

## INTELIGENCIA ARTIFICIAL Y DERECHOS HUMANOS: UNA REGULACIÓN ÉTICA PARA LA PROTECCIÓN DEL ESTADO DE DERECHO

### **Coordinadores**

Alejandra Cedallin Martínez-Sánchez  
David Espinosa-González

### **Introducción**

Ma. Lilia López López

### **Autores**

Albano Torres-Gómez  
Alejandra Cedallin Martínez-Sánchez  
Bryan Azcué Hoyos  
Claudia Lorena Morffi Collado  
David Espinosa-González  
Francisco Raúl Matus Ramírez  
Gabriel Pérez-Galmiche  
Johanna Mildred Pinto Garcia  
Marcos Cruz González  
María de Jesús Méndez Verduzco  
Omar Fernando Becerra Partida



REVISTA **IUS**

REVISTA DE INVESTIGACIÓN JURÍDICA

nº **57**

**NUEVA ÉPOCA, VOLUMEN 21, No. 57, ENERO-OCTUBRE 2026.**



Revista IUS, del Centro Internacional de Estudios sobre Ley y Derecho, CIELDE A.C.

**DIRECTOR EDITORIAL**

Cristian Badillo Gutiérrez

Centro Internacional de Estudios Sobre Ley y Derecho, A.C.

**EDITOR ACADÉMICO**

Agustín Herrera Fragoso

Universidad Popular Autónoma del Estado de Puebla-UPAEP

**CONSEJO ACADÉMICO**

Agustín Antonio Herrera Fragoso

Alberto Patiño Reyes

Antonio Casciano

Francesco Tommaso Trabalza Contessini

Javier Dondé Matute

Jorge Ulises Carmona Tinoco

José Antonio Guevara Bermúdez

José Gabriel González Merlano

Max Silva Abbott

Úrsula Cristina Basset

Víctor Rojas Amandi

**COORDINADORES NÚMERO 57**

Alejandra Cedallín Martínez-Sánchez

David Espinosa-González

ISSN versión impresa: 2594-2816. ISSN versión electrónica: 1870-2147. Reserva de Derechos al Uso Exclusivo 04-2018-060313134100-203, otorgados por el Instituto Nacional del Derecho de Autor de la Secretaría de Cultura.

Indizada en: CONAHCYT, Redalyc, Biblat, Latindex, Scielo Citation Index, Scielo México, UNAM, vlex.

REVISTA IUS, vol. 21. no. 57, enero-octubre 2026, es una difusión periódica semestral editada por el Centro Internacional de Estudios sobre Ley y Derecho, CIELDE A.C., con domicilio en Priv. de los Industriales #111, Int. 102b, col. Jurica, Querétaro, Qro. C.P. 76100. Tel.: 442 241 9413, [contacto@revistaius.com](mailto:contacto@revistaius.com).

Las opiniones expresadas por los autores no necesariamente reflejan la postura del editor de la publicación.

[www.revistaius.com](http://www.revistaius.com)



Revista IUS está distribuido bajo una Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional.

## ÍNDICE

### Introducción | 7

*Ma. Lilia López López*

### Empleo de la IA por estudiantes universitarios: En búsqueda del empleo ético-regulado para la contribución al derecho humano a la educación | 11

*Use of AI by university students: In search of ethical and regulated use for the contribution to the human right to education*

*Alejandra Cedallin Martínez-Sánchez, David Espinosa-González y Gabriel Pérez-Galmiche*

### Reconfiguración de libertades fundamentales en la era de la inteligencia artificial:

### Vigilancia, poder algorítmico y crisis del estado de derecho | 29

*Reconfigurations of fundamental freedoms in the age of artificial intelligence: Surveillance, algorithmic power, and the crisis of the rule of law*  
*Bryan Azcué Hoyos*

### Sesgos algorítmicos en sistemas educativos con inteligencia artificial: Implicaciones éticas, jurídicas y pedagógicas desde la literatura reciente | 49

*Algorithmic bias in educational systems with artificial intelligence: Ethical, legal, and pedagogical implications from recent literature*

*Albano Torres-Gómez*

### Desarrollo y regulación de la inteligencia artificial en México: Un análisis jurídico de sus avances y retos con base en los derechos humanos | 65

*Development and regulation of artificial intelligence in Mexico: A legal analysis of its advances and challenges based on human rights*  
*Marcos Cruz González*

### Los derechos humanos de la nación mexicana: Rompiendo paradigmas teóricos para el pleno ejercicio de derechos mediante la inteligencia artificial | 87

*The human rights of the Mexican nation: Breaking theoretical paradigms for the full exercise of rights through artificial intelligence*

*David Espinosa-González, Alejandra Cedallin Martínez-Sánchez y Gabriel Pérez-Galmiche*

### Predecir no es juzgar: Sistemas algorítmicos de riesgo penal y presunción de inocencia en América Latina | 101

*Predicting is not judging: Algorithmic systems of criminal risk and presumption of innocence in Latin America*

*Claudia Lorena Morffi Collado*

### Traducción normativa en sistemas algorítmicos: Insuficiencias del paradigma de derechos humanos por diseño y bases para una teoría fuerte | 123

*Normative translation in algorithmic systems: Structural shortcomings of the human rights by design paradigm and foundations for a strong theory*

*Johanna Mildred Pinto Garcia*

### Inteligencia artificial y la protección del patrimonio neuro-genómico: Hacia una regulación ética frente a los nuevos desafíos del estado de derecho | 143

*Artificial intelligence and the protection of neuro-genomic heritage: Towards ethical regulation in the face of the new challenges to the rule of law*

*Omar Fernando Becerra Partida y María de Jesús Méndez Verduzco*

### Identidad narrativa y derechos humanos: Una aproximación hermenéutica a la migración forzada y la inteligencia artificial | 155

*Narrative identity and human rights: A hermeneutical approach to forced migration and artificial intelligence*

*Francisco Raúl Matus Ramírez*



## INTRODUCCIÓN

Esta obra no pudo haber llegado en mejor momento, dada la presencia de la inteligencia artificial en diferentes contextos de nuestra vida profesional. Actualmente, estamos en contacto con diferentes artefactos y aplicaciones que utilizan IA; dichas herramientas evolucionan cada vez más rápido y su uso desregulado puede generar debates éticos, jurídicos y pedagógicos. Durante la lectura de la presente obra, nos adentraremos en diferentes ideas, posturas y sugerencias respecto al uso desmedido de la IA. En el primer capítulo, los autores nos invitan a reflexionar sobre la ausencia de marcos normativos en el uso de la IA, lo cual incide en la formación académica y en los derechos del estudiantado en México. Con la información recabada, los autores argumentan que, tanto por la falta de normativas institucionales como federales, se vulnera el derecho humano a la educación.

En el segundo capítulo, la incorporación de la IA en las decisiones sobre crédito, empleo o programas sociales es una realidad que reta las garantías jurídicas tradicionales, lo que implicaría una reconfiguración estructural de las libertades fundamentales, alterando la autonomía, la privacidad y la igualdad. Por ello, los autores se plantean una reconfiguración normativa basada en derechos humanos que demanda centralizar a la persona, la explicabilidad de los modelos, la rendición de cuentas y auditorías independientes para garantizar la equidad. Esto conduce a una perspectiva donde la IA no es un trámite técnico, sino una decisión ética y política sobre la protección de la dignidad y la libertad humana.

Desde la perspectiva ética, en el capítulo 3, el autor nos invita a reflexionar sobre la integración de la IA en la educación, por medio de plataformas adaptativas y sistemas predictivos, cuestionando si la automatización amplía derechos o perpetúa desigualdades. Ante esta situación, se nos muestra una revisión crítica de las implicaciones éticas, jurídicas y pedagógicas de la integración de la IA en la educación. El planteamiento que se propone se centra en el sesgo algorítmico que surge en la educación y que no es precisamente un error en la estadística o la parte técnica, sino que se percibe como un fenómeno sociotécnico multidimensional en relación con los derechos humanos, la no discriminación, la protección de datos personales, la transparencia y la justicia educativa. Finalmente, se nos propone un marco de gobernanza algorítmica articulado en siete criterios clave, apelando al diseño de una IA alineada con el Estado de derecho y la equidad social.

Sin duda, con los innumerables avances tecnológicos se vislumbran profundos retos éticos y jurídicos que requieren un modelo normativo enfocado en la dignidad humana; el caso del estudio del autor Cruz González nos presenta la postura del Estado mexicano frente a esta vertiginosa transformación. En este apartado, nos menciona que los derechos públicos subjetivos deben ser la base que guíe la regulación, así como el desarrollo soberano de la tecnología. Mediante un enfoque metodológico mixto, el trabajo examina el panorama normativo en México, resaltando las iniciativas parlamentarias y los marcos éticos. Asimismo, evalúa la posición del país en América Latina basándose en indicadores de la CEPAL, mostrando tanto las fortalezas como las áreas de oportunidad para alcanzar

la soberanía tecnológica. El autor nos muestra una visión contextualizada y soberana que impulsa la economía basada en el conocimiento. Sin duda, esta obra constituye una lectura imprescindible para juristas, legisladores y académicos interesados en el desarrollo tecnológico con la protección de la dignidad humana ante la IA.

Frente a la recurrente vulneración de los derechos humanos en el país, el capítulo 5 de esta obra ofrece una aportación dogmática al derecho constitucional en México. Se propone considerar a la Nación mexicana como un nuevo sujeto de derecho, entendido como una persona moral, que, al igual que las personas físicas, puede ser titular de derechos humanos. A través de una metodología analítico-sintética de carácter documental, los autores articulan una argumentación que enlaza el concepto sociológico y político de la nación con el ordenamiento jurídico mexicano. El capítulo concluye con un análisis sobre la IA como una herramienta clave para garantizar y potenciar los derechos civiles, políticos, económicos, sociales y culturales, permitiendo el debate respecto al rol del Estado en el acceso equitativo a tecnologías emergentes.

Siguiendo esta misma vertiente, el capítulo subsecuente nos ofrece un análisis de los casos de Argentina, Colombia, Brasil y Chile, en donde las tecnologías no han permeado una modernización garantista, sino la reproducción de desigualdades estructurales. Se introduce el concepto de «positivismo tecnificado», donde la autoridad del dato cuantitativo reemplaza paulatina y progresivamente a la argumentación jurídica. Ante los riesgos de los errores algorítmicos, el estudio propone tres ejes normativos indispensables: explicabilidad, auditabilidad continua y supervisión humana efectiva. De esta manera, la obra se consolida como un referente fundamental para subordinar el desarrollo tecnológico a los principios del Estado de derecho.

Asimismo, esta obra nos expone un análisis sobre las limitaciones conceptuales de integrar los derechos fundamentales en el desarrollo de la IA. Sostiene que el enfoque que impera por diseño adolece de una fragmentación y reduce los conceptos jurídicos a simples métricas computacionales. Por ello, para superar esta brecha entre el lenguaje cualitativo del derecho y el código formalizado, la obra plantea una teoría de la traducción normativa articulada en cuatro niveles: principios jurídicos, normas operativas, especificaciones técnicas e implementación sistémica. Además, mantiene una mirada crítica desde el Sur Global que alerta sobre la colonialidad algorítmica y la imposición de estándares del Norte Global en realidades institucionales asimétricas. Con esta propuesta procedimental, este estudio se posiciona como un aporte imprescindible para garantizar la trazabilidad y la constitucionalización efectiva de la IA en el ciclo de vida tecnológico.

Por otra parte, el procesamiento de datos por medio de la IA ha puesto de manifiesto problemas evidentes para las garantías del Estado de derecho. En uno de los capítulos finales, se nos invita a reflexionar sobre la importancia de proteger la integridad biológica y mental frente a los avances de los algoritmos de aprendizaje. Se sostiene que el marco normativo actual de protección de datos resulta insuficiente ante la capacidad predictiva y de reidentificación de la IA, lo cual afecta a la autonomía y a la intimidad individual. Por ello, ante estos riesgos, se propone reconocer los neuroderechos a nivel constitucional e integrar el concepto de neurodato en la legislación.

Finalmente, la presente obra nos lleva a tomar conciencia del papel de la IA ante su uso desmedido y no normado. Hay que reconocer que la IA presenta ventajas y bondades

en nuestro trabajo profesional; sin embargo, también presenta riesgos y situaciones complejas; por ende, esta obra es una invitación para reflexionar sobre lo mencionado y prepararnos para futuros contextos.

En sus manos, estimado lector, se presentan diversos puntos de vista acerca de la IA a la luz de la multidisciplinariedad, buscando contribuir a la solución de problemáticas identificadas desde un enfoque científico y académico.

**Dra. Ma. Lilia López López.**



## Empleo de la IA por estudiantes universitarios: En búsqueda del empleo ético-regulado para la contribución al derecho humano a la educación\*

*Use of AI by university students: In search of ethical and regulated use for the contribution to the human right to education*

Alejandra Cedallin Martínez – Sánchez\*\* David Espinosa – González\*\*\* Gabriel Pérez – Galmiche\*\*\*\*

### RESUMEN

La Inteligencia Artificial (IA) representa una herramienta de gran aportación para el ámbito académico e investigativo. Empero, la inexistencia de bases regulatorias basadas en la ética y principios axiológicos que guíen a las (os) estudiantes, docentes e investigadoras (es), deja a la libre consideración, interpretación y posturas de quienes emplean la información emanada por las plataformas de libre acceso de IA. El presente artículo se fundamenta en una metodología cuantitativa, el método analítico-sintético y la encuesta, diseñando un cuestionario que se dirige a una muestra de 129 estudiantes universitarios de la licenciatura en Administración de Empresas del Complejo Regional Mixteca (Benemérita Universidad Autónoma de Puebla). Se obtiene como principal resultado que la inexistencia de un reglamento federal, estatal y/o institucional implicaría una violación al derecho humano a la educación del estudiantado e incluso podría impactar en sus trayectorias escolares. También se genera la posibilidad de recaer en plagio ante el desconocimiento del tratamiento de información emanada de IA.

### PALABRAS CLAVE

Inteligencia Artificial; Derechos Humanos; Reglamento.

### ABSTRACT

Artificial Intelligence (AI) represents a highly significant tool for the academic and research fields. However, the lack of regulatory frameworks based on ethics and axiological principles to guide students, faculty, and researchers leaves the use of information generated by open-access AI platforms to the individual discretion, interpretation, and stance of the user. This article is based on a quantitative methodology, employing the analytical-synthetic method and the survey technique. A questionnaire was designed and administered to a sample of 129 undergraduate Business Administration students from the Complejo Regional Mixteca (Benemérita Universidad Autónoma de Puebla). The primary finding indicates that the absence of federal, state, and/or institutional regulations would constitute a violation of the students' human right to education and could even impact their academic trajectories. Furthermore, it creates the potential for plagiarism due to a lack of knowledge regarding the proper handling of information sourced from AI.

### KEYWORDS

Artificial Intelligence; Human Rights; Regulation.

\* Artículo de Investigación

\*\* Doctora en Educación y Profesora investigadora en la Benemérita Universidad Autónoma de Puebla (alejandra.martinezsa@correo.buap.mx). <https://orcid.org/0000-0002-7798-7670>.

\*\*\* Doctor en Derecho y Educación, maestro en derecho, y abogado, notario y actuario por la Benemérita Universidad Autónoma de Puebla. (david.espinosa@correo.buap.mx). <https://orcid.org/0000-0002-6993-3349>

\*\*\*\* Doctor en derecho y profesor investigador de la Benemérita Universidad Autónoma de Puebla (gabriel.galmiche@correo.buap.mx). <https://orcid.org/0000-0003-0843-1887>.

## SUMARIO

1. Introducción
2. Generalidades teórico-conceptuales de la investigación
3. Metodología
4. Resultados
5. Discusión
6. Conclusiones
7. Referencias bibliográficas

### 1. INTRODUCCIÓN

La presente investigación expone un tema de importancia jurídica y ética con dirección al empleo de la información emanada por la Inteligencia Artificial, ya que actualmente, únicamente existen propuestas de leyes y reglamentos para el acceso a las plataformas, empleo y tratamiento de la información emanada de Inteligencia Artificial (IA), ya que al basarse en la postura individual, no olvidando las posturas de los diversos formatos de cita, puede existir desde plagio por desconocimiento (indirecto) hasta la identificación de violación de derechos humanos del estudiantado, docentes e investigadoras (es) ante la omisión de establecer lineamientos normativos que regulen éticamente el empleo de esta información.

En este sentido, la presente investigación plantea el objetivo principal de "analizar el impacto de la inexistencia de reglamentos dirigidos al acceso, empleo y tratamiento de información emanada de IA desde la perspectiva constitucional y educativa", formulando la pregunta de investigación: "¿Cuál es el impacto de la inexistencia de un reglamento en territorio mexicano que regule el acceso, empleo y tratamiento de información emanada de la IA desde la postura estudiantil y constitucional?", y adoptando la postura hipotética principal de que "la inexistencia de una reglamentación en el territorio mexicano, contribuye a la violación del derecho humano a la educación por omisión de la autoridad mexicana, generando brechas de diversa índole e influenciando negativamente al cumplimiento satisfactorio de trayectorias escolares del estudiantado universitario".

De igual forma, la presente investigación, se estructura en cuatro partes, donde la primera, expone la base teórica conceptual; en la segunda, se menciona la metodología empleada; continuando, con el tercer apartado, se expresan los resultados entrelazados con bases teóricas sustanciales, y finalmente, se expone la discusión final.

El lector tiene en sus manos un tema que se pone en la mesa de discusión, donde el ámbito jurídico y educativo, una vez más exigen medios reglamentarios suficientes y viables para el ejercicio pleno del Derecho a la Educación.

### 2. GENERALIDADES TEÓRICO-CONCEPTUALES DE LA INVESTIGACIÓN

La presente investigación aborda una problemática prioritaria, el empleo de la Inteligencia Artificial (IA) sin una regulación o enfoque ético. Lo que impacta de manera directa e indirecta al Derecho Humano a la Educación en México.

