institucionales, jurisprudencia y literatura especializada sobre la relación entre tecnología, poder judicial y garantías procesales. Metodológicamente, se desarrolla como estudio comparado de casos con selección intencional de México, Brasil y Colombia, países que representan distintas modalidades de adopción tecnológica: herramientas incipientes en el control del amparo, automatización a gran escala en el Supremo Tribunal Federal y priorización de acciones constitucionales en la Corte Constitucional, lo que permite identificar tanto beneficios administrativos como riesgos para la función jurisdiccional.

Diseño comparado

El diseño comparado se construye sobre la base de una comparación funcional, no meramente descriptiva. Esto significa que los casos no se analizan como realidades aisladas, sino como respuestas institucionales a problemas semejantes: congestión judicial, sobrecarga de trabajo, necesidad de clasificación masiva de expedientes y búsqueda de eficiencia en el procesamiento de información jurídica. A partir de ello, el estudio identifica semejanzas y diferencias en la forma en que cada sistema judicial

incorpora inteligencia artificial, qué tareas delega a la automatización y cuáles mantiene bajo control humano.

La comparación también permite examinar cómo cada jurisdicción enfrenta la tensión entre eficiencia y garantía. Mientras en algunos contextos la tecnología se presenta como una solución para agilizar trámites y reducir tiempos de espera, en otros se advierte el riesgo de que el uso de algoritmos altere el equilibrio procesal y reduzca la posibilidad de contradicción efectiva. Esta dimensión resulta especialmente importante en América Latina, donde las asimetrías institucionales, la desigualdad digital y la debilidad de ciertos mecanismos de control incrementan la probabilidad de que la tecnología se utilice sin suficientes salvaguardas.

La investigación se apoya en un corpus documental integrado por fuentes normativas, jurisprudenciales, institucionales y académicas. En el plano normativo e institucional se revisan directrices y lineamientos emitidos por organismos como CEJA, la OCDE, la UNESCO y el Consejo de Europa, por tratarse de documentos que establecen principios de gobernanza, transparencia, no discriminación, supervisión humana y uso responsable de IA en sistemas judiciales. Estas fuentes ofrecen un marco de referencia útil para evaluar los criterios con los que los poderes judiciales latinoamericanos están incorporando herramientas algorítmicas.

En el plano jurisprudencial y de política pública se examinan decisiones y proyectos vinculados con el uso de sistemas automatizados en la justicia, así como documentos institucionales que describen experiencias concretas en México, Brasil y Colombia. A ello se suma literatura académica especializada proveniente de repositorios como Scielo, Redalyc y publicaciones universitarias o institucionales de acceso abierto, especialmente aquellas centradas en algoritmos, derechos humanos, acceso a la justicia y administración digital. La combinación de estas fuentes permite construir una base analítica suficientemente sólida para articular una discusión de carácter teórico y empírico.

Cuadro 1. Criterios de comparación

Criterio/País	México	Brasil	Colombia
Institución analizada	Consejo de la Judicatura Federal / juicio de amparo	Supremo Tribunal Federal	Corte Constitucional
Herramienta principal	Sistemas de apoyo y eventual filtrado	VICTOR	PretorIA
Función crítica	Procedencia del amparo y defensa técnica	Repercusión general y clasificación masiva	Priorización de tutelas
Riesgo principal	Opacidad algorítmica	Hipernormatización artificial	Sesgo de clase
Derecho en tensión	Acceso a la justicia y debido proceso	Igualdad, motivación y evolución jurisprudencial	Igualdad procesal y tutela judicial efectiva

Fuente: Elaboración propia con base en CEJA (2025) y CEPEJ (2018/2019).

