

42

TEMAS PROCESALES

Vanessa Franco Ramírez
Editora



RED

— Proceso y Justicia —

2025-2 ISSN 2619-3655

Injusticia algorítmico-epistémica y valoración probatoria: desafíos probatorios de la injusticia testimonial y lineamientos para minimizar sus riesgos en la decisión judicial¹

Algorithmic-epistemic injustice and evidential assessment: evidential challenges of testimonial injustice and guidelines to minimize its risks in judicial decisions

Injustiça Algorítmico-Epistêmica e Valoração Probatória: Desafios Probatórios da Injustiça Testemunhal e Diretrizes para Minimizar seus Riscos na Decisão Judicial

Eddison David Castrillón García

Universidad Pontificia Bolivariana.

eddison.castrillon@upb.edu.co.

Orcid: <https://orcid.org/0000-0002-1252-8887>

Recibido: 12/07/2025 / Aceptado: 12/09/2025 / Publicado: 12/12/2025

DOI: 10.63865/temasp.v42n42a8

Resumen

Este artículo tiene como propósito formular lineamientos epistemológicos, jurídicos y técnicos que contribuyan al fortalecimiento de la administración de justicia, partiendo de la necesidad de mitigar el riesgo inherente a la injusticia algorítmico-epistémica. Con el fin de lograr este objetivo propuesto, se desarrolló una investigación de carácter cualitativo, sustentada en un enfoque analítico-deductivo. Esta metodología integró el análisis documental con la aplicación del enfoque fenomenológico propio de la tradición filosófica, concretando un abordaje interdisciplinario del problema objeto de estudio, el cual se ha denominado injusticia algorítmico-epistémica. La injusticia algorítmica se puede conjugar con la injusticia epistémica en detrimento de personas y grupos históricamente excluidos y vulnerables frente al acceso a la justicia y la tutela judicial efectiva, por lo que se requiere su adecuada comprensión para minimizar los riesgos y lograr el fortalecimiento de la tutela judicial y la observancia de las garantías procesales en escenarios en los cuales se aplican sistemas de inteligencia artificial en la administración de justicia.

Palabras clave: injusticia algorítmico-epistémica, injusticia epistémica, injusticia hermenéutica, injusticia testimonial, inteligencia artificial.

¹ Artículo resultado de la investigación doctoral que lleva por nombre «Fenomenología ética del testimonio en la justicia digital: Lineamientos éticos para la implementación de la IA en la decisión judicial frente a la prueba testimonial», inscrita en el Grupo de Investigación Epimeleia de la Universidad Pontificia Bolivariana y desarrollada en el marco de la formación doctoral en Filosofía de la misma Universidad.

Abstract

This article aims to formulate epistemological, legal, and technical guidelines that contribute to strengthening the administration of justice, based on the need to mitigate the risk inherent in algorithmic-epistemic injustice. To achieve this proposed objective, a qualitative research project was developed, based on an analytical-deductive approach. This methodology integrated documentary analysis with the application of the phenomenological approach characteristic of the philosophical tradition, concretizing an interdisciplinary approach to the problem under study, which has been called algorithmic-epistemic injustice. Algorithmic injustice can be combined with epistemic injustice, to the detriment of individuals and groups historically excluded and vulnerable to access to justice and effective judicial protection. Therefore, a proper understanding of these injustices is required to minimize their risks and strengthen judicial protection and the observance of procedural guarantees in scenarios where AI systems are applied in the administration of justice.

Keywords: algorithmic-epistemic injustice, epistemic injustice, hermeneutical injustice, testimonial injustice, artificial intelligence.

Resumo

Este artigo tem como propósito formular diretrizes epistemológicas, jurídicas e técnicas que contribuam para o fortalecimento da administração da justiça, partindo da necessidade de mitigar o risco inerente à injustiça algorítmico-epistêmica. A fim de alcançar este objetivo proposto, desenvolveu-se uma pesquisa de caráter qualitativo, sustentada em uma abordagem analítico-dedutiva. Esta metodologia integrou a análise documental com a aplicação da abordagem fenomenológica própria da tradição filosófica, concretizando uma abordagem interdisciplinar do problema objeto de estudo, o qual foi denominado injustiça algorítmico-epistêmica. A injustiça algorítmica pode se conjugar com a injustiça epistêmica em detrimento de pessoas e grupos historicamente excluídos e vulneráveis no que se refere ao acesso à justiça e à tutela jurisdicional efetiva, razão pela qual se requer sua adequada compreensão para minimizar os riscos e lograr o fortalecimento da tutela judicial e a observância das garantias processuais em cenários nos quais se aplicam sistemas de inteligência artificial na administração da justiça.

Palavras-chave: injustiça algorítmico-epistêmica, injustiça epistêmica, injustiça hermenêutica, injustiça testemunhal, inteligência artificial.

1. Introducción

La implementación de sistemas de inteligencia artificial (IA) en el ámbito de la administración de justicia representa grandes desafíos, pues la neutralidad de los sistemas de IA es cada vez más cuestionada, como lo prueba la reiterada evidencia sobre los sesgos discriminatorios que pueden reproducir dichos sistemas.

Con el auge de la IA en todos los campos del conocimiento, se evidencian situaciones en las cuales se toman decisiones apoyadas en IA que causan perjuicios injustos en las personas. A este fenómeno se le denomina injusticia algorítmica (Castrillón García, 2023), y puede presentarse también en conjunto con otro tipo de injusticia, denominado por Fricker (2017) *injusticia epistémica* y en tal caso, en el ámbito judicial, se puede profundizar la situación desfavorable de grupos socialmente vulnerables frente al acceso a la justicia y la tutela judicial efectiva.

El estudio de este fenómeno anteriormente descrito resulta importante, toda vez que no ha sido suficientemente estudiado. Por ello, en el presente artículo se

realiza su abordaje desde una perspectiva interdisciplinaria, con el fin de proponer lineamientos epistemológicos, jurídicos y técnicos que permitirán fortalecer la administración de justicia teniendo como premisa el control o la eliminación del riesgo de injusticia algorítmico-epistémica.

Para lograr este propósito se llevó a cabo una investigación de tipo cualitativo, desde un enfoque analítico-deductivo que combinó análisis documental y aplicación del enfoque fenomenológico propio de la filosofía.

En este orden de ideas, en un primer apartado se estudiará la injusticia algorítmico-epistémica en el ámbito de la administración de justicia; en un segundo apartado, se analizarán los desafíos probatorios de la injusticia testimonial y la injusticia algorítmica en el ámbito judicial; en un tercer apartado, se propondrán lineamientos para minimizar el riesgo de injusticia algorítmico-epistémica en la administración de justicia; y finalmente, se presentarán unas conclusiones.

2. Injusticia algorítmica en diferentes ámbitos

La injusticia algorítmica puede manifestarse en diferentes ámbitos como el médico, el laboral, el financiero y el jurídico.

En el escenario laboral se han identificado ciertos casos en los que los algoritmos empleados para seleccionar personal han generado exclusión frente a grupos minoritarios, en especial tratándose de aspectos como la raza o el género. Estas prácticas discriminatorias evidencian cómo los sistemas automatizados pueden perpetuar desigualdades estructurales preexistentes que terminan afectando negativamente la igualdad en el acceso al empleo (Corrales Aznar, 2022, p. 418).

Una situación similar a la anterior se ha presentado en el ámbito financiero, en donde se ha evidenciado la presencia de sesgos algorítmicos en los procesos de evaluación crediticia, en los cuales los sistemas de IA incorporaron variables no asociadas a la solvencia económica de una persona, reproduciendo criterios discriminatorios como el origen étnico o el género (Eguíluz Castañeira, 2020, p. 357).

