

JUSTICIA DIGITAL EN EL SUR GLOBAL Y EUROPA: NEUTRALIDAD TECNOLÓGICA, DESIGUALDAD Y GOBERNANZA DEMOCRÁTICA CON MIRADAS PARA CHILE

DIGITAL JUSTICE IN THE GLOBAL SOUTH AND EUROPE: TECHNOLOGICAL NEUTRALITY, INEQUALITY, AND DEMOCRATIC GOVERNANCE WITH PERSPECTIVES FOR CHILE

ISNEL MARTÍNEZ MONTENEGRO^{*}

VÍCTOR ANTONIO NAVARRETE ACUÑA^{**}

SANTIAGO ROJAS BAUVIGNET^{***}

RESUMEN

El artículo analiza críticamente la noción de neutralidad tecnológica en la justicia digital, comparando marcos regulatorios de Europa, países del Sur Global y Chile. Se sostiene que los sistemas automatizados no son neutrales, sino que reproducen y amplifican desigualdades estructurales; cada decisión técnica, desde el diseño de interfaces hasta la ponderación de variables en algoritmos, implica elecciones normativas con consecuencias distributivas. El análisis europeo evidencia avances en protección de datos y en los procedimientos de control, pero su enfoque formalista y procedimental resulta insuficiente para abordar cuestiones sustantivas de justicia, especialmente en contextos del Sur Global caracterizados por brechas digitales, exclusión social y debilidad institucional. En Chile, la adopción acrítica de modelos europeos podría perpetuar inequidades, evidenciando la necesidad de enfoques contextualizados. Se propone un modelo de gobernanza democrática de la justicia digital que integre Privacy by Design, Equidad por Diseño, Diseño Participativo y Evaluación

ABSTRACT

This article critically examines the notion of technological neutrality in digital justice by comparing regulatory frameworks in Europe, the Global South, and Chile. It argues that automated systems are not neutral but rather reproduce and amplify structural inequalities; every technical decision—from interface design to the weighting of variables in algorithms—entails normative choices with distributive consequences. The European analysis highlights progress in data protection and oversight procedures, but its formalistic and procedural approach proves insufficient to address substantive issues of justice, particularly in Global South contexts marked by digital divides, social exclusion, and institutional weakness. In Chile, the uncritical adoption of European models could perpetuate inequities, underscoring the need for context-sensitive approaches. The article proposes a model of democratic governance of digital justice that incorporates Privacy by Design, Equity by Design, Participatory Design, and Algorithmic Impact Assessment, combined

* Doctor en Derecho, Ciencia Política y Criminología de la Universidad de Valencia, España. Profesor e académico de la Universidad Católica de Temuco, Chile.
E-mail: imartinez@uct.cl. ORCID: <https://orcid.org/0000-0003-0322-1071>.

** Cientista Político Universidad Católica de Temuco, Chile. Dirección de Investigación, Universidad Católica de Temuco, Chile.
E-mail: victor.navarrete@uct.cl. ORCID: <https://orcid.org/0009-0009-6162-6481>.

*** Licenciado en Ciencias Jurídicas de la Universidad Católica de Temuco, Chile. Centro de Innovación en Aprendizaje, Docencia y Tecnología Educativa CINAP, Universidad Católica de Temuco, Chile.
E-mail: santiago.rojas@uct.cl. ORCID: <https://orcid.org/0009-0006-2767-965X>.

de Impacto Algorítmico, acompañado de supervisión humana significativa y deliberación ciudadana. La alfabetización tecnopolítica y la consideración de la tecnología como un espacio de disputa política resultan esenciales para construir sistemas judiciales inclusivos, transparentes y responsables. En definitiva, el artículo plantea que la justicia digital no debe concebirse como un destino inevitable, sino como una elección normativa, ofreciendo un camino hacia modelos más equitativos y democráticos, especialmente relevantes para los desafíos del Sur Global y el contexto chileno.

PALABRAS CLAVE: Justicia Algorítmica. Neutralidad Tecnológica. Discriminación Algorítmica. Gobernanza de la IA

with meaningful human oversight and civic deliberation. Techno-political literacy and the recognition of technology as a site of political contestation are essential for building inclusive, transparent, and accountable judicial systems. Ultimately, the article contends that digital justice should not be conceived as an inevitable destiny but as a normative choice, offering a pathway toward more equitable and democratic models, particularly relevant for the challenges of the Global South and the Chilean context.

KEYWORDS: *Algorithmic Justice. Technological Neutrality. Algorithmic Discrimination. AI Governance.*

INTRODUCCIÓN

En las últimas décadas, la digitalización ha penetrado de manera progresiva e inexorable en el campo jurídico. Desde la implementación de sistemas para gestionar expedientes judiciales hasta complejos algoritmos para decisiones como la evaluación de riesgos de reincidencia o la asignación de beneficios sociales, la narrativa tecnocrática sostiene que la justicia será más eficiente, objetiva y accesible¹. Sin embargo, este relato que presenta la tecnología como una solución neutral y desprovista de problemas históricos de demora, sesgos y desigual, oculta tensiones críticas que requieren un análisis profundo. La introducción de estas herramientas implica más que una actualización técnica; representa una intervención política que modifica las relaciones de poder, los modos de producción de la verdad y las formas de exclusión social².

Este artículo cuestiona el supuesto fundamental de que la tecnología aplicada al derecho es neutra. La tesis central que se defiende es que la "neutralidad tecnológica" es una construcción ideológica que despolitiza decisiones esencialmente políticas, encubriendo los valores, prejuicios y estructuras de poder integrados en el código, los algoritmos y la infraestructura digital³. Lejos de ser meras herramientas, estos sistemas son artefactos políticos que pueden reproducir, amplificar y automatizar desigualdades estructurales preexistentes, generando nuevas formas de discriminación y exclusión. Sin una gobernanza democrática y un control sustantivo, la "justicia digital" corre el grave riesgo de convertirse en un sistema de exclusión eficiente, opaco y socialmente corrosivo.

1 EGUÍLUZ CASTAÑEIRA, Juan Antonio. Desafíos y retos que plantean las decisiones automatizadas y los perfiles para los derechos fundamentales. *Estudios de Deusto*, Bilbao, v. 68, n. 2, p. 325–367, 2020. DOI: [https://doi.org/10.18543/ed-68\(2\)-2020pp325-367](https://doi.org/10.18543/ed-68(2)-2020pp325-367).

2 GAO , Qifan; AN , Hongyu. Technology-neutral Illusion: The Ethical and Social Challenges in the Age of Artificial Intelligence. *Sociology, Philosophy and Psychology*, [S. l.], v. 1, n. 2, p. 33–40, 2024. DOI: [10.70267/frygj645](https://doi.org/10.70267/frygj645).

3 GAO; AN, 2024, p. 39.

Para sostener este argumento, la investigación se apoya en una metodología cualitativa de tipo teórico-interpretativa. A través del análisis conceptual y la revisión de una literatura crítica interdisciplinaria, busca problematizar el mito de la neutralidad digital en el ámbito jurídico. Entre sus fundamentos se encuentran aportes de la Teoría Crítica del Derecho, que interpreta el derecho como un campo de disputa política⁴; los Estudios de Ciencia, Tecnología y Sociedad (STS), que evidencian cómo los artefactos tecnológicos encarnan política⁵, y la filosofía de la tecnología, que alerta sobre la lógica de dominación inherente a la racionalidad puramente instrumental⁶.

La idea ampliamente difundida de que la tecnología consiste en instrumentos neutrales, cuyo valor moral depende solo del uso que les den los humanos es una creencia profundamente equívoca⁷. En el ámbito de la justicia, esta noción sirve como un potente legitimador para la implementación de sistemas automatizados, presentados como soluciones objetivas a problemas sociales complejos. Sin embargo, un análisis crítico desde la filosofía y la sociología de la tecnología revela que los artefactos tecnológicos nunca son neutrales, de hecho, están intrínsecamente cargados de valores, sesgos y decisiones políticas.

El discurso predominante defiende a la tecnología como un mero medio para alcanzar fines, desvinculándola de valores y política. No obstante, esta visión instrumentalista ha sido desafiada de manera contundente por teóricos como Langdon Winner y Lawrence Lessig. Winner, en su influyente ensayo "Do Artifacts Have Politics?", sostiene que los artefactos técnicos poseen propiedades políticas de dos formas⁸ primero, porque su invención, diseño o configuración pueden servir para resolver problemas comunitarios, estableciendo jerarquías de poder; y segundo porque hay tecnologías que son intrínsecamente políticas, cuyo funcionamiento requiere ciertas organizaciones sociales y políticas específicas. El clásico ejemplo de Winner son los puentes bajos de Robert Moses en Long Island, diseñados para impedir el acceso a las playas de los autobuses y, por ende, de sectores sociales populares y minorías raciales⁹.

En el ámbito judicial, un sistema informático para la gestión de casos o un algoritmo de valoración de riesgos no se limita a procesar datos. Constituye una arquitectura que regula el comportamiento de jueces, abogados y usuarios de manera en ocasiones más efectiva e invisible que la legislación procesal mis-

4 CORREAS, Óscar. Introducción a la crítica del derecho moderno. México: Fontamara, 2006.

5 WINNER, Langdon. *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. Chicago: University of Chicago Press, 1986.

6 MARCUSE, Herbert. One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society. Boston: Beacon Press, 1964.

7 GAO; AN, 2024, p. 33–34.

8 WINNER, 1986, p. 19–39.

9 WINNER, 1986, p. 23–24.

ma¹⁰. Las decisiones sobre qué datos se recogen, cómo se ponderan variables, qué funciones ofrece la interfaz o qué resultados se priorizan son decisiones políticas que condicionan resultados y distribuyen poder¹¹.

Complementando a Winner, Lawrence Lessig popularizó la expresión "Code is Law"¹² ("el código es ley"). Lessig sostiene que el ciberespacio está regido por cuatro fuerzas: la ley, las normas sociales, el mercado y la arquitectura¹³. De estas, la arquitectura del código es a menudo la más eficaz y menos visible forma de regulación, pues no persuade, sino que obliga; determinando lo que está permitido o prohibido en el entorno digital¹⁴. Por ello, un sistema judicial digital no es solo una plataforma para aplicar las leyes; su código constituye una forma de legislación que impone restricciones y habilita posibilidades, habitualmente sin debate público ni supervisión democrática¹⁵. La falacia de la neutralidad tecnológica radica en ignorar que cada línea de código y decisión de diseño representa una forma de legislación que incorpora valores y ejerce poder¹⁶.

La investigación se centra en cómo la supuesta neutralidad tecnológica en la justicia digital tiende a reproducir relaciones de poder y desigualdades estructurales. En este contexto, surge la pregunta de investigación: ¿de qué manera la neutralidad tecnológica en la justicia digital perpetúa estas relaciones y desigualdades, y cómo pueden compararse críticamente los marcos regulatorios de países del Sur Global, Europa y Chile para diseñar un modelo de justicia digital democrático adaptado a países en desarrollo? El objetivo general es analizar críticamente el concepto de neutralidad tecnológica en el ámbito judicial y examinar cómo los sistemas digitales reflejan decisiones políticas y distribuciones de poder¹⁷. Los objetivos específicos incluyen el estudio desde la teoría crítica del derecho, la sociología de la tecnología y la filosofía tecnológica.

Además, se empleó el método de comparación jurídica de corte funcional y crítico, que confronta normas, regulaciones, sus contextos de aplicación, efecto

10 WACHTER, Sandra; MITTELSTADT, Brent; RUSSELL, Chris. Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. Oxford Internet Institute, University of Oxford, 2020, p. 7–8. DOI: 10.2139/ssrn.3547922.