La selección de casos y documentos respondió a tres criterios principales. El primero fue la relevancia institucional, entendida como la capacidad de la experiencia analizada para incidir de manera significativa en la administración de justicia o en la garantía de derechos fundamentales. El segundo criterio fue la disponibilidad de información pública suficiente para reconstruir el funcionamiento general de la herramienta, su finalidad y sus riesgos. El tercero fue la diversidad funcional, es decir, la posibilidad de incluir casos que representaran distintos usos de inteligencia artificial: filtrado de demandas, clasificación de recursos, priorización de expedientes, apoyo a la toma de decisiones o análisis de grandes volúmenes documentales.

Este criterio de diversidad funcional es particularmente útil porque evita reducir el fenómeno a una sola forma de automatización. No todas las herramientas de IA impactan de la misma manera el sistema de justicia. Algunas se orientan a tareas de gestión y archivo, mientras otras intervienen en procesos de valoración, interpretación o decisión. Por ello, la investigación diferencia entre usos administrativos y jurisdiccionales, y dentro de estos últimos distingue entre herramientas de apoyo, herramientas de selección y herramientas que pueden llegar a sustituir materialmente la voluntad judicial.

La estrategia analítica se organiza mediante una matriz comparativa construida a partir de cuatro dimensiones: jurisdicción analizada, herramienta utilizada, función crítica desempeñada y riesgo principal identificado. Esta matriz permite ordenar la información de manera sistemática y facilita la lectura transversal de los casos. Además, posibilita identificar patrones de riesgo que se repiten en distintos contextos, tales como la opacidad algorítmica, el sesgo de automatización, la pérdida de control sobre los datos y la tendencia a privilegiar la eficiencia por encima de la motivación jurídica.

El análisis se complementa con una lectura crítica de los principios de transparencia, no discriminación, supervisión humana y rendición de cuentas, especialmente desarrollados por la Carta Ética Europea sobre el uso de la IA en los sistemas judiciales y por los lineamientos de CEJA para los poderes judiciales de las Américas. A partir de este marco, la investigación no se limita a valorar si las herramientas funcionan, sino que examina si su funcionamiento es compatible con los estándares constitucionales y convencionales de justicia.

México: amparo y defensa técnica

En el caso mexicano, el análisis se centra en el juicio de amparo como mecanismo de control constitucional y espacio privilegiado para observar los efectos de la digitalización judicial sobre la defensa técnica. El amparo

ocupa una posición singular en el constitucionalismo mexicano al fungir como garantía frente a actos de autoridad que pueden afectar derechos fundamentales, por lo que cualquier herramienta algorítmica que intervenga en su trámite, clasificación o filtrado debe examinarse con especial cautela.

La digitalización judicial ha sido impulsada por la necesidad de atender la carga de trabajo del Consejo de la Judicatura Federal y ampliar el acceso remoto a los servicios de justicia; sin embargo, también puede generar nuevas asimetrías si la opacidad técnica impide a la defensa conocer los criterios con los que un sistema automatizado filtra una demanda o sugiere una improcedencia. El riesgo principal no es solo técnico, sino procesal: que el algoritmo opere como un guardián invisible que limite el acceso a la jurisdicción constitucional sin una posibilidad real de contradicción.

Brasil: automatización masiva y repercusión general

Brasil constituye un caso especialmente relevante por el alto grado de digitalización de su Poder Judicial y la implementación del sistema VICTOR en el Supremo Tribunal Federal, utilizado para identificar temas de repercusión general y clasificar masivamente recursos extraordinarios. La magnitud del volumen procesal convierte a esta herramienta en un laboratorio institucional para analizar la relación entre automatización y justicia constitucional, ya que no solo agiliza tareas, sino que organiza el acceso mismo a la decisión judicial. Esto puede condicionar la evolución jurisprudencial y reducir la sensibilidad del tribunal frente a nuevos contextos sociales o cambios fácticos no previstos en la base de datos. A ello se suma el problema del racismo estructural y la eventual ausencia de auditorías que detecten sesgos de clase o raza, de modo que, en un sistema marcado por desigualdades históricas, la automatización sin controles puede reforzar exclusiones bajo una apariencia de neutralidad técnica.