Por lo que es indispensable partir de conocer el término de Inteligencia Artificial (IA), considerado por el Diccionario Panhispánico como la "disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico"<sup>1</sup>.

De igual forma, puede considerarse como "el medio por el cual las computadoras, los robots y otros dispositivos realizan tareas que normalmente requieren de la inteligencia humana"<sup>2</sup>.

En este sentido, se le preguntó a ChatGPT cómo podría definirse, por lo que emitió diversos enfoques:

Técnico: Sistema computacional basado en redes neuronales profundas y aprendizaje automático. Funcional: Herramienta capaz de interpretar lenguaje humano y generar respuestas útiles. Jurídico-tecnológico: Agente digital automatizado que procesa datos y emite resultados algorítmicos sin personalidad jurídica propia. Filosófico: Simulación computacional de ciertas capacidades cognitivas humanas, particularmente relacionadas con el lenguaje y el razonamiento.<sup>3</sup>

Considerando otro término valioso, es posible considerar "reglamento" como una "norma que rige la organización y funcionamiento de cualquier establecimiento o institución, públicos o privados"<sup>4</sup>. También, puede ser definido como "una regla obligatoria impuesta por una autoridad pública que no es el Parlamento"<sup>5</sup>.

Continuando con el hilo conceptual, los Derechos Humanos, parte esencial de la presente investigación, pueden definirse como los "derechos inherentes a todos los seres humanos, sin distinción alguna de nacionalidad, lugar de residencia, sexo, origen nacional o étnico, color, religión, lengua, o cualquier otra condición"<sup>6</sup>.

No obstante, es posible retomar una postura más actual teniendo la mirada en los precedentes que impulsan la actual teoría de Derechos Humanos:

Principios humanos con fundamento internacional ubicados en diversos instrumentos u objetos, que representan la responsabilidad de protección de la existencia del ser y sus necesidades básicas de existencia. Adecuados al tiempo y espacio.<sup>7</sup>

Finalmente, se considera oportuno mencionar lo comprendido por "Derecho humano a la educación":

La educación es, en sí misma, una herramienta esencial para el desarrollo humano individual y social. La educación es indispensable para el ejercicio de otros derechos humanos, tales como el acceso al trabajo, el derecho a la participación en los asuntos públicos, el cuidado de la salud, etc. Mientras, la dimensión social de la educación es determinante para la igualdad de género, la integración social, promover la paz y el bienestar colectivo.<sup>8</sup>

<sup>1</sup> REAL ACADEMIA ESPAÑOLA. Diccionario panhispánico del español jurídico, voz «Inteligencia artificial». España, 2025.

<sup>2</sup> GÓMEZ HERRERA, Renato. La inteligencia artificial ¿Hacia dónde nos lleva? ¿Cómo ves? Divulgación de la ciencia. México: UNAM, 7 de mayo de 2026.

<sup>3</sup> OPENAI. ChatGPT 4. Modelo de lenguaje grande. Contenido generado el 7 de mayo de 2026. Prompt: «¿Cómo podrías definirte?».

<sup>4</sup> REAL ACADEMIA ESPAÑOLA, op. cit., nota 1, voz «reglamento».

<sup>5</sup> DE PINA, Rafael; DE PINA VARA, Rafael. Diccionario de Derecho. México: Porrúa, p. 436.

<sup>6</sup> NACIONES UNIDAS. ¿Qué son los derechos humanos? Oficina del Alto Comisionado. México, 2026.

<sup>7</sup> PÉREZ-GALMICHE, Gabriel; ESPINOSA-GONZÁLEZ, David. La crisis de los Derechos Humanos: Estudio comparativo México-Italia. Ciudad de México: Montiel & Soriano Editores y BUAP, 2021, p. 257.

<sup>8</sup> NACIONES UNIDAS. El derecho humano a la educación. Oficina del Alto Comisionado, América del Sur, 2022, p. 1.

Recuperando estos elementos conceptuales, se ha construido un andamiaje teórico que sustenta el argumento fundamental de esta investigación: el acceso a las plataformas de Inteligencia Artificial por parte de los universitarios sugiere un riesgo constante ante la falta de reglamentos federales, estatales e institucionales, basándose en la postura ética de las (os) estudiantes y docentes. Claramente, dejando a la libre interpretación ética del empleo de las plataformas, información y su tratamiento.

Esta situación ha llegado a ser confusa e, incluso, llegar a afectar la formación y desarrollo de las (os) universitarios, impactando en su trayectoria escolar.

Por otra parte, en el fondo, visualizando una violación al Derecho Humano a la Educación por omisión de la Autoridad, ante la inexistencia de reglamentaciones viables y efectivas que conduzcan a un ejercicio pleno del Derecho Humano a la Educación, como lo establece el artículo 103 fracción I:

Los Tribunales de la Federación resolverán toda controversia que se suscite:

I. Por normas generales, actos u omisiones de la autoridad que violen los derechos humanos reconocidos y las garantías otorgadas para su protección por esta Constitución, así como por los tratados internacionales de los que el Estado Mexicano sea parte.<sup>9</sup>

Recordando que en el párrafo 3º del artículo 1º constitucional, determina la obligación de las Autoridades para procurar el respeto y goce de los Derechos Humanos:

Todas las autoridades, en el ámbito de sus competencias, tienen la obligación de promover, respetar, proteger y garantizar los derechos humanos de conformidad con los principios de universalidad, interdependencia, indivisibilidad y progresividad. En consecuencia, el Estado deberá prevenir, investigar, sancionar y reparar las violaciones a los derechos humanos, en los términos que establezca la ley.<sup>10</sup>

Por lo que, a continuación, se visualizarán los puntos de vista de 129 estudiantes del Complejo Regional Mixteca de la Benemérita Universidad Autónoma de Puebla (México), para comprender la importancia de formular una Regulación sobre el empleo de estas herramientas y la necesidad de bases éticas orientadoras para el empleo y tratamiento de información de las Plataformas de IA.

### 3. METODOLOGÍA

La presente investigación se fundamenta en un enfoque cuantitativo, ya que en "esta clase de investigaciones resulta común en las ciencias y el mundo profesional, sobre todo en situaciones en las que existen pocos datos"<sup>11</sup>, impulsando el estudio de la problemática del empleo de la IA sin regulaciones enfocadas al ámbito académico universitario.

De igual forma, se emplea el método analítico – sintético para conocer la problemática desde la perspectiva universitaria; considerando oportuno emplear la encuesta como técnica principal y el cuestionario como instrumento digital (Microsoft Forms) aplicado a la muestra.

La población asciende a 228 estudiantes que se encontraban matriculados hasta el periodo otoño 2025 en la licenciatura en Administración de Empresas del Complejo

<sup>9</sup>CONGRESO CONSTITUYENTE DE 1917. Constitución Política de los Estados Unidos Mexicanos, art. 103. México: H. Congreso de la Unión, 6 de mayo de 2026, p. 107.

<sup>10</sup>Ibidem, art. 1º, p. 1.

<sup>11</sup>HERNÁNDEZ-SAMPIERI, Roberto; MENDOZA TORRES, Christian Paulina. Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta. México: McGraw-Hill, 2018, p. 107.

Regional Mixteca de la Benemérita Universidad Autónoma de Puebla (Atlixco, Puebla; México) y la muestra fue integrada por 129 universitarios que atendieron a la aplicación del instrumento digital, considerando desde el 1er semestre hasta 9º (y otros).

Cabe señalar que el instrumento fue piloteado con cinco estudiantes de tercer semestre de la propia licenciatura.

La selección de la muestra obedece a que los autores del presente artículo ofrecen mayoritariamente cátedra en esta licenciatura de las seis ofertadas en el municipio de Atlixco de este Complejo Regional (Licenciatura en Administración de Empresas, en Administración Turística, Contaduría Pública, Gastronomía, Procesos Educativos y Derecho); teniendo mayor control sobre el estudio y aplicación del instrumento.

Finalmente, el siguiente análisis de datos partió de la agregación estadística y visualización automática realizada por el software Microsoft Forms, que se caracteriza por un proceso de recopilación de respuestas, clasificación de los datos, cálculo de estadísticas básicas y creación de gráficas; para posteriormente, emprender un proceso de análisis humano especializado por parte de los autores de la presente investigación.

#### **4. RESULTADOS**

La propia investigación conduce a la exposición de resultados del instrumento aplicado a las (os) estudiantes de la Licenciatura en Administración de Empresas del Complejo Regional Mixteca (Campus Atlixco y Unidad de Vinculación Social).

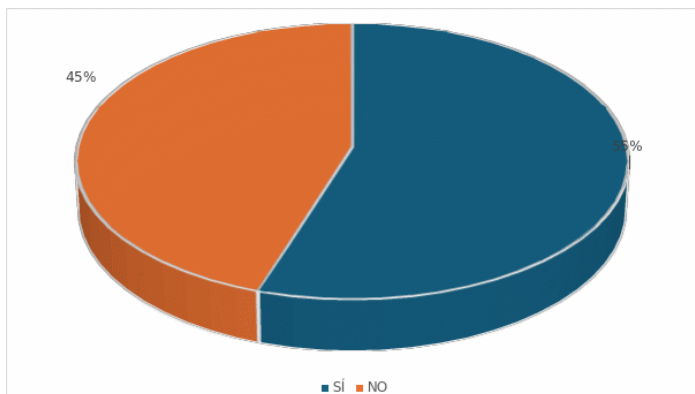
Como se mencionó en el anterior apartado, la muestra asciende a 129 estudiantes, considerando importante el alcance de respuesta porque se retoman alumnas (os) desde primer semestre hasta el 9º (y otros).

Uno de los principales propósitos de esta investigación es tratar de comprobar las posturas hipotéticas que originan este tema de investigación: Las (os) estudiantes cometen plagio digital por desconocimiento y/u orientación en el tratamiento de información emanada de Inteligencia Artificial.

Por otra parte, existe la estigmatización del empleo de las plataformas de inteligencia artificial, al considerarse de tajo como un acto falto de ética, desde algunas posturas conservadoras. Por lo que es importante conocer el punto de vista de los universitarios para identificar el área de oportunidad y lograr fortalecer las habilidades investigativas al aprovechar un medio tecnológico valioso como lo es la IA.

En primer lugar, se partió de la formulación de la presente pregunta: ¿Las (os) docentes conocen y emplean la IA dentro de las actividades científicas y académicas de la asignatura?, con la intención de conocer si las (os) docentes conocen, emplean e, incluso, dominan las plataformas de Inteligencia Artificial (ver gráfica 1).

Gráfica 1. Empleo y conocimiento de la IA por parte de los docentes



Fuente: elaboración propia.

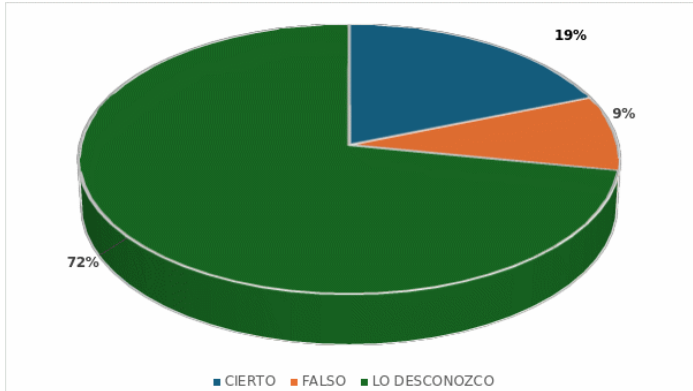
Observando que el 55% (71 individuos) de las (os) universitarios, consideran que sus docentes conocen y emplean la IA, y un 45% (58 individuos) mencionaron que no. Es un primer acercamiento valioso ante la necesidad de incorporar los recursos tecnológicos digitales innovadores en las asignaturas, como lo expresa la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (por sus siglas en inglés: UNESCO):

La IA proporciona nuevas vías que permiten explorar y buscar ayuda sobre determinados temas, aunque también proporciona accesos rápidos. (...) En otras palabras, se trata de replantear la manera en que aprendemos y enseñamos y, en consecuencia, cómo se enseña a los educandos y qué es lo que éstos deberán priorizar.<sup>12</sup>

Continuando con el hilo de esta idea, la orientación es indispensable para dejar de estigmatizar el empleo de estas plataformas, siendo de vital importancia hacerles comprender al estudiantado que el empleo de información proveniente de IA, debe de ser considerada como información digital, por lo que la siguiente cuestión se dirige a conocer el punto de vista de los individuos encuestados para develar si consideran, desde su origen, el empleo de esta información como un acto de deshonestidad académica, es decir, plagio (ver gráfica 2).

<sup>12</sup>NACIONES UNIDAS. El uso de la IA en la educación: decidir el futuro que queremos. París: UNESCO, 17 de mayo de 2024.

Gráfica 2. El empleo de información emanada de la IA es considerado plagio.



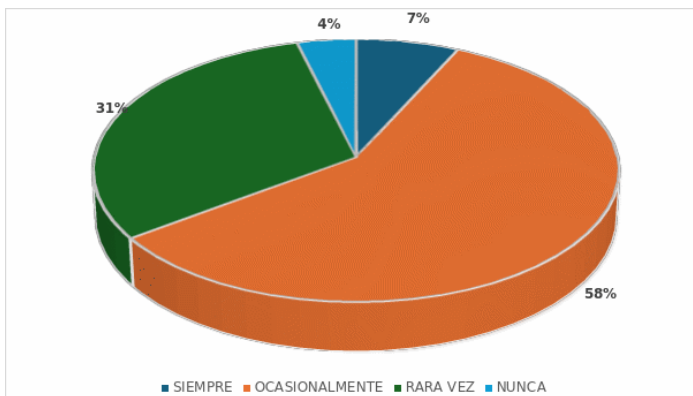
Fuente: elaboración propia.

Se obtiene que la opción "Lo desconozco" prevalece en este ítem con 93 individuos, que representan el 72 %, continuando con la opción "Cierto", con 24 encuestados (19 %) y cerrando con la opción "Falso" (9 %). De ello se obtiene que el desconocimiento de las (os) universitarios, podría provocar una conducta deshonesta como el plagio, una situación que podría ser limitada con la orientación docente:

El plagio involuntario consiste en la apropiación accidental del trabajo de otros por desconocimiento de las normas de documentación. Sin embargo, este uso indebido de las fuentes se considera una violación de la integridad académica.<sup>13</sup>

Prosiguiendo, la tecnología educativa ha llegado a las aulas y a las (os) estudiantes, lamentablemente, en ocasiones, el estudiantado depende forzosamente de la IA para desarrollar e incluso brindarse una idea sobre las tareas, trabajos e investigación escolar. Por lo que la siguiente pregunta, es de gran interés para el presente artículo científico (ver gráfica 3):

Gráfica 3. La IA me ha ayudado en los trabajos académicos.



Fuente: elaboración propia.

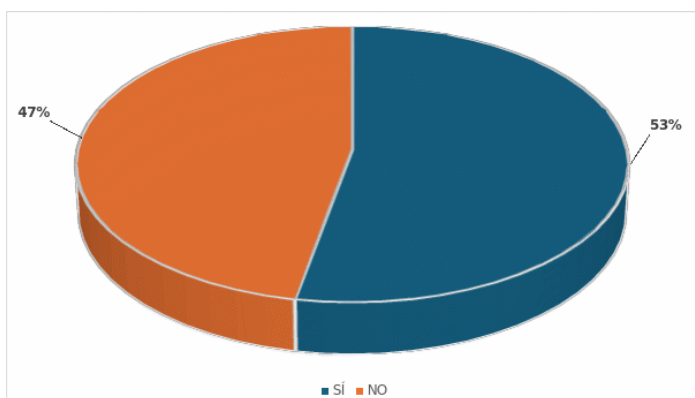
<sup>13</sup>NATIONAL UNIVERSITY. Inteligencia artificial e investigación académica: Integridad académica / Los dos tipos de plagio. Library National University, EE. UU., 23 de abril de 2026.

Como se identificó, 9 individuos de la investigación seleccionaron la opción "siempre" (7 %), 75 eligieron "ocasionalmente" (58 %), 40 optaron por "rara vez" (31 %) y 5 retomaron la respuesta "nunca" constitutiva del 4 %. Un tema de interés que impulsa eventos académicos de debate universitario, como los exponen los diálogos generados en la Universidad Anáhuac:

De acuerdo con un estudio realizado por la Universidad Anáhuac Mayab, ocho de cada 100 alumnos universitarios tienen un apego 'alto' a la tecnología. Así lo señalaron los especialistas Dr. José J. Karam Espósitos (Escuela de Psicología) y del Dr. José Gabriel Urzaiz Lares (División de Ingeniería).<sup>14</sup>

Conduciendo al cuestionamiento base, sobre si las (os) docentes han orientado para el empleo correcto, viable y ético de la información emanada de la IA (ver gráfica 4).

Gráfica 4. Muestra de empleo ético de la IA por parte de los docentes



Fuente: elaboración propia.

Conociendo que el 53 %, es decir, 69 encuestadas (os), respondieron "sí" y el 47 % expresaron que "no" (60 individuos). Contemplando un importante porcentaje de compromiso en la guía del estudiantado. Por lo que es necesario trabajar a la par, docente-estudiante y orientar en su empleo:

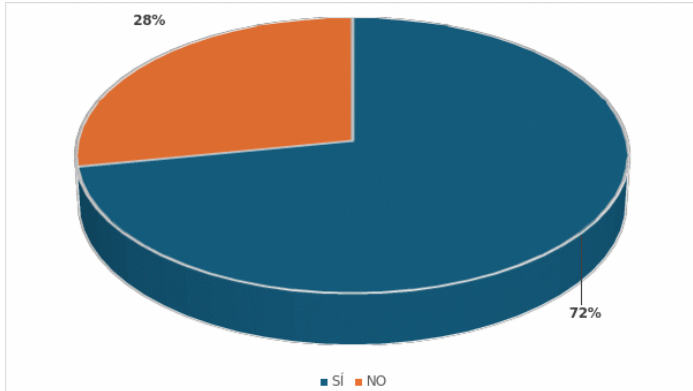
No podemos quedarnos en el entusiasmo tecnológico. La verdadera innovación ocurre cuando el maestro y el estudiante comprenden las implicaciones de lo que están haciendo con la tecnología.<sup>15</sup>

Consecutivamente, una postura fundamental, desde qué momento el estudiantado universitario considera plagio durante el empleo de información emanada de la IA (ver gráfica 5).