11 FLORES RIVAS, Juan Carlos; AGUIRREZABAL GRÜNSTEIN, Maite. Análisis de algunos aspectos procesales del sistema de resolución de controversias en concesiones de obra pública. *Revista Chilena de Derecho y Ciencia Política*, Santiago, v. 16, n. 1, p. 1–28, ene. 2025. DOI: <https://doi.org/10.7770/rchdcv16n1-art412>

12 LESSIG, Lawrence. *Code and Other Laws of Cyberspace*. New York: Basic Books, 1999.

13 LESSIG, 1999, p. 235–236.

14 LESSIG, 1999, p. 235–236.

15 WACHTER; MITTELSTADT; RUSSELL, 2020, p. 7–8.

16 GAO; AN, 2024, p. 39

17 MARTÍNEZ MONTENEGRO, Isnel. El problema de investigación jurídica: claves metodológicas para una formulación pertinente. *Revista Chilena de Derecho y Ciencia Política*, Temuco, v. 16, n. 1, p. 1–7, jun. 2025. DOI: <https://doi.org/10.7770/rchdcv16n1-art503INNN>

y fundamentos sociopolíticos¹⁸. Este enfoque pone en evidencia los límites que implica copiar sin adaptación modelos europeos en contextos del Sur Global, donde las condiciones sociales, institucionales y tecnológicas varían y requieren de participación ciudadana y la deliberación judicial adaptadas para evitar la reproducción de desigualdades y garantizar un acceso efectivo a la justicia digital.

El análisis documental abarcó fuentes normativas, proyectos de ley, documentos oficiales de política pública y literatura académica crítica, complementando con matrices comparativas que sistematicen principios, enfoques y vacíos de cada marco regulatorio, lo que permite identificar cómo los sistemas judiciales digitales no son neutrales, sino artefactos políticos que configuran relaciones de poder. De este modo, se propone un modelo de justicia digital para el contexto del Sur Global basado en los principios de Privacy by Design, Equidad por Diseño, Diseño Participativo y Evaluación de Impacto Algorítmico, con control humano significativo y promoviendo la justicia democrática.

1. LA ILUSIÓN DE LA NEUTRALIDAD TECNOLÓGICA EN LA JUSTICIA: ARTEFACTOS, CÓDIGO Y ALGORITMOS COMO FORMAS DE PODER POLÍTICO

La crítica a la neutralidad tecnológica se profundiza con el análisis de Herbert Marcuse sobre la "racionalidad tecnológica". En su obra "El Hombre Unidimensional", Marcuse sostiene que la sociedad contemporánea, caracterizada por su desarrollo industrial, ha establecido la racionalidad técnica e instrumental como forma dominante de pensar, la cual limita la capacidad de cuestionamiento y la oposición¹⁹. Según el autor la racionalidad tecnológica estructura un entorno donde solo se consideran válidas las preguntas vinculadas a la eficiencia y el control, dejando de lado cualquier otra perspectiva que no se ajuste a dichos parámetros²⁰.

Cuando esta lógica se inserta en el sistema de justicia, sus efectos resultan especialmente problemáticos. El derecho se ocupa de valores irrenunciables como equidad, deliberación, contexto, empatía, dignidad y la posibilidad de redención. Sin embargo, la racionalidad tecnológica tiende a transformar aspectos humanos y cualitativos en elementos susceptibles de medición y procesamiento, o directamente a descartar aquello que se percibe como ineficaz²¹. Los algoritmos de sentencia pueden optimizar la gestión de datos y producir puntuaciones de riesgo, pero son incapaces de captar el contexto personal, la trayectoria vital

18 MARTÍNEZ MONTENEGRO, Isnel. Sobre los métodos de la investigación jurídica. *Revista Chilena de Derecho y Ciencia Política*, Temuco, v. 14, n. 1, p. 1–4, jun. 2023. DOI: <https://doi.org/10.7770/rchdcp-V14N1-art312>

19 MARCUSE, 1964, p. 17.

20 MARCUSE, 1964, p. 17–18.

21 MARCUSE, 1964, p. 17–18

o el potencial de cambio del individuo. En este sentido, la labor judicial requiere que las decisiones se fundamenten adecuadamente, ya que la justificación de los fallos debe involucrar tanto criterios objetivos previstos por la norma como una valoración integral de la prueba y la búsqueda de justicia sustantiva²². El supuesto central implica una exigencia de humanidad en la administración de justicia, entendida como el análisis integral del individuo y de su conducta en su totalidad. Esta tarea ya representa un desafío para la labor jurisdiccional; por lo tanto, asignar dicha función a sistemas algorítmicos o de inteligencia artificial no implicaría, en principio, excluirlos de dicho desafío.

Marcuse advierte que el peligro no reside en la tecnología per se, sino en la ideología que la acompaña: la creencia de que todos los problemas admiten soluciones técnicas²³. En el ámbito jurídico, esta premisa desemboca en la aceptación acrítica de sistemas automatizados por percibirlos como superiores, debido a su rapidez o capacidad de procesamiento. Así, la justicia corre el riesgo de ser reducida a un problema de gestión informacional, subordinando su propósito esencial —la realización de la justicia en casos concretos— a los imperativos de eficiencia y predicción. De esta manera, la tecnología redefine incluso qué se considera un problema o una solución legítima, desplazando formas de razonamiento que escapan a su lógica unidimensional²⁴.

La Teoría Crítica del Derecho, impulsada por juristas como Óscar Correas, aporta un marco adicional para comprender la ausencia de neutralidad tecnológica en el ámbito legal. Correas sostiene que el derecho es un campo de disputa social y política, donde el lenguaje jurídico se utiliza para legitimar intereses y proyectos de poder²⁵.

Desde esta perspectiva, la introducción de la inteligencia artificial y la automatización en el sistema judicial no puede interpretarse como una mera modernización técnica, sino como la irrupción de nuevos actores ideológicos. Los algoritmos y plataformas digitales —dotados de una aparente objetividad matemática— pueden consolidar relaciones de poder, naturalizar decisiones discretionales y marginar perspectivas de colectivos desfavorecidos²⁶²⁷. Pueden servir para solidificar las relaciones de poder existentes, para hacer que las decisiones discretionales y políticas parezcan inevitables y objetivas, y para marginar las voces y perspectivas de los grupos ya desfavorecidos.

22 SQUELLA NARDUCCI, Agustín. Justificar decisiones jurídicas y justificar decisiones judiciales. *Revista de Derecho* (Valdivia), Valdivia, v. XIX, n. 1, p. 277–292, jul. 2006, p. 280–282.

23 MARCUSE, 1964, p. 16–18

24 GAO; AN, 2024, p. 35–36.

25 CORREAS, 2006, p. 18–20.

26 CRUZ RIVERO, Diego. Disputabilidad y equidad en los mercados digitales: una visión de Derecho europeo. *Revista Chilena de Derecho y Ciencia Política*, Temuco, v. 15, n. 1, p. 1–31, jul. 2024. DOI: <https://doi.org/10.7770/rchdcv-15n1-art393>

27 GAO; AN, 2024, p. 39

La tecnología, por tanto, no es un árbitro imparcial, sino que actúa como agente activo en la lucha por la definición de la justicia. La elección de qué datos utilizar para entrenar un algoritmo, qué variables considerar relevantes o qué resultados optimizar son decisiones que benefician a ciertos grupos y perjudican a otros. Al presentar estas decisiones como puramente técnicas, el discurso de la neutralidad tecnológica oculta su carácter político y las sustraе del debate democrático. La verdadera tarea, desde una perspectiva crítica, es desvelar la política inscrita en la tecnología y someterla a un escrutinio democrático riguroso.

El mito de la neutralidad tecnológica opera como mecanismo ideológico de legitimación con profundas consecuencias. Presentar los sistemas automatizados como apolíticos favorece su adopción y aplicación sin debate crítico, tanto por operadores jurídicos como por la sociedad. La población tiende a mostrar mayor disposición a aceptar transformaciones profundas en su vida cotidiana cuando estas se presentan bajo el rótulo de la innovación tecnológica, en lugar de enmarcarse como reformas políticas explícitas²⁸. Este fenómeno crea una peligrosa dinámica: el discurso de la neutralidad actúa como un caballo de Troya que facilita la implementación de cambios estructurales en la administración de justicia —quién decide, con qué criterios y con qué consecuencias— sin un debate político transparente. El resultado es una delegación de facto de la soberanía judicial desde foros públicos y deliberativos, como los tribunales y los parlamentos, hacia espacios opacos y privados, como las empresas de tecnología y los equipos de desarrolladores, que operan según lógicas de mercado y diseño técnico, fuera del alcance del control democrático.

En síntesis, la tecnología aplicada a la justicia nunca es neutral ni exclusivamente instrumental. Cada decisión de diseño y cada algoritmo incorporado en los sistemas judiciales introduce una forma de regulación que moldea el ejercicio del derecho y redefine el sentido mismo de lo justo. La ilusión de neutralidad tecnológica funciona como un velo que encubre relaciones de poder y desplaza el debate político hacia terrenos aparentemente técnicos y resta transparencia y participación democrática en la toma de decisiones. De ahí la necesidad de asumir una mirada crítica que reconozca a la tecnología como un actor político y promueva modelos de desarrollo orientados a la equidad, la deliberación y la dignidad humana. El desafío no radica en rechazar la innovación, sino en orientar su utilización hacia una justicia verdaderamente democrática e inclusiva.

28 WINNER, 1986, p. 22-24

2. ENTRE LA TRANSPARENCIA FORMAL Y EL CONTROL EFECTIVO: LOS DILEMAS REGULATORIOS DE LA JUSTICIA DIGITAL EUROPEA CON ESPECIAL REFERENCIA A ESPAÑA

Frente a los crecientes riesgos de discriminación y exclusión derivados de la justicia automatizada, han surgido diversos marcos regulatorios, especialmente en la Unión Europea, con el objetivo de establecer salvaguardias y garantizar la protección de los derechos fundamentales. Sin embargo, un análisis crítico de estas normativas revela limitaciones: aunque supone un avance, se sustenta en una lógica procedural y de gestión del riesgo que no siempre atiende a las dimensiones sustantivas de poder y justicia.

La Unión Europea ha adoptado un enfoque pionero y multifacético para la regulación de la tecnología digital, creando un ecosistema normativo compuesto principalmente por el Reglamento General de Protección de Datos (RGPD)²⁹, la Ley de Servicios Digitales (DSA)³⁰ y la Ley de Inteligencia Artificial (Ley de IA)³¹.

El Reglamento General de Protección de Datos (RGPD), en su Artículo 22, reconoce el derecho de toda persona a no quedar sometida a decisiones basadas exclusivamente en procesos automatizados —incluida la elaboración de perfiles— que tengan consecuencias jurídicas o un impacto significativo en su vida³². En apariencia, esta disposición parece una defensa robusta contra la tiranía algorítmica. Sin embargo, sus excepciones y ambigüedades limitan severamente su eficacia práctica. Entre ellas destaca la excepción que permite tales decisiones cuando resultan necesarias para celebrar o ejecutar un contrato, lo que, bajo una interpretación excesivamente amplia permite la automatización de decisiones en ámbitos como la concesión de créditos, seguros o contrataciones sin garantías efectivas. Según Davis y Schwemer, esta redacción favorece una especie de automatización “legitimada a medias”, en la que los sistemas funcionan sin mecanismos sustantivos de revisión humana, debilitando la protección de derechos fundamentales³³.

29 UNIÓN EUROPEA. Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (Reglamento General de Protección de Datos). Diario Oficial de la Unión Europea, L 119, 4 mayo 2016.