Colombia: tutela y priorización de casos

Colombia ofrece una experiencia distinta, pero igualmente valiosa, mediante la herramienta PretorIA, diseñada para gestionar el elevado número de acciones de tutela presentadas ante la Corte Constitucional. Su relevancia metodológica reside en que permite analizar un uso de IA orientado a priorizar casos con presunta vulneración de derechos fundamentales, especialmente en salud, pensiones y seguridad social. No obstante, también plantea el riesgo de reproducir sesgos de clase o capital jurídico: si el sistema se entrena con tutelas históricamente exitosas, podría privilegiar formatos discursivos más técnicos, asociados a litigantes de

grandes centros urbanos, y relegar reclamaciones provenientes de zonas rurales o de personas con menor formación jurídica. Así, la herramienta puede convertirse en un mecanismo de triage judicial que, sin regulación adecuada, amplíe la brecha entre quienes saben formular mejor sus derechos y quienes enfrentan mayores obstáculos para hacerlo.

Tabla 2. Comparativa

Sistema judicial	Herramienta	Función crítica	Riesgo principal identificado
México (CJF)	Diversas herramientas para gestión y posible filtrado	Filtrado de demandas de amparo y apoyo a la tramitación	Opacidad algorítmica, afectación a la defensa técnica y riesgo de rechazo invisible de casos
Brasil (STF)	VICTOR	Clasificación de recursos por repercusión general	Hipernormatización artificial, rigidez categorial y posible reproducción de sesgos estructurales
Colombia (Corte Constitucional)	PretorIA	Priorización y selección de acciones de tutela	Sesgo de clase, exclusión de narrativas rurales y ventaja para lenguajes jurídicos más técnicos

Fuente: Elaboración propia con base en CEJA (2025) y CEPEJ (2018/2019).

La comparación entre México, Brasil y Colombia muestra que los riesgos de la IA en la justicia no se distribuyen de manera uniforme, sino que dependen de contextos institucionales y sociales específicos. En México predomina la preocupación por la opacidad y la afectación de la defensa técnica; en Brasil, por la rigidez jurisprudencial y la reproducción de desigualdades históricas; y en Colombia, por la priorización sesgada de ciertas narrativas jurídicas en el acceso a la tutela judicial efectiva. En consecuencia, el riesgo debe evaluarse según la desigualdad social, la calidad institucional y la capacidad de supervisión. La cuestión central no es si la IA puede utilizarse en la justicia, sino bajo qué condiciones normativas, técnicas y democráticas resulta compatible con el debido proceso, los derechos fundamentales y la preservación del juicio humano, la motivación judicial y el acceso igualitario a la justicia.

Discusión

La comparación sistemática de los casos de México, Brasil y Colombia revela una tensión estructural que atraviesa la incorporación de la IA en la administración de justicia: la primacía de la eficiencia técnica sobre la preservación de la justicia como proceso humano, deliberativo y orientado por principios de derechos humanos. En los tres contextos analizados, herramientas como los sistemas de filtrado para amparos en México, VICTOR en el Supremo Tribunal Federal brasileño o PretorIA en la corte constitucional colombiana han sido implementadas con el propósito

explícito de agilizar trámites, reducir tiempos de espera y gestionar volúmenes masivos de expedientes judiciales. Si bien estos objetivos responden a problemas reales de congestión y sobrecarga laboral, la narrativa de la eficiencia ha desplazado una discusión más profunda sobre la calidad sustantiva de la justicia y sus efectos sobre la igualdad procesal. Como advierte la Carta Ética Europea sobre el uso de la IA en los sistemas judiciales, el procesamiento algorítmico tiende a generar una “norma cuantitativa” donde el resultado estadísticamente más probable se convierte en estándar de decisión, lo cual es particularmente peligroso en América Latina, donde el derecho debe operar como contrapeso de desigualdades históricas y proteger a grupos vulnerables que no se ajustan a promedios estadísticos (CEPEJ, 2018/2019; CEJA, 2025).