<sup>14</sup>REDACCIÓN ANÁHUAC. Incrementa dependencia de la IA entre jóvenes. Universidad Anáhuac Campus Mayab. México, 2-3 de julio de 2025.

<sup>15</sup>HERNÁNDEZ, Jorge; GARDUÑO MUÑOZ, Thania Corina. La IA, una extensión del trabajo docente. Gaceta UNAM. México: Universidad Nacional Autónoma de México, 20 de noviembre de 2025.

Gráfica 5. Valoración de plagio en el empleo de información emanada de la IA y presentarla como propia



Fuente: elaboración propia.

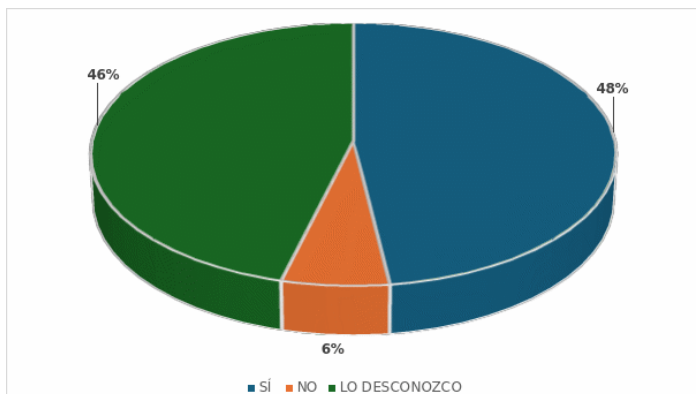
De la anterior interrogante, se identificó que la prevalencia de la respuesta "sí", es visible con 93 estudiantes que equivale al 72% y 36 individuos, respondieron que "no", es decir, un 28 %. Teniendo presente analizar el momento o consideración de esta conducta al tratar la información:

El plagio, en esencia, consiste en presentar las ideas, el trabajo o las palabras de otra persona como propias sin la debida atribución (p. ej., McCuen, 2008). Sin embargo, la cuestión de a quién se le roba la propiedad intelectual cuando se presenta contenido generado por IA como obra original en trabajos o tareas complica el concepto de plagio (Eke, 2023). Por lo tanto, los investigadores argumentan que las fuentes generadas por IA deben citarse adecuadamente en las tareas académicas, y que el papel de la IA en el proceso de escritura debe definirse explícitamente; de lo contrario, enfatizan que esto conllevará un riesgo de plagio (Jarrah, Wardat y Fidalgo, 2023; Perkins, 2023).<sup>16</sup>

Al hilo del argumento y análisis, se continúa con la cuestión dirigida al tratamiento y empleo de la información obtenida de plataformas de IA (ver gráfica 6).

<sup>16</sup>BAŞER, Miraç; KOZAK, Metin; ERDOĞAN, İbrahim. ¿Nos preocupa el uso de la inteligencia artificial y el plagio? El comportamiento de los estudiantes con la IA a través del triángulo del fraude. *The Internet and Higher Education*, vol. 69, 2026.

Gráfica 6. ¿Puedo emplear IA sin recaer en plagio?



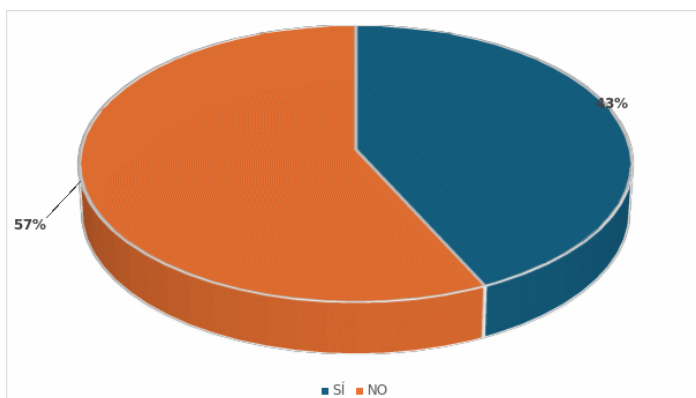
Fuente: elaboración propia.

Los resultados de la presente cuestión, exponen la postura de 62 sujetos de muestra que eligieron que "sí", reflejando un 48 %; la opción "no" con un 6 % (8 individuos), y con 59 encuestados (46 %) "lo desconozco". Teniendo presente que se recomendará citar la información obtenida, sin brindar coautoría a la IA, por lo que es un reto constante y ético:

Si bien el marco regulador en desarrollo pretende requerir a los proveedores de IA Generativa que reconozcan y aseguren la protección de los derechos de propiedad de los autores del contenido usado para desarrollar un modelo, el reto se está haciendo cada vez más complicado dada la inmensa cantidad de trabajos generado.<sup>17</sup>

Vinculándose con la anterior cuestión, es indispensable conocer si el alumnado sabe cómo citar la información emitida por la IA (ver gráfica 7).

Gráfica 7. Sé citar la información obtenida de IA.



Fuente: elaboración propia.

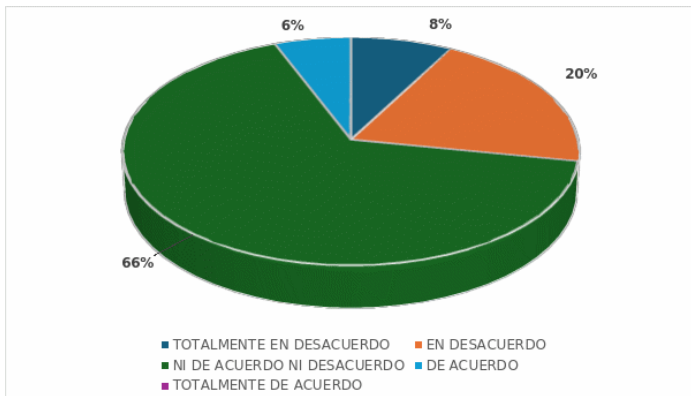
De la anterior gráfica, se obtiene que 55 sujetos de la muestra (43 %) seleccionaron la opción "sí" y 74 individuos eligieron "no", representando un 57 %; no obstante, la

<sup>17</sup>GARCÍA TEJEDOR, Nieves. Uso de la Inteligencia Artificial Generativa sin caer en el plagio y respetando la integridad académica. Biblioteca UNED, 16 de abril de 2024.

falsedad o la inexistencia, pueden significar una problemática seria: "las referencias bibliográficas son un elemento clave en la validación del conocimiento, ya que permiten rastrear las fuentes y verificar la información"<sup>18</sup>, por lo que su inexistencia impactaría al proceso de investigación y resultados.

Continuando el desarrollo de este estudio, se tiene que considerar el punto de vista de las (os) estudiantes encuestados, para conocer si consideran "indispensable" el empleo de la IA, para la realización de documentos (ver gráfica 8).

Gráfica 8. Considero indispensable la IA para comenzar un trabajo de investigación



Fuente: elaboración propia.

Lo anterior, ofrece una perspectiva interesante, ya que el 8 % (10 sujetos) eligieron la opción "totalmente en desacuerdo"; el 20 %, que son 26 individuos, seleccionaron "en desacuerdo"; prevaleciendo "ni de acuerdo ni desacuerdo" con un 66 % representando 85 encuestadas (os); "de acuerdo" fueron el 6 % considerándose 8 estudiantes, y, finalmente, "totalmente de acuerdo:" no fue elegido por ningún miembro de la muestra (0 %; 0 individuos). Como se observa, la muestra selecciona una opción "neutra", dando cabida a la interpretación de los autores, donde se percibe su empleo "orientador", empero, no es "siempre" necesario para comenzar, brindando una importante área de oportunidad para impulsar otras herramientas tecnológicas y didácticas; coincidiendo con Newsroom Infobae:

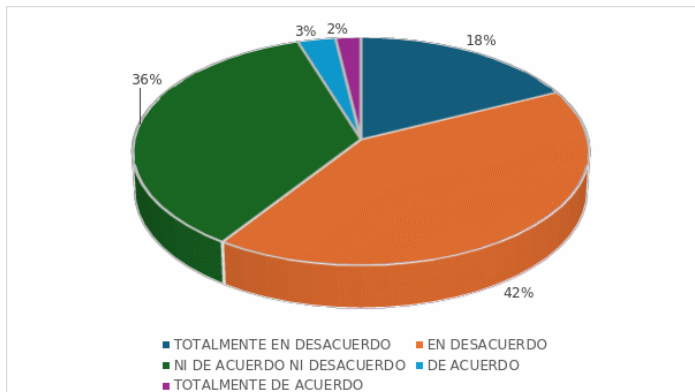
El 90 % de los universitarios utiliza la inteligencia artificial (IA) para estudiar y casi cuatro de cada diez la usa a diario. (...) Son datos que se desprenden del estudio "Uso y percepción de la IA en el entorno universitario" realizado por la Fundación Conocimiento y Desarrollo (CYD) a través de dos encuestas sobre una base de casi 1 000 estudiantes y 12 universidades públicas y 8 privadas, lo que representa al 24 % y al 27 % de los centros públicos y privados [en Madrid].<sup>19</sup>

Vinculándose con el siguiente cuestionamiento, elevando el trabajo académico a investigación científica dentro del aula (ver gráfica 9).

<sup>18</sup>NORIEGA, Pedro. Miles de publicaciones científicas podrían incluir citas inventadas por inteligencia artificial, según la revista Nature. Infobae. México, 6 de mayo de 2026.

<sup>19</sup>NEWSROOM INFOBAE. El 90 % de los universitarios usa la IA para estudiar y sus profesores para crear contenido. Infobae. Madrid, 17 de mayo de 2025.

Gráfica 9. Tengo problemas si no cuento con el acompañamiento de IA para mis trabajos de investigación.



Fuente: elaboración propia.

El presente ítem del instrumento, ofrece un cuestionamiento que busca profundizar en el individuo, ya que posiblemente, al no contar con una orientación y guía en el proceso indagatorio, les sea más viable a las (os) universitarias (os) acceder a estas plataformas de acceso libre y dejarse llevar, como se visualiza con la selección de respuestas.

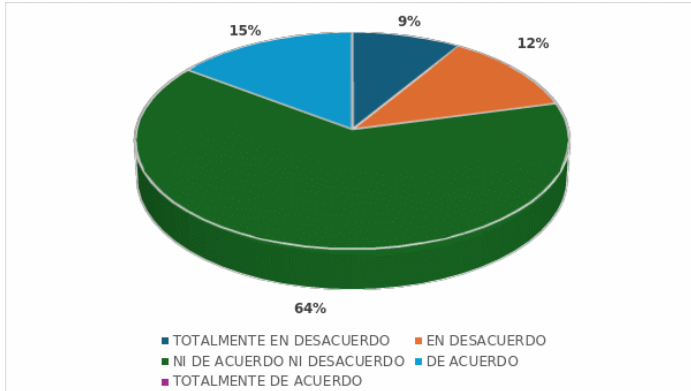
Siendo 23 individuos que eligieron la opción "totalmente en desacuerdo" (18%); con 54 sujetos se ubica la respuesta "en desacuerdo" (42%); "Ni de acuerdo ni desacuerdo" resultaron 46 encuestados (36%); con 4 universitarios (3%) se considera "de acuerdo", y, finalmente, "totalmente de acuerdo" cuenta con 2 integrantes de la muestra, es decir, el 2%. Estos resultados reflejan una negación y una postura neutra que se inclina por una posibilidad eventual más que constante. Coincidiendo con gran parte de la comunidad científica, donde cabe la posibilidad de que la IA, con orientación y bases éticas mostradas en praxis investigativa a las (os) estudiantes, podrían retomar herramientas valiosas en el ámbito investigativo:

El reto de la educación superior es la formación de docentes que puedan aplicar la tecnología en la investigación y desarrollo de métodos de enseñanza aprendizaje que favorezcan el formar profesionales con una visión humanista y ética, que mejore la calidad y eficiencia en la educación, facilite la comprensión de contenidos, que genere experiencias de aprendizaje personalizados y favorezca el pensamiento crítico, así como regular su uso, tomando en cuenta que la IA no sustituye a las personas y su capacidad, es solo una herramienta que ayuda a optimizar y agilizar las tareas en todas las áreas del conocimiento.<sup>20</sup>

Al continuar con la presentación de la siguiente cuestión graficada, es posible identificar diversos puntos de vista, por demás enriquecedores, no solo en la siguiente gráfica, sino incluso en la comunidad científica (ver gráfica 10).

<sup>20</sup> PRIETO GARCÍA, Brenda; ANCHONDO CHAVARRÍA, Guadalupe; MARTÍNEZ ARANDA, María Rebeca; GONZÁLEZ GONZÁLEZ, Javier Antonio; TINTORI LARA, Manuel. El uso de la IA para la realización de tareas de los alumnos de la carrera de DN de la UTCH. Revista Politécnica de Aguascalientes, vol. 4, 2024, año 3, p. 5.

Gráfica 10. Si yo le emito todas las indicaciones, directrices y especificaciones a la IA, no es plagio



Fuente: elaboración propia.

De este ítem, se desprenden los siguientes resultados: "totalmente en desacuerdo": 9% (11 individuos de muestra); "en desacuerdo": 12 % (16 individuos de muestra); "ni de acuerdo ni desacuerdo": 64 % (82 individuos de muestra); "de acuerdo": 15 % (19 individuos de muestra), y "totalmente de acuerdo": 1 % (1 individuo).

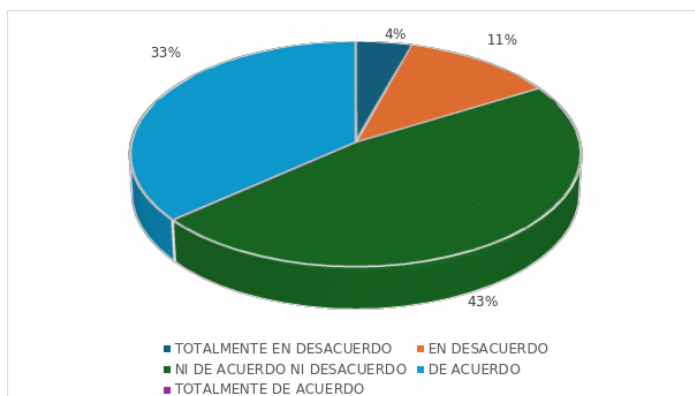
Sin embargo, no es que exista una respuesta errónea, sino parámetros éticos establecidos, impuestos o institucionalizados lo que nos brindará certeza y una dirección unificada sobre estas acciones, ya que la lógica radica en que como todo "software", un humano emite las indicaciones, por ende, la información generada, creada y sesgada por el ser humano, no es información del software, ¿o sí? En este sentido, se identifica una postura interesante que apoya su sentido lógico del empleo de la información de la IA, resultante de indicaciones humanas:

Cuando se usa un chatbot para realizar algún escrito y solo se copia literalmente lo que ha dado como resultado, no se está haciendo plagio, al menos no de forma directa. Cuando éste se ha cometido, es porque se ha quitado la autoría de una obra y no se reconoce al verdadero autor, pero en el caso de una herramienta de inteligencia artificial (IA), no se puede decir que sea la autora de esa obra, pues el usuario está dando las indicaciones de lo que quiere, por tanto, también ha puesto de su parte, explicó Gabriela Elisa Sued, becaria posdoctoral del Instituto de Investigaciones Sociales (IIS) de la Universidad Nacional Autónoma de México (UNAM).<sup>21</sup>

Se finaliza con un punto provechoso para la presente investigación, al cuestionar a las (os) estudiantes si consideran necesaria la regulación del empleo de información emitida por las plataformas de IA (ver gráfica 11).

<sup>21</sup>ZÁRATE GÓMEZ, Itahi. El uso de chatbots en la elaboración de textos no implica cometer plagio. Instituto de Investigaciones Sociales, Departamento de Difusión del IIS-UNAM. México, 23 de junio de 2023.

Gráfica 11. Valores necesaria la regulación del empleo de información emanada de la IA dentro del ámbito académico



Fuente: elaboración propia.

Finalmente, se retoma el resultado del presente ítem graficado para hacer expreso que la muestra universitaria adopta una postura neutra, "ni de acuerdo ni en desacuerdo" con un 43 % (56 individuos de muestra); continuando con "de acuerdo" con un 33 % (43 individuos de muestra); prosiguiendo con "en desacuerdo" con un 11 % o 14 individuos; siguiendo con "totalmente de acuerdo" identificando un 9 % (11 individuos de muestra), y "totalmente en desacuerdo" representando un 4 % (5 individuos).

Logrando identificar la malinterpretación de "regular" y "prohibir", en otras palabras, en el momento de la aplicación del instrumento y sin buscar sesgar el resultado de este (no se emitió comentario alguno), las (os) estudiantes comentaron que la reglamentación implicaba prohibir el empleo de la IA.

Si bien las instituciones educativas han buscado impulsar propuestas de reglamentación, el propio Senado de la República Mexicana ha emitido su postura para comenzar a regular a la IA:

La Comisión de Ciencia y Tecnología del Senado presenta un nuevo marco normativo para regular el uso de la inteligencia artificial (IA) en México, con el objetivo de establecer garantías para los derechos humanos, la privacidad y la protección de datos personales. (...) La Comisión del Senado reporta que esta medida responde al crecimiento vertiginoso de la IA y la urgencia de adaptar la legislación nacional en materia de salud, educación, brechas digitales y otras áreas prioritarias.<sup>22</sup>

Culminando este apartado de Resultados, es viable continuar con la discusión de esta investigación.