En el ámbito médico, la problemática asociada a la injusticia algorítmica se manifiesta cuando se emplean sistemas de IA que se alimentan con bases de datos que no reflejan de manera adecuada el contexto específico de los pacientes que son objeto de estudio o cuando dichos sistemas no han sido entrenados adecuadamente (Hao, 2021). En este escenario se pueden derivar consecuencias perjudiciales e injustas para el paciente, al emplear los resultados generados por la IA con el fin de establecer un diagnóstico o un tratamiento (Castrillón García, 2025). Un ejemplo es el programa *Medicaid* en el Estado de New Jersey, donde miles de personas con discapacidad reciben servicios médicos, cuya asignación está mediada por el algoritmo NJCAT (Reisman, 2022). En este escenario se identificaron múltiples casos en los que personas mayores con discapacidad estaban en

situaciones de riesgo para su vida e integridad personal, pues el sistema de IA presentó errores significativos, interpretando de manera incorrecta las necesidades de los pacientes con discapacidad y generando resultados considerados «catastróficos» (Reisman, 2022). Estos perjuicios se atribuyeron a la presencia de sesgos en la construcción del algoritmo NJCAT, específicamente por no contemplar de forma adecuada las necesidades diferenciadas de atención médica de las personas con discapacidad en etapas avanzadas de la vida.

3. Injusticia algorítmica en la justicia predictiva

Se han documentado casos de discriminación algorítmica en sistemas de IA predictivos diseñados para calcular la probabilidad de reincidencia delictiva. Es así como se ha podido detectar que estos sistemas han incorporado variables estrechamente vinculadas con la raza y el nivel socioeconómico, lo que ha producido en la práctica decisiones sesgadas que refuerzan dinámicas estructurales de exclusión social (Barona Vilar, 2021). Por este motivo, es fundamental someter a un riguroso escrutinio los algoritmos que, pese a su aparente neutralidad, constituyen el núcleo operativo de la IA aplicada al ámbito de la administración de justicia. En este sentido, Agudelo Londoño (2022) advierte, por ejemplo, sobre el uso del sistema experto «Fiscal Watson» usado en Colombia, señalándolo de utilizar solo una fracción de la limitada información disponible sobre los delitos reportados para estimar la probabilidad de atribución de responsabilidad penal frente a una persona determinada. Y aun así, en este sistema las oportunidades para que una persona pueda impugnar o cuestionar el resultado probabilístico generado por el algoritmo son muy limitadas, toda vez que se trata de una herramienta protegida por patente y desarrollada por una empresa privada (Agudelo Londoño, 2022).

El auge en la implementación de la justicia predictiva implica analizar los casos de injusticia algorítmica, lo cual es un riesgo potencial en tanto que la transparencia y explicabilidad de los algoritmos de los sistemas de IA son todo un desafío. Esto debido a que dichos algoritmos pueden operar como auténticas cajas negras, puesto que el proceso mediante el cual se generan las decisiones puede ser inteligible. Esta falta de claridad en la lógica subyacente dificulta de manera significativa la identificación y corrección de posibles sesgos discriminatorios que se reproducen en los resultados que arrojan estos sistemas.

A lo anterior se suma el hecho de que cada vez somos más dependientes de la tecnología y en ocasiones llegamos a depositar tanta confianza en ella que resulta siendo indispensable para la toma de ciertas decisiones. Este fenómeno lleva consigo la configuración de una vigilancia disminuida frente a aquellas situaciones que le delegamos a la tecnología, so pretexto de que ella lo hace mejor que nosotros, por lo que tendemos a relegar nuestras funciones de monitoreo, verificación de la información e incluso el ejercicio del juicio propio (Hardré, 2016, citado por Arévalo Ortiz, 2024).

Es precisamente en este escenario en el cual no solo se pone en juego una forma atenuada de vigilancia, sino que también se manifiesta una modalidad específica de injusticia epistémica: la injusticia testimonial. De acuerdo con Fricker (2017), esta ocurre cuando los prejuicios del oyente conducen a una atribución injustamente reducida de credibilidad a las afirmaciones del hablante. Es así como confiar excesivamente en sistemas de IA aplicados a la justicia predictiva tiene el potencial de que los modelos asociados a dichos sistemas terminen siendo priorizados en la toma de decisiones sin cuestionar los criterios de diseño o entrenamiento o sin cotejar los resultados con otras fuentes de verificación.

Esto último sin mencionar que la implementación de la justicia predictiva puede terminar privilegiando la eficiencia sobre los principios éticos y jurídicos. Un ejemplo para ilustrar esta situación es el caso de *PredPol*, un sistema de IA creado para predecir las zonas de una ciudad en las que se tiene más probabilidad de que se cometa un delito y que únicamente se alimenta de tres datos básicos: el tipo de crimen, el lugar donde se cometió dicho crimen y la fecha exacta en la que tuvo lugar. La experiencia asociada con la aplicación de dicho algoritmo ha evidenciado una criminalización y sobrevigilancia de comunidades históricamente marginadas, desatendiendo de forma paralela las zonas privilegiadas y reproduciendo los ciclos de exclusión y pobreza. Se ha concluido que estos efectos son consecuencia de sesgos implícitos en el diseño del modelo predictivo producto de la percepción o perspectiva de los programadores. En este orden de ideas, se puede dar el caso de que el testimonio de una persona que pertenezca a dichas comunidades sobrevigiladas llegue a ser desestimado debido al sesgo tecnológico, sin importar las pruebas que pudieran existir sobre el buen comportamiento del sujeto.

Para el caso de *PredPol*, entonces, es imperativo habilitar espacios deliberativos en los que las voces de las comunidades afectadas sean no solo escuchadas, sino también reconocidas como fuentes legítimas de conocimiento en el debate sobre las políticas de seguridad y las prácticas policiales. Esto, toda vez que la aplicación de *PredPol* «evidenció que en muchos casos los testimonios de las personas judicializadas fueron desestimados a favor de las decisiones algorítmica, lo cual también puede ser observado como un problema de vigilancia disminuida» (Arévalo Ortiz, 2024, p.52). Esta sugerencia, por ejemplo, sería una forma de generar una apertura epistemológica para comenzar a confrontar la injusticia testimonial que las tecnologías de vigilancia tienden a perpetuar (Arévalo Ortiz, 2024).

La falta de transparencia y de mecanismos claros de rendición de cuentas en el diseño y la implementación de algoritmos de IA constituye un elemento crítico que favorece la perpetuación de la discriminación algorítmica. En numerosas ocasiones, estos sistemas son desarrollados y operados por entidades privadas sin que exista una supervisión adecuada o un marco regulatorio claro. Todo esto dificulta también la comprensión de los procesos internos de los algoritmos y de los criterios que se emplean para la toma de decisiones automatizadas.

En consecuencia, la implementación de sistemas de IA en el ámbito judicial plantea desafíos significativos en términos de transparencia y comprensión de los procesos decisionales. La naturaleza opaca de ciertos sistemas algorítmicos compromete la rendición de cuentas y dificulta la interpretación de las decisiones judiciales, lo cual puede minar la confianza pública en la administración de justicia. Por ello, la adecuación de los marcos normativos, el fomento de la transparencia en los procesos algorítmicos, la prevención de sesgos sistémicos y la capacitación permanente de los actores judiciales constituyen pilares fundamentales para garantizar que el uso de tecnologías en el ámbito jurídico se alinee con los principios rectores de la justicia y, en especial, con el respeto por el debido proceso.