30 UNIÓN EUROPEA. Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo, de 19 de octubre de 2022, relativo a un mercado único de servicios digitales (Ley de Servicios Digitales). Diario Oficial de la Unión Europea, L 277, 27 oct. 2022.

31 UNIÓN EUROPEA. Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de marzo de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial). Diario Oficial de la Unión Europea, L 148, 30 mayo 2024.

32 UNIÓN EUROPEA. Reglamento (UE) 2016/679, art. 22.

33 DAVIS, Peter Alexander Earls; SCHWEMER, Sebastian Felix. Rethinking decisions under Ar-

Asimismo, la norma únicamente se aplica a aquellas decisiones basadas de manera plenamente automatizada, lo que permite a las organizaciones evadir su alcance mediante la incorporación de una supuesta "intervención humana significativa". Este concepto carece de una definición precisa y, en la práctica, puede quedar reducido a una revisión meramente simbólica o *rubber-stamping*³⁴ ³⁵. En este mismo sentido, el disputado derecho a la explicación —que se desprende de los artículos 13, 14 y 15 del RGPD y cuyo alcance ha generado un amplio debate académico— constituye otro reflejo de esta debilidad estructural³⁶.

Aunque el enfoque del DSA es innovador al imponer obligaciones de transparencia y responsabilidad a las grandes plataformas digitales, su ámbito de aplicación se restringe casi exclusivamente al entorno privado y comercial. No aborda de forma directa ni específica el uso de sistemas algorítmicos por parte de las administraciones públicas, incluidos los utilizados en ámbitos tan sensibles como la justicia o la seguridad social. Esta omisión ha sido destacada por el Parlamento Europeo, que advierte que el uso de decisiones algorítmicas en la administración pública plantea riesgos significativos para los derechos fundamentales³⁷.

En comparación, la Ley de Inteligencia Artificial (Ley de IA) constituye la iniciativa más ambiciosa y directa. Su enfoque se basa en la gestión del riesgo e identifica como de alto riesgo a los sistemas de IA empleados en la administración de justicia y procesos democráticos, así como en la aplicación de la ley³⁸. A partir de esta clasificación se impone obligaciones estrictas a los desarrolladores y usuarios: desde la obligación de realizar evaluaciones de conformidad, garantizar la calidad de los datos para reducir sesgos, asegurar la transparencia hacia los usuarios e implementar mecanismos de supervisión humana adecuados, hasta la obligación de robustez técnica y ciberseguridad³⁹. Sin embargo, pese a su aparente rigor, este marco ha sido criticado por su enfoque predominantemente procedimental. En la práctica, el cumplimiento puede terminar traduciéndose en

ticle 22 of the GDPR: implications for semi-automated legal decision-making. In: Proceedings of the Third International Workshop on Artificial Intelligence and Intelligent Assistance for Legal Professionals in the Digital Workplace (LegalAIIA 2023, ICAIL 2023, Braga, Portugal, 19 jun. 2023. CEUR Workshop Proceedings, 2023, p. 1–2.

34 LEVITINA, Anna. Humans in automated decision-making under the GDPR and AI Act. Barcelona: CIDOB, 2024. Disponible en: <<https://www.cidob.org/en/publications/humans-automated-decision-making-under-gdpr-and-ai-act>>

35 EGÚÍLUZ CASTAÑEIRA, 2020, p. 341–343.

36 EGÚÍLUZ CASTAÑEIRA, 2020, p. 344–346

37 EUROPEAN PARLIAMENTARY RESEARCH SERVICE. *Understanding algorithmic decision-making: opportunities and challenges*. Brussels: European Parliament, 2019, p. 13–15.

38 UNIÓN EUROPEA. Reglamento (UE) 2024/1689, art. 6.

39 UNIÓN EUROPEA. Reglamento (UE) 2024/1689, arts. 8–15.

trámites formales —como completar formularios o superar auditorías— sin que ello garantice un funcionamiento realmente justo del sistema⁴⁰.

Tabla 2. Marco Regulatorio de la UE para Sistemas de IA en la Justicia

Instrumento legal	Disposición clave	Requisitos principales (alto riesgo en justicia)	Críticas y limitaciones
Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (RGPD)	Art. 22: derecho a no ser objeto de decisiones basadas únicamente en tratamientos automatizados; derecho a intervención humana y a impugnar.	Garantía formal contra decisiones 100% automatizadas; derecho a explicación derivado de Arts. 13–15.	Excepciones amplias (ej. contractual); fácil elusión al añadir “intervención humana” simbólica; derecho a explicación ambiguo.
Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo, de 19 de octubre de 2022, relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE (Ley de Servicios Digitales, DSA)	Arts. 27, 34: transparencia en sistemas de recomendación y evaluación de riesgos sistémicos.	Obligaciones de información y auditoría de grandes plataformas en línea.	Enfoque en el sector privado y moderación de contenidos; no cubre decisiones algorítmicas públicas (justicia, seguridad social).

⁴⁰ AUDE CEFALIELLO; MIRIAM KULLMANN. Offering false security: How the draft artificial intelligence act undermines fundamental workers' rights. **European Labour Law Journal**, v. 13, n. 4, p. 542–562, 2022. DOI: <https://doi.org/10.1177/20319525221114474>

Instrumento legal	Disposición clave	Requisitos principales (alto riesgo en justicia)	Críticas y limitaciones
Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de IA)	Art. 6 y Anexo III: clasificación como “alto riesgo” de sistemas usados en justicia y procesos democráticos.	Evaluación de conformidad; requisitos de calidad de datos, transparencia, supervisión humana y robustez.	

La tabla sintetiza el ecosistema normativo europeo en materia de decisiones automatizadas en la justicia, mostrando tanto sus avances como sus límites. En primer lugar, el RGPD (2016/679) constituye la norma fundacional en materia de protección de datos y fue pionero en establecer el derecho a no ser objeto de decisiones basadas exclusivamente en tratamientos automatizados. Sin embargo, su eficacia práctica queda debilitada por dos factores: las amplias excepciones (como la relativa a contratos) y la facilidad con la que se puede eludir la prohibición mediante la inclusión de una intervención humana mínima o simbólica. Además, el célebre “derecho a la explicación” carece de un alcance jurídico preciso, lo que ha generado un debate sobre su verdadera capacidad para garantizar transparencia.

En segundo lugar, la Ley de Servicios Digitales (Reglamento 2022/2065) representa un paso importante en términos de responsabilidad y transparencia en el ecosistema digital, pero su foco está puesto en las grandes plataformas privadas y la moderación de contenidos. Al no aplicarse de forma directa a los sistemas algorítmicos del sector público, su impacto en la justicia digital es, en la práctica, marginal.

En tercer lugar, la Ley de IA (Reglamento 2024/1689) introduce por primera vez un marco específico de control para sistemas algorítmicos de “alto riesgo”, dentro de los cuales se incluyen aquellos aplicados a la justicia. Sus obligaciones son las más estrictas: calidad de datos, transparencia, supervisión humana y robustez técnica. No obstante, diversos autores advierten que este modelo se basa en un enfoque procedural y de gestión de riesgos. En consecuencia, el cumplimiento puede quedar reducido a auditorías y certificaciones formales sin que ello garantice decisiones sustantivamente justas.

En conjunto, la tabla revela que la regulación europea ha avanzado en crear mecanismos de control preventivo y de transparencia, pero sigue enfrentando una tensión estructural: mientras el derecho opera sobre principios normativos e interpretativos, la regulación de la IA se orienta hacia la estandarización técnica y procedimental. El riesgo, por tanto, es que la defensa de los derechos fundamentales quede atrapada en una lógica burocrática de cumplimiento sin asegurar un control judicial y social real.

Esta lógica regulatoria, si bien busca gestionar riesgos con buenas intenciones, resulta paradójica. Al intentar controlar la racionalidad tecnológica desde parámetros de cuantificación y cálculo. El peligro radica en que la defensa de los derechos fundamentales podría transformarse en una mera gestión burocrática de la ética. Así, sistemas perjudiciales podrían legitimarse si cumplen los procedimientos, lo que vaciaría de contenido el control judicial y social efectivo.

La limitación más profunda del enfoque regulatorio actual, y del proyecto de justicia algorítmica en general, reside en una tensión fundamental entre la naturaleza del derecho y la lógica de la automatización. Como señalan Wachter y sus colegas, el derecho antidiscriminatorio europeo, y el razonamiento jurídico en general, se fundamenta en la "igualdad contextual"⁴¹: determinar una discriminación no es un cálculo mecánico, sino una práctica interpretativa y ca-suística, sensible al contexto social, político y normativo. En este marco, jueces y tribunales evalúan situaciones de manera flexible, decidiendo qué constituye un trato desigual, quién puede considerarse un grupo de comparación legítimo y si una diferencia de trato está objetivamente justificada⁴².

Este enfoque es, por diseño, incompatible con la lógica de los sistemas de IA. La automatización a escala requiere reglas claras, estáticas, cuantificables y homogéneas. Un algoritmo no puede deliberar sobre el "contexto" o hacer juicios normativos matizados ya que opera aplicando reglas predefinidas a conjuntos de datos. Así, cualquier intento de traducir la justicia en métricas de fairness supone una simplificación que contradice la esencia del derecho como práctica interpretativa. La justicia europea, por diseño, no puede automatizarse sin perder sus atributos fundamentales: flexibilidad, sensibilidad al contexto y apertura a la deliberación⁴³.

Ante las limitaciones de la regulación y la opacidad de los sistemas algorítmicos, el poder judicial ha comenzado a emerger como un contrapeso crucial, reafirmando la primacía de los derechos fundamentales sobre la eficiencia tecnológica. Dos casos europeos recientes son particularmente ilustrativos.

El caso SyRI en los Países Bajos es un ejemplo paradigmático. El gobierno holandés implementó el *Systeem Risico Indicatie* (SyRI), un sistema algorítmico

⁴¹ WACHTER; MITTELSTADT; RUSSELL, 2020, p. 12-14.

⁴² WACHTER; MITTELSTADT; RUSSELL, 2020, p. 13.

⁴³ WACHTER; MITTELSTADT; RUSSELL, 2020, p. 15.

de perfilado de riesgos para detectar posibles fraudes en materia de seguridad social e impuestos⁴⁴. En febrero de 2020, el Tribunal de Distrito de La Haya declaró ilegal el uso de SyRI⁴⁵. La decisión es notable porque no se basó en un fallo técnico del algoritmo, sino en una violación del Artículo 8 del Convenio Europeo de Derechos Humanos (CEDH), que protege el derecho a la vida privada⁴⁶. El tribunal consideró que el sistema, debido a su opacidad y a la falta de transparencia sobre el modelo de riesgo utilizado, no lograba un "justo equilibrio" entre el interés legítimo del Estado en combatir el fraude y el derecho fundamental a la privacidad de los ciudadanos⁴⁷. El tribunal destacó además que la opacidad del sistema, sumada a la ausencia de mecanismos que permitieran verificar su funcionamiento, lo hacía intrínsecamente desproporcionado e incompatible con los principios de un Estado de derecho democráticos. Como argumentan Appelman, Ó Fathaigh y van Hoboken, la sentencia se apoyó en una interpretación estricta del Art. 8 CEDH, enfatizando la necesidad de transparencia y límites explícitos para evitar un desequilibrio injustificado entre el interés estatal y los derechos individuales⁴⁸.