## 5. DISCUSIÓN

Como es posible observar a lo largo de esta investigación y exposición de resultados, es imperante migrar a una postura innovadora en cuanto al empleo de la información de IA con una base ética institucionalizada, donde las y los estudiantes conozcan y dominen el manejo y tratamiento ético de esta información, impulsando y contribuyendo con

<sup>22</sup>ROF, Diana. Desde Comisión del Senado impulsan iniciativa de ley para regular el uso de inteligencia artificial en México. Infobae. México, 30 de abril de 2026.

su formación académica y limitar los rezagos y brechas tecnológicas, que vulneran su derecho humano a la educación.

En este sentido, los organismos internacionales, como la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (por sus siglas en inglés: UNESCO) han priorizado este tema buscando una regulación que permita el empleo de este avance tecnológico basado en una base axiológica.

El 23 de noviembre de 2021, en París (Francia), se celebró la 41.ª Conferencia General de la UNESCO, durante la cual se generó la Recomendación sobre la ética de la inteligencia artificial siendo adoptada por 193 Estados miembros de esta Organización, considerándose el primer marco normativo en esta materia.

Postulando oportuno el exhorto que plasma en el propio preámbulo del instrumento:

Recomienda que los Estados Miembros apliquen, de manera voluntaria, las disposiciones de la presente Recomendación mediante la adopción de las medidas adecuadas, en particular las medidas legislativas o de otra índole que puedan ser necesarias, de acuerdo con la práctica constitucional y las estructuras de gobierno de cada Estado, con el fin de dar efecto en sus respectivas jurisdicciones a los principios y normas enunciados en la Recomendación, de conformidad con el derecho internacional, incluido el derecho internacional de los derechos humanos.<sup>23</sup>

Buscando generar ese impulso en los Estados Miembros, adopta una serie de valores que las (os) autoras (es) de la presente investigación consideran viables para emprender una reglamentación Nacional, Estatal e Institucional:

- Proporcionalidad e inocuidad;
- Seguridad y protección;
- Equidad y no discriminación;
- Sostenibilidad;
- Derecho a la intimidad y protección de datos;
- Supervisión y decisión humanas;
- Transparencia y explicabilidad;
- Responsabilidad y rendición de cuentas;
- Sensibilización y educación, y
- Gobernanza y colaboración adaptativas y de múltiples partes interesadas.<sup>24</sup>

Por lo que las propuestas y, en concreto, los instrumentos "deberían" existir armonizados con estos principios que bien podrían funcionar como ejes rectores.

Dentro de las propuestas existentes y por generarse, en el ámbito educativo, buscando ejercitar el Derecho Humano a la Educación y buscar "borrar" brechas tecnológicas, es viable considerar el párrafo número 107, dentro del Ámbito de Actuación 8: Educación e Investigación:

Los Estados Miembros deberían promover y apoyar las investigaciones sobre la IA, en particular las investigaciones sobre la ética de la IA, por ejemplo mediante inversiones en este tipo de investigaciones o creando incentivos para que los sectores público y privado inviertan en este ámbito, reconociendo que las investigaciones contribuyen de manera significativa al desarrollo y la mejora ulteriores de las tecnologías de la IA con miras a promover el derecho internacional y los valores y principios enunciados en la presente Recomendación.

<sup>23</sup>ORGANIZACIÓN DE LAS NACIONES UNIDAS PARA LA EDUCACIÓN, LA CIENCIA Y LA CULTURA. Recomendación sobre la ética de la inteligencia artificial. París: UNESCO, 23 de noviembre de 2021, p. 7.

<sup>24</sup>Ibidem, pp. 20-23.

Los Estados Miembros también deberían promover públicamente las mejores prácticas de los investigadores y las empresas que desarrollan la IA de forma ética y cooperar con ellos<sup>25</sup>.

No obstante, en la actualidad únicamente existen propuestas de ley que buscan reglamentar el empleo de estas plataformas de acceso libre, por las bancadas de MORENA y el PVEM en el año 2025 y, previamente, en 2024, se presentó una propuesta en el Senado Mexicano.

Considerando una probable violación de derechos humanos por omisión, ya que, al no garantizar el acceso, empleo, regulación y fomento de los adelantos tecnológicos, se vulnera el contenido de la fracción V del artículo 3º:

Toda persona tiene derecho a gozar de los beneficios del desarrollo de la ciencia y la innovación tecnológica. El Estado apoyará la investigación e innovación científica, humanística y tecnológica, y garantizará el acceso abierto a la información que derive de ella, para lo cual deberá proveer recursos y estímulos suficientes, conforme a las bases de coordinación, vinculación y participación que establezcan las leyes en la materia; además alentará el fortalecimiento y difusión de nuestra cultura.<sup>26</sup>

Ergo, no lograr concluir la trayectoria escolar de los universitarios del país, y como es posible observar, del territorio poblano.

Por lo que contribuir a la formación y desarrollo universitario a través del empleo ético de la IA y regulado, contribuirá al cumplimiento de los objetivos escolares y académicos de cada estudiante de nivel superior.

## 6. CONCLUSIONES

El desarrollo de la presente investigación permite expresar que se lograron cumplir con los elementos metodológicos que guiaron al presente artículo científico-jurídico, de igual forma, se han recuperado las siguientes conclusiones:

Primera. – Se requiere un reglamento gubernamental e institucional que busque conducir y guiar el empleo de las plataformas de Inteligencia Artificial para tener un instrumento para la formación de las (os) estudiantes universitarios.

Segunda. – Existe falta de certeza y desconocimiento en el acceso, empleo y tratamiento de información por parte del estudiantado universitario (CRM-BUAP) emanada por las plataformas de Inteligencia Artificial por lo que puede existir plagio y conductas poco éticas por desconocimiento.

Tercero. – Las diversas posturas éticas y de interpretación pueden conducir a emitir posturas por parte de las (os) docentes e investigadores que conduzcan al plagio directo e indirecto.

Cuarto. – Se identifica una violación del derecho a la educación ante la omisión de reglamentos federales, estatales e institucionales para el tratamiento de la información por parte de las (os) universitarios, docentes e investigadoras (es).

Quinta. – Se coincide con la postura de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (por sus siglas en inglés: UNESCO), sobre todo en retomar los valores expuestos en la Recomendación sobre la ética de la inteligencia artificial, considerando oportuno retomarlos como ejes de acción que podrían guiar los Reglamentos en esta materia en territorio mexicano.

---

<sup>25</sup>Ibidem, pp. 35-36.

<sup>26</sup>CONGRESO CONSTITUYENTE DE 1917, op. cit., nota 9, art. 3.º, p. 11.

Sexta. – La falta de regulación jurídico-ética sobre el empleo de la información emanada por la IA vulnera los Derechos Humanos de las (os) universitarios e incluso afecta a su trayectoria escolar.

## **7. REFERENCIAS BIBLIOGRÁFICAS**

- BAŞER, Miraç; KOZAK, Metin; ERDOĞAN, İbrahim. ¿Nos preocupa el uso de la inteligencia artificial y el plagio? El comportamiento de los estudiantes con la IA a través del triángulo del fraude. En línea. *The Internet and Higher Education*, vol. 69 (2026). [Disponible en: <https://bsu.buap.mx/ggT>].
- CONGRESO CONSTITUYENTE DE 1917. Constitución Política de los Estados Unidos Mexicanos. En línea. México: H. Congreso de la Unión, 6 de mayo de 2026. [Disponible en: <https://bsu.buap.mx/gig>].
- DE PINA, Rafael; DE PINA VARA, Rafael. *Diccionario de Derecho*. México: Porrúa.
- GARCÍA TEJEDOR, Nieves. Uso de la Inteligencia Artificial Generativa sin caer en el plagio y respetando la integridad académica. En línea. Biblioteca UNED, 16 de abril de 2024. [Disponible en: <https://bsu.buap.mx/ggU>].
- GÓMEZ HERRERA, Renato. La inteligencia artificial ¿Hacia dónde nos lleva? En línea. ¿Cómo ves? Divulgación de la ciencia. México: UNAM, 7 de mayo de 2026. [Disponible en: <https://bsu.buap.mx/gil>].
- HERNÁNDEZ, Jorge; GARDUÑO MUÑOZ, Thania Corina. La IA, una extensión del trabajo docente. En línea. *Gaceta UNAM*. México: Universidad Nacional Autónoma de México, 20 de noviembre de 2025. [Disponible en: <https://bsu.buap.mx/ggS>].
- HERNÁNDEZ-SAMPIERI, Roberto; MENDOZA TORRES, Christian Paulina. *Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta*. México: McGraw-Hill, 2018.
- NACIONES UNIDAS. ¿Qué son los derechos humanos? En línea. Oficina del Alto Comisionado. México, 2026. [Disponible en: <https://bsu.buap.mx/gip>].
- NACIONES UNIDAS. El uso de la IA en la educación: decidir el futuro que queremos. En línea. París: UNESCO, 17 de mayo de 2024. [Disponible en: <https://bsu.buap.mx/geS>].
- NACIONES UNIDAS. El derecho humano a la educación. En línea. Oficina del Alto Comisionado, América del Sur, 2022. [Disponible en: <https://bsu.buap.mx/giq>].
- NATIONAL UNIVERSITY. Inteligencia artificial e investigación académica: Integridad académica / Los dos tipos de plagio. En línea. Library National University, EE. UU., 23 de abril de 2026. [Disponible en: <https://bsu.buap.mx/geT>].
- NEWSROOM INFOBAE. El 90 % de los universitarios usa la IA para estudiar y sus profesores para crear contenido. En línea. Infobae. Madrid, 17 de mayo de 2025. [Disponible en: <https://bsu.buap.mx/ghF>].
- NORIEGA, Pedro. Miles de publicaciones científicas podrían incluir citas inventadas por inteligencia artificial, según la revista Nature. En línea. Infobae. México, 6 de mayo de 2026. [Disponible en: <https://bsu.buap.mx/ghY>].
- OPENAI. ChatGPT 4. Modelo de lenguaje grande. En línea. Contenido generado el 7 de mayo de 2026. [Disponible en: <https://bsu.buap.mx/gio>. Prompt: «¿Cómo podrías definirte?»].
- ORGANIZACIÓN DE LAS NACIONES UNIDAS PARA LA EDUCACIÓN, LA CIENCIA Y LA CULTURA. Recomendación sobre la ética de la inteligencia artificial. En línea. París: UNESCO, 23 de noviembre de 2021. [Disponible en: <https://bsu.buap.mx/ghW>].
- PÉREZ-GALMICHE, Gabriel; ESPINOSA-GONZÁLEZ, David. *La crisis de los Derechos Humanos: Estudio comparativo México-Italia*. Ciudad de México: Montiel Et Soriano Editores y BUAP, 2021.
- PRIETO GARCÍA, Brenda; ANCHONDO CHAVARRÍA, Guadalupe; MARTÍNEZ ARANDA, María Rebeca; GONZÁLEZ GONZÁLEZ, Javier Antonio; TINTORI LARA, Manuel. El uso de la IA para la realización de tareas de los alumnos de la carrera de DN de la UTCH. En línea. *Revista Politécnica de Aguascalientes*, vol. 4 (2024), año 3, pp. 1-5. [Disponible en: <https://bsu.buap.mx/ghG>].

- REAL ACADEMIA ESPAÑOLA. Diccionario panhispánico del español jurídico. En línea. España, 2025. [Disponible en: <https://bsu.buap.mx/gii>].
- REDACCIÓN ANÁHUAC. Incrementa dependencia de la IA entre jóvenes. En línea. Universidad Anáhuac Campus Mayab. México, 2-3 de julio de 2025. [Disponible en: <https://bsu.buap.mx/ggR>].
- ROF, Diana. Desde Comisión del Senado impulsan iniciativa de ley para regular el uso de inteligencia artificial en México. En línea. Infobae. México, 30 de abril de 2026. [Disponible en: <https://bsu.buap.mx/ghV>].
- ZÁRATE GÓMEZ, Itahi. El uso de chatbots en la elaboración de textos no implica cometer plagio. En línea. Instituto de Investigaciones Sociales, Departamento de Difusión del IIS-UNAM. México, 23 de junio de 2023. [Disponible en: <https://bsu.buap.mx/ghH>].

# Reconfiguración de libertades fundamentales en la era de la inteligencia artificial: Vigilancia, poder algorítmico y crisis del estado de derecho\*

*Reconfigurations of fundamental freedoms in the age of artificial intelligence: Surveillance, algorithmic power, and the crisis of the rule of law*

Bryan Azcué Hoyos\*\*

## RESUMEN

El presente artículo sostiene que la incorporación de sistemas de inteligencia artificial no solo introduce riesgos para las libertades fundamentales, sino que reconfigura su ejercicio en términos de vigilancia, control y poder algorítmico. Desde un enfoque teórico analítico, anclado en el marco de los derechos humanos y una perspectiva ético-crítica, se examinan tres dimensiones: la erosión de la autonomía frente a decisiones automatizadas, la intensificación de la vigilancia digital y la reproducción de desigualdades mediante sesgos algorítmicos. Se argumenta que estos procesos no son contingentes, sino estructurales, y evidencian la insuficiencia de los marcos jurídicos tradicionales para garantizar el Estado de derecho en entornos algorítmicos. Como aporte, se plantea la necesidad de una reconfiguración normativa que incorpore una lectura crítica del poder tecnológico desde los derechos humanos, reconociendo que la dignidad humana no puede quedar subordinada a la lógica de la optimización técnica.

## PALABRAS CLAVE

Inteligencia artificial; libertades fundamentales; derechos humanos; poder algorítmico; vigilancia digital; sesgos algorítmicos.

## ABSTRACT

This article argues that the incorporation of artificial intelligence systems not only introduces risk to fundamental freedoms but also reconfigures their exercise in terms of surveillance, control, and algorithmic power. From a theoretical-analytical approach grounded in human rights and a critical ethical perspective, three dimensions are examined. The erosion of autonomy in the face of automated decision-making, the intensification of digital surveillance, and the reproduction of inequalities through algorithmic bias. It is argued that these processes are not contingent but structural, revealing the insufficiency of traditional legal frameworks to guarantee the rule of law in algorithmic environments. As a contribution, the article proposes the need for a normative reconfiguration that incorporates a critical understanding of technological power from a human rights perspective, recognizing that human dignity cannot be subordinated to the logic of technical optimization.

## KEYWORDS

Artificial intelligence; fundamental freedom; human rights; algorithmic power; digital surveillance; algorithmic bias.

\* Artículo de investigación

\*\* Maestro en Administración y doctorando en educación, docente y coordinador del Laboratorio de Gastronomía en el Complejo Regional Mixteca de la Benemérita Universidad Autónoma de Puebla (bryan.azcue@correo.buap.mx). <https://orcid.org/009-0001-2513-9071>.

## SUMARIO

1. Introducción
2. Marco conceptual: libertades fundamentales e inteligencia artificial en contextos contemporáneos
3. Reconfiguración de las libertades fundamentales en entornos algorítmicos
4. Poder algorítmico y crisis estructural del Estado de derecho
5. Hacia una reconfiguración normativa desde los derechos humanos
6. Conclusiones
7. Referencias Bibliográficas

### 1. INTRODUCCIÓN

Vivimos en un momento en el que las decisiones que inciden sobre la vida de millones de personas (quién accede a un crédito, quién es considerado candidato apto para un empleo, quién recibe atención prioritaria en un servicio público, quién es identificado como sospechoso por un sistema de vigilancia) son adoptadas, en proporciones crecientes, por sistemas de inteligencia artificial. Esta transformación no ha ocurrido de manera abrupta ni espectacular, sino a través de una incorporación gradual, casi silenciosa, de tecnologías algorítmicas en los procesos que organizan la vida social, económica e institucional.

La novedad no reside únicamente en la automatización de tareas rutinarias. Lo que distingue a los sistemas de inteligencia artificial contemporáneos es su capacidad para procesar volúmenes masivos de datos, identificar patrones no evidentes para el ojo humano y producir inferencias con consecuencias jurídicas y sociales concretas. Esta capacidad, que en principio podría orientarse hacia la eficiencia y la mejora de servicios, introduce al mismo tiempo un conjunto de tensiones estructurales que afectan el ejercicio de derechos y libertades fundamentales.

El debate académico y normativo en torno a la inteligencia artificial ha estado dominado, durante largo tiempo, por un enfoque centrado en la gestión de riesgos: identificar los posibles daños que estas tecnologías podrían causar y establecer salvaguardias técnicas o regulatorias para mitigarlos. Sin desconocer la utilidad de este enfoque, el presente artículo propone una lectura diferente: la inteligencia artificial no solo genera riesgos adicionales sobre libertades previamente establecidas, sino que transforma las condiciones mismas bajo las cuales dichas libertades son ejercidas. En otros términos, el problema no es únicamente que los algoritmos puedan vulnerar la privacidad o producir discriminación, sino que reconfiguran la estructura del ejercicio de la autonomía, la igualdad y la privacidad en el entorno contemporáneo.

Esta distinción tiene consecuencias relevantes para el análisis jurídico. Si los efectos de la inteligencia artificial fueran meramente contingentes (fallas técnicas, errores corregibles, riesgos evitables con mayor regulación, etc.) bastaría con ampliar y actualizar los marcos normativos existentes. Pero si, como se sostiene en este trabajo, las transformaciones son estructurales, entonces la respuesta jurídica no puede limitarse a ajustes superficiales: requiere una revisión de los fundamentos conceptuales sobre

los cuales se ha construido el sistema de protección de derechos de las sociedades contemporáneas.

El artículo se organiza de la siguiente manera. En primer lugar, se establece un marco conceptual que articula la noción de libertades fundamentales con las características específicas de los sistemas algorítmicos. A continuación, se analizan tres dimensiones concretas de reconfiguración: la autonomía frente a decisiones automatizadas, la transformación de la privacidad en el contexto del capitalismo de la vigilancia y la reproducción de desigualdades mediante sesgos algorítmicos. Posteriormente, se examina el concepto de poder algorítmico y sus implicaciones para el Estado de derecho. Finalmente, se propone una reconfiguración normativa orientada por los derechos humanos y se formulan las conclusiones.