4. Injusticia algorítmico-epistémica e injusticia epistémica generativa

A partir de los planteamientos de Fricker (2017), cabe destacar la manera como Kay et al. (2024) se aproximan al problema de la injusticia algorítmico-epistémica, pero a través del concepto de «injusticia epistémica generativa». Como se muestra en la Tabla 1, con este concepto se pueden establecer cuatro variaciones o configuraciones de la injusticia epistémica generativa.

Tabla 1.
Configuraciones de la injusticia epistémica generativa

Denominación	Descripción de la configuración
Injusticia testimonial amplificada generativa	Cuando la IA generativa magnifica y produce puntos de vista socialmente sesgados a partir de sus datos de entrenamiento
Injusticia testimonial manipulativa generativa	Cuando los humanos fabrican injusticias testimoniales con la IA generativa
Ignorancia hermenéutica generativa	Cuando la IA generativa carece de los marcos interpretativos para comprender las experiencias humanas
Injusticia de acceso hermenéutico generativa	Cuando la IA generativa facilita un acceso desigual a la información y el conocimiento

Fuente. elaboración propia con base en Kay et al. (2024).

De acuerdo con lo anterior, la injusticia algorítmico-epistémica no es una mera conjetura teórica, sino que es un problema real cuya relevancia se ha intensificado a raíz de los avances recientes en IA; de ahí que es posible traer a colación casos empíricos específicos que dan cuenta de la necesidad de un abordaje interdisciplinario.

A continuación, se presenta cada una de las configuraciones antes mencionadas por Kay et al. (2024), con el fin de aclarar mejor el fenómeno aquí abordado.

La injusticia testimonial ampliada se concreta cuando los sistemas de IA reproducen los sesgos existentes en los datos con los que son entrenados, lo que implica que se perpetúen formas de injusticia testimonial al desacreditarse

sistemáticamente con ello los grupos marginados. Esta situación se agrava por la desigual representación en los conjuntos de datos, lo que privilegia las voces dominantes y silencia las perspectivas de los más vulnerables, dando lugar a un ciclo que refuerza las narrativas hegemónicas y profundiza las inequidades epistémicas dentro de un grupo poblacional.

La injusticia testimonial manipulativa implica la manipulación consciente de la información que alimenta el sistema de IA. Esto significa que la injusticia epistémica no solo surge de prejuicios inconscientes, sino también de la manipulación deliberada de los datos que la soportan o de los cuales se sirve el algoritmo, afectando grupos marginados. En este contexto, dicho tipo de injusticia se manifiesta cuando se usan sistemas de IA para desacreditar a estos grupos. Tal es el caso del uso de los *deepfakes* para sembrar dudas sobre testimonios auténticos, lo cual termina perpetuando el silenciamiento y la exclusión desde el punto de vista epistémico (Kay et al., 2024).

La ignorancia hermenéutica generativa, por su parte, se presenta cuando los sistemas de inteligencia artificial terminan representando de forma errónea o excluyendo a grupos marginados como consecuencia de una comprensión limitada del contexto cultural o experiencial que rodea a dichos grupos. Este fenómeno implica no solo deficiencias en los datos de entrenamiento, sino también en los mecanismos de supervisión y ajuste fino, que tienen el riesgo de pasar por alto las voces sobrerrepresentadas (Kay et al., 2024).

A diferencia de la injusticia hermenéutica tradicional (que surge de la falta de recursos conceptuales compartidos entre humanos), la ignorancia hermenéutica generativa es propia de las limitaciones tecnológicas de la IA, incapaz de incorporar vivencias individuales y saberes culturales. Esto la convierte en una forma única de malinterpretación, con el potencial de generar nuevas distorsiones más allá de los sesgos presentes en los datos.

Por último, la ignorancia de acceso hermenéutico es una forma específica de injusticia hermenéutica que ocurre en sistemas de IA generativa y que parte de la idea de Fricker (2017), según la cual la injusticia hermenéutica surge cuando las personas no pueden interpretar o expresar sus experiencias por falta de recursos conceptuales. En el contexto de la inteligencia artificial, este fenómeno se manifiesta en el control sesgado del acceso a la información, afectando especialmente a grupos marginados. Un ejemplo documentado por Kay et al. (2024) es el caso de las dificultades de los usuarios afroamericanos con las tecnologías de reconocimiento de voz, con lo que se concretan barreras al acceso equitativo al conocimiento, aislamiento informativo y fragmentación epistémica, lo que intensifica las desigualdades (Kay et al., 2024).

5. Injusticia algorítmico-epistémica en el ámbito judicial

Fricker (2017) establece una distinción fundamental entre dos formas de injusticia epistémica: la testimonial y la hermenéutica. Por un lado, la injusticia testimonial se manifiesta cuando el testimonio de un individuo es desacreditado injustamente debido a prejuicios arraigados en su identidad social. Esta forma de injusticia revela cómo los sesgos sociales pueden erosionar la credibilidad de quienes pertenecen a grupos marginados.

Por otro lado, la injusticia hermenéutica emerge cuando existe una brecha entre las vivencias personales y los marcos interpretativos disponibles en el imaginario colectivo. El término «hermenéutica» alude entonces al proceso de interpretación del conocimiento, y en este contexto se refiere a los recursos conceptuales compartidos que permiten comprender y comunicar experiencias. Cuando dichos recursos son insuficientes o están sesgados por la perspectiva de los grupos dominantes, las experiencias de ciertos individuos o comunidades quedan malinterpretadas o, peor aún, invisibilizadas.

En este punto cabe resaltar que la IA también pueden perpetuar las dinámicas de injusticia epistémica, en tanto sus algoritmos configuran y condicionan la producción de conocimiento (Kay et al., 2024). Por esto se reitera aquí la posibilidad latente de que se presenten riesgos de injusticia algorítmico-epistémica.

En el campo emergente que explora la confluencia entre la injusticia epistémica y los sistemas algorítmicos de IA ha sido acuñada la noción «injusticia algorítmico-epistémica», término que fue usado por primera vez por Byrnes y Spears (2023) y que ha sido profundizado en campos como la medicina en investigaciones relacionadas con la atención de pacientes con discapacidad auditiva (Castrillón García, 2025) y en el ámbito jurídico con el fin de situar un entramado conceptual más amplio para su abordaje sistemático (Kay et al., 2024).

El fenómeno de la injusticia epistémica se ha identificado en sus dos manifestaciones (injusticia testimonial e injusticia hermenéutica), lo cual también implica variaciones al concepto de injusticia algorítmico-epistémica. En primer lugar, la injusticia testimonial se manifiesta cuando los algoritmos son investidos de una autoridad epistémica superior a la de los agentes humanos, lo cual puede ampliar y perpetuar sesgos sociales preexistentes. En segundo lugar, las formas de injusticia hermenéutica emergen cuando los algoritmos producen de manera autónoma marcos de sentido e interpretaciones, especialmente en entornos automatizados que operan sin mediación humana directa, lo que plantea desafíos ontológicos y prácticos en torno a su comprensión y su administración (Kay et al., 2024).

Sirve para ilustrar lo anterior el estudio de Symons y Alvarado (2022), quienes examinaron un caso paradigmático en el que una persona fue injustamente privada de la libertad condicional como resultado de las evaluaciones realizadas por el algoritmo de predicción de reincidencia Compas (*Correctional Offender*

Management Profiling for Alternative Sanctions). Se destaca en dicho estudio que a pesar de que el individuo presentó evidencia que respaldaba su solicitud ante autoridades humanas, su situación no fue reconsiderada (Wexler, 2018, citado por Kay et al., 2024). Esto indica la existencia de dificultades epistémicas enmarcadas en la falta de transparencia inherente a la IA que limita gravemente la posibilidad de impugnar las decisiones erróneas, contribuyendo a la producción de injusticia algorítmico-epistémica.