Esta eficiencia técnica no es meramente retórica, sino que se manifiesta en prácticas que erosionan la agencia judicial y la legitimidad institucional. El sesgo de automatización describe el fenómeno por el cual operadores judiciales exhaustos confían excesivamente en las sugerencias de los sistemas, reducen la revisión crítica y convierten al juez en validador formal de outputs técnicos. En Brasil, por ejemplo, VICTOR procesa decenas de miles de recursos anuales para identificar temas de repercusión general, acelera la clasificación, pero induce a una hipernormatización artificial que fuerza los casos a encajar en categorías preexistentes, limita la sensibilidad del tribunal frente a transformaciones sociales o matices fácticos no previstos en sus datos de entrenamiento. De manera similar, en Colombia, PretorIA prioriza tutelas basándose en patrones lingüísticos y temáticos, lo que puede favorecer reclamos con lenguaje técnico urbano sobre narrativas rurales o menos formalizadas, reproduce así brechas de capital jurídico. En México, la eventual automatización del análisis de procedencia en amparos agrava esta dinámica, porque el algoritmo podría rechazar demandas sin considerar la suplencia de la deficiencia de la queja, un principio protector esencial del control constitucional. Estudios recientes sobre IA en procesos judiciales destacan que esta confianza acrítica no solo acelera trámites, sino que debilita la motivación de las decisiones y la posibilidad real de impugnación, transforma la justicia en una operación administrativa opaca.

El problema se agrava cuando la eficiencia se logra a expensas de la atención personalizada a casos complejos. En contextos latinoamericanos marcados por racismo estructural, desigualdad territorial y brechas digitales, los algoritmos entrenados con datos judiciales históricos sesgados pueden amplificar exclusiones bajo apariencia de neutralidad técnica. Investigaciones sobre justicia algorítmica en la región advierten que sistemas como los analizados tienden a reproducir patrones discriminatorios de vigilancia selectiva, sentencias más severas para poblaciones

pobres o marginación de minorías, legitimándolos estadísticamente en lugar de corregirlos. La falacia radica en equiparar mayor velocidad con mayor equidad, cuando la realidad evidencia que la IA puede consolidar una “justicia *low-cost*” accesible para algunos, pero inaccesible o sesgada para otros, y erosiona la esencia del debido proceso como garantía de contradicción efectiva y motivación comprensible.

Frente a esta racionalidad algorítmica expansiva, el control de convencionalidad surge como límite normativo ineludible y herramienta de resistencia institucional. En los sistemas jurídicos latinoamericanos, este control obliga a los jueces a confrontar prácticas tecnológicas con la Convención Americana sobre Derechos Humanos y la jurisprudencia interamericana, desaplica cualquier uso de IA que vulnere el derecho a un proceso justo, la igualdad de armas o la tutela judicial efectiva. Plataformas como THEMIS IA de la Corte IDH ilustran cómo la tecnología puede fortalecer el Estado de Derecho si se alinea con principios de transparencia, rendición de cuentas y dignidad humana, pero también advierten que la soberanía tecnológica es una exigencia jurídica: los poderes judiciales no pueden delegar conocimiento jurídico en cajas negras comerciales protegidas por secretos industriales, porque ello compromete la independencia judicial y la rendición de cuentas democrática. Es importante el derecho a conocer la lógica de decisión en un lenguaje comprensible, no solo técnico es condición *sine qua non* para la impugnación efectiva, convirtiéndose en una garantía procesal frente a la opacidad inherente a modelos complejos.