Desde el punto de vista metodológico, el trabajo adopta un enfoque teórico-analítico que articula aportes de la teoría jurídica, la filosofía política y los estudios críticos de tecnología. Si bien no se desarrolla un estudio de caso en particular, el análisis se apoya en referencias a experiencias documentadas y a instrumentos normativos internacionales que permiten contrastar los postulados teóricos con las tendencias regulatorias y sociales observables.

El objetivo central de este trabajo es demostrar que la incorporación de sistemas de inteligencia artificial en los procesos de toma de decisiones reconfigura estructuralmente el ejercicio de las libertades fundamentales, evidenciando una crisis en la capacidad del Estado de derecho para garantizar su protección efectiva, y proponer, desde una perspectiva de derechos humanos, los principios orientadores de una reconfiguración normativa que responda a esa crisis.

## **2. MARCO CONCEPTUAL: LIBERTADES FUNDAMENTALES E INTELIGENCIA ARTIFICIAL EN CONTEXTOS CONTEMPORÁNEOS**

El análisis de las implicaciones de la inteligencia artificial sobre las libertades fundamentales exige partir de un marco conceptual que articule, por un lado, la traducción jurídica de los derechos humanos y, por otro, las transformaciones derivadas de la digitalización y automatización de los procesos de decisión. Este esfuerzo no es trivial, porque implica tender un puente entre dos tradiciones que han evolucionado, en gran medida, de manera paralela y con escasos puntos de contacto: la teoría jurídica de los derechos y la teoría crítica de la tecnología.

Las libertades fundamentales han sido históricamente entendidas como garantías esenciales para la protección de la dignidad humana frente a la intervención arbitraria del poder. En la tradición liberal clásica, estas libertades operaban principalmente como límites negativos frente al Estado: la libertad de expresión, la libertad de movimiento, la inviolabilidad del domicilio y la correspondencia, la igualdad ante la ley. Con el desarrollo del Estado social y la consolidación de los sistemas internacionales de derechos humanos durante el siglo XX, este catálogo se amplió para incluir dimensiones positivas, exigibles al Estado no solo como abstención, sino como acción.

Antonio Enrique Pérez Luño, en su análisis clásico sobre los Derechos Humanos, el Estado de derecho y la Constitución, caracteriza las libertades fundamentales como posiciones subjetivas que encuentran su fundamento en la dignidad humana y que

requieren, para su efectividad, tanto de garantías normativas como de condiciones institucionales que hagan posible su ejercicio real<sup>1</sup>. Esta perspectiva resulta especialmente relevante en el contexto actual, porque sugiere que la protección de las libertades no puede reducirse a su reconocimiento formal, sino que depende de las condiciones materiales (y hoy, también, tecnológicas) en las que se desenvuelve la vida de las personas.

La emergencia de sistemas de inteligencia artificial introduce un desplazamiento significativo en esas condiciones. A diferencia de los modelos tradicionales de toma de decisiones, donde la acción humana era identificable y jurídicamente imputable, los sistemas algorítmicos operan mediante procesos automatizados que, en muchos casos, resultan opacos incluso para quienes los diseñan o implementan. Esta opacidad no es un defecto técnico superable, sino una característica estructural de los sistemas de aprendizaje automático, en los que los modelos ajustan sus parámetros de manera autónoma a partir de los datos de entrenamiento, generando representaciones internas que no siempre admiten una interpretación clara en términos humanos.

Desde esta perspectiva, la inteligencia artificial puede ser entendida como un conjunto de sistemas capaces de procesar grandes volúmenes de datos, identificar patrones y tomar decisiones o recomendaciones con distintos grados de autonomía. No obstante, su relevancia jurídica radica en su creciente utilización en ámbitos sensibles (la administración pública, los procesos laborales, el sistema de justicia penal, los servicios de salud y los servicios financieros) donde sus decisiones pueden afectar de manera directa el acceso a derechos y oportunidades. En estos contextos, la automatización no solo incrementa la eficiencia: también desplaza la responsabilidad, dificulta la impugnación y normaliza formas de exclusión que resultan difíciles de visibilizar bajo las categorías jurídicas tradicionales.

Diversos organismos internacionales han advertido que la incorporación de sistemas algorítmicos en la toma de decisiones plantea riesgos específicos para las libertades fundamentales. La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (por sus siglas en inglés: UNESCO), en su Recomendación sobre la ética de la inteligencia artificial de 2021, subraya la necesidad de que el desarrollo y uso de estos sistemas se rija por principios de respeto a la dignidad humana, la justicia, la transparencia y la responsabilidad<sup>2</sup>.

La Organización para la Cooperación y el Desarrollo Económico (OCDE), por su parte, adoptó en 2019 sus Principios sobre Inteligencia Artificial, que establecen estándares orientadores para los gobiernos y las empresas en relación con la transparencia, la explicabilidad y el respeto a los derechos humanos<sup>3</sup>.

El Consejo de Europa, que desde 2019 había identificado ámbitos de preocupación sobre IA y derechos humanos, culminó ese proceso con la adopción, el 17 de mayo de 2024, del Convenio Marco sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de Derecho (el primer tratado internacional jurídicamente vinculante en la

---

<sup>1</sup>PÉREZ LUÑO, Antonio Enrique. *Derechos Humanos, Estado de Derecho y Constitución*. Madrid: Tecnos, 1999.

<sup>2</sup>ORGANIZACIÓN DE LAS NACIONES UNIDAS PARA LA EDUCACIÓN, LA CIENCIA Y LA CULTURA (UNESCO). *Recomendación sobre la ética de la inteligencia artificial*. París: UNESCO, 2021.

<sup>3</sup>ORGANIZACIÓN PARA LA COOPERACIÓN Y EL DESARROLLO ECONÓMICOS (OCDE). *Principios sobre Inteligencia Artificial*. París: OCDE, 2019.

materia), cuyo objetivo central es garantizar que las actividades dentro del ciclo de vida de los sistemas de inteligencia artificial sean plenamente coherentes con los derechos humanos, la democracia y el Estado de derecho<sup>4</sup>.

Este instrumento representa el intento más ambicioso, hasta la fecha, de articular una respuesta jurídica comprehensiva a los desafíos que plantean estas tecnologías, aunque también ha sido objeto de críticas por su énfasis en el riesgo técnico en detrimento de un enfoque más robusto en derechos.

Esta brecha conceptual constituye uno de los principales obstáculos para la protección efectiva de las libertades fundamentales en entornos algorítmicos, y ha comenzado a ser abordada desde la doctrina jurídica más reciente, que analiza de manera sistemática las formas en que los algoritmos condicionan el ejercicio legítimo de los derechos fundamentales<sup>5</sup>.

No obstante, más allá de estos esfuerzos normativos, persiste una tensión fundamental que ninguno de estos instrumentos resuelve plenamente: los marcos jurídicos tradicionales fueron diseñados para regular relaciones entre sujetos identificables y conductas deliberadas, mientras que los sistemas de inteligencia artificial operan a través de lógicas probabilísticas, aprendizaje automático y procesamiento masivo de datos que desafían las categorías jurídicas existentes. La noción de responsabilidad presupone la identificación de un agente; la noción de discriminación presupone la intencionalidad o, al menos, la identificabilidad de la conducta discriminatoria; la noción de privacidad presupone la distinción entre lo público y lo privado. Los sistemas algorítmicos ponen en cuestión cada una de estas presunciones.

Esta brecha conceptual no es un problema técnico-jurídico menor, sino uno de los principales obstáculos para la protección efectiva de las libertades fundamentales en entornos algorítmicos. Abordarla exige no solo reinterpretar los alcances de estas libertades, sino también reconocer que su ejercicio se encuentra en proceso de reconfiguración, en un contexto donde el poder ya no se ejerce exclusivamente desde estructuras institucionales visibles, sino también a través de sistemas tecnológicos que median, condicionan y, en ocasiones, determinan decisiones con impacto directo en la vida de las personas.

### **3. RECONFIGURACIÓN DE LAS LIBERTADES FUNDAMENTALES EN ENTORNOS ALGORÍTMICOS**

El análisis de la reconfiguración de las libertades fundamentales en entornos algorítmicos exige descender del plano conceptual hacia las manifestaciones concretas en que dicha reconfiguración opera. No se trata de riesgos abstractos o hipotéticos, sino de transformaciones que se producen de manera diferenciada en distintas dimensiones del ejercicio de los derechos: en la capacidad de los individuos para tomar decisiones propias, en las condiciones bajo las cuales se desenvuelve su vida privada y en las posibilidades reales de acceder a oportunidades en condiciones de igualdad.

---

<sup>4</sup>CONSEJO DE EUROPA. Convenio Marco sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de Derecho (CETS n.º 225). Adoptado por el Comité de Ministros el 17 de mayo de 2024. Abierto a la firma en Vilna el 5 de septiembre de 2024.

<sup>5</sup>MARTÍN-HERRERA, David (ed.). La inteligencia artificial y el control algorítmico de los derechos fundamentales. Cizur Menor: Aranzadi, 2024. ISBN 978-84-1162-477-0.

Con ese propósito, el presente apartado examina tres dimensiones que, si bien son analíticamente distinguibles, se encuentran profundamente interrelacionadas en la práctica: la erosión de la autonomía frente a decisiones automatizadas, la transformación de la privacidad en el contexto de la vigilancia digital y la reproducción de desigualdades estructurales mediante sesgos algorítmicos. En cada una de ellas es posible identificar no solo la influencia de los marcos jurídicos tradicionales, sino la forma específica en que la lógica algorítmica desplaza, condiciona o vacía de contenido el ejercicio efectivo de las libertades fundamentales.

### 3.1 Autonomía y toma de decisiones automatizadas

La autonomía individual, entendida como la capacidad de las personas para tomar decisiones libres e informadas sobre su propia vida, constituye uno de los pilares fundamentales de los derechos humanos y, en particular, del derecho a la libertad personal en su dimensión más amplia. Desde Kant hasta las formulaciones contemporáneas del liberalismo político, la autonomía ha sido concebida como el atributo que distingue a las personas como sujetos de derecho: la capacidad de gobernarse a uno mismo, de ser autor de las propias decisiones y de no ser tratado meramente como un medio para los fines ajenos<sup>6</sup>.

Esta noción, sin embargo, fue constituida en un contexto en que las decisiones que afectaban la vida de las personas eran tomadas, en última instancia, por otros seres humanos: funcionarios públicos, jueces, empleadores, instituciones financieras. Incluso en los casos en que esas decisiones eran arbitrarias o discriminatorias, la arquitectura jurídica del Estado de derecho proporcionaba, al menos en teoría, mecanismos para identificar al responsable, exigir una justificación y obtener reparación. La automatización de decisiones introduce una ruptura en esta arquitectura que no puede ser ignorada.

A diferencia de las decisiones tradicionales, donde es posible identificar a un agente con responsabilidad jurídica, los sistemas algorítmicos operan a partir de modelos probabilísticos que procesan volúmenes inconmensurables de datos para generar predicciones o recomendaciones. El resultado es una forma de mediación tecnológica que, si bien puede incrementar la eficiencia o reducir ciertos sesgos individuales, también reduce la capacidad de los individuos para comprender, cuestionar o impugnar las decisiones que los afectan. La autonomía no desaparece, pero se ve progresivamente condicionada por sistemas cuyo funcionamiento resulta, en muchos casos, opaco incluso para sus propios diseñadores.

Diversos estudios han señalado que esta forma de toma de decisiones tiende a consolidar lo que se ha denominado una "racionalidad algorítmica": una forma de legitimación de las decisiones basada no en su justificación deliberativa, sino en su aparente objetividad técnica<sup>7</sup>. Como advierte Cathy O'Neil en su análisis sobre los efectos sociales de los modelos matemáticos, estos sistemas pueden operar como instrumentos de amplificación de desigualdades bajo la apariencia de neutralidad, ocultando los criterios,

---

<sup>6</sup>KANT, Immanuel. *Fundamentación de la metafísica de las costumbres*. Madrid: Alianza Editorial, 2002, pp. 103-115.

<sup>7</sup>DANAHER, John. *Algorithmic governmentality and the death of politics*. En: DANAHER, John y McARTHUR, Neil (eds.). *Robot Rules: Regulating Artificial Intelligence*. Cham: Palgrave Macmillan, 2018, pp. 71-94.

supuestos y valores que los sustentan detrás de una fachada de rigor estadístico<sup>8</sup>. La objetividad del dato y la sofisticación del modelo se convierten así en mecanismos de legitimación que dificultan el cuestionamiento político y jurídico de las decisiones automatizadas.

Esta dinámica tiene consecuencias particularmente graves en ámbitos donde están en juego derechos fundamentales: En el sistema de justicia penal, varios países han implementado herramientas de evaluación de riesgo algorítmica (como COMPAS en los Estados Unidos de Norteamérica) para informar decisiones sobre libertad bajo fianza o condicional. Estas herramientas han sido objeto de severas críticas por parte de investigadores que documentaron cómo sus predicciones presentaban tasas de error diferenciadas según la raza de los evaluados, asignando calificaciones de alto riesgo a personas negras con mayor frecuencia que a personas blancas con perfiles comparables<sup>9</sup>. El problema no reside únicamente en el sesgo del resultado, sino en la forma en que la autoridad del algoritmo desplaza la deliberación judicial y convierte en técnica lo que es, en esencia, una decisión normativa sobre la libertad de una persona.

La propia UNESCO, en su Kit de herramientas sobre inteligencia artificial y Estado de derecho, publicado en 2023 y dirigido específicamente a operadores de justicia, advierte que la opacidad de los sistemas algorítmicos aplicados al ámbito judicial compromete principios esenciales como el debido proceso y la independencia judicial, subrayando la necesidad de que los jueces mantengan un papel activo y no deleguen en los sistemas automatizados la sustancia de sus decisiones.<sup>10</sup>

En el ámbito del empleo, los sistemas de selección automatizada utilizan técnicas de procesamiento de lenguaje natural y análisis de patrones para filtrar currículos, evaluar entrevistas grabadas y clasificar candidatos. Más allá de los problemas de sesgo que se examinarán en el apartado siguiente, estos sistemas plantean una cuestión de fondo sobre la autonomía: cuando una persona no obtiene un empleo porque un algoritmo descartó su candidatura sin ninguna revisión humana significativa, ¿frente a quién ejerce su derecho a la defensa? ¿Qué mecanismo le permite conocer los criterios utilizados y cuestionar su aplicación?

Virginia Eubanks, en su documentación de casos de automatización en políticas públicas estadounidenses, muestra cómo las personas en situación de vulnerabilidad social quedan atrapadas en lo que denomina "redes digitales de vigilancia y control": sistemas automatizados que determinan el acceso a la vivienda, los beneficios sociales y los servicios de salud, sin que los afectados dispongan de información suficiente sobre los criterios aplicados ni de recursos efectivos para impugnarlos.<sup>11</sup> La conclusión que emerge de su análisis no es que la tecnología sea necesariamente perjudicial, sino que, desplegada sin las salvaguardias adecuadas, puede institucionalizar formas de exclusión

---

<sup>8</sup>O'NEIL, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Nueva York: Crown Publishing Group, 2016.

<sup>9</sup>ANGWIN, Julia, LARSON, Jeff, MATTU, Surya y KIRCHNER, Lauren. *Machine bias*. ProPublica [en línea]. 23 de mayo de 2016.

<sup>10</sup>ORGANIZACIÓN DE LAS NACIONES UNIDAS PARA LA EDUCACIÓN, LA CIENCIA Y LA CULTURA (UNESCO). *Inteligencia Artificial y Estado de Derecho: Kit de herramientas para el poder judicial*. París: UNESCO, 2023.

<sup>11</sup>EUBANKS, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. Nueva York: St. Martin's Press, 2018.

que resultan en especial difíciles de combatir de manera precisa porque se presentan como decisiones técnicas y no políticas.

Frente a este panorama, la autonomía se ve erosionada no solo por la sustitución de la decisión humana, sino por la dificultad estructural de los individuos para ejercer control sobre procesos que inciden directamente en su acceso a derechos y oportunidades. La imposibilidad de comprender los criterios utilizados por los sistemas, la ausencia de mecanismos efectivos de rendición de cuentas y la asimetría de información entre quienes diseñan los algoritmos y quienes son afectados por ellos configuran un escenario en que el derecho a la defensa y el principio de motivación de las decisiones (pilares del debido proceso) se convierten en garantías formalmente vigentes, pero materialmente vaciadas de contenido<sup>12</sup>.

En términos conceptuales, este fenómeno puede describirse como una reconfiguración de las condiciones de ejercicio de la autonomía: los individuos siguen siendo formalmente reconocidos como sujetos de derecho, pero las decisiones que estructuran sus posibilidades vitales son tomadas en entornos donde su agencia está sistemáticamente limitada. En lugar de decisiones plenamente deliberadas, los individuos se enfrentan a entornos donde sus opciones están preconfiguradas por sistemas que operan con lógicas de optimización y predicción, desplazando progresivamente el centro de la decisión desde el sujeto hacia la infraestructura tecnológica<sup>13</sup>.

### 3.2 Vigilancia digital y transformación de la privacidad

La privacidad ha sido reconocida como un derecho fundamental en los principales instrumentos internacionales de derechos humanos, desde el artículo 12 de la Declaración Universal hasta el artículo 17 del Pacto Internacional de Derechos Civiles y Políticos<sup>14</sup>. En su formulación clásica, este derecho protege a las personas frente a la intromisión arbitraria en su vida privada, su familia, su domicilio y su correspondencia. Se trata, en esencia, de un derecho de exclusión: la persona tiene el derecho de mantener ciertos espacios de su vida a salvo de la mirada y la intervención de terceros, incluido el Estado.

Esta concepción fue construida en un contexto en que la recopilación de información sobre las personas era costosa, fragmentada y requería de actos deliberados de intromisión. La vigilancia estatal requería agentes, archivos físicos y procedimientos formales; la vigilancia privada era limitada por las mismas restricciones tecnológicas y económicas. En ese contexto, el modelo jurídico de protección basado en la prohibición de la intromisión injustificada resultaba adecuado<sup>15</sup>.