En España, el caso de la reforma de la LOREG (Ley Orgánica del Régimen Electoral General) ofrece otra lección importante. En 2018, se introdujo el Artículo 58 bis, que permitía a los partidos políticos recopilar datos personales relativos a las opiniones políticas de los ciudadanos a partir de fuentes de acceso público para realizar actividades de propaganda electoral en periodo de elecciones⁴⁹. Esta medida, que abría la puerta al microtargeting y al perfilado ideológico masivo, fue recurrida ante el Tribunal Constitucional. En su sentencia de 2019, el Tribunal declaró nulo el precepto⁵⁰. El argumento central fue que la ley era excesivamente vaga y carecía de las garantías adecuadas para proteger derechos fundamentales como la protección de datos⁵¹ y la libertad ideológica⁵².

44 EGUÍLUZ CASTAÑEIRA, 2020, p. 351–353.

45 PAÍSES BAJOS. Rechtbank Den Haag. Sentencia de 5 de febrero de 2020. ECLI:NL:RBDHA:2020:1878 (Asunto SyRI). La Haya, 2020

46 CONSEJO DE EUROPA. Convenio Europeo para la Protección de los Derechos Humanos y de las Libertades Fundamentales, Roma, 4 nov. 1950. Art. 8.

47 EGUÍLUZ CASTAÑEIRA, 2020, p. 352–353.

48 NAOMI APPELMAN; RONAN Ó FATHAIGH; JORIS VAN HOBOKEN. Social Welfare, Risk Profiling and Fundamental Rights: The Case of SyRI in the Netherlands. *JIPITEC – Journal of Intellectual Property, Information Technology and E-Commerce Law*, v. 12, p. 257-271, 2021. Disponible en: <<https://ssrn.com/abstract=3984935>>.

49 EGUÍLUZ CASTAÑEIRA, 2020, p. 348–349

50 ESPAÑA. Tribunal Constitucional. Sentencia 76/2019, de 22 de mayo. Madrid: Tribunal Constitucional, 2019.

51 ESPAÑA. Constitución Española. Boletín Oficial del Estado, 29 de diciembre de 1978. Art. 18.4.

52 ESPAÑA. Constitución Española. Boletín Oficial del Estado, 29 de diciembre de 1978. Art. 16.

⁵³. El tribunal determinó que una habilitación tan genérica para tratar datos especialmente sensibles, sin establecer límites claros ni condiciones precisas, constituyía una injerencia desproporcionada e inconstitucional⁵⁴.

Ambos casos demuestran que, incluso cuando los marcos regulatorios específicos sobre IA son incipientes o insuficientes, los tribunales pueden y deben recurrir a los principios fundamentales de los derechos humanos y constitucionales para actuar como un dique de contención contra los excesos de la vigilancia y la automatización estatal. Estas sentencias reafirman que la eficiencia y la innovación tecnológica no pueden prevalecer sobre los derechos y las garantías que constituyen el núcleo del Estado de derecho.

Si bien la sección anterior analizó los desafíos estructurales de acceso a la justicia digital en el Sur Global, en el Norte Global el debate ha sido recientemente reconfigurado por la irrupción de una tecnología disruptiva: los modelos de lenguaje generativo a gran escala (LLMs). A diferencia de los sistemas predictivos tradicionales, que procesan variables para clasificar o tomar decisiones, la IA generativa tiene la capacidad de producir directamente textos jurídicos: demandas, contratos, alegatos e incluso sentencias simuladas. Ello genera riesgos epistémicos y normativos de una magnitud sin precedentes. Desde una perspectiva crítica, esta tecnología puede interpretarse como una forma extrema de lo que Herbert Marcuse denominó "racionalidad unidimensional": la reducción de procesos complejos, como la deliberación jurídica, a cadenas formales que oscurecen su dimensión política y valorativa.⁵⁵. Los modelos generativos aplican una lógica probabilística al lenguaje legal, anticipando patrones textuales en lugar de interpretar normas con sentido. Este tránsito de la interpretación a la predicción encarna una forma de justicia "unidimensional", donde la complejidad contextual se ve sustituida por fórmulas lingüísticas entrenadas estadísticamente.

Entre los riesgos más significativos de esta transformación se encuentran la "alucinación algorítmica" es decir, la generación de información inexistente o ficticia, como referencias jurisprudenciales inventadas. Estudios estiman que entre el 17% y el 33% de las respuestas de sistemas de IA jurídica contienen este tipo de errores⁵⁶.

El caso Mata v. Avianca (2023) resultó paradigmático: abogados presentaron ante un tribunal de Nueva York una demanda con múltiples citas jurisprudenciales generadas por ChatGPT, todas completamente falsas. El tribunal no solo rechazó el documento, sino que sancionó a los profesionales involu-

⁵³ EGUÍLUZ CASTAÑEIRA, 2020, p. 348–349

⁵⁴ EGUÍLUZ CASTAÑEIRA, 2020, p. 352–353.

⁵⁵ MARCUSE, 1964, p. 16–18

⁵⁶ MAGESH, Anant; et al. Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools. *Journal of Empirical Legal Studies*, Hoboken, 2025, p. 14–16.

crados por violaciones éticas graves⁵⁷ ⁵⁸. Este episodio expone la magnitud del riesgo estructural que implica delegar tareas jurídicas a sistemas que, por diseño, carecen de sentido normativo y verificación ontológica. Ante estos desafíos, organismos profesionales han comenzado a establecer directrices específicas. En julio de 2024, la American Bar Association (ABA) emitió la Opinión Formal 512, donde establece que los abogados deben obtener consentimiento informado del cliente antes de utilizar IA generativa, verificar exhaustivamente la veracidad del contenido generado, y asumir responsabilidad plena sobre los documentos presentados⁵⁹.

Este panorama demuestra que el problema trasciende lo técnico y pone en cuestión la idea misma de supervisión humana como garantía suficiente. La noción de "intervención significativa", tal como aparece en algunos marcos europeos, resulta insuficiente cuando no se acompaña de condiciones estructurales que garanticen comprensión y responsabilidad real. En este contexto, el concepto de Control Humano Significativo (MHC), originado en el ámbito de armas autónomas, emerge como alternativa robusta. Sus tres requisitos son: proporcionar a las personas información comprensible, garantizar una capacidad efectiva de intervención o voto, y asignar una responsabilidad legal clara⁶⁰. Este marco se está ampliando a contextos civiles de alto riesgo, como la justicia, con el fin de asegurar una supervisión sustantiva.

En suma, la IA generativa no solo introduce nuevos riesgos técnicos, sino que exige repensar las condiciones de legitimidad y responsabilidad en el uso de sistemas automatizados en el derecho. En lugar de asumir que basta con una supervisión superficial, es necesario articular marcos institucionales que garanticen un control humano sustutivo, deliberativo y contextualizado, especialmente en decisiones que afectan derechos fundamentales.

La regulación europea sobre justicia digital ha avanzado considerablemente en los últimos años, con instrumentos como el RGPD, la Ley de Servicios Digitales (DSA) y la Ley de Inteligencia Artificial (Ley de IA). El RGPD establece derechos como el de no ser objeto de decisiones basadas únicamente en tratamientos automatizados, pero su efectividad práctica se ve limitada por amplias excepciones y la posibilidad de eludirlo mediante intervenciones humanas sim-

57 BRODERICK, P. Kevin. Order to Show Cause, Sanctions. Documento judicial en el caso Mata v. Avianca, Inc., No. 22-CV-1461 (PKC). New York: United States District Court, S.D.N.Y., 2023.

58 UNITED STATES DISTRICT COURT (Southern District of New York). Mata v. Avianca, Inc., No. 22-CV-1461 (PKC), 2023 WL 4114965, 22 jun. 2023.

59 AMERICAN BAR ASSOCIATION. Formal Opinion 512 on Lawyers' Use of Generative Artificial Intelligence. Chicago: ABA, 2024.

60 SANTONI DE SIO, Filippo; VAN DEN HOVEN, Jeroen. Meaningful Human Control over Autonomous Systems: A Philosophical Account. *Frontiers in Robotics and AI*, Lausanne, v. 5, art. 15, 2018. DOI: <https://doi.org/10.3389/frobt.2018.00015>

bólicas⁶¹. La DSA promueve transparencia y responsabilidad en plataformas digitales privadas, aunque no aborda directamente decisiones algorítmicas públicas⁶². La Ley de IA introduce obligaciones estrictas para sistemas de alto riesgo, incluyendo la justicia, sin embargo, su enfoque procedural y basado en gestión de riesgos puede reducirse a formalidades, sin garantizar decisiones sustantivamente justas, generando una tensión entre la regulación técnica y los principios normativos del derecho, que dependen de interpretación contextual y sensibilidad social⁶³.

Esta brecha se hace visible en la práctica judicial: casos como SyRI en Países Bajos y la reforma de la LOREG en España muestran cómo los tribunales han actuado en defensa de los derechos fundamentales frente a sistemas opacos o normas excesivamente amplias, subrayando que la innovación tecnológica nunca puede imponerse sobre las garantías legales. La aparición de la IA generativa, con su capacidad de elaborar textos jurídicos completos, introduce riesgos inéditos, entre ellos la llamada “alucinación algorítmica”, que consiste en la producción de referencias o jurisprudencia inexistentes, y la insuficiencia de una supervisión humana meramente formal. Ante esto, conceptos como el Control Humano Significativo (MHC) ofrecen un marco más sólido para garantizar responsabilidad y comprensión efectiva, demostrando que la legitimidad de la justicia automatizada depende no solo del cumplimiento formal, sino de mecanismos de control humano sustantivo, deliberativo y contextualizado.

3. JUSTICIA DIGITAL Y EXCLUSIÓN: EL LADO OSCURO DE LA AUTOMATIZACIÓN EN LA ERA DE LA IA EN EL SUR GLOBAL

La crítica teórica a la neutralidad tecnológica tiene especial interés cuando se examinan sus manifestaciones concretas. La implementación de sistemas automatizados en la justicia no es una promesa futura, sino una realidad presente que ya está generando formas tangibles de discriminación y exclusión. Esta sección analiza los mecanismos específicos a través de los cuales la tecnología, lejos de ser un ecualizador, se convierte en un multiplicador de desigualdades.

Los sesgos algorítmicos no derivan necesariamente de intenciones deliberadas, sino de elecciones técnicas de diseño que reproducen desigualdades sociales ya existentes. Según Eguílez Castañeira, la discriminación no se limita al uso de datos sesgados, aunque este aspecto sea relevante, sino que puede introducirse en múltiples etapas críticas del diseño de un sistema de aprendizaje automático, muchas veces de manera inadvertida⁶⁴.

61 UNIÓN EUROPEA. Reglamento (UE) 2016/679, art. 22

62 UNIÓN EUROPEA. Reglamento (UE) 2022/2065, art. 34.

63 UNIÓN EUROPEA. Reglamento (UE) 2024/1689, arts. 6 y 8–15.

64 EGUILÉZ CASTAÑEIRA, 2020, p. 334–336

Un ejemplo central es la definición de la llamada "variable objetivo". Cuando un algoritmo de predicción se crea, los desarrolladores deben decidir qué resultado se va a optimizar. En un sistema de evaluación de riesgos de reincidencia, por ejemplo, podría elegirse "nueva detención" en lugar de "nueva condena" como métrica principal. Tal decisión, aparentemente técnica, tiene profundas implicaciones: mientras las detenciones están influidas por sesgos en las prácticas policiales, las condenas exigen un estándar probatorio más elevado. Como señala Eguílez Castañeira, si se optimiza utilizando una variable previamente sesgada, el modelo simplemente replicará y ampliará ese sesgo⁶⁵.

Del mismo modo, la definición de las "etiquetas de clase" también es un momento de riesgo. Los sistemas de Machine Learning (ML) funcionan clasificando datos en categorías preestablecidas. Sin embargo, el modo en que se construyen esas etiquetas puede introducir prejuicios. Por ejemplo, considerar como indicador de un "buen empleado" la continuidad laboral ininterrumpida resulta discriminatorio hacia mujeres, quienes estadísticamente interrumpen más sus carreras por motivos de cuidado familiar⁶⁶.