6. Desafíos probatorios de la injusticia testimonial y la injusticia algorítmica en el ámbito judicial

Con el fin de establecer los aspectos clave a ser considerados para la construcción de lineamientos de justicia algorítmico-epistémica, es necesario profundizar en los alcances y desafíos del empleo de la IA en el escenario probatorio.

6.1. El empleo de la IA en la valoración de la prueba pericial

En el ámbito de la prueba pericial, la IA ofrece múltiples aplicaciones. Esta prueba permite que el juez acceda a conocimientos especializados ajenos a su formación jurídica, a través del perito. El juez evalúa críticamente los resultados del peritaje, considerando el perfil del experto y los criterios sustantivos, como los criterios *Daubert*. Estos criterios, originados en la jurisprudencia estadounidense, han influido en normas como el artículo 702 de las *Federal Rules of Evidence* y también en la doctrina y jurisprudencia española, más allá de lo establecido en la Ley de Enjuiciamiento Civil (Castellano, 2021). Con respecto a la evaluación del perfil del perito es claro que los sistemas de IA pueden estimar el nivel de experiencia y examinar las contribuciones previas de un experto para valorar su competencia profesional (Nieva Fenoll, 2018).

No obstante, la automatización de los criterios *Daubert* plantea desafíos significativos, especialmente en lo que respecta a su implementación mediante inteligencia artificial. En este contexto, se propone centrar la automatización en los cinco elementos centrales de los criterios *Daubert*: validez científica, revisión por pares, tasa de error conocida, estándares de control y aceptación general en la comunidad científica. Estos aspectos son susceptibles de ser integrados en sistemas expertos diseñados específicamente, como lo sugiere Nieva Fenoll (2018) y con base en ellos es posible configurar un algoritmo útil para evaluar el perfil del perito (Castellano, 2021).

Se resalta que en todos los casos los sistemas de IA para la valoración probatoria se proponen como apoyo a la labor del juez y en ningún modo deben sustituirlo (Castellano, 2021; Nieva Fenoll, 2028). Y esto tiene sentido, pues en el ámbito judicial los sistemas de IA implican una «caja negra» para el juez. No obstante, aunque se piense en una solución como lo sería la creación de un organismo público que

asegure la implementación y el cumplimiento de los principios de publicidad, transparencia, trazabilidad, auditabilidad y aplicabilidad del código algorítmico, esto no aseguraría la eliminación de sesgos, de ahí que este es otro problema que se debe enfrentar. Sin embargo, se destaca la importancia de que al interior de la Rama Judicial se pudiera establecer un organismo de control que garantice la aplicación efectiva de dichos principios, lo que implicaría incluir a los actores del sistema judicial desde el diseño hasta la evaluación y la mejora continua del sistema asegurando una participación activa.

En Colombia, lo anterior va de la mano de los progresos que a nivel jurisprudencial se han tenido en los últimos años en materia de reconocimiento del derecho a la explicación de las decisiones algorítmicas. Tal es el avance de la Corte Constitucional frente al derecho a conocer los algoritmos usados en decisiones administrativas, que se traduce en un derecho a la explicación de las decisiones tomadas en el ámbito público con base en algoritmos (Corte Constitucional de Colombia, Sentencia T-067, 2025).

6.2. El empleo de la IA en la valoración de la prueba testimonial

Si bien los jueces valoran la credibilidad de los testimonios basándose en el contexto y su experiencia, se debe considerar que la memoria humana es limitada y carece de precisión, lo que conduce a que los análisis estadísticos realizados por seres humanos frecuentemente presenten sesgos de confirmación o incluso de discriminación. Frente a esto, cabe resaltar que las variables que inciden en la credibilidad de los testimonios poseen un grado de objetividad parcial (Nieva Fenoll, 2018), por lo que es posible sistematizarlas y traducirlas a lenguaje de programación. De hecho, ya se han tenido experiencias con programas que pretenden evaluar la credibilidad de los testigos, como el caso de *Advokate*, con el cual se evalúa la credibilidad de los testigos mediante factores objetivos como la distancia y las condiciones de observación, el vínculo previo con la persona observada y las características personales del testigo (edad, educación, estado mental), la precisión del relato y el lapso entre los hechos y la declaración. Lo anterior ejemplifica los criterios que pudieran ofrecer una base metodológica para valorar la fiabilidad del testimonio en el ámbito judicial (Bromby & Hall, 2002).

Las variables mencionadas pueden ser implementadas mediante algoritmos que, apoyados en bases estadísticas, perfeccionan la intuición humana y permiten una aproximación más precisa a los hechos. Según Nieva Fenoll (2018), estas fórmulas deben considerar la coherencia interna de la declaración, la capacidad del interrogado para contextualizar, la corroboración con otras pruebas y la detección de comentarios adicionales oportunistas o adicionales que encubran la intención de influir en la percepción del juzgador.

Es así como la inteligencia artificial pudiera ser una herramienta de gran utilidad para abordar las corroboraciones, pues supera las limitaciones humanas

para verificar rápidamente la compatibilidad entre las declaraciones y las pruebas (Castellano, 2021).

Por último, respecto a la prueba testimonial, Nieva Fenoll (2018) propone que la IA puede utilizarse para detectar y eliminar preguntas no neutrales en los interrogatorios, es decir, aquellas que inducen al declarante a aceptar premisas implícitas, lo cual contribuiría a preservar la objetividad e imparcialidad del testimonio.

7. Inteligencia artificial y desnaturalización de las actuaciones encaminadas a generar y controvertir la prueba en el proceso judicial

El empleo de herramientas tecnológicas de investigación ha generado transformaciones en el derecho penal, pasando de ser reactivo a adoptar un enfoque preventivo. Esta evolución de la mano de la implementación de sistemas de IA en la justicia predictiva difumina las fronteras jurídicas tradicionales, plantea riesgos para los derechos fundamentales y genera una regresión hacia un derecho penal simbólico centrado en el control y la seguridad, pues acentúa la aceptación social de mayor vigilancia a costa de las libertades individuales.

Así, las tecnologías asociadas a la investigación afectan no solo a los involucrados directos, sino también a terceros sin relación con los hechos delictivos, debido al uso de tecnologías que permiten una vigilancia preventiva por parte de la policía. Un ejemplo que ilustra lo anterior es el impacto de Compas, herramienta aludida con antelación y empleada para calcular la probabilidad de que una persona cometa delitos en el futuro. Con respecto al empleo de Compas cabe traer a colación aquí el caso controversial de Eric Loomis, a quien en 2013 se le impuso una pena más severa basada en esta herramienta algorítmica que predecía una alta probabilidad de reincidencia. Se destaca en este caso el hecho de que al apelar la decisión se solicitó acceso al funcionamiento del algoritmo para ejercer el derecho de defensa, pero la empresa lo negó por razones de propiedad intelectual (Barona, 2024). En este caso se evidencia, por tanto, el conflicto entre el derecho al debido proceso y la protección de los derechos de autor en el ámbito judicial, lo cual implica sobreponer intereses privados por encima de principios procesales que además tienen naturaleza de derechos fundamentales, poniendo de relieve una desnaturalización de la actuación judicial.

A lo anterior se suma el hecho de que la incorporación de sistemas de IA en la generación de fuentes de prueba, entendidas como los informes o documentos producidos mediante sistemas algorítmicos, transforman su naturaleza al integrarse como medios probatorios dentro del proceso. Sin mencionar que la algoritmización de las fuentes de prueba y su integración al acervo probatorio terminan alterando los principios y condicionando las decisiones judiciales, de tal modo que «convergen una serie de instrumentos que inciden en la función que desempeñan los jueces, a saber, en la *Judge Craft*, esto es, tanto en la *Judicial Decision* como en su proceso de elaboración a través de la constatación probatoria» (Barona Vilar, 2024, p.108).