La digitalización ha transformado radicalmente ese contexto. La recopilación de datos se ha convertido en una actividad continua, automatizada y prácticamente invisible que acompaña cada interacción de las personas con los entornos digitales. Cada búsqueda en internet, cada desplazamiento geolocalizado, cada compra, cada

<sup>12</sup>EUBANKS, Virginia. *Automating Inequality...*, op. cit.; MARTÍN-HERRERA, David (ed.). *La inteligencia artificial y el control algorítmico de los derechos fundamentales*, op. cit.

<sup>13</sup>ZUBOFF, Shoshana. *The Age of Surveillance Capitalism...*, op. cit. EUBANKS, Virginia. *Automating Inequality...*, op. cit.

<sup>14</sup>ORGANIZACIÓN DE LAS NACIONES UNIDAS. *Declaración Universal de Derechos Humanos*. París: ONU, 1948, art. 12; ORGANIZACIÓN DE LAS NACIONES UNIDAS. *Pacto Internacional de Derechos Civiles y Políticos*. Nueva York: ONU, 1996, art. 17.

<sup>15</sup>ZUBOFF, Shoshana. *The Age of Surveillance Capitalism...*, op. cit.

comunicación a través de plataformas digitales, cada patrón de sueño registrado por un dispositivo portátil. Produce datos que son capturados, almacenados y procesados por sistemas que, en su mayor parte, operan fuera del conocimiento y del control de los propios usuarios.

Shoshana Zuboff denomina este fenómeno como "capitalismo de la vigilancia": una lógica económica en la que las experiencias humanas son transformadas en datos, los datos son procesados para producir predicciones sobre el comportamiento futuro, y esas predicciones se venden en mercados que trafican con certezas sobre lo que las personas harán, pensarán o desaparecerán<sup>16</sup>. En este modelo, la privacidad no es simplemente violada; es convertida en un recurso productivo. El valor económico no reside en la información en sí misma, sino en la capacidad predictiva que esa información habilita, y esa capacidad predictiva es utilizada no solo para anticiparse al comportamiento, sino para modificarlo.

Este diagnóstico tiene implicaciones jurídicas de primera magnitud. Si la privacidad es transformada en un insumo productivo dentro de un sistema económico, entonces su protección no puede limitarse a la prohibición de intromisiones puntuales: requiere una reconfiguración del marco normativo que reconozca la dimensión estructural de la recopilación de datos y sus efectos sobre la autonomía y la dignidad humana<sup>17</sup>.

El Reglamento General de Protección de Datos de la Unión Europea representa un avance significativo en esta dirección, al establecer principios como la minimización de datos, la limitación de la finalidad y el consentimiento informado<sup>18</sup>. No obstante, su efectividad ha sido cuestionada por quienes señalan que el modelo de consentimiento es fácilmente eludible mediante interfaces diseñadas para obtener un acuerdo superficial sin una comprensión real de sus implicaciones<sup>19</sup>.

Byung-Chul Han ofrece una perspectiva complementaria al señalar que la vigilancia contemporánea no se ejerce únicamente desde fuera, sino que se integra en las prácticas cotidianas de los propios individuos a través de lo que denomina "psicopolítica"<sup>20</sup>. En la sociedad de la transparencia que describe Han, los sujetos no solo son vigilados: participan activamente en su propia vigilancia, exponiendo voluntariamente aspectos de su vida privada en redes sociales, aceptando condiciones de uso sin leerlas, consintiendo el rastreo de su ubicación a cambio de servicios gratuitos. La vigilancia se naturaliza y se convierte en el precio implícito de la participación en la vida digital.

Esta dimensión subjetiva de la vigilancia tiene consecuencias importantes para el análisis jurídico: si las personas participan voluntariamente en dinámicas que erosionan su propia privacidad, ¿puede el derecho asumir que el consentimiento expresado es libre e informado? ¿O debe reconocer que ese consentimiento opera en condiciones

---

<sup>16</sup>ZUBOFF, Shoshana. *The Age of Surveillance Capitalism...* op. cit.

<sup>17</sup>ZUBOFF, Shoshana. *The Age of Surveillance Capitalism...*, op. cit. PÉREZ LUÑO, Antonio Enrique. *Derechos Humanos, Estado de Derecho y Constitución*, op. cit.

<sup>18</sup>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. Diario Oficial de la Unión Europea [en línea]. 4 de mayo de 2016, L 119, pp. 1-88.

<sup>19</sup>ZUBOFF, Shoshana. *The Age of Surveillance Capitalism...*, op. cit.

<sup>20</sup>HAN, Byung-Chul. *Psicopolítica: Neoliberalismo y nuevas técnicas de poder*. Barcelona: Herder, 2014.

de asimetría estructural de información y poder que lo invalidan como fundamento suficiente de legitimidad?<sup>21</sup> La respuesta a estas preguntas no es meramente técnica: implica una toma de posición sobre el alcance de la autonomía y las obligaciones del Estado en relación con la protección de condiciones que hagan posible su ejercicio real.

En términos concretos, la vigilancia digital se manifiesta en la capacidad de las plataformas y los Estados para rastrear ubicaciones, analizar comportamientos en línea, inferir preferencias, orientaciones políticas o estados emocionales, y construir perfiles detallados que determinan los contenidos que se muestran, los precios que se ofrecen, las noticias que se reciben y las oportunidades que se presentan. El reconocimiento facial, el análisis de sentimientos, la inferencia de rasgos de personalidad a partir de patrones de comportamiento digital: todas estas tecnologías amplían las capacidades de vigilancia a dimensiones que los marcos jurídicos tradicionales no estaban equipados para enfrentar<sup>22</sup>.

En el ámbito estatal, la vigilancia algorítmica ha encontrado justificación en el discurso de la seguridad pública y la prevención del terrorismo. China ha desarrollado uno de los sistemas de vigilancia más extensos del mundo, que incluye reconocimiento facial en espacios públicos, sistemas de puntuación social y monitoreo de comunicaciones digitales. Pero la tendencia no se limita a regímenes autoritarios: en democracias liberales, los programas de vigilancia masiva revelados por Edward Snowden pusieron de manifiesto que la recopilación de datos a gran escala por agencias de inteligencia había adquirido dimensiones incompatibles con el derecho a la privacidad tal como lo reconocen los instrumentos internacionales de derechos humanos<sup>23</sup>.

La transformación de la privacidad en entornos algorítmicos no radica únicamente en la posibilidad de su vulneración, sino en su redefinición como elemento funcional dentro de sistemas de vigilancia y control. Cuando la privacidad se convierte en la excepción y la exposición en la norma, el marco jurídico que la protege pierde su capacidad de operar sobre la realidad social que pretende regular. La inteligencia artificial, al potenciar estas dinámicas mediante el análisis automatizado de datos, contribuye a consolidar un entorno en que la vigilancia se normaliza y se vuelve prácticamente imperceptible, dificultando su problematización desde los marcos normativos existentes.

### 3.3 Sesgos algorítmicos y reproducción de desigualdades

Uno de los argumentos más frecuentes en favor de la utilización de sistemas algorítmicos en la toma de decisiones es precisamente su aparente capacidad para superar las limitaciones y sesgos del juicio humano. Un juez, un empleador o un funcionario público pueden estar influenciados por prejuicios conscientes o inconscientes; un algoritmo, en cambio, trataría a todos los casos de la misma manera, aplicando criterios uniformes y objetivos. Esta intuición, sin embargo, descansa sobre una premisa errónea ampliamente cuestionada por la investigación empírica: los algoritmos no son neutrales, porque no

---

<sup>21</sup>HAN, Byung-Chul. *Psicopolítica...* op. cit.; ZUBOFF, Shoshana. *The Age of Surveillance Capitalism...*, op. cit.

<sup>22</sup>ZUBOFF, Shoshana. *The Age of Surveillance Capitalism...*, op. cit. HAN, Byung-Chul. *Psicopolítica...*, op. cit. GREENWALD, Glenn. *No Place to Hide...*, op. cit.

<sup>23</sup>GREENWALD, Glenn. *No Place to Hide: Edward Snowden, the NSA, and the U.S. Surveillance State*. Nueva York: Metropolitan Books, 2014.

operan en el vacío, sino sobre datos que reflejan el mundo tal como ha sido y no necesariamente como debería ser<sup>24</sup>.

Los sesgos algorítmicos surgen, en primer lugar, de los datos de entrenamiento. Si un sistema de selección de personal es entrenado con datos históricos de contratación de una empresa que, durante décadas, ha contratado predominantemente hombres para posiciones de liderazgo, el algoritmo aprenderá a asociar el perfil masculino con la idoneidad para esos roles, replicando y potencialmente amplificando el sesgo histórico.

Si un sistema de evaluación crediticia es entrenado con datos de poblaciones que han tenido acceso desigual al sistema financiero formal, sus predicciones reflejarán y consolidarán esas desigualdades de acceso. Amazon tuvo que discontinuar en 2018 un sistema de selección automatizada de candidatos precisamente porque su modelo había aprendido a penalizar currículos que incluían palabras asociadas con el género femenino<sup>25</sup>.

En segundo lugar, los sesgos pueden derivarse de la elección de las variables utilizadas por el modelo. En el ámbito del crédito, por ejemplo, el uso del código postal como variable predictiva puede funcionar como un proxy de la raza o el origen étnico en sociedades donde la segregación residencial ha sido históricamente pronunciada, produciéndose efectos discriminatorios indirectos que son jurídicamente tan problemáticos como la discriminación directa. En el ámbito de la justicia penal, las variables utilizadas para predecir la reincidencia (como el nivel educativo, el historial de empleo o el vecindario de residencia) reproducen las condiciones de desventaja estructural de los grupos más vulnerables, convirtiendo la desigualdad social en un factor de riesgo individual<sup>26</sup>.

Cathy O'Neil caracteriza estos sistemas como "armas de destrucción matemática": modelos que se presentan como objetivos e infalibles, pero que en realidad codifican valores, amplían desigualdades y afectan desproporcionadamente a las personas con menos recursos para impugnarlos<sup>27</sup>. La perversidad del fenómeno reside, en buena medida, en el mecanismo de retroalimentación: un sistema que predice que ciertos individuos son más propensos a incumplir sus obligaciones crediticias genera que esas personas reciban condiciones más onerosas o sean excluidas del sistema, lo que incrementa su vulnerabilidad financiera y confirman, en apariencia, la predicción original.

Virginia Eubanks documenta casos concretos en que herramientas automatizadas aplicaba a políticas públicas han producido exclusiones sistemáticas de poblaciones en situación de vulnerabilidad<sup>28</sup>. En Indiana (EE.UU.), un sistema automatizado de gestión de beneficios sociales generó miles de cancelaciones de ayudas por supuestos errores o fraudes que, en la mayoría de los casos, eran atribuibles a fallas del propio sistema o a condiciones estructurales de las personas afectadas (dificultades para acceder a documentación, barreras idiomáticas, falta de conectividad digital). En Los Ángeles (EE.UU.),

---

<sup>24</sup>O'NEIL, Cathy. *Weapons of Math Destruction...*, op. cit. EUBANKS, Virginia. *Automating Inequality...*, op. cit.

<sup>25</sup>DASTIN, Jeffrey. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters [en línea]. 10 de octubre de 2018.

<sup>26</sup>O'NEIL, Cathy, *Weapons of Math Destruction...*, op. cit.

<sup>27</sup>O'NEIL, Cathy, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Nueva York: Crown Publishing Group, 2016.

<sup>28</sup>EUBANKS, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. Nueva York: St. Martin's Press, 2018. ISBN 978-1-250-07431-1.

un sistema predictivo de identificación de personas en riesgo de quedarse sin hogar fue cuestionado por producir perfiles que penalizaban a las personas más marginalizadas, precisamente aquellas que más necesitaban los servicios.

A ello se suma el problema de la opacidad. Muchos de los sistemas de inteligencia artificial utilizados en contextos de alto impacto operan como cajas negras: sus procesos internos no son accesibles ni para los usuarios afectados ni, en muchos casos, para las propias instituciones que los utilizan. Esta opacidad no solo dificulta la rendición de cuentas, sino que restringe el acceso a mecanismos efectivos de defensa: si una persona no puede conocer los criterios que determinaron una decisión adversa, difícilmente puede impugnarla de manera efectiva. El principio de igualdad ante la ley presupone, entre otras cosas, la posibilidad de cuestionar que afectan las propias decisiones que afectan los propios derechos; un sistema opaco hace que esa posibilidad sea meramente formal<sup>29</sup>.

Desde la perspectiva de los derechos humanos, los sesgos algorítmicos plantean un desafío directo al principio de no discriminación, que prohíbe tanto la discriminación directa (el trato diferente basado explícitamente en una característica protegida) como la discriminación indirecta (trato aparentemente neutro que produce efectos desproporcionadamente adversos sobre un grupo protegido). Los sistemas algorítmicos son especialmente propensos a la discriminación indirecta, porque pueden incorporar variables que funcionan como “proxies” de características protegidas sin utilizarlas explícitamente, produciendo efectos discriminatorios que resultan difíciles de identificar y probar con los instrumentos procesales disponibles<sup>30</sup>.

En consecuencia, los sesgos algorítmicos no deben entenderse como fallas técnicas aisladas, corregibles mediante mejoras en los modelos o en los datos. Son, más bien, manifestaciones de problemas estructurales que preexisten la tecnología y que esta amplifica y formaliza. La inteligencia artificial no crea la desigualdad social, pero puede darle una nueva forma de legitimidad al convertirla en el resultado de un proceso técnico aparentemente objetivo<sup>31</sup>. Esta conversión es jurídicamente inaceptable cuando ocurre en ámbitos donde están en juego derechos fundamentales, y exige incorporar criterios de equidad y justicia no solo en la regulación, sino en el propio diseño de los sistemas.

#### 4. PODER ALGORÍTMICO Y CRISIS ESTRUCTURAL DEL ESTADO DE DERECHO

El análisis de la autonomía, la vigilancia y los sesgos algorítmicos permite advertir que los impactos de la inteligencia artificial en las libertades fundamentales no son fenómenos aislados ni contingentes. Son expresiones de una transformación más profunda en las formas de ejercicio del poder. En este contexto, resulta necesario articular un concepto que permita nombrar y analizar esta transformación: el poder algorítmico.

Por poder algorítmico se entiende, en el marco de este trabajo, la capacidad de los sistemas de inteligencia artificial para condicionar, orientar y determinar decisiones con

---

<sup>29</sup>O'NEIL, Cathy. *Weapons of Math Destruction...*, op. cit.; MARTÍN-HERRERA, David (ed.). *La inteligencia artificial y el control algorítmico de los derechos fundamentales*, op. cit.

<sup>30</sup>O'NEIL, Cathy. *Weapons of Math Destruction...*, op. cit. EUBANKS, Virginia. *Automating Inequality...*, op. cit.; MARTÍN-HERRERA, David (ed.). *La inteligencia artificial y el control algorítmico de los derechos fundamentales*, op. cit.

<sup>31</sup>O'NEIL, Cathy. *Weapons of Math Destruction...*, op. cit.; EUBANKS, Virginia. *Automating Inequality...*, op. cit.

efectos jurídicos y sociales relevantes, operando a través de lógicas automatizadas que se integran en las estructuras institucionales de manera progresiva y, en muchos casos, imperceptible. A diferencia de las formas tradicionales de poder (que se ejercían a través de mandatos explícitos, instituciones identificables y procedimientos formales) el poder algorítmico se caracteriza por su opacidad, su escala, su velocidad y su capacidad de integrarse en la vida cotidiana sin suscitar las resistencias que suele generar el poder visible.

Zuboff, en su análisis del capitalismo de la vigilancia, describe cómo las grandes plataformas tecnológicas han desarrollado capacidades de influencia sobre el comportamiento humano que no dependen de la coerción ni del convencimiento deliberativo, sino de la modificación del entorno en que se producen las decisiones<sup>32</sup>.

Mediante el control de los flujos de información, la personalización de los contenidos y la anticipación de las preferencias, estos sistemas ejercen una forma de poder que opera por debajo del umbral de la conciencia deliberativa, moldeando las condiciones en que los individuos forman sus preferencias y toman sus decisiones. Esta forma de poder es, en cierto sentido, más eficaz que el poder coercitivo: no requiere reprimir la voluntad, sino darle forma.

En el ámbito estatal, el poder algorítmico se manifiesta en la utilización de sistemas de inteligencia artificial para decisiones de alto impacto en políticas públicas, gestión de servicios sociales, administración de justicia y vigilancia. Cuando un sistema algorítmico determina quién recibe una prestación social, quién es sometido a una medida cautelar o quién es identificado como sospechoso, el Estado ejerce su poder de una manera que es formalmente atribuible a sus instituciones, pero materialmente mediada por sistemas cuyo funcionamiento no siempre comprende ni controla; Esta mediación tecnológica introduce una forma de responsabilidad difusa que resulta problemática desde la perspectiva del Estado de derecho<sup>33</sup>.

El Estado de derecho, en su formulación clásica, descansa sobre un conjunto de principios que constituyen su fundamento: la legalidad, la separación de poderes, la garantía de derechos fundamentales, la transparencia y la rendición de cuentas. Cada uno de estos principios se ve tensionado por las dinámicas del poder algorítmico<sup>34</sup>.

El principio de legalidad exige que las decisiones que afectan derechos individuales tengan un fundamento normativo claro y sean adoptadas conforme a procedimientos establecidos. Cuando las decisiones son tomadas por sistemas algorítmicos cuyos criterios no han sido establecidos por normas accesibles al público, o cuando esos criterios se modifican automáticamente mediante el aprendizaje continuo del sistema, la cadena entre la norma y la decisión se rompe de manera que dificulta el control jurídico<sup>35</sup>.

---

<sup>32</sup>ZUBOFF, Shoshana. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Nueva York: PublicAffairs, 2019.

<sup>33</sup>ZUBOFF, Shoshana. *The Age of Surveillance Capitalism...*, op. cit.; MARTÍN-HERRERA, David (ed.). *La inteligencia artificial y el control algorítmico de los derechos fundamentales*, op. cit.