La gobernanza ética de la IA judicial demanda medidas concretas y contextualizadas al escenario latinoamericano, superando regulaciones abstractas por enfoques específicos adaptados a los riesgos identificados. En primer término, las evaluaciones de impacto algorítmico (AIA/EIA) deben ser obligatorias en fases previas y continuas de implementación, permite detectar sesgos discriminatorios, evaluar la calidad y representatividad de los datos, y prever afectaciones a derechos fundamentales como la no discriminación o la privacidad. Estos instrumentos, ya regulados en marcos europeos y recomendados por la OCDE, son esenciales en justicia porque permiten mitigar riesgos como la discriminación indirecta o la falta de trazabilidad, transformándose en mecanismos de responsabilidad proactiva. En segundo lugar, el derecho a la explicabilidad debe institucionalizarse como garantía procesal robusta: toda persona afectada por decisión influida por IA tiene derecho a reconstruir, en términos accesibles, la lógica algorítmica, los datos ponderados y su integración en el razonamiento judicial, evitando que la motivación se reduzca a validaciones formales. Finalmente, el diseño centrado en derechos humanos exige integrar desde la programación principios como

proporcionalidad, equidad, supervisión humana y protección de datos sensibles, alineando la IA con normas internacionales y evitar la reproducción de desigualdades estructurales.

En síntesis, la discusión comparada muestra que la IA judicial no es inherentemente benéfica ni perjudicial; su legitimidad depende de una arquitectura institucional subordinada al control democrático y a los estándares interamericanos. La amenaza principal no radica en la tecnología en sí, sino en su uso acrítico para privilegiar eficiencia sobre equidad, opacidad sobre motivación y automatización sobre deliberación humana. Por ello, una gobernanza situada exige evaluaciones de impacto, explicabilidad efectiva y diseño normativamente anclado en DDHH, a fin de que la justicia digital fortalezca, y no erosione, el constitucionalismo latinoamericano. De lo contrario, el riesgo es una administración judicial eficiente pero injusta, donde la velocidad técnica sustituya la profundidad ética (CEJA, 2025).

Conclusiones

La incorporación de la IA en los sistemas judiciales de México, Brasil y Colombia responde a la lógica del capitalismo digital y a la necesidad de enfrentar la mora judicial regional. Aunque estas herramientas permiten procesar grandes volúmenes de información, agilizar trámites y mejorar la gestión, también revelan tensiones estructurales en materia de justicia social, debido proceso y derechos humanos. El análisis comparado muestra que la eficiencia técnica no puede legitimarse como fin autónomo cuando entra en conflicto con garantías esenciales como la motivación comprensible de las decisiones, la igualdad procesal y la contradicción efectiva (CEJA, 2025; CEPEJ, 2018/2019).

En México, la opacidad de los algoritmos aplicados al filtrado de demandas de amparo puede convertir una garantía constitucional en un mecanismo de exclusión técnica, dificultando la defensa y erosionando la esencia protectora del amparo. En Brasil, la clasificación masiva de recursos mediante VICTOR corre el riesgo de rigidizar la jurisprudencia y reproducir racismo estructural y desigualdades de clase al privilegiar patrones históricos sobre nuevas realidades sociales. En Colombia, PreorIA representa un avance en la priorización de tutelas, pero enfrenta el desafío de no marginar reclamos rurales, indígenas o populares cuyos formatos discursivos suelen estar subrepresentados en bases de datos urbanas y técnicas.

Estas experiencias evidencian que la IA no es neutral: aprende de datos históricos cargados de sesgos sociales, culturales y estructurales, y puede reproducirlos bajo apariencia de objetividad. En contextos latinoamericanos, ello puede traducirse en vigilancia selectiva sobre

poblaciones pobres, decisiones más severas para minorías racializadas o exclusión de reclamos formulados en lenguajes no estandarizados. El sesgo de automatización agrava el problema cuando operadores judiciales sobrecargados confían acríticamente en resultados técnicos y reducen la revisión humana.