8. Dificultades probatorias de la injusticia algorítmico-epistémica en el ámbito jurídico

Las dificultades para evidenciar los sesgos algorítmicos que conducen al fenómeno de la injusticia algorítmica radican en lo difícil que resulta comprender o interpretar el algoritmo, pues en el ámbito judicial los sistemas de IA constituyen una «caja negra» para el juez y el operador jurídico en general, más aún cuando se trata de sistemas basados en aprendizaje autónomo, pues en tales casos no es posible garantizar la transparencia y publicidad del código algorítmico, el cual en dichos sistemas es cambiante. De ahí que no sería posible, por ejemplo, en el ámbito penal, que un imputado o acusado con base en una decisión apoyada en IA se pueda defender por el desconocimiento de las reglas lógicas que son difusas por la opacidad del algoritmo.

Adicionalmente, el hecho de que para el juez un sistema de inteligencia artificial permanezca como un enigma o una caja negra, porque sus procesos algorítmicos no son entendidos, significa que a su vez el sistema no puede ser supervisado por el juez para poder intervenir sobre los sesgos que puede tener el resultado que entrega el algoritmo. Esto si se considera que un sistema de IA se fundamenta en una arquitectura algorítmica de elevada complejidad, concebida y desarrollada por seres humanos, ya sea un programador individual, un diseñador especializado en IA o un equipo interdisciplinario conformado para dicho propósito, y aunque estos actores puedan contar con el apoyo de otras herramientas automatizadas, incluida la propia IA, el acto de programar sigue siendo esencialmente una actividad humana. Por ello, no está exenta de la posibilidad de incorporar sesgos cognitivos, culturales o biológicos propios de los diseñadores del sistema de IA.

Si bien ante lo anterior se pueden proponer soluciones encaminadas a disminuir y controlar los sesgos en la programación, el diseño o el entrenamiento del algoritmo y, así mismo, dirigidas a asegurar la aplicabilidad y transparencia del algoritmo, el reto mayor tiene que ver con la injusticia testimonial y su dificultad para probarse en el ámbito jurídico.

Una de las manifestaciones más frecuentes y menos analizadas de la parcialidad subjetiva en el ámbito judicial es la reducción de la confianza que se le otorga a un testigo, motivada por un prejuicio de identidad implícito por parte del juez. Este fenómeno ha sido abordado en el campo de la epistemología social con el concepto denominado aquí «injusticia testimonial». Aspectos como la raza, el género, la orientación sexual, la condición de discapacidad o de desplazamiento de un testigo pueden generar una disminución en la credibilidad que se le otorga, con ocasión de prejuicios que son inherentes a la persona que recibe el testimonio.

No obstante, aunque se trata de un fenómeno real en su ocurrencia, constituye una de las manifestaciones de parcialidad más complejas de probar, debido a que la mayoría de los sistemas judiciales suelen operar según el principio de presunción de imparcialidad, salvo prueba en contrario. Esto significa que a menos

que en la decisión judicial o en la transcripción de un interrogatorio se evidencie de manera explícita la utilización indebida de estereotipos o prejuicios, resulta muy complejo establecer de forma concluyente la existencia de una parcialidad subjetiva en relación con la valoración de la credibilidad de un testigo (Páez & Arcila-Valenzuela, 2023).

Cabe aclarar que el propósito de reconocer casos de injusticia testimonial en el contexto de la administración de justicia no radica en endilgarle responsabilidad disciplinaria a los operadores judiciales, sino más bien en formular estrategias que permitan prevenir y corregir este tipo de injusticia, tanto en el plano individual como institucional. En este sentido, se reitera que, según Fricker (2017), una persona es víctima de injusticia testimonial cuando su credibilidad es menospreciada como consecuencia de un prejuicio identitario negativo por parte de quien recibe y valora su testimonio. Esta noción como tal resalta el impacto epistémico adverso derivado del sesgo del oyente sin discernir de manera explícita que la disminución de credibilidad del hablante es producto de una acción intencional o consciente.

De acuerdo con lo anterior, una acción intencional o consciente encaminada a cuestionar la veracidad del testimonio del individuo o debilitar su confianza no implica necesariamente un error en la evaluación del estatus epistémico del interlocutor. Esta situación que es frecuentemente identificable en el ámbito judicial, y en especial en las dinámicas entre fiscales y testigos, está revestida de intencionalidad por generar un impacto psicológico para la valoración del testimonio (Páez & Arcila-Valenzuela, 2023).

Por el contrario, la ausencia de una intención consciente de desacreditar al testigo da lugar a una forma de injusticia testimonial que se manifiesta como consecuencia de una evaluación errónea y no deliberada de la competencia epistémica del hablante. Esta variante de injusticia marcada por su carácter inadvertido introduce un reto epistemológico significativo distinto a la modalidad previamente analizada, pues el error de juicio no proviene de una acción voluntaria, consciente o intencional.

Para establecer de manera fundamentada la existencia de un caso de injusticia testimonial, es necesario establecer tres elementos clave: en primer lugar, identificar si el oyente, en este caso el juez, presenta un prejuicio de identidad negativo frente al hablante, independientemente de si es consciente o no de ello; en segundo lugar, establecer si dicho sesgo ha operado como causa directa del menoscabo de credibilidad otorgada al testigo; y, por último, constatar si efectivamente se produjo una merma injustificada en la credibilidad del testimonio (Páez & Arcila-Valenzuela, 2023).

De acuerdo con lo anterior, identificar casos individuales de injusticia testimonial es todo un reto, más aún cuando las dinámicas de exclusión y discriminación que se pueden evidenciar en el ámbito judicial no suelen derivarse de la adopción deliberada de sesgos por parte de los operadores jurídicos. Por el contrario, suelen

manifestarse como consecuencias de prejuicios identitarios implícitos y factores contextuales que, de manera inadvertida, inciden en los procesos de decisión y valoración dentro del ámbito jurídico (Páez y Arcila-Valenzuela, 2023).

A continuación, se presentan algunos ejemplos de este fenómeno. El primero se documenta en una investigación realizada por Tuerkheimer (2017), quien identificó una tendencia sistemática en el sistema legal estadounidense a desacreditar injustificadamente los testimonios de víctimas de violencia sexual, fenómeno que cabe dentro de la denominación de «injusticia testimonial». Adicionalmente, Tuerkheimer (2017) señala que este fenómeno se acentúa con el requerimiento de corroboración, de acuerdo con el cual se exigen pruebas independientes al testimonio de la víctima con el supuesto de que las denuncias pueden ser malintencionadas o no ser sinceras. A pesar de su carácter discriminatorio, este requisito había sido históricamente defendido por las cortes, como se evidencia en el informe del Comité de Revisión de Derecho Penal de Inglaterra y Gales (1972), que le recomendaba al jurado tratar con especial cautela dichos testimonios de víctimas de violencia sexual, atribuyendo como posibles motivaciones de las denunciadas los celos, las fantasías, los deseos de venganza o la vergüenza de haber consentido un acto sexual (Ho, 2013, p.254, citado por Páez y Arcila-Valenzuela, 2023).

En el contexto colombiano, la Jurisdicción Especial para la Paz (JEP) ha adoptado el concepto de injusticia epistémica en su protocolo de comunicación con víctimas de violencia sexual (JEP Unidad de Investigación y Acusación, 2018), con el propósito de contrarrestar los efectos de la revictimización y el estigma que históricamente han recaído sobre mujeres y personas LGTBI en el marco del conflicto armado interno. Este enfoque busca transformar las prácticas institucionales que han reproducido formas de exclusión y deslegitimación de los testimonios de estas poblaciones en Colombia (Páez y Arcila-Valenzuela, 2023).