La recopilación de datos de entrenamiento es otro punto crítico. Como advierte el mismo autor, si los datos provienen de una sociedad desigual, el algoritmo no hará más que aprender y codificar esa desigualdad. El sistema de reclutamiento creado por Amazon ilustra bien esto: al ser entrenado con currículums históricos dominados por perfiles masculinos, el algoritmo terminó penalizando solicitudes que incluían la palabra "femenino". En la misma línea, los sistemas de "policía predictiva" entrenados con datos de detenciones en zonas sobrevigiladas generan bucles de retroalimentación: recomiendan vigilar esas áreas, lo que provoca más arrestos y, en consecuencia, refuerza la "validez" de las predicciones⁶⁷.

Las denominadas "variables proxy" constituyen otro mecanismo problemático. Incluso si se prohíbe explícitamente el uso de variables protegidas como la raza o el género, los algoritmos pueden aprender a utilizar otras variables altamente correlacionadas para hacer las mismas discriminaciones. El código postal, por ejemplo, puede funcionar como un sustituto altamente correlacionado con raza o clase social. Tal como explica Eguílez Castañeira⁶⁸, un algoritmo que deniega créditos en función del código postal, aunque no incluya la variable "raza", puede generar discriminaciones estructurales. En esta línea, O'Neil (2017) describe estos sistemas como "armas de destrucción matemática" por su alcance masivo, su opacidad y su capacidad de producir daños significativos bajo una apariencia de neutralidad.

65 EGUÍLUZ CASTAÑEIRA, 2020, p. 336–338.

66 EGUÍLUZ CASTAÑEIRA, 2020, p. 338–340.

67 EGUÍLUZ CASTAÑEIRA, 2020, p. 340–341.

68 EGUÍLUZ CASTAÑEIRA, 2020, p. 339–340.

Un caso paradigmático ha sido el algoritmo COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), empleado en varios estados de EE. UU. para predecir la probabilidad de reincidencia. La investigación de ProPublica en 2016 reveló que, aunque la precisión global del modelo era similar para acusados negros y blancos (cerca del 61%), los patrones de error mostraban disparidades significativas por raza⁶⁹⁷⁰. Concretamente, los acusados negros que no reincidieron fueron clasificados erróneamente como de “alto riesgo” casi el doble de veces que los acusados blancos (45% frente a 23%). Por otro lado, los acusados blancos que sí reincidieron fueron etiquetados con mucha más frecuencia como de “bajo riesgo” en relación con los acusados negros (48% frente a 28%)⁷¹. En términos prácticos, el sistema imponía una carga desigual de errores sobre la población negra, sometiéndola a un escrutinio más severo y a consecuencias jurídicas más gravosas.

La empresa desarrolladora, Northpointe (hoy Equivant), argumentó que su modelo cumplía con la llamada “paridad predictiva”: para un mismo nivel de riesgo, la probabilidad real de reincidencia resultaba equivalente entre personas negras y blancas⁷². Sin embargo, como subraya la literatura contemporánea, cuando las tasas iniciales de un fenómeno son distintas entre grupos, es matemáticamente imposible cumplir al mismo tiempo la calibración, la igualdad en tasas de falsos positivos y la paridad demográfica⁷³. Este hallazgo pone en cuestión modelos como COMPAS, que, si bien buscan preservar lo que Dieterich et al. (2016) denominan “paridad predictiva”, pueden hacerlo a costa de aumentar la disparidad en los errores de predicción entre distintos grupos raciales⁷⁴. La investigación empírica de Angwin et al. (2016) evidenció precisamente esto: personas negras clasificadas como de alto riesgo reincidieron con menor frecuencia que personas blancas con la misma puntuación⁷⁵. Estudios posteriores han reforzado estas críticas, señalando que herramientas como estas exigen tomar decisiones normativas explícitas sobre qué métrica priorizar, y que no basta con

69 ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren. Machine Bias: There’s Software Used Across the Country to Predict Future Criminals. And it’s Biased Against Blacks. *ProPublica*, 23 may. 2016. Disponible en: <<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>>.

70 KLEINBERG, Jon; MULLAINATHAN, Sendhil; RAGHAVAN, Manish. Inherent Trade-Offs in the Fair Determination of Risk Scores. arXiv preprint arXiv:1609.05807v2, 2016. Disponible en: <<https://arxiv.org/abs/1609.05807>>.

71 ANGWIN et al., 2016.

72 DIETERICH, William; MENDOZA, Christina; BRENNAN, Tim. COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity. Traverse City (MI): Northpointe Inc., 2016. Disponible en: <<https://www.documentcloud.org/documents/2998391-ProPublica-Commentary-Final-070616.html>>.

73 KLEINBERG; MULLAINATHAN; RAGHAVAN, 2016, p. 1-3.

74 DIETERICH; MENDOZA; BRENNAN, 2016, p. 4.

75 ANGWIN et al., 2016, p. 3-5.

adherirse a un estándar estadístico si se ignora el contexto social y jurídico en el que se aplican⁷⁶.

El debate subsiguiente entre ProPublica, Northpointe y la comunidad académica reveló una verdad matemática fundamental: cuando las tasas base de un resultado (en este caso, la reincidencia) difieren entre dos grupos poblacionales, es matemáticamente imposible que un algoritmo satisfaga simultáneamente la paridad predictiva y el equilibrio en las tasas de error (como la igualdad de tasas de falsos positivos y falsos negativos)⁷⁷. La "justicia" algorítmica, por lo tanto, no es un problema técnico que pueda resolverse con mejores datos o algoritmos más sofisticados. Autores como Wachter, Mittelstadt y Russell (2020) señalan que el problema no puede resolverse solo con datos más amplios o modelos más potentes; lo que está en juego son decisiones valorativas acerca de la distribución de riesgos y beneficios⁷⁸.

Tabla 1. Comparativa de Métricas de Justicia Algorítmica en el caso COMPAS

Métrica de justicia	Definición sencilla	Resultado en COMPAS	Implicación normativa
Paridad predictiva (calibración)	Para una puntuación de riesgo dada, la probabilidad real de reincidencia es la misma en todos los grupos.	Satisfiecha (Northpointe/Equivant) ⁷⁹	Mantiene la consistencia en el significado de las puntuaciones, pero permite que los errores se distribuyan de manera desigual.
Equilibrio en falsos positivos	La probabilidad de ser clasificado como “alto riesgo” sin reincidir es igual entre grupos.	No satisfecha (tasa mayor en acusados negros) ⁸⁰	Evita la sobrepenalización injusta, pero puede entrar en conflicto con la calibración del modelo.

76 WACHTER; MITTELSTADT; RUSSELL, 2020, p. 46–47.

77 KLEINBERG; MULLAINATHAN; RAGHAVAN, 2016, p. 1–3.

78 WACHTER; MITTELSTADT; RUSSELL, 2020, p. 46–47.

79 DIETERICH; MENDOZA; BRENNAN, 2016.

80 ANGWIN et al., 2016.

Métrica de justicia	Definición sencilla	Resultado en COMPAS	Implicación normativa
Equilibrio en falsos negativos	La probabilidad de ser clasificado como “bajo riesgo” y reincidir es igual entre grupos.	No satisfecha (tasa mayor en acusados blancos) ⁸¹	Protege a la sociedad de reincidencias no detectadas, pero aumenta la carga de falsos positivos en otros grupos.

Esta tabla visualiza el conflicto central: la "justicia" no es un término unívoco. Obliga al lector a confrontar que la elección de una métrica es una decisión política con consecuencias distributivas, desmitificando la supuesta objetividad del algoritmo.

Ahora bien, más allá del sesgo algorítmico, la digitalización de la justicia genera un nivel aún más radical de exclusión, que puede comprenderse a través del concepto de "pensamiento abisal" del sociólogo Boaventura de Sousa Santos. Este autor sostiene que el pensamiento occidental moderno establece una línea divisoria invisible que separa un el "lado metropolitano" habitado por sujetos plenos de derechos y resueltos los conflictos mediante regulación y emancipación, y un y el "lado colonial", donde los sujetos son despojados de humanidad y tratados bajo lógicas de apropiación y violencia⁸². Según Santos, el derecho moderno funciona fundamentalmente en el ámbito metropolitano, pero al mismo tiempo produce y legitima esa exclusión del ámbito colonial⁸³. El despliegue de tecnologías sin una mirada crítica puede reforzar nuevas divisiones abisales. La justicia digital separa así a quienes tienen internet, dispositivos y competencias técnicas, de quienes carecen de ellos. Los primeros viven como "ciudadanos metropolitanos", que sí experimentan mejoras en eficiencia y accesibilidad; los segundos representan a los "ciudadanos coloniales" del entorno digital —adultos mayores, comunidades rurales, migrantes o sectores empobrecidos— para quienes sistemas que exigen correo electrónico o cargar documentos en línea se vuelven imposibles⁸⁴.

81 ANGWIN et al., 2016.

82 Boaventura de Sousa Santos, *Para além do pensamento abissal: Das linhas globais a uma ecologia de saberes*, Revista Crítica de Ciências Sociais, n.º 78, outubro 2007, pp. 3-46. Disponível en: <http://journals.openedition.org/rccs/753>

83 SANTOS, 2007, p. 13-15.

84 AMAO; OGUNKUNLE; AGU; OMONIYI; SAKPERE, 2025, p. 2-5.

Esta exclusión digital no es simplemente una brecha que pueda resolverse con talleres de alfabetización tecnológica o acceso a Wi-Fi. Se trata de una exclusión estructural en la que el sistema judicial deja de ser accesible para parte de la población. En términos de Santos (2007), constituye un fenómeno abisal: el acceso mismo a los mecanismos de justicia es negado⁸⁵. Para el excluido digital, el tribunal “desaparece”. Así, la tecnología no solo replica desigualdades ya existentes, sino que las cristaliza en la arquitectura del propio Estado de derecho. La “exclusión abisal” (tal como se desarrolló anteriormente) describe una lógica de división entre una zona metropolitana en donde se aplican plenamente los derechos, la legalidad y el funcionamiento de las instituciones y una zona colonial, donde predomina la arbitrariedad y la invisibilidad jurídica. Esta dinámica se distingue especialmente en la adopción de tecnologías digitales en los sistemas judiciales del Sur Global. A diferencia de los debates del Norte Global centrados en la imparcialidad estadística de los algoritmos, en América Latina, África e India el problema central es aún más elemental: quién puede siquiera participar del sistema de justicia digital.

En India, el proyecto e-Courts busca digitalizar el acceso judicial, pero se ha mostrado que al contrario este sistema ha profundizado las desigualdades en regiones rurales, donde la conectividad es limitada y gran parte de la población no obtiene una asesoría técnica adecuada. De esta manera la digitalización lejos de resolver barreras reconfigura la exclusión, desplazándola del espacio físico al virtual sin eliminarla⁸⁶.

Por otra parte, en Brasil, el programa “Justiça 4.0” fue lanzado con la intención de ampliar el acceso a la justicia mediante herramientas de inteligencia artificial y automatización⁸⁷. No obstante, estudios han mostrado que estas soluciones digitales tienden a beneficiar principalmente a quienes ya cuentan con buena infraestructura tecnológica y competencias digitales. Por el contrario, muchas comunidades en favelas o zonas rurales, donde el acceso a internet es limitado y la alfabetización digital es baja, enfrentan dificultades reales para interactuar con los servicios judiciales en línea⁸⁸. Así, en lugar de reducir desigual-

85 SANTOS, 2007, p. 7–15.

86 SUBHAJIT BASU; CHITRA JHA. Evaluating ICT Adoption in the Indian Judiciary: Challenges, Opportunities, and the Impact of the eCourts Project. *Indian Journal of Law and Justice*, v. 15, n. 1, p. 1–25, 2024. DOI: 10.36745/ijca.343

87 WEUDER MARTINS CÂMARA; PATRÍCIA VILAR BORBA GUIMARÃES. A inovação e o uso de inteligência artificial no judiciário brasileiro: uma análise a partir do open justice e do programa justiça 4.0. *Cadernos de Direito Actual*, n.º 25 (extraordinario), p. 384-401, 2024.