<sup>34</sup>PÉREZ LUÑO, Antonio Enrique. *Derechos Humanos, Estado de Derecho y Constitución*, op. cit.

<sup>35</sup>PÉREZ LUÑO, Antonio Enrique. *Derechos Humanos, Estado de Derecho y Constitución*, op. cit.; MARTÍN-HERRERA, David (ed.). *La inteligencia artificial y el control algorítmico de los derechos fundamentales*, op. cit.

El principio de transparencia exige que las decisiones que afectan derechos sean explicables y accesibles a quienes las sufren. Los sistemas de inteligencia artificial basados en aprendizaje profundo son paradigmáticamente opacos: incluso sus diseñadores pueden ser incapaces de explicar por qué un modelo específico produjo un resultado específico. El llamado "derecho a la explicación" introducido por el RGPD europeo representa un intento de responder a esta exigencia, pero su efectividad ha sido cuestionada por quienes señalan que las explicaciones producidas por los sistemas pueden ser post hoc, simplificadas o directamente incompatibles con la estructura técnica del modelo que las genera.<sup>36</sup>

El principio de responsabilidad exige que quien ejerce poder en nombre del Estado o con autorización del Estado sea identificable y responda por sus decisiones. La cadena de responsabilidad en el despliegue de sistemas de inteligencia artificial es, con frecuencia, extraordinariamente compleja: involucra a los diseñadores del modelo, a quienes recopilaban los datos de entrenamiento, a quienes decidieron implementar el sistema, a quienes lo utilizan en casos concretos y a quienes lo auditan o supervisan. Esta complejidad puede derivar en situaciones de responsabilidad difusa en que nadie es efectivamente responsable, aunque el daño sea real y cuantificable.

La creciente concentración de capacidades tecnológicas en manos de un número reducido de grandes corporaciones introduce una dimensión adicional a la crisis del Estado de derecho. Estas entidades no solo desarrollan y controlan los sistemas de inteligencia artificial más avanzados, sino que también gestionan los datos que los alimentan y las infraestructuras sobre las que operan, lo que les otorga una posición de poder que trasciende la esfera económica para incidir directamente en la gobernanza de ámbitos que, en el paradigma del Estado de derecho, eran considerados de competencia pública. La pregunta sobre si los marcos regulatorios nacionales e internacionales son capaces de limitar y supervisar el ejercicio de este poder es, hoy, una de las más urgentes de la teoría jurídica contemporánea.

En América Latina, estas tensiones adquieren una dimensión particular. La región combina una alta penetración de plataformas digitales con marcos regulatorios débiles, instituciones con capacidades técnicas limitadas para supervisar el uso de la inteligencia artificial y poblaciones con condiciones de vulnerabilidad estructural que las hacen especialmente expuestas a los riesgos de exclusión algorítmica. La CEPAL ha advertido que la gobernanza de datos y algoritmos es una cuestión estratégica para el desarrollo sostenible e inclusivo de la región, y que, sin marcos adecuados, la digitalización puede profundizar las brechas existentes en lugar de reducirlas<sup>37</sup>.

En conjunto, estos elementos configuran una crisis estructural del Estado de derecho: no en el sentido de que este haya desaparecido o colapsado, sino en el sentido de que los principios que lo sustentan (legalidad, transparencia, responsabilidad, control del poder) se ven sistemáticamente tensionados por dinámicas tecnológicas que escapan a los mecanismos de control diseñados para un entorno analógico. La inteligencia artificial

<sup>36</sup>UNIÓN EUROPEA. Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, Op. Cit. Art. 22 y considerando 71-72.

<sup>37</sup>COMISIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE (CEPAL). Datos, algoritmos y políticas públicas. Santiago de Chile: CEPAL, 2022.

no solo introduce nuevos desafíos regulatorios: cuestiona las premisas sobre las cuales se han construido los mecanismos de protección de las libertades fundamentales, y obliga a preguntarse si esos mecanismos son capaces de operar efectivamente en un entorno radicalmente diferente al que les dio origen.

## **5. HACIA UNA RECONFIGURACIÓN NORMATIVA DESDE LOS DERECHOS HUMANOS**

Las tensiones identificadas a lo largo de este análisis evidencian que los marcos jurídicos tradicionales resultan insuficientes para responder a los desafíos que plantea la inteligencia artificial. Esta insuficiencia no es una cuestión de voluntad política ni de falta de regulación en términos cuantitativos: es una cuestión de categorías. Los instrumentos normativos disponibles fueron diseñados para operar sobre supuestos (la identificabilidad de los actores, la deliberación humana de las decisiones, la distinción entre lo público y lo privado) que los entornos algorítmicos ponen sistemáticamente en cuestión.

Frente a este diagnóstico, diversas voces han propuesto respuestas que van desde la simple extensión de las categorías existentes hasta la creación de regímenes jurídicos completamente nuevos. En este trabajo se sostiene que ninguno de estos extremos resulta satisfactorio. La extensión acrítica de las categorías existentes corre el riesgo de producir una falsa sensación de protección sin modificar las condiciones estructurales que generan la vulneración de derechos. La creación de regímenes completamente nuevos, por su parte, puede desconectarse de los estándares de derechos humanos laboriosamente construidos durante décadas y abrir espacios para la legitimación de restricciones de derechos bajo el pretexto de la innovación tecnológica.

Lo que se propone, en cambio, es una reconfiguración normativa: un proceso de revisión y reformulación de los marcos jurídicos existentes que, desde una perspectiva centrada en los derechos humanos, incorpore las particularidades de los sistemas algorítmicos sin abandonar los estándares que hacen de esos marcos un instrumento de protección de la dignidad humana. Esta reconfiguración no implica necesariamente la creación inmediata de nuevas disposiciones legales altamente especializadas, sino el desarrollo de principios orientadores capaces de guiar tanto el diseño como la implementación y la regulación de la inteligencia artificial.

El primero de estos principios es el de centralidad de la persona. Este principio establece que los sistemas de inteligencia artificial deben estar al servicio del desarrollo humano y no a la inversa. En términos concretos, esto implica que cuando un sistema automatizado entra en conflicto con la dignidad, la autonomía o los derechos fundamentales de una persona, debe prevalecer la protección de estos sobre la eficiencia del sistema. La UNESCO, en su Recomendación sobre la ética de la inteligencia artificial, establece explícitamente que el bienestar humano debe ser el objetivo central de las políticas de inteligencia artificial, y que ninguna consideración de eficiencia o progreso tecnológico puede justificar la subordinación de la dignidad humana a la lógica de la optimización técnica<sup>38</sup>.

---

<sup>38</sup>ORGANIZACIÓN DE LAS NACIONES UNIDAS PARA LA EDUCACIÓN, LA CIENCIA Y LA CULTURA (UNESCO). Recomendación sobre la ética de la inteligencia artificial. París: UNESCO, 2021.

El segundo principio es el de transparencia y explicabilidad. Este principio exige que los sistemas algorítmicos utilizados en contextos de alto impacto sobre derechos fundamentales sean comprensibles y cuestionables por quienes resultan afectados por sus decisiones. Esta exigencia no se limita a la publicación del código fuente o a la descripción técnica del modelo: requiere una explicación inteligible de los criterios utilizados, las variables consideradas y los mecanismos de ponderación empleados, en un lenguaje accesible para las personas no especializadas. La Comisión Europea ha subrayado, desde sus directrices éticas de 2019 y de manera vinculante a través del AI Act de 2024, que la explicabilidad es un requisito fundamental que debe acompañar el desarrollo técnico de los sistemas y no ser tratada como una adición posterior<sup>39</sup>.

El tercero es el principio de responsabilidad y rendición de cuentas. Este principio establece que deben existir mecanismos claros para identificar a los actores responsables del diseño, la implementación y el uso de sistemas de inteligencia artificial, y para exigirles responder por los daños que esos sistemas puedan causar. La dilución de la responsabilidad en estructuras tecnológicas complejas no puede ser utilizadas como escudo frente a las reclamaciones de quienes sufren daños derivados del funcionamiento de esos sistemas. En este sentido, resulta necesario desarrollar conceptos jurídicos de responsabilidad adaptados a las características de los entornos algorítmicos, que permitan atribuir responsabilidades de manera proporcional a los distintos actores involucrados en el ciclo de vida de un sistema de inteligencia artificial.

El cuarto principio es el de equidad y no discriminación. Este principio exige que el diseño, el entrenamiento, la implementación y la auditoría de los sistemas de inteligencia artificial incorporen criterios de justicia orientados a prevenir la reproducción de desigualdades estructurales. Esto implica, entre otras cosas, que los datos de entrenamiento sean evaluados críticamente desde una perspectiva de derechos, que los modelos sean sometidos a pruebas de impacto diferencial sobre grupos en situación de vulnerabilidad, y que los resultados sean auditados periódicamente por organismos independientes con capacidad técnica y autoridad jurídica para intervenir cuando se detecten patrones discriminatorios<sup>40</sup>.

Estos cuatro principios no deben entenderse como elementos aislados, sino como componentes de un enfoque integral que reconoce la interdependencia entre la innovación tecnológica y la protección de los derechos humanos. Su implementación efectiva requiere, además, el fortalecimiento de las capacidades institucionales para supervisar el uso de la inteligencia artificial en los distintos sectores: no basta con establecer principios en instrumentos normativos si las instituciones encargadas de aplicarlos carecen de los recursos técnicos y jurídicos para hacerlo.

En el contexto latinoamericano, esta reconfiguración normativa enfrenta desafíos específicos. La heterogeneidad regulatoria entre los países de la región, la dependencia tecnológica respecto de empresas y plataformas con sede en otros continentes, y las

---

<sup>39</sup>COMISIÓN EUROPEA. Ethics Guidelines for Trustworthy AI. Bruselas: High-Level Expert Group on Artificial Intelligence, 2019.

<sup>40</sup>UNIÓN EUROPEA. 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. Diario Oficial de la Unión Europea [en línea]. 12 de julio de 2024, L 2024/1689.

asimetrías estructurales en el acceso a la tecnología plantean interrogantes sobre la capacidad de los Estados de la región para implementar marcos regulatorios efectivos de manera unilateral. Esto sugiere la necesidad de avanzar también en mecanismos de cooperación regional que permitan desarrollar estándares comunes, compartir capacidades institucionales y negociar con mayor poder frente a los actores tecnológicos globales.

En última instancia, el desafío de la reconfiguración normativa no radica únicamente en regular la inteligencia artificial, sino en garantizar que su incorporación en la vida social se realice bajo condiciones que respeten y fortalezcan el Estado de derecho. Esto exige reconocer que las transformaciones tecnológicas actuales no solo demandan respuestas jurídicas puntuales, sino también una revisión crítica de los fundamentos sobre los cuales se han construido los sistemas de protección de derechos en las sociedades contemporáneas. Una revisión que, lejos de abandonar esos fundamentos, los reafirme y los adapte a las condiciones de un entorno radicalmente diferente al que los vio nacer.

## **6. CONCLUSIONES**

El análisis desarrollado a lo largo de este artículo permite sostener, con fundamento, que la inteligencia artificial no constituye únicamente un desafío técnico o regulatorio. Es un fenómeno que reconfigura de manera profunda las condiciones bajo las cuales se ejercen las libertades fundamentales, afectando la autonomía de los individuos, transformando las estructuras de la privacidad y amplificando desigualdades históricas mediante mecanismos que se presentan como objetivos pero que reproducen los sesgos del mundo social del que emergen.

La primera conclusión que se desprende del análisis es que la distinción entre riesgos contingentes y transformaciones estructurales tiene una importancia decisiva para el enfoque jurídico. Si los efectos de la inteligencia artificial fueran meramente contingentes (errores corregibles, fallas superables con mayor regulación), bastaría con ampliar y actualizar los marcos normativos existentes. Pero si, como se ha argumentado, las transformaciones son estructurales, entonces la respuesta jurídica no puede limitarse a ajustes superficiales: requiere una revisión de las premisas conceptuales sobre las cuales se ha construido el sistema de protección de derechos.

La segunda conclusión es que el poder algorítmico representa una forma cualitativamente nueva de ejercicio del poder, que no se limita a los actores estatales, sino que involucra también a grandes corporaciones tecnológicas capaces de incidir, a través de sus plataformas y sistemas, en ámbitos de la vida social que el paradigma del Estado de derecho reservaba a la regulación pública. Esta dispersión del poder entre actores estatales y privados, mediada por infraestructuras tecnológicas complejas, plantea interrogantes que las categorías jurídicas tradicionales no están equipadas para responder de manera adecuada.

La tercera conclusión es que la crisis estructural del Estado de derecho que genera el poder algorítmico no es un estado final sino un proceso: una erosión progresiva de los principios de legalidad, transparencia, responsabilidad y control del poder que, si no se atiende con la seriedad que merece, puede derivar en un vaciamiento material de las garantías jurídicas. Las normas seguirán existiendo, los tribunales seguirán funcionando,

las declaraciones de derechos seguirán siendo invocadas, pero su capacidad de operar efectivamente sobre las condiciones reales en las que se desenvuelve la vida de las personas quedará sistemáticamente limitada.

La cuarta conclusión es que la reconfiguración normativa propuesta en este trabajo, orientada por los principios de centralidad de la persona, transparencia, responsabilidad y equidad, no es una propuesta utópica sino una exigencia que se deriva de los compromisos internacionales de derechos humanos ya asumidos por los Estados. El desafío no consiste en inventar nuevos fundamentos, sino en tomarse en serio los que ya existen y aplicarlos con coherencia a un entorno tecnológico que los pone a prueba de maneras nuevas y complejas.

Finalmente, cabe señalar que la discusión sobre inteligencia artificial y libertades fundamentales no es un debate académico sin consecuencias prácticas. Es una conversación sobre el tipo de sociedad que estamos construyendo y sobre las condiciones en que las personas podrán (o no) ejercer su autonomía, proteger su privacidad y acceder a un trato igualitario en un entorno cada vez más mediado por sistemas tecnológicos. La pregunta que subyace no es únicamente cómo regular la inteligencia artificial, sino hasta qué punto estamos dispuestos a replantear las bases del orden jurídico para evitar que la protección de los derechos se convierta en una promesa vacía frente al avance del poder algorítmico. Esa pregunta no admite respuestas técnicas: exige una decisión política y ética sobre los valores que queremos que orienten el desarrollo tecnológico de nuestras sociedades.

## 7. REFERENCIAS BIBLIOGRÁFICAS

- ANGWIN, Julia; LARSON, Jeff; MATTU, Surya y KIRCHNER, Lauren. Machine bias. ProPublica [en línea]. 23 de mayo de 2016. [Consulta: 29 de marzo de 2026]. [Disponible en: <https://www.propublica.org/article/machine-bias-risk-assessment-in-criminal-sentencing>].
- COMISIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE (CEPAL). Datos, algoritmos y políticas públicas. Santiago de Chile: CEPAL, 2022. [Consulta: 31 de marzo de 2026]. [Disponible en: <https://repositorio.cepal.org>].
- COMISIÓN EUROPEA. Ethics Guidelines for Trustworthy AI. Bruselas: High-Level Expert Group on Artificial Intelligence, 2019. [Consulta: 06 de abril de 2026]. [Disponible en: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>].
- CONSEJO DE EUROPA. Convenio Marco sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de Derecho (CETS n.º 225). Adoptado el 17 de mayo de 2024. [Consulta: 25 de marzo de 2026]. [Disponible en: <https://www.coe.int/es/web/conventions/full-list/-/conventions/treaty/225>].
- DANAHER, John. Algorithmic governmentality and the death of politics. En: DANAHER, John y McARTHUR, Neil (eds.). Robot Rules: Regulating Artificial Intelligence. Cham: Palgrave Macmillan, 2018, pp. 71-94.
- DASTIN, Jeffrey. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters [en línea]. 10 de octubre de 2018. [Consulta: 30 de marzo de 2026]. [Disponible en: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>].
- EUBANKS, Virginia. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. Nueva York: St. Martin's Press, 2018.
- GREENWALD, Glenn. No Place to Hide: Edward Snowden, the NSA, and the U.S. Surveillance State. Nueva York: Metropolitan Books, 2014.

- HAN, Byung-Chul. *Psicopolítica: Neoliberalismo y nuevas técnicas de poder*. Barcelona: Herder, 2014.
- KANT, Immanuel. *Fundamentación de la metafísica de las costumbres*. Madrid: Alianza Editorial, 2002.
- MARTÍN-HERRERA, David (ed.). *La inteligencia artificial y el control algorítmico de los derechos fundamentales*. Cizur Menor: Aranzadi, 2024.
- O'NEIL, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Nueva York: Crown Publishing Group, 2016.
- ORGANIZACIÓN DE LAS NACIONES UNIDAS. *Declaración Universal de Derechos Humanos*. París: ONU, 1948.
- ORGANIZACIÓN DE LAS NACIONES UNIDAS. *Pacto Internacional de Derechos Civiles y Políticos*. Nueva York: ONU, 1966.
- ORGANIZACIÓN DE LAS NACIONES UNIDAS PARA LA EDUCACIÓN, LA CIENCIA Y LA CULTURA (UNESCO). *Inteligencia Artificial y Estado de Derecho: Kit de herramientas para el poder judicial*. París: UNESCO, 2023. [Consulta: 29 de marzo de 2026]. [Disponible en: <https://www.unesco.org/es/artificial-intelligence/rule-law>].
- ORGANIZACIÓN DE LAS NACIONES UNIDAS PARA LA EDUCACIÓN, LA CIENCIA Y LA CULTURA (UNESCO). *Recomendación sobre la ética de la inteligencia artificial*. París: UNESCO, 2021. [Consulta: 03 de abril de 2026]. [Disponible en: [https://unesdoc.unesco.org/ark:/48223/pf0000381137\\_spa](https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa)].
- ORGANIZACIÓN PARA LA COOPERACIÓN Y EL DESARROLLO ECONÓMICOS (OCDE). *Principios sobre Inteligencia Artificial*. París: OCDE, 2019. [Consulta: 25 de marzo de 2026]. [Disponible en: <https://oecd.ai/en/ai-principles>].
- PÉREZ LUÑO, Antonio Enrique. *Derechos Humanos, Estado de Derecho y Constitución*. Madrid: Tecnos, 1999.
- UNIÓN EUROPEA. Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016. *Diario Oficial de la Unión Europea* [en línea]. 4 de mayo de 2016, L 119, pp. 1-88. [Consulta: 29 de marzo de 2026]. [Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32016R0679>].
- UNIÓN EUROPEA. 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. *Diario Oficial de la Unión Europea* [en línea]. 12 de julio de 2024, L 2024/1689. [Consulta: 18 de abril de 2026]. [Disponible en: [https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=OJ:L\\_202401689](https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=OJ:L_202401689)].
- ZUBOFF, Shoshana. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Nueva York: PublicAffairs, 2019.