9. Lineamientos para minimizar el riesgo de injusticia algorítmico-epistémica en la administración de justicia

La injusticia testimonial en el contexto de la toma de decisiones mediadas por el uso de sistemas de IA se fundamenta en el cumplimiento de cinco condiciones esenciales que se pueden considerar al momento de evaluar el testimonio de una persona que hace parte de un grupo poblacional históricamente excluido o discriminado. Tales condiciones incluyen el prejuicio identitario, que se manifiesta en una predisposición a desestimar la credibilidad del testimonio de la persona en función de sus características personales; la condición de desventaja, que evidencia la posición desigual del paciente en su capacidad de comunicar eficazmente sus experiencias; la condición de parte interesada, que subraya la afectación directa o indirecta de la persona por las decisiones tomadas; la condición epistémica, que reconoce la relevancia del conocimiento experiencial de la persona en el proceso de toma de decisiones; y la condición de justicia social, que vincula la

discriminación epistémica con las desigualdades estructurales y sociales que históricamente han afectado al grupo poblacional al cual pertenece la persona. La violación conjunta de estas condiciones da lugar a una situación de injusticia epistémica, y tal como lo han señalado Fricker (2017) y Byskov (2021), ninguna de las condiciones mencionadas resulta suficiente por sí sola para determinar que una discriminación epistémica constituya una injusticia.

Desde una perspectiva técnica, resulta preocupante la limitada capacidad del servidor judicial para detectar la presencia de sesgos en los sistemas de IA aplicados a la administración de justicia. Esta dificultad compromete tanto la posibilidad de implementar mecanismos de corrección como la fiabilidad de las decisiones fundamentadas en los resultados generados por dichos sistemas. En consecuencia, se torna necesario establecer un principio de precaución que regule el uso de la IA en el ámbito judicial. Esto implica un enfoque que exige mantener una actitud de escucha activa y deliberativa frente a la persona que rinde un testimonio, reafirmando el papel central del juicio humano en la toma de decisiones. Tal como lo plantea Fricker (2017), esta empatía se manifiesta en la capacidad del juez para convertirse en un «oyente virtuoso», es decir, en alguien que ha logrado revisar críticamente sus propios prejuicios y ajustar sus esquemas interpretativos a partir de la experiencia acumulada. Esta virtud epistémica exige del juez una vigilancia constante sobre sus propios sesgos, así como una disposición reflexiva que equilibre la espontaneidad con la deliberación, en aras de una comunicación respetuosa y eficaz.

La injusticia epistémica, desde su concepción originaria, representa un reto ético que exige ser enfrentado. Respecto a dicha injusticia, Fricker (2017) propone que el cultivo de virtudes epistémicas puede orientar la construcción de una cultura más justa en términos del conocimiento. En este proceso, los grupos con poder social tienen una responsabilidad crucial: deben fomentar la justicia testimonial mediante la reflexión crítica sobre sus propios prejuicios y promover la justicia hermenéutica mediante una apertura sensible hacia las voces marginadas. Una salida similar tendríamos que entretener para el ámbito de la administración de justicia a partir del rol del juez en el proceso. Sin embargo, resulta más complejo su abordaje cuando se trata del fenómeno de la injusticia algorítmico-epistémica, dada la necesidad de su abordaje interdisciplinario y multidimensional, lo que a fin de cuentas recae en el mismo problema que denota la teoría de la injusticia epistémica.

Tal como se ha indicado con antelación, la injusticia testimonial en el ámbito jurídico surge cuando los prejuicios identitarios por parte del funcionario judicial afectan negativamente la evaluación de la credibilidad del testigo. Frente a esto Fricker (2017) argumenta que los prejuicios identitarios negativos son epistémicamente censurables debido a su persistencia frente a la evidencia que es contraria. Estos prejuicios se consolidan tanto en el plano epistémico como emocional, lo que explica su durabilidad o persistencia en el tiempo; de ahí la solución propuesta por Fricker (2017) de mitigar la injusticia epistémica mediante el

cultivo de virtudes individuales por parte del oyente. En consecuencia, de acuerdo con Fricker (2017) y Páez y Arcila-Valenzuela (2023), los servidores judiciales pueden desarrollar virtudes epistémicas para mejorar su capacidad de evaluar testimonios de forma justa y precisa, cumpliendo con sus responsabilidades epistémicas en el ámbito judicial.

No es posible establecer un grado exacto y universal de credibilidad para un hablante, ya que depende de la percepción del oyente y de la evidencia disponible. Factores no epistémicos — como prejuicios, aversión al riesgo o influencias culturales— afectan esta valoración, por lo que la credibilidad siempre está mediada por contextos personales, sociales e institucionales, lo que impide una evaluación neutral y objetiva del testimonio (Páez y Arcila-Valenzuela, 2023).

La finalidad práctica de identificar situaciones de injusticia testimonial en el ámbito judicial consiste en verificar la existencia de una parcialidad judicial subjetiva y reconocer la presencia de sesgos en el juicio del oyente, de tal modo que a partir de allí se puedan aplicar soluciones como la propuesta por Fricker (2017) basada en la teoría de la virtud, que implica que el oyente sea capaz de identificar los episodios en los que ha incurrido en un injusticia testimonial, con el propósito de corregir su conducta.

Aplicando la fórmula de Fricker (2017) en el ámbito judicial, para evitar la injusticia testimonial, los servidores judiciales deben desarrollar una disposición virtuosa que contrarreste los prejuicios al evaluar la credibilidad de los testigos. Esta virtud busca una evaluación imparcial de la evidencia no solo mediante una reflexión deliberada, sino también a través de una sensibilidad testimonial, es decir, con una capacidad de juzgar espontáneamente con justicia. Esta sensibilidad se cultiva, según Fricker (2017), mediante una virtud correctiva, que implica un proceso constante de automonitoreo y ajuste de juicios por parte del oyente. Sin embargo, dicho proceso se ve obstaculizado por el fenómeno conocido como «sesgo del punto ciego» (Pronin et al., citado por Páez y Arcila-Valenzuela, 2023). Este sesgo se refiere a la tendencia de las personas a reconocer los sesgos cognitivos y motivaciones más fácilmente en los demás que en sí mismas. Esto revela una limitada capacidad para evaluar la imparcialidad propia, ya que los mismos sesgos que se intentan corregir interfieren en ese proceso.

A pesar de la relevancia del enfoque de justicia epistémica propuesto por Fricker (2017), diversas críticas han señalado sus limitaciones desde una perspectiva sociopolítica. Particularmente, se argumenta que la atención centrada en los prejuicios implícitos puede desviar el análisis de las estructuras sociales y los sesgos explícitos que perpetúan las injusticias testimoniales (Alcoff, 2010). Además, es cuestionable la estabilidad de las disposiciones virtuosas individuales, al demostrar que las variaciones mínimas en el contexto pueden inducir comportamientos contrarios a dichas disposiciones (Doris, 2002). No obstante, es preciso resaltar que Fricker (2017) no considera que las virtudes personales constituyan una solución exhaustiva frente a la injusticia testimonial. Si bien reconoce su relevancia, también

sostiene que estas virtudes representan únicamente un componente dentro de una respuesta más amplia.