88 Zaida Espinosa Zárate, Celia Camilli Trujillo y Jesús Plaza-de-la-Hoz, *Digitalization in Vulnerable Populations: A Systematic Review in Latin America, Social Indicators Research*, 170, 2023, pp. 1183-1207, <https://doi.org/10.1007/s11205-023-03239-x>

dades, la digitalización puede profundizar las brechas existentes, reproduciendo la exclusión bajo el discurso del progreso tecnológico⁸⁹.

En Colombia, el uso de plataformas digitales para la asignación de tutelas, la radicación de demandas y la realización de audiencias virtuales ha incrementado la eficiencia judicial, pero también ha generado nuevos obstáculos. Organizaciones de la sociedad civil han advertido que personas en zonas rurales, comunidades indígenas o sectores empobrecidos enfrentan barreras tecnológicas y de alfabetización que transforman el derecho a la justicia en una promesa meramente formal⁹⁰. Un estudio sobre la implementación de infraestructura tecnológica en el país reveló que esta no logró reducir la brecha digital en las áreas rurales⁹¹.

En Nigeria, el *Judiciary Information Technology Policy* ha sido presentado como una solución estructural al atraso judicial, con el objetivo de delinear una política para la aplicación de la tecnología en los tribunales y así mejorar la productividad y la confianza pública⁹². Sin embargo, fuera de las capitales regionales, la falta de infraestructura básica como el suministro eléctrico inestable, la conectividad de red deficiente y la escasez de instalaciones TIC adecuadas, impide que amplios sectores se beneficien del sistema digital⁹³. De hecho, la justicia digital tiende a consolidar un sistema dual: uno moderno y funcional para las élites y zonas con mejor infraestructura, y otro precario e inaccesible para el resto de la población, dado que muy pocas oficinas de abogados están conectadas a internet y el sistema judicial en general carece de un marco tecnológico moderno⁹⁴.

Estos ejemplos ponen de manifiesto que la digitalización judicial, cuando se aplica sin medidas correctivas, en lugar de democratizar el acceso puede generar nuevas fronteras abisales entre “ciudadanos conectados” y “ciudadanos invisibles”.

La digitalización de la justicia no constituye un proceso neutral ni meramente técnico, sino un terreno en el que se reproducen y profundizan desigualdades sociales preexistentes. Los sesgos algorítmicos se infiltran en distintas etapas del diseño y aplicación de los sistemas —desde la definición de variables y etiquetas hasta la selección de datos o el uso de variables proxy—, generando

⁸⁹ ESPINOSA ZÁRATE et al., 2023, p. 1203–1205.

⁹⁰ ESPINOSA ZÁRATE et al., 2023, p. 1198–1201.

⁹¹ ESPINOSA ZÁRATE et al., 2023, p. 1190–1193.

⁹² ESTHER E. ARCHIBONG; KENE-UNWANA TRENCHARD IBIA; GEORGE SUSANA KENNEDY; GEORGE WILLIAMS KENNEDY. Invigoration of E-Judicial System in Nigeria: Challenges and Improvement Strategies to Achieve SDG 2030. *Research Journal of Mass Communication and Information Technology*, v. 8, n. 1, p. 101–108, 2022. DOI: 10.56201/rjmcit.v8.no1.2022.pg101.108

⁹³ ARCHIBONG; IBIA; KENNEDY; KENNEDY, 2022, p. 105.

⁹⁴ ARCHIBONG; IBIA; KENNEDY; KENNEDY, 2022, p. 103.

efectos discriminatorios que afectan de manera desproporcionada a ciertos grupos sociales. Casos como el algoritmo COMPAS muestran que incluso cuando se cumplen determinadas métricas de equidad, los errores en la clasificación se distribuyen de forma desigual entre poblaciones, revela que la “justicia algorítmica” no es un problema que pueda resolverse únicamente con mejores datos o modelos más sofisticados, sino que implica decisiones normativas y éticas sobre qué tipo de injusticia se está dispuesto a tolerar.

Más allá de los sesgos técnicos, la justicia digital abre un frente de exclusión aún más radical: la exclusión estructural del acceso mismo al sistema jurídico. A la luz del “pensamiento abisal” de Boaventura de Sousa Santos, puede observarse cómo la tecnología traza nuevas líneas divisorias entre quienes cuentan con conectividad, alfabetización digital y recursos, y quienes quedan fuera de esa esfera, invisibilizados como “ciudadanos coloniales” de la era digital. Para estos últimos, el tribunal deja de existir como instancia a la que acudir, transformando la justicia en un privilegio reservado a los sujetos metropolitanos de lo digital. De este modo, la tecnología no solo reproduce desigualdades, sino que las cristaliza en la propia arquitectura del Estado de derecho, volviendo la exclusión más profunda, opaca y difícil de cuestionar.

4. HACIA UNA JUSTICIA DIGITAL DEMOCRÁTICA: PRINCIPIOS Y PROPUESTAS PARA CHILE

La crítica a la neutralidad tecnológica y la constatación de los límites de la regulación actual exigen ir más allá del diagnóstico y proponer marcos alternativos para la gobernanza de la tecnología en la justicia. No se trata de rechazar la tecnología, sino de reorientar su desarrollo e implementación para que sirva a los fines de una justicia democrática, inclusiva y sustantiva. Esto requiere un cambio de paradigma: de la optimización técnica a la deliberación política, y de la neutralidad formal a la justicia relacional.

Un punto de partida erróneo para la regulación es el principio de "neutralidad legal de la IA". Este sostiene que la ley debería tratar de forma equivalente los comportamientos de los sistemas automatizados y los de los humanos⁹⁵. Sin embargo, aunque esta postura busca maximizar beneficios de la IA, tropieza con el riesgo de una falsa equivalencia: se centra excesivamente en el individuo —ya sea persona o sistema— y depende de razonamientos analógicos, sin atender suficientemente a las diferencias sustantivas y al contexto social⁹⁶.

En contraposición, se ha planteado un enfoque relacional de la regulación tecnológica⁹⁷. Este enfoque no parte de una supuesta equivalencia, sino que

95 CARYS J. CRAIG. *The Relational Robot: A Normative Lens for AI Legal Neutrality*. *Jerusalem Review of Legal Studies*, 2022 (versión en SSRN, DOI: 10.2139/ssrn.4118849), p. 3.

96 CRAIG, 2022, p. 3–4.

97 CRAIG, 2022, p. 2, 13–16.

evalúa la tecnología en función de cómo configura y transforma las relaciones: entre el ciudadano y el Estado, entre el justiciable y el sistema judicial, y entre los propios ciudadanos. La pregunta clave no es si un algoritmo es "tan bueno como un juez", sino qué tipo de relación de poder establece un sistema algorítmico entre el Estado y la persona sometida a su juicio. ¿Es una relación de transparencia o de opacidad? ¿De deliberación o de imposición? ¿De dignidad o de cosificación?

Este enfoque relacional se complementa con la noción de derechos de neutralidad de la plataforma, desarrollada por Hannibal Travis⁹⁸. Aplicado al contexto judicial, esto implica ir más allá de los derechos negativos (protección contra la interferencia) para afirmar derechos positivos. Según Travis, debe asegurarse no solo la posibilidad formal de acceso al sistema, sino también un "derecho a ser escuchado" en condiciones de equidad: que los ciudadanos no sean invisibilizados, distorsionados o degradados por la lógica de un algoritmo. Ello implica un derecho efectivo a la contestación, a la revisión humana significativa y a una reparación frente a decisiones algorítmicas injustas, garantizando que la voz individual no quede silenciada por el código⁹⁹.

El principio de *Justicia por Diseño* propone una transformación en la forma en que se desarrollan y regulan los sistemas algorítmicos en el ámbito judicial. En lugar de limitarse constatar ex post si una herramienta vulnera derechos, se busca que los valores fundamentales del Estado de derecho —equidad, debido proceso, no discriminación, transparencia y rendición de cuentas— se integren desde el inicio del ciclo de vida tecnológico^{100 101}.

Para pasar del nivel declarativo a un marco operativo concreto, se propone articular *Justicia por Diseño* en torno a cuatro componentes complementarios y verificables:

1. Privacy by Design (PbD).

Como explica Cavoukian (2009), este enfoque sostiene que la protección de derechos debe incorporarse de manera estructural en el sistema, no añadirse como medida correctiva posterior. Sus siete principios fundacionales (proactividad, privacidad por defecto, integración sistémica, funcionalidad positiva, seguridad integral, visibilidad y respeto al usuario) funcionan también en el

⁹⁸ TRAVIS, Hannibal. *Platform Neutrality Rights: AI Censors and the Future of Freedom*. Abingdon; New York: Routledge, 2025. (Routledge Research in IT and e-Commerce Law), p 15.

⁹⁹ TRAVIS, 2025, p. 15–16.

¹⁰⁰ CAVOUKIAN, Ann. *Privacy by Design: The 7 Foundational Principles*. Toronto: Information & Privacy Commissioner of Ontario, 2009. Disponible en: <https://privacy.ucsc.edu/resources/privacy-by-design--foundational-principles.pdf>

¹⁰¹ ADA LOVELACE INSTITUTE. *The AI Act: A User Guide*. London: Ada Lovelace Institute, 2022. Disponible en: <https://www.adalovelaceinstitute.org/resource/aia-user-guide/>

ámbito judicial como guía para que los sistemas incorporen salvaguardas desde su concepción técnica¹⁰².

2. Equidad por Diseño (EbD).

Green (2022) propone superar la simple auditoría posterior y avanzar hacia la “equidad estructural”, lo que implica orientar el diseño de manera estratégica a corregir desigualdades históricas. Esto incluye definir métricas que identifiquen impactos diferenciales sobre poblaciones marginadas y prever escenarios que eviten exclusiones indirectas o sistémicas¹⁰³.

3. Diseño Participativo (PD).

El Ada Lovelace Institute (2022) muestra que incorporar la voz de múltiples actores —particularmente de comunidades afectadas— en etapas de diseño y evaluación genera legitimidad social y ayuda a identificar riesgos normativos invisibles desde una perspectiva puramente técnica¹⁰⁴.

4. Evaluación de Impacto Algorítmico (EIA).

El Gobierno de Canadá (2019) ha implementado marcos regulatorios que hacen obligatoria la EIA ex ante en servicios públicos, mediante metodologías estandarizadas y transparentes¹⁰⁵. Este mecanismo permite documentar, deliberar y revisar periódicamente los efectos sociales, jurídicos y éticos de cada sistema, bajo estándares claros y con participación multisectorial. En conjunto, estos componentes conforman un ciclo de gobernanza sustantiva: *PbD* y *EbD* definen los principios normativos que deben guiar el desarrollo; *PD* establece los actores que deben participar activamente en el proceso; y *EIA* ofrece una herramienta verificable para auditar su cumplimiento. Así, *Justicia por Diseño* no es un ideal abstracto, sino una arquitectura institucional replicable.