Desde una perspectiva ética, estos riesgos comprometen principios como equidad, transparencia y proporcionalidad. La falta de explicabilidad impide impugnar decisiones, detectar errores y sostener la confianza ciudadana en la justicia. Por ello, una implementación legítima exige diversidad en los datos, auditorías independientes, mitigación proactiva de sesgos y diseños orientados por la dignidad humana y la justicia social.

Se concluye que la autoridad algorítmica debe subordinarse a un control humano significativo y a un control de convencionalidad difuso, alineado con la Convención Americana y estándares interamericanos. La eficiencia técnica no puede ocultar la erosión de garantías fundamentales; para que la IA sea progreso, debe evolucionar de herramienta burocrática a instrumento de transparencia y empoderamiento ciudadano, garantizando que la justicia latinoamericana permanezca humana, ética y equitativa (CEJA, 2025).

Referencias

- Agencia de los Derechos Fundamentales de la Unión Europea (FRA) (2020). *Getting the future right: artificial intelligence and fundamental rights*. Viena. <https://fra.europa.eu/en/publication/2020/getting-future-right-artificial-intelligence-and-fundamental-rights>
- BBVA *OpenMind* (2025). Innovación y tecnología para combatir los sesgos de la inteligencia artificial. BBVA. <https://www.bbva.com/es/innovacion/innovacion-y-tecnologia-para-combatir-los-sesgos-de-la-inteligencia-artificial/>
- Castets-Renard, Céline (2021). Human rights and algorithmic impact assessment for predictive policing. En *Constitutional challenges in the algorithmic society* (pp. 123-145). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108914891.007>
- Centro de Estudios de Justicia de las Américas (CEJA). (2025). IA en los poderes judiciales: reflexiones y lineamientos para las Américas. Santiago. https://cejamericas.org/wp-content/uploads/2025/12/Inteligencia-Artificial-en-los-Poderes-Judiciales-de-las-Américas_.pdf
- Ciccolella, Pablo (2024). Capitalismo digital, modo de desarrollo informacional y territorio en América Latina. *Geo UERJ*, (46), 1-25.

- <https://www.e-publicacoes.uerj.br/geouerj/article/download/87488/51901/329040>
- Comisión Europea para la Eficiencia de la Justicia (CEPEJ). (2018). *Carta ética europea sobre el uso de la IA en los sistemas judiciales*. Estrasburgo: Consejo de Europa. <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>
- Consiglio di Stato (2019). Sentenza n. 8472 del 13 dicembre 2019. Roma. <https://images.go.wolterskluwer.com/WebWoltersKluwer-7Bae89464-6ad9-479d-ac0b-ce43a7f696127Dconsiglio-di-stato-sentenza-8472-2019.pdf>
- Corte Constitucional de Colombia (2020). PretorIA: un ejemplo de incorporación de tecnologías de punta en el sector justicia. Bogotá. <https://www.corteconstitucional.gov.co/relatoria/autos/2020/a1280-20.htm>
- Corte Interamericana de Derechos Humanos (2024). Estándares interamericanos de DDHH y uso estatal de la IA. San José. https://www.eff.org/files/2024/10/30/estandares_interamericanos_de_ddhh_y_uso_estatal_de_la_ia_final.pdf
- De Lara-García, José.(2022). IA y justicia: experiencias en América Latina. *DIVULGARE*, 9(17), 45-67. <https://dialnet.unirioja.es/descarga/articulo/10097225.pdf>
- Gluj, Anabella. (2024). La concepción ampliada del capitalismo de Nancy Fraser: reflexiones sobre su conceptualización y periodización. *Rey Desnudo. Revista de Libros*, 13(25), 194-218. <https://dialnet.unirioja.es/descarga/articulo/10194615.pdf>
- Goethe-Institut (2024). Sesgo y error: los prejuicios de la IA. *Goethe Institut*. Ciudad de México. <https://www.goethe.de/ins/mx/es/kul/wir/22842955.html>
- Gordo López, Ángel, y Rubio Martín, Ma. José (2024). Incertidumbres algorítmicas en torno a las violencias de género. El caso del sistema VioGén y otros sistemas de predicción del riesgo. *Revista Española De Sociología*, 33(2), a225.
- Greimas, Algirdas Julien. (1966). *Sémantique structurale: recherche de méthode*. Paris: Larousse.
- Gutiérrez, Juan D. (2024). Critical appraisal of large language models in judicial decision-making. En Regine, Paul Emma Carmel y Jennifer Cobbe (eds.), *Handbook on public policy and AI*. Cheltenham: Edward Elgar Publishing. En prensa. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4667572
- Habermas, Jürgen (1984). *The theory of communicative action: reason and the rationalization of society* (Vol. 1, T. McCarthy, trad.). Boston: Beacon Press.