De acuerdo con lo anterior, cabe resaltar las múltiples propuestas que se han formulado y que merecerían ser combinadas de manera complementaria para garantizar la minimización de riesgos de injusticia algorítmica asociadas al fenómeno aquí en estudio:

Una primera medida es la representación fidedigna de las comunidades históricamente discriminadas a través de modelos incluyentes con enfoques de género y diferenciales, así como con esquemas de compensación ajustados a las particularidades de los datos que alimentan los sistemas de IA. En segundo lugar, los desarrolladores de inteligencia artificial deben analizar críticamente los sesgos en los datos de entrenamiento, considerando que la injusticia hermenéutica aumenta cuando los grupos dominantes controlan la definición de los valores que guían la respuesta de estos sistemas de IA. Por último, se debe resaltar que la falta de transparencia en los sistemas de inteligencia artificial contribuye a la injusticia epistémica al impedir que los grupos afectados comprendan y cuestionen las decisiones injustas. De ahí que, para contrarrestar esta situación, se debe fomentar la transparencia mediante interfases accesibles y el avance tecnológico encaminado a que el usuario de los sistemas de IA pueda comprender los procesos internos del modelo algorítmico empleado.

Es prioritario promover la adopción de algoritmos cuya estructura y lógica operativa sean comprensibles y transparentes. Esta orientación exige privilegiar modelos que posibiliten el análisis crítico frente a otros sistemas de naturaleza opaca, como las redes neuronales profundas, particularmente en escenarios sensibles como el judicial. En aquellos casos en los que se recurra a modelos de mayor complejidad, se vuelve imperativo incorporar mecanismos explicativos que permitan desentrañar los factores que inciden en los resultados, garantizando así una interpretación responsable de los resultados (Chaverra Mena, 2023).

Es menester establecer consideraciones legales encaminadas a garantizar el debido proceso y, en especial, el derecho de defensa y contradicción en el escenario de la actividad probatoria en el proceso judicial. Es así como los sistemas jurídicos deben establecer normas que regulen la viabilidad técnica y las condiciones para integrar herramientas tecnológicas en los procesos judiciales, garantizando mecanismos que permitan verificar los resultados de los sistemas de IA empleados y proteger el derecho de defensa de las partes.

Adicionalmente, sería importante excluir como prueba los resultados de sistemas algorítmicos cuando son la única base de una condena, para proteger la presunción de inocencia (Barona Vilar, 2024), si se considera que el uso de estos sistemas puede inducir una presunción de culpabilidad y una alteración de la carga probatoria, afectando el principio de *in dubio pro reo*. Por ello, se requiere

una regulación procesal que mitigue estos riesgos y garantice decisiones judiciales debidamente motivadas y transparentes.

Por último, es importante reconocer que la incorporación de sistemas inteligentes en el ámbito judicial no debe entenderse como un fin en sí mismo, sino como medios técnicos de carácter instrumental, siempre supeditados a los principios rectores de la justicia y al cumplimiento riguroso del ordenamiento jurídico garantizando la tutela judicial efectiva. En este orden de ideas, la función auxiliar o de apoyo que estas tecnologías pueden desempeñar en la resolución de controversias adquiere especial relevancia, particularmente en lo que concierne a su utilización por parte de los órganos judiciales como apoyo en la fundamentación de sus decisiones (San Miguel Caso, 2023). Por lo anterior, resulta de vital importancia reforzar la exigencia de la motivación de las decisiones judiciales en aras de brindar transparencia en la mismas y permitir un control de la decisión judicial y, por tanto, garantizar el derecho de contradicción y defensa, de modo que se salvaguarden las garantías procesales al implementar el uso de la IA en la administración de justicia. Las estrategias institucionales para prevenir este fenómeno y mitigar los efectos de los sesgos y prejuicios implican transformaciones que requieren de la agencia de sujetos virtuosos capaces de impulsarlas. En este caso sentido, es fundamental el quehacer de los servidores judiciales, independientemente de que al interior de los ordenamientos jurídicos existan suficientes criterios de corrección y control de la decisión judicial, pues en el campo judicial es bastante complejo identificar casos individuales de injusticia testimonial, aunque sea innegable que existen grupos poblacionales histórica y sistemáticamente discriminados y que, como consecuencia, han tenido efectos negativos sobre su tutela judicial efectiva.

10. Conclusiones

A lo largo de este artículo se establecen unas bases para erigir el concepto de justicia algorítmico-epistémica como un principio ético cardinal que debe ser considerado de manera rigurosa no solo en el diseño de tecnologías que inciden en las estructuras de poder en la sociedad, sino también en la implementación de los sistemas de IA en campos tan sensibles y de amplio impacto como la administración de justicia. Desde el marco teórico aquí abordado se propone en la labor del juez la reflexión sobre la incorporación estratégica de virtudes epistémicas, especialmente en aquellos escenarios vinculados a la valoración probatoria, ámbitos en los cuales la estimación del acervo probatorio en su conjunto puede implicar no solo resultados predictivos con empleo de la IA, sino diversos tipos de pruebas como la testimonial, que pudieren distar del resultado que arroja el sistema de IA con incidencia en la decisión judicial. Todo esto con el fin de enfrentar y atenuar las múltiples manifestaciones de injusticia epistémica y algorítmico-epistémica en el ámbito de la administración de justicia.

Con respecto a la regulación de la IA en el ámbito público, es importante garantizar la transparencia del algoritmo de modo que se le permita a la ciudadanía

conocer los ámbitos específicos en los cuales se están tomando decisiones fundamentadas total o parcialmente en procesos algorítmicos; igualmente, es fundamental establecer el impacto de dichas decisiones algorítmicas sobre los derechos y libertades de los ciudadanos y los criterios de procedimiento con los cuales la administración de justicia o las instituciones públicas adoptan sus decisiones.

Adicionalmente al aspecto regulatorio, es posible emplear la misma IA para la eliminación de sesgos algorítmicos, tal como ocurre con la técnica de «reducción de sesgos», la cual consiste en intervenir los datos de entrenamiento o la estructura misma de los modelos con el propósito de mitigar distorsiones sistemáticas. A esto se suman metodologías como la selección de características relevantes y la aplicación de técnicas de regularización. En conjunto, estas prácticas contribuyen a mejorar la precisión de los algoritmos y promueven una mayor equidad en sus resultados, lo que puede incidir directamente en la disminución de fenómenos de discriminación algorítmica y, por ende, de los casos de injusticia algorítmica. No obstante, resulta fundamental reconocer que estas técnicas por sí solas no constituyen una solución integral al desafío que representa la discriminación algorítmica. Por esto, su implementación debe ir acompañada de estrategias complementarias que impliquen la promoción de la transparencia y de la responsabilidad en el diseño y la aplicación de sistemas de IA, además de la implementación de la evaluación de la calidad de los algoritmos que incluya auditorías y estudios de impacto.

Referencias

Agudelo Londoño, S. (2021). Inteligencia artificial en el proceso penal: análisis a la luz del Fiscal Watson. *Pensamiento Jurídico*, 1(54), 147-164. <https://revistas.unal.edu.co/index.php/peju/article/%20view/96091>

Alcoff, L. M. (2010). Epistemic identities. *Episteme*, 7(2), 128-137. <https://doi.org/10.3366/epi.2010.0003>

Arévalo Ortiz, A. V. (2024). *Sesgos cognitivos e injusticia testimonial en la inteligencia artificial* [tesis de grado, Universidad El Bosque]. Repositorio Unbosque. <https://hdl.handle.net/20.500.12495/13082>

Barona Vilar, S. (ed.) (2021). *Justicia algorítmica y neuroderecho. Una mirada multidisciplinar*. Editorial Tirant lo Blanch.