Un componente esencial de este marco es superar la noción difusa de “supervisión humana significativa” establecida en marcos como la Ley de IA de la Unión Europea. La literatura en psicología cognitiva ha demostrado que, en condiciones de alta carga de trabajo, los humanos tienden a aceptar las decisiones de los sistemas automatizados sin cuestionamiento (automation bias). Por tanto, se requiere una redefinición más exigente del rol humano en estos procesos.

102 CAVOUKIAN, 2009, p. 1–5.

103 BEN GREEN. Escaping the Impossibility of Fairness: From Formal to Substantive Algorithmic Fairness. *Philosophy & Technology*, v. 35, n. 90, 2022. p. 17–20. DOI: <https://doi.org/10.1007/s13347-022-00584-6>.

104 ADA LOVELACE INSTITUTE, 2022, p. 12–15.

105 CANADÁ. Government of Canada. Directive on Automated Decision-Making. Ottawa: Treasury Board of Canada Secretariat, 2019 (última mod. 2025). Disponible en: <<https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>>

El concepto de Control Humano Significativo (*Meaningful Human Control*, MHC), originalmente formulado en el contexto de los sistemas de armas autónomas, ofrece una base más sólida. Según Santoni de Sio y van den Hoven (2018), este control requiere tres condiciones fundamentales: acceso a información clara y comprensible; autoridad real para intervenir, modificar o detener la decisión automatizada; y una asignación jurídica explícita de responsabilidad. Aplicado al ámbito judicial, esto significa que jueces y funcionarios deben contar no solo con la facultad, sino también con la obligación institucional de auditar críticamente los sistemas, evitando un rol meramente ceremonial o de validación¹⁰⁶.

El Proyecto de Ley de Inteligencia Artificial en Chile (Boletín 16821-19), presentado en mayo de 2024, constituye un esfuerzo pionero para regular los sistemas de IA en clave de derechos fundamentales, innovación responsable y control institucional. Entre sus mecanismos incluye una clasificación por nivel de riesgo definida por un Consejo Técnico Asesor, sanciones de hasta 20.000 UTM y facultades de fiscalización a la futura Agencia de Protección de Datos Personales¹⁰⁷.

Desde una perspectiva crítica, este marco legal chileno puede interpretarse como un espacio propicio para la implementación de los principios de Justicia por Diseño. La clasificación de riesgos y la prohibición de ciertos usos —como la vigilancia biométrica remota o la manipulación subliminal— revelan un intento por prevenir daños desde el diseño. Igualmente, la exigencia de explicabilidad y el recurso a entornos regulatorios de tipo sandbox se alinean con las lógicas de EIA, al ofrecer herramientas para auditar anticipadamente los efectos sociales y jurídicos de la IA (Ministerio de Ciencia, 2024).

La supervisión humana obligatoria y la inclusión de sanciones legales claras sientan las bases para trascender la noción formal de supervisión, en favor de un auténtico marco de Control Humano Significativo, en la línea planteada por Santoni de Sio y van den Hoven (2018)¹⁰⁸ Muñoz Valdebenito (2024) ofrece un examen sistemático del texto legal, reconociendo fortalezas y debilidades, incluyendo la complejidad de la clasificación de riesgos, potencial rigidez frente a la rápida evolución tecnológica y falta de claridad en temas clave como la IA generativa y la propiedad intelectual. El contraste con marcos internacionales permite identificar mejores prácticas y recomendaciones para alcanzar una regu-

¹⁰⁶ SANTONI DE SIO, Filippo & VAN DEN HOVEN, Jeroen. *Meaningful Human Control over Autonomous Systems: A Philosophical Account*. *Frontiers in Robotics and AI*, vol. 5, art. 15, 2018, p. 1-15. DOI: <https://doi.org/10.3389/frobt.2018.00015>

¹⁰⁷ CHILE. MINISTERIO DE CIENCIA, TECNOLOGÍA, CONOCIMIENTO E INNOVACIÓN. *Proyecto de Ley que regula los sistemas de inteligencia artificial* [Boletín 16821-19]. Santiago de Chile, 7 mayo 2024. Disponible en: <https://minciencia.gob.cl/areas/inteligencia-artificial/Inteligencia-Artificial/Proyecto-Ley-regula-sistemas-IA/>

¹⁰⁸ SANTONI DE SIO; VAN DEN HOVEN, 2018, p. 8-10.

lación más equilibrada y adaptativa¹⁰⁹. En consecuencia, Chile tiene una oportunidad estratégica: este proyecto puede servir como un laboratorio normativo en el que converjan innovación, justicia estructural y control democrático. Sin embargo, para evitar replicar los límites de la regulación europea—marcada por el formalismo procedimental—debe reforzarse la participación ciudadana y la deliberación judicial, integrando plenamente los principios de Privacy by Design, Equidad por Diseño, Diseño Participativo y Evaluación de Impacto Algorítmico.

En el Sur Global, y en particular en Chile, el debate sobre justicia digital no puede resolverse mediante la mera importación del modelo europeo. Si bien este ha generado avances relevantes en materia de protección de datos e inteligencia artificial, su orientación procedural responde a realidades institucionales distintas, con mayores recursos regulatorios y capacidades técnicas. Aplicar estas mismas fórmulas en contextos de desigualdad estructural y limitada capacidad de fiscalización corre el riesgo de convertir la regulación en un ejercicio meramente declarativo, sin eficacia práctica para proteger derechos.

Sin embargo, para evitar replicar los límites de la regulación europea —marcada por el formalismo procedimental— debe reforzarse la participación ciudadana y la deliberación judicial, integrando plenamente los principios de Privacy by Design, Equidad por Diseño, Diseño Participativo y Evaluación de Impacto Algorítmico. Esta exigencia se explica porque el modelo europeo, aunque avanzado en términos técnicos, ha mostrado su tendencia a reducir la protección de derechos a una serie de auditorías, certificaciones y obligaciones de cumplimiento que no siempre garantizan resultados sustantivos en materia de justicia¹¹⁰. En contextos como el chileno, donde la brecha digital, la desigualdad socioeconómica y la limitada capacidad institucional de supervisión son factores estructurales, la mera transposición de ese modelo corre el riesgo de generar una “ética de papel”, incapaz de controlar realmente los impactos de la automatización en la administración de justicia.

Por ello, una aproximación contextualizada resulta indispensable. El principio de Kasirzadeh¹¹¹ *Privacy by Design*, asegura que los derechos fundamentales no se concibían como controles ex post, sino como garantías integradas en el diseño tecnológico¹¹². La *Equidad por Diseño* exige corregir desigual-

109 MUÑOZ VALDEBENITO, Claudio. *Proyecto de ley que regula la inteligencia artificial en Chile: análisis crítico hacia una innovación responsable*. Revista Chilena de Derecho y Ciencia Política, v. 15, n. 1, p. 95-120, 2024.

110 WACHTER; MITTELSTADT; RUSSELL, 2020, p. 42-47.

111 KASIRZADEH, Atoosa. *Algorithmic Fairness and Structural Injustice: Insights from Feminist Political Philosophy*. In: AAAI/ACM Conference on AI, Ethics, and Society (AIES'22), Oxford, UK, 1-3 ago. 2022. New York: ACM, 2022. p. 349-356. DOI: <https://doi.org/10.1145/3514094.3534188>

112 CAVOUKIAN, 2009, p. 1-4.

dades históricas al momento de definir métricas y objetivos algorítmicos, lo que resulta crucial en sociedades con fuertes asimetrías como las del Sur Global¹¹³. El *Diseño Participativo* amplía la legitimidad democrática al incorporar a comunidades afectadas en las fases de desarrollo y supervisión, evitando que las decisiones técnicas se tomen de forma unilateral desde el Estado o desde corporaciones tecnológicas¹¹⁴. Finalmente, la *Evaluación de Impacto Algorítmico* proporciona un mecanismo institucional de control continuo y verificable, permitiendo anticipar y corregir riesgos antes de que las vulneraciones de derechos se materialicen¹¹⁵.

En suma, la experiencia europea puede servir de referencia inicial, pero no puede adoptarse acríticamente. En el caso chileno, la clave está en aplicar estos principios de manera efectiva y sustantiva para avanzar hacia una justicia digital democrática, sensible a la desigualdad estructural y abierta al control ciudadano.

5. CONCLUSIÓN

El análisis comparativo muestra que, aunque los marcos regulatorios europeos —y en particular el español— han logrado avances relevantes la formalización de procedimientos, la protección de datos y evaluación de riesgos algorítmicos, su enfoque predominantemente procedural limita la atención a cuestiones sustantivas de poder, la equidad y participación ciudadana. En el Sur Global, la replicación acrítica de estos modelos pueden profundizar desigualdades estructurales, debido a la heterogeneidad social, la persistencia de la brecha digital y la fragilidad institucional. En el caso de Chile, la incorporación de tecnologías judiciales enfrenta el desafío de equilibrar eficiencia técnica con justicia sustantiva, lo que exige integrar los principios de Privacy by Design, Equidad por Diseño, Diseño Participativo y Evaluación de Impacto Algorítmico, al tiempo que se consolida la deliberación judicial y se amplía la participación ciudadana.

En todos los casos, la investigación subraya que la neutralidad tecnológica es una falacia: cada decisión de diseño—desde la elección de variables hasta la interfaz de usuario—constituye un acto político con consecuencias distributivas. En consecuencia, justicia digital no debe reducirse a una lógica de optimización técnica, sino orientarse hacia criterios de inclusión, transparencia, responsabilidad y deliberación democrática. Asumir esta perspectiva requiere un cambio de paradigma: priorizar derechos sustantivos sobre eficiencia, implementar auditorías vinculantes, supervisión humana significativa y diseñar sistemas que reflejen los valores sociales y democráticos del contexto específico.

¹¹³ KASIRZADEH, 2022, p. 351–353.

¹¹⁴ ADA LOVELACE INSTITUTE, 2022, p. 12–15.

¹¹⁵ CANADÁ, 2019, última mod. 2025, p. 4–6.

La tecnología no es un destino inevitable, sino una elección regulatoria y política. El modo en que se implemente, determinará si la justicia digital reproduce desigualdades existentes o si, por el contrario, contribuye a fortalecer la equidad y la democracia en cada región. Europa y España ofrecen referencias normativas útiles, aunque limitadas; los países del Sur Global deben desarrollar modelos adaptados a sus realidades socio institucionales; y Chile tiene, ante sí, la oportunidad de liderar un enfoque de justicia digital democrática que combine control técnico, participación ciudadana y principios de equidad contextual.

REFERENCIAS

ADA LOVELACE INSTITUTE. The AI Act: A User Guide. London: Ada Lovelace Institute, 2022. Disponible en: <https://www.adalovelaceinstitute.org/resource/aia-user-guide/>. Acceso en: 2 set. 2025.

AMAO, D. K.; OGUNKUNLE, O. O.; AGU, J. N.; OMONIYI, A. D.; SAKPERE, W. Role of technology in a social, economic and ethical environment: jurisprudence approaches to legal reasoning. African Journal of Law, Political Research and Administration, v. 8, n. 2, p. 1-17, 2025. Disponible en: https://www.researchgate.net/publication/391365746_Role_of_Technology_in_a_Social_Economic_and_Ethical_Environment_Jurisprudence_Approaches_to_Legal_Reasoning. Acceso en: 2 set. 2025.

AMERICAN BAR ASSOCIATION (ABA). Formal Opinion 512 on Lawyers' Use of Generative Artificial Intelligence. Chicago: American Bar Association, 2024.

ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren. Machine bias: risk assessments in criminal sentencing. ProPublica, 23 May 2016. Disponible en: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acceso en: 2 set. 2025.