- Illich, Iván (1974). *La convivencialidad*. Barcelona: Barral Editores.
- Ledesma, José O. (2022). Algoritmos y género: IA al servicio de la violencia simbólica. *Revista Llapanchikpaq: Justicia*, 4(5), 209-236. <https://doi.org/10.51197/lj.v4i5.659>
- Miranda Bonilla, Haideer (2021). Algoritmos y derechos humanos. *Revista de la Facultad de Derecho de México*, 71(280), 123-145.
- Neptune Rivera, Vivian I. (2015). *La desigualdad digital y el acceso a la justicia*. New Haven: Yale Law School. https://law.yale.edu/sites/default/files/documents/pdf/SELA15_Neptune_CV_Sp.pdf
- Ferrante, Enzo (2021). Inteligencia artificial y sesgos algorítmicos. ¿Por qué deberían importarnos? *Nueva Sociedad*. <https://nuso.org/articulo/inteligencia-artificial-y-sesgos-algoritmicos/>
- OCDE. (2025). IA en la administración y acceso a la justicia: gobernar con la inteligencia artificial. París: OCDE. https://www.oecd.org/es/publications/gobernar-con-la-inteligencia-artificial_dc00e56a-es/full-report/ai-in-justice-administration-and-access-to-justice_f0cbe651.html
- Quipildor Rocha, Pilar (2025). Los riesgos del sesgo algorítmico en la IA y su impacto en el derecho del consumidor. *Sistema Argentino de Información Jurídica*. Buenos Aires. <https://www.saij.gov.ar/IdSAIJ: DACF250037>
- Pérez-Ugena Coromina, María (2024). Sesgo de género (en IA). *Eunomía: Revista en Cultura de la Legalidad*, 311-330.
- Sánchez Caparrós, Mariana y Ross, Zarina (2024). *Violencia de género digital en la era de la inteligencia artificial*. Ushuaia: Poder Judicial de Tierra del Fuego. <https://www.justierradelfuego.gov.ar/wp-content/uploads/2024/04/Violencia-de-genero-digital-en-la-era-de-la-inteligencia-artificial.pdf>
- Soto, Antonio (2024). Desenmascarando los sesgos algorítmicos: ética y transparencia en la IA. *LinkedIn*. Sunnyvale. <https://es.linkedin.com/pulse/desenmascarando-los-sesgos-algor%C3%ADtmicos-%C3%A9tica-y-en-la-antonio-soto-o29wf>
- Supremo Tribunal Federal de Brasil (STF) (2024). *Impactos da inteligência artificial no constitucionalismo contemporâneo*. Brasília. https://www.stf.jus.br/arquivo/cms/codi/anexo/Impactos_da_IA_no_Constitucional_Contemporaneo_2024_2703251.pdf
- UNESCO (2023). *Kit de herramientas sobre la IA y el Estado de derecho para el poder judicial*. París: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000387331_spa

Vásquez Pérez, María Nely (2024). Ética y género en la IA: identificar sesgos de género en IA mediante pensamiento complejo. *Revista Iberoamericana De Complejidad Y Ciencias Económicas*, 2(2), 49-62. <https://doi.org/10.48168/ricce.v2n2p49>