# Sesgos algorítmicos en sistemas educativos con inteligencia artificial: Implicaciones éticas, jurídicas y pedagógicas desde la literatura reciente\*

*Algorithmic bias in educational systems with artificial intelligence: Ethical, legal, and pedagogical implications from recent literature*

Albano Torres-Gómez\*\*

## RESUMEN

Este artículo analiza la literatura reciente sobre sesgos algorítmicos en sistemas educativos con inteligencia artificial (IA), con énfasis en sus implicaciones éticas, jurídicas y pedagógicas. Mediante una revisión crítica de fuentes académicas especializadas y documentos normativos internacionales, se examinan los modos en que los modelos predictivos, la analítica del aprendizaje, la evaluación automatizada, los tutores inteligentes y la supervisión remota pueden reproducir o intensificar desigualdades previas. El argumento central sostiene que el sesgo algorítmico no constituye solo un problema de precisión técnica, sino una cuestión social y jurídica vinculada con derechos humanos, no discriminación, protección de datos, transparencia, rendición de cuentas y justicia educativa. El trabajo identifica vacíos de investigación en contextos iberoamericanos y propone criterios para una gobernanza educativa basada en auditoría, explicabilidad situada, participación de la comunidad y supervisión humana.

## PALABRAS CLAVE

Inteligencia artificial educativa, sesgo algorítmico, derechos humanos, no discriminación, gobernanza algorítmica.

## ABSTRACT

This article examines recent literature on algorithmic bias in educational systems with artificial intelligence (AI), emphasizing its ethical, legal, and pedagogical implications. Through a critical review of specialized academic sources and international policy documents, it analyzes how predictive models, learning analytics, automated assessment, intelligent tutoring systems, and remote monitoring may reproduce or intensify prior inequalities. The central argument is that algorithmic bias is not merely a matter of technical accuracy, but a social and legal issue connected to human rights, non-discrimination, data protection, transparency, accountability, and educational justice. The article identifies research gaps in Iberoamerican contexts and proposes criteria for educational governance based on auditing, situated explainability, community participation and human oversight.

## KEYWORDS

Artificial intelligence in education, algorithmic bias, human rights, non-discrimination, algorithmic governance.

\* Artículo de investigación

\*\* Doctor en educación y maestro en comunicación por la Benemérita Universidad Autónoma de Puebla, con una estancia posdoctoral en la Universidad Nacional Autónoma de México (albano.torres@comunidad.unam.mx). <https://orcid.org/0000-0002-8451-031X>.

## SUMARIO

1. Introducción
2. Metodología de revisión
3. Del sesgo técnico al perjuicio educativo
4. Implicaciones éticas y jurídicas
5. Implicaciones pedagógicas
6. Hacia una gobernanza educativa de la justicia algorítmica
7. Conclusiones
8. Referencias Bibliograficas

### 1. INTRODUCCIÓN

La inteligencia artificial aplicada a la educación ha dejado de ser una promesa del futuro para convertirse en infraestructura cotidiana de administración escolar, diagnóstico, evaluación, tutoría y personalización del aprendizaje; en donde los diversos sistemas tecnológicos emergentes puede ser usados por las instituciones como, plataformas adaptativas, evaluadores automáticos, asistentes generativos, tutores inteligentes y herramientas de supervisión remota que operan sobre datos de desempeño, interacción, tiempo de conexión, comportamiento y, en ocasiones, rasgos biométricos. Esa expansión no solo transforma la técnica educativa; también altera la distribución de oportunidades, apoyos, reconocimientos y cargas institucionales; ante este escenario, el sesgo algorítmico se vuelve una posibilidad real que sirve para analizar si la automatización educativa amplía derechos o consolida desigualdades ya existentes.<sup>1</sup>

La literatura sobre inteligencia artificial educativa muestra un crecimiento sostenido en los últimos años, pero también se ha dado de forma irregular, gran parte de la investigación ha sido impulsada desde ciencias de la computación, ingeniería y minería de datos, mientras que las preguntas pedagógicas, éticas y jurídicas aparecen con frecuencia como consecuencias posteriores, no como condiciones de diseño. Esta brecha importa porque la educación no es un dominio neutral, en ella se decide quién recibe apoyo, quién es clasificado como estudiante en riesgo, quién es considerado competente, quién es vigilado con mayor intensidad y quién puede quedar excluido por datos incompletos o categorías mal definidas.<sup>2</sup>

El problema no se reduce a preguntar si un modelo acierta; en el ámbito educativo, una predicción correcta para el promedio puede resultar injusta para grupos específicos; una intervención técnicamente eficiente puede desplazar la responsabilidad institucional hacia el estudiante; y una métrica de equidad puede ser incompatible con otra cuando existen desigualdades estructurales en las tasas de acceso, permanencia o logro. La justicia algorítmica en educación exige considerar todos los pasos del proceso formativo; es decir, medición, selección de datos, definición de etiquetas, entrenamiento,

---

<sup>1</sup>BAKER, RYAN S. Y HAWN, AARON, «Algorithmic Bias in Education», *International Journal of Artificial Intelligence in Education*, vol. 32, 2022, pp. 1052-1092.

<sup>2</sup>ZAWACKI-RICHTER, OLAF; MARÍN, VICTORIA I.; BOND, MELISSA y GOUVERNEUR, FRANZISKA, «Systematic Review of Research on Artificial Intelligence Applications in Higher Education: Where Are the Educators?», *International Journal of Educational Technology in Higher Education*, vol. 16, art. 39, 2019.

implementación, interpretación docente, reacción institucional y efectos acumulativos sobre trayectorias educativas.<sup>3</sup>

El objetivo de este artículo es construir un estado del arte crítico sobre sesgos algorítmicos en sistemas educativos con inteligencia artificial (IA) y derivar sus implicaciones éticas, jurídicas y pedagógicas. La pregunta que guía el análisis es ¿qué muestra la literatura reciente sobre las fuentes, manifestaciones y riesgos del sesgo algorítmico en educación, y qué criterios de gobernanza pueden orientar su prevención desde una perspectiva de derechos humanos y justicia educativa?

La premisa que este trabajo sostiene es que el sesgo algorítmico educativo debe entenderse como un fenómeno sociotécnico, ya que sus causas no se localizan únicamente en datos incompletos o modelos defectuosos, sino en la interacción entre historias de desigualdad, decisiones de diseño, finalidades institucionales, opacidad comercial, regulaciones insuficientes y prácticas pedagógicas que pueden asumir el resultado de un algoritmo como si fuera evidencia objetiva. Por ello, las respuestas puramente técnicas, aunque necesarias, son insuficientes, es por esto que se requieren mecanismos de auditoría, documentación, participación, protección de datos, supervisión humana y deliberación pedagógica.

## **2. METODOLOGÍA DE REVISIÓN**

El artículo se desarrolla mediante un proceso de revisión crítica de literatura, en la que se pretende realizar un análisis integrador de fuentes científicas y normativas que permita identificar consensos, tensiones, vacíos y criterios de acción. La recuperación bibliográfica privilegió artículos indexados o publicados por editoriales académicas reconocidas en las bases de datos Google Academic, SciELO, Redalyc y Dialnet, además de documentos normativos de organismos internacionales y fuentes oficiales.

La estrategia de búsqueda se organizó en los siguientes grupos conceptuales; inteligencia artificial educativa, sesgo algorítmico en educación, justicia o equidad algorítmica, indicadores de aprendizaje, predicción de riesgo académico, evaluación automatizada, supervisión remota, gobernanza digital, derechos humanos, auditoría y documentación de sistemas de IA. Se priorizaron fuentes de alto impacto por citación y centralidad conceptual, así como trabajos recientes que conectan explícitamente justicia algorítmica con sistemas educativos con IA.

Los criterios de inclusión fueron: a) pertinencia directa para educación, algoritmos, equidad, derechos o gobernanza; b) disponibilidad con acceso abierto, página editorial, repositorio institucional u organismo oficial; c) contribución conceptual, empírica o normativa al estado del arte; y d) preferencia por artículos revisados por pares, libros académicos o documentos de política pública con autoridad institucional. Se excluyeron textos sin datos bibliográficos claros, fuentes duplicadas y materiales que sólo repetían hallazgos ya consolidados. Al final del proceso de búsqueda se identificaron 26 fuentes pertinentes para este trabajo.

La revisión reconoce como limitación relevante, que la literatura disponible está geográficamente concentrada, con predominio de estudios de Estados Unidos y Europa,

---

<sup>3</sup>KIZILCEC, RENÉ F. Y LEE, HANSOL, «Algorithmic Fairness in Education», en Holmes, Wayne y Porayska-Pomsta, Kaska (eds.), *The Ethics of Artificial Intelligence in Education*, Londres, Routledge, 2022, pp. 174-202.

mientras que América Latina e Iberoamérica aparecen subrepresentadas. Esta diferencia reconoce la existencia de sistemas educativos con desigualdades, categorías de análisis, infraestructuras digitales y marcos de protección distintos; lo que enfatiza la necesidad de investigación local contextualizada.<sup>4</sup>

Para ordenar la revisión se construyó una tipología de análisis con cuatro niveles; el primer nivel corresponde a sesgos de entrada, tales como, datos incompletos, muestras no representativas y registros producidos por plataformas que capturan actividad digital pero no necesariamente aprendizaje. El segundo nivel corresponde a sesgos de modelado; en particular, selección de objetivos, optimización por precisión promedio, etiquetas históricas y métricas de justicia no fundamentadas pedagógicamente. El tercer nivel corresponde a sesgos de implementación; por ejemplo, interpretación acrítica, automatización de decisiones, opacidad contractual y ausencia de canales de reclamación. El cuarto nivel corresponde a daños educativos; entre los que se encuentra, exclusión, estigmatización, vigilancia desproporcionada, pérdida de autonomía, disminución de oportunidades y reproducción de desigualdades.

Esta tipología permite una lectura reflexiva que busca responder la pregunta de cómo el sesgo se formula en términos de trayectoria; es decir, cómo un dato se convierte en predicción, cómo una predicción se convierte en intervención y cómo esa intervención altera la experiencia educativa de una persona o grupo. En consecuencia, un sistema puede ser problemático, aunque no discrimine explícitamente en su código; basta con que sus insumos de información, objetivos o usos institucionales produzcan efectos diferenciados injustificados.

### 3. DEL SESGO TÉCNICO AL PERJUICIO EDUCATIVO

Luego de la lectura de los materiales de información identificados y recolectados, un primer consenso es que el sesgo algorítmico no se origina en un único punto del sistema; este puede introducirse en la definición del problema, en la recolección de datos, en la medición de variables, en la elección de etiquetas, en el aprendizaje del modelo, en la evaluación, en el despliegue y en el modo en que las personas interpretan y actúan frente a los resultados. En el ámbito educativo, estos puntos son especialmente sensibles porque muchas variables usadas como indicadores de aprendizaje son en realidad un reflejo de condiciones sociales específicas; por ejemplo, conectividad, disponibilidad de tiempo, familiaridad con el lenguaje académico, capital cultural, escuela de procedencia o trayectorias previas.<sup>5</sup>

La literatura crítica también advierte que el término sesgo suele usarse de manera ambigua, ya que en algunos textos designa error estadístico; en otros, disparidad de desempeño entre grupos; en otros más, daño distributivo, estereotipación, exclusión o reproducción de jerarquías. Esta ambigüedad no es menor, si una universidad afirma que su sistema no tiene sesgos porque mantiene una precisión promedio, puede ignorar

<sup>4</sup>OECD, «Algorithmic Bias: The State of the Situation and Policy Recommendations», en OECD Digital Education Outlook 2023, París, OECD Publishing, 2023.

<sup>5</sup>SURESH, HARINI y GUTTAG, JOHN, «A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle», Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization, 2021

que comete más errores con estudiantes indígenas, hablantes no nativos, personas con discapacidad o estudiantes de bajo nivel socioeconómico; de esta manera, la definición de sesgo es, por tanto, una decisión normativa.<sup>6</sup>

La teoría formal de la equidad algorítmica aporta herramientas importantes, como paridad demográfica, igualdad de oportunidades, calibración, equidad individual o restricciones de error por grupo. Sin embargo, la misma literatura demuestra que estas métricas pueden entrar en conflicto, tratar de igual manera a estudiantes que en apariencia son similares exige una métrica de similitud que no es evidente en educación; garantizar tasas iguales de intervención puede reducir precisión; y calibrar un modelo puede no eliminar disparidades cuando los grupos parten de condiciones estructurales distintas; por ello, la equidad educativa no puede resolverse eligiendo una métrica aislada.<sup>7,8</sup>

En cuanto a indicadores de aprendizaje, la preocupación central se ubica en los sistemas de predicción de riesgo académico. Estos modelos prometen identificar tempranamente a estudiantes que podrían reprobado, abandonar o requerir apoyo. Su atractivo institucional es evidente, porque permite anticipar problemas, asignar recursos y mejorar retención; pero la literatura muestra que los algoritmos pueden funcionar de forma desigual entre grupos y que la optimización por precisión puede invisibilizar errores diferenciados. Un falso positivo puede estigmatizar; un falso negativo puede privar de apoyo; y una intervención basada en riesgo puede convertirse en una meta institucional errónea si docentes o tutores reducen expectativas.<sup>9,10</sup>

Los estudios sobre predicción de éxito estudiantil muestran que los ajustes posteriores para mejorar una dimensión de equidad pueden empeorar otra, esto obliga a abandonar la idea de que existe una solución universal de justicia educativa. En educación, la selección de la métrica debe justificarse con base en el daño que se busca evitar; por ejemplo, exclusión del acceso, pérdida de becas, vigilancia injustificada, ausencia de apoyo, clasificación estigmatizante, invisibilidad de necesidades o disminución de autonomía. La pregunta correcta no es solo si el modelo es justo en abstracto, sino justo respecto de qué derecho, qué grupo, qué decisión y qué consecuencia.<sup>11</sup>

Los tutores inteligentes y plataformas adaptativas plantean otro tipo de riesgo, dado que, al personalizar contenidos, ritmos o retroalimentación, pueden asignar experiencias educativas distintas según patrones históricos de desempeño. Si los datos reflejan

<sup>6</sup>BLODGETT, SU LIN; BAROCAS, SOLON; DAUMÉ III, HAL y WALLACH, HANNA, «Language (Technology) is Power: A Critical Survey of 'Bias' in NLP», *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, pp. 5454-5476.

<sup>7</sup>DWORK, CYNTHIA; HARDT, MORITZ; PITASSI, TONIANN; REINGOLD, OMER y ZEMEL, RICHARD, «Fairness through Awareness», *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, 2012, pp. 214-226.

<sup>8</sup>KLEINBERG, JON; MULLAINATHAN, SENDHIL y RAGHAVAN, MANISH, «Inherent Trade-Offs in the Fair Determination of Risk Scores», *8th Innovations in Theoretical Computer Science Conference*, LIPIcs, vol. 67, 2017, art. 43.

<sup>9</sup>BAYER, VACLAV; HLOSTA, MARTIN y FERNANDEZ, MIRIAM, «Learning Analytics and Fairness: Do Existing Algorithms Serve Everyone Equally?», en *Artificial Intelligence in Education. AIED 2021, Lecture Notes in Computer Science*, vol. 12749, Cham, Springer, 2021, pp. 71-75.

<sup>10</sup>RIAZY, SHIRIN; SIMBECK, KATHARINA y SCHRECK, VANESSA, «Fairness in Learning Analytics: Student At-risk Prediction in Virtual Learning Environments», *Proceedings of the 12th International Conference on Computer Supported Education*, vol. 1, 2020, pp. 15-25.

<sup>11</sup>LEE, HANSOL y KIZILCEC, RENÉ F., «Evaluation of Fairness Trade-Offs in Predicting Student Success», *arXiv*, 2020.

desigualdades previas, el sistema puede ofrecer tareas menos desafiantes a estudiantes que requieren más oportunidades, no menos. Es por esto que, la personalización puede convertirse en segmentación silenciosa, tal es el caso de estudiantes considerados de bajo rendimiento que reciben trayectorias empobrecidas, mientras quienes ya cuentan con ventajas reciben retos de mayor complejidad; lo que implica que la justicia pedagógica debe exigir una revisión de la exactitud del diagnóstico, aunado a la calidad de la experiencia que se deriva de él.

Por otra parte, el monitoreo remoto de exámenes muestra con claridad la convergencia entre sesgo, vigilancia y derechos, a causa de las capacidades de las herramientas de vigilancia que pueden detectar rostros, movimientos, sonidos, miradas o patrones de navegación para inferir conductas sospechosas. Estos sistemas han sido cuestionados por riesgos de privacidad, autonomía, libertad, transparencia, confianza y no discriminación. En términos pedagógicos, introducen un clima de sospecha que puede afectar de manera diferenciada a estudiantes con discapacidad, estudiantes neurodivergentes, personas con condiciones domésticas precarias o usuarios de equipos y conexiones inestables.<sup>12</sup>

De forma similar, la evaluación automatizada, en particular la calificación de ensayos, respuestas abiertas o producción oral concentra riesgos adicionales porque traduce expresiones complejas de aprendizaje en puntuaciones. Si un sistema