Barona Vilar, S. (2024). Justicia con algoritmos e inteligencia artificial, ¿acuerpando garantías y derechos procesales o liquidándolos? *Derechos y libertades*, 51(2), 83-116. <https://hdl.handle.net/10550/104058>

Bromby, M. C., & Hall, M. J. (2002). The Development and Rapid Evolution of the Knowledge Model of ADVOKATE: An Advisory System to Assess the Credibility of Eyewitness Testimony. En T. J. M. Bench-Capon, A. Daskalopulu, & R. G. F. Winkels (eds.), *Legal Knowledge and Information Systems. The Fifteenth Annual Conference* (pp. 143-152). IOS Press.

Byskov, M. F. (2021). What makes epistemic injustice an “injustice”? *Journal of Social Philosophy*, 52(1), 114-131. <https://doi.org/10.1111/josp.12348>

Castellano, P. S. (2021). Inteligencia Artificial y valoración de la prueba: las garantías Jurídico-Constitucionales del órgano de control. *THEMIS Revista de Derecho*, (79), 283-297. <https://doi.org/10.18800/themis.202101.015>

Castrillón García, E. D. (2023). Lineamientos éticos para la implementación de la inteligencia artificial en la decisión judicial: la política de modernización de la justicia civil y los retos éticos frente al riesgo de injusticia algorítmica. En M. M. Bustamante Rúa, A. del P. Henao Ochoa, & D. M. Ramírez Carvajal (coords.), *La justicia en la era de la revolución tecnológica* (pp. 206-228). Fondo Editorial Institución Universitaria de Envigado. https://www.iue.edu.co/fondo_editorial/la-justicia-en-la-era-de-la-revolucion-tecnologica/

Castrillón García, E. D. (2025). *Injusticia Algorítmico-Epistémica. Fenomenología ética del testimonio en la discapacidad auditiva frente a la implementación de IA*. Fondo Editorial Red para el Estudio del Proceso y la Justicia. https://www.procesalyjusticia.org/_files/ugd/0e0037_953aeae9678c4a3da591060db0d01cbb.pdf

Chaverra Mena, Y. E. (2025). *El Sesgo Algorítmico en Sistemas Judiciales y su Impacto en Víctimas de Género* [paper]. SSRN. <http://dx.doi.org/10.2139/ssrn.5339543>

Corrales Aznar, J. M. (2022). Derecho al trabajo y discriminación: Las políticas públicas de empleo como instrumentos de igualdad. En E. García-Antón Palacios (dir.), *Los derechos humanos en la inteligencia artificial: su integración en los ODS de Agenda 2030* (pp. 399-422). Thomson Reuters Aranzadi.

Corte Constitucional de Colombia. (2025, 26 de febrero). *Sentencia T-067 de 2025* [M. P.: Cabo, N. Á.]. <https://www.corteconstitucional.gov.co/relatoria/2025/t-067-25.htm>

Doris, J. M. (2002). *Lack of character: Personality and moral behavior*. Cambridge University Press.

Eguíluz Castañeira, J. A. (2020). Desafíos y retos que plantean las decisiones automatizadas y los perfilados para los derechos fundamentales. *Estudios de Deusto: Revista de Derecho Público*, 68(2), 325-367. [https://doi.org/10.18543/ed-68\(2\)-2020pp325-367](https://doi.org/10.18543/ed-68(2)-2020pp325-367)

Fricker, M. (2017). *Injusticia epistémica. El poder y la ética del conocimiento* (R. García Pérez, trad.). Herder.

Hao, K. (2021). Una IA descubre racismo involuntario en una prueba médica estándar. *MIT Technology Review*. <https://technologyreview.es/article/una-ia-descubre-racismo-involuntario-en-una-prueba-medica-estandar/>

JEP Unidad de Investigación y Acusación. (2018). *Protocolo de comunicación de la Unidad de Investigación y Acusación con víctimas de violencia sexual*. Justicia Especial para la Paz-Unidad de Investigación y Acusación.

Kay, J., Kasirzadeh, A., & Mohamed, S. (2024). Epistemic injustice in generative AI. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 7(1), 684-697. <https://doi.org/10.1609/aies.v7i1.31671>

Nieva Fenoll, J. (2018). *Inteligencia artificial y proceso judicial*. Marcial Pons.

Páez, A., & Arcila-Valenzuela, M. (2023). Los problemas probatorios de la injusticia testimonial en el derecho. *Isonomía-Revista de teoría y filosofía del Derecho*, (59). <https://doi.org/10.5347/isonomia.59/2023.653>

Reisman, D. (2022, 9 de agosto). Cómo el gobierno se apoya en algoritmos para designar los beneficios de la atención médica y por qué estas fórmulas secretas amenazan los derechos fundamentales de los pacientes. *ACLU New Jersey*. <https://www.aclu-nj.org/es/news/como-el-gobierno-se-apoya-en-algoritmos-para-designar-los-beneficios-de-la-atencion-medica-y>

San Miguel Caso, C. (2023). Inteligencia Artificial y algoritmos: la controvertida evolución de la tutela judicial efectiva en el proceso penal. *Estudios Penales y Criminológicos*, 44(Ext.), 1-23. <https://doi.org/10.15304/epc.44.8859>

Symons, J., & Alvarado, R. (2022). Epistemic injustice and data science technologies. *Synthese*, 200(2), 87. <https://doi.org/10.1007/s11229-022-03631-z>

Tuerkheimer, D. (2017). Incredible women: Sexual violence and the credibility discount. *Penal Law Review*, 166(1). https://scholarship.law.upenn.edu/penn_law_review/vol166/iss1/1

42 | TEMAS PROCESALES

2025-2

Foro Internacional
Tutela Judicial Efectiva y Prueba
2025



RED

— Proceso y Justicia —

La presente edición de Temas Procesales reúne un conjunto de investigaciones que reflejan la diversidad, complejidad y actualidad del debate procesal contemporáneo. Con aportes provenientes de España, Colombia, Brasil e Italia, esta revista ofrece al lector un recorrido por problemáticas emergentes y enfoques renovados que dialogan entre la teoría, la práctica judicial y los desafíos tecnológicos que atraviesan el derecho en la actualidad.

Abrimos con un análisis sobre trastornos del lenguaje y pruebas personales, una reflexión necesaria para comprender cómo las condiciones comunicativas inciden en la credibilidad, la percepción judicial y las garantías procesales. A continuación, un estudio sobre los fundamentos teóricos y normativos de las pruebas digitales aborda su creciente centralidad en los sistemas de justicia y los retos que plantean para la autenticidad, integridad y cadena de custodia.

Italia aporta un texto sobre la valoración de las pruebas y su control por la Corte di Cassazione, que permite observar cómo este tribunal ha construido criterios de racionalidad y límites para el juez de mérito. En materia tecnológica, el artículo sobre prueba científica y tecnologías de registro distribuido profundiza en la fiabilidad, trazabilidad y potencial probatorio de sistemas como blockchain. Se suma un estudio sobre lingüística forense y su utilidad para la identificación y atribución de mensajes, seguido de un análisis del criminal compliance program y la prueba en el proceso penal español, especialmente relevante para organizaciones sujetas a responsabilidad penal.

La edición continúa con una reflexión sobre la prueba en la determinación de la filiación, así como un aporte teórico sobre injusticia algorítmico-epistémica y valoración probatoria, tema crucial ante el avance de sistemas automatizados de decisión.

Finalmente, dos estudios inspirados en Taruffo cierran este número: la cientificación del proceso en lo contencioso administrativo colombiano y el principio de precaución ambiental como argumento en la creación judicial del derecho. Esta revista invita a pensar, comparar y transformar nuestras prácticas procesales desde una perspectiva plural y rigurosa.