APPELMAN, Naomi; Ó FATHAIGH, Ronan; VAN HOBOKEN, Joris V. J. Social welfare, risk profiling and fundamental rights: the case of SyRI in the Netherlands. Journal of Intellectual Property, Information Technology and E-Commerce Law (JIPITEC), v. 12, n. 4, 2021. Disponible en: <https://ssrn.com/abstract=3984935>. Acceso en: 2 set. 2025.

ARCHIBONG, Esther E.; IBIA, Kene-Unwana Trenchard; KENNEDY, George Susana; KENNEDY, George Williams. Invigoration of e-judicial system in Nigeria: challenges and improvement strategies to achieve SDG 2030. Research Journal of Mass Communication and Information Technology, v. 8, n. 1, p. 101-108, 2022.

BASU, Subhajit; JHA, Chitra. Evaluating ICT adoption in the Indian judiciary: challenges, opportunities, and the impact of the eCourts project. *Indian Journal of Law and Justice*, v. 15, n. 1, p. 1-26, 2024.

CÂMARA, Weuder Martins; GUIMARÃES, Patrícia Vilar Borba. A inovação e o uso de inteligência artificial no judiciário brasileiro: uma análise a partir do open justice e do programa Justiça 4.0. *Cadernos de Dereito Actual*, n. 25 (extraordinario), p. 384-401, 2024. DOI: 10.5281/zenodo.13926000.

CAVOUKIAN, Ann. Privacy by design: the 7 foundational principles. Toronto: Information & Privacy Commissioner of Ontario, 2009. Disponible en: <https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf>. Acceso en: 2 set. 2025.

CEFALIELLO, Aude; KULLMANN, Miriam. Offering false security: how the draft Artificial Intelligence Act undermines fundamental workers' rights. *European Labour Law Journal*, v. 13, n. 4, p. 542-562, 2022. DOI: <https://doi.org/10.1177/20319525221114474>

CORREAS, Óscar. Introducción a la crítica del derecho moderno. México: Fontamara, 2006.

CRAIG, Carys J. The relational robot: a normative lens for AI legal neutrality. *Jerusalem Review of Legal Studies*, 2022. Disponible en: <https://ssrn.com/abstract=4118849>. Acceso en: 2 set. 2025.

CRUZ RIVERO, Diego. Disputabilidad y equidad en los mercados digitales: una visión de Derecho europeo. *Revista Chilena de Derecho y Ciencia Política*, v. 15, n. 1, art. 17, 2024. Epub 17 jul. 2024. DOI: <https://dx.doi.org/10.7770/rchdcv15n1-art393>

DAVIS, Peter Alexander Earls; SCHWEMER, Sebastian Felix. Rethinking decisions under Article 22 of the GDPR: implications for semi-automated legal decision-making. In: INTERNATIONAL WORKSHOP ON ARTIFICIAL INTELLIGENCE AND INTELLIGENT ASSISTANCE FOR LEGAL PROFESSIONALS IN THE DIGITAL WORKPLACE (LEGALAIIA 2023), 3., 2023, Braga. Proceedings.... Braga: ICAIL, 2023. p. 1-9. Disponible en: <https://ceur-ws.org/Vol-3420/paper1.pdf>. Acceso en: 2 set. 2025.

DIETERICH, William; MENDOZA, Christina; BRENNAN, Tim. COMPAS risk scales: demonstrating accuracy, equity and predictive parity. Northpointe Inc., 2016. Disponible en: <https://www.equivant.com/compass-risk-scales-demonstrating-accuracy-equity-and-predictive-parity/>. Acceso en: 2 set. 2025.

EGUÍLUZ CASTAÑEIRA, Juan Antonio. Desafíos y retos que plantean las decisiones automatizadas y los perfilados para los derechos fundamentales. Estu-

dios de Deusto, v. 68, n. 2, p. 325-367, 2020. DOI: [https://doi.org/10.18543/ed-68\(2\)-2020pp325-367](https://doi.org/10.18543/ed-68(2)-2020pp325-367)

ESPINOSA ZÁRATE, Zaida; CAMILLI TRUJILLO, Celia; PLAZA-DE-LA-HOZ, Jesús. Digitalization in vulnerable populations: a systematic review in Latin America. *Social Indicators Research*, v. 170, p. 1183-1207, 2023. DOI: <https://doi.org/10.1007/s11205-023-03239-x>

FLORES RIVAS, Juan Carlos; AGUIRREZABAL GRÜNSTEIN, Maite. Análisis de algunos aspectos procesales del sistema de resolución de controversias en concesiones de obra pública. *Revista Chilena de Derecho y Ciencia Política*, v. 16, n. 1, art. 06, 2025. Epub 15 abr. 2025. DOI: <https://dx.doi.org/10.7770/rchdcv16n1-art412>

GAO, Qifan; AN, Hongyu. Technology-neutral illusion: the ethical and social challenges in the age of artificial intelligence. *Sociology, Philosophy and Psychology*, v. 1, n. 2, p. 33-40, 2024. DOI: 10.70267/f, Ben. Escaping the impossibility of fairness: from formal to substantive algorithmic fairness. *Philosophy & Technology*, v. 35, n. 90, 2022. DOI: <https://doi.org/10.1007/s13347-022-00584-6>

CANADÁ. Directive on Automated Decision-Making. Ottawa: Treasury Board of Canada Secretariat, 2019. Atualizada em jun. 2025. Disponível en: <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>. Acceso en: 2 set. 2025.

KASIRZADEH, Atoosa. Algorithmic fairness and structural injustice: insights from feminist political philosophy. In: AAAI/ACM CONFERENCE ON AI, ETHICS, AND SOCIETY (AIES'22), 2022, Oxford. Proceedings.... New York: Association for Computing Machinery, 2022. p. 349-356. DOI: <https://doi.org/10.1145/3514094.3534188>

KLEINBERG, Jon; MULLAINATHAN, Sendhil; RAGHAVAN, Manish. Inherent trade-offs in the fair determination of risk scores. arXiv preprint, arXiv:1609.05807, 2016. Disponível en: <https://arxiv.org/abs/1609.05807> . Acceso en: 2 set. 2025.

LESSIG, Lawrence. Code and other laws of cyberspace. New York: Basic Books, 1999.

LEVITINA, Anna. Humans in automated decision-making under the GDPR and AI Act. CIDOB Policy Report, 2024. Disponível en: <https://www.cidob.org/en/publications/humans-automated-decision-making-under-gdpr-and-ai-act> . Acceso en: 2 set. 2025.

MAGESH, Varun; SURANI, Faiyaz; DAHL, Matt; SUZGUN, Mirac; MANNING, Christopher D.; HO, Daniel E. Hallucination-free? assessing the reliability of leading AI legal research tools. *Journal of Empirical Legal Studies*, 2025. DOI: <https://doi.org/10.1111/jels.12413>

MARCUSE, Herbert. One-dimensional man: studies in the ideology of advanced industrial society. Boston: Beacon Press, 1964.

MARTÍNEZ MONTENEGRO, Isnel. Sobre los métodos de la investigación jurídica. Revista Chilena de Derecho y Ciencia Política, v. 14, n. 1, art. 01, 2023. Epub 30 jun. 2023. DOI: <https://dx.doi.org/10.7770/rchdcv14n1-art312>

MARTÍNEZ MONTENEGRO, Isnel. El problema de investigación jurídica: claves metodológicas para una formulación pertinente. Revista Chilena de Derecho y Ciencia Política, v. 16, n. 1, art. 09, 2025. Epub 28 jun. 2025. DOI: <https://dx.doi.org/10.7770/rchdcv16n1-art503>

CHILE. MINISTERIO DE CIENCIA, TECNOLOGÍA, CONOCIMIENTO E INNOVACIÓN. Proyecto de ley que regula los sistemas de inteligencia artificial [Boletín 16821-19]. Santiago de Chile, 7 mayo 2024. Disponible en: <https://minciencia.gob.cl/areas/inteligencia-artificial/Inteligencia-Artificial/Proyecto-Ley-regula-sistemas-IA/>. Acceso en: 2 set. 2025.

O'NEIL, Cathy. Armas de destrucción matemática. Madrid: Capitán Swing, 2017. Disponible en: <https://www.capitanswing.com/libros/armas-de-destrucion-matematica/>. Acceso en: 2 set. 2025.

SANTONI DE SIO, Filippo; VAN DEN HOVEN, Jeroen. Meaningful human control over autonomous systems: a philosophical account. *Frontiers in Robotics and AI*, v. 5, art. 15, p. 1-15, 2018. DOI: <https://doi.org/10.3389/frobt.2018.00015>

SOUSA SANTOS, Boaventura de. Más allá del pensamiento abisal: de las líneas globales a una ecología de saberes. *Revista Crítica de Ciências Sociais*, n. 78, p. 3-46, 2007. DOI: <https://doi.org/10.4000/rccs.753>

SQUELLA NARDUCCI, Agustín; VALENZUELA CORI, Rodrigo. Seminario “La justificación de las decisiones judiciales”, Facultad de Ciencias Jurídicas y Sociales, Universidad Austral de Chile, 23 mar. 2005. *Revista de Derecho (Valdivia)*, v. 19, n. 1, p. 277-292, jul. 2006. Disponible en: http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-09502006000100015-&lng=en&nrm-iso. Acceso en: 2 set. 2025. DOI: <http://dx.doi.org/10.4067/S0718-09502006000100015>

TRAVIS, Hannibal. Platform neutrality rights: AI censors and the future of freedom. Abingdon; New York: Routledge, 2025. (Routledge Research in IT and e-Commerce Law). DOI: <https://doi.org/10.4324/9781003496106>

WACHTER, Sandra; MITTELSTADT, Brent; RUSSELL, Chris. Why fairness cannot be automated: bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review*, v. 41, art. 105567, 2021. Preprint disponible en SSRN (2020). DOI: <https://doi.org/10.2139/ssrn.3547922>

WINNER, Langdon. *The whale and the reactor: a search for limits in an age of high technology*. Chicago: University of Chicago Press, 1986.

JURISPRUDENCIA

BRODERICK, P. Kevin. Order to Show Cause, Sanctions. Documento judicial en el caso Mata v. Avianca, Inc., No. 22-CV-1461 (PKC). New York: United States District Court, S.D.N.Y., 2023.

UNITED STATES DISTRICT COURT (Southern District of New York). Mata v. Avianca, Inc., No. 22-CV-1461 (PKC), 2023 WL 4114965, 22 jun. 2023.

ESPAÑA. TRIBUNAL CONSTITUCIONAL. Sentencia 76/2019, de 22 de mayo. Madrid: Tribunal Constitucional, 2019.

TRIBUNAL DE DISTRITO DE LA HAYA. Sentencia de 5 de febrero de 2020, ECLI:NL:RBDHA:2020:1878 (Asunto SyRI). La Haya: Tribunal de Distrito, 2020.

NORMAS Y TRATADOS

CONSTITUCIÓN ESPAÑOLA. Madrid: Jefatura del Estado, 1978.

CONVENIO EUROPEO DE DERECHOS HUMANOS (CEDH). Roma, 1950.

UNIÓN EUROPEA. Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (Reglamento general de protección de datos). Diario Oficial de la Unión Europea, L 119, 4 may. 2016.

UNIÓN EUROPEA. Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo, de 19 de octubre de 2022, relativo a un mercado único de servicios digitales (Ley de Servicios Digitales). Diario Oficial de la Unión Europea, L 277, 27 oct. 2022.

UNIÓN EUROPEA. Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de marzo de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial). Diario Oficial de la Unión Europea, L 148, 30 mayo 2024.

EUROPEAN PARLIAMENTARY RESEARCH SERVICE. Understanding algorithmic decision-making: opportunities and challenges. Brussels: European Parliament, Mar. 2019. Disponible en: [https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU\(2019\)624261_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf). Acceso en: 2 set. 2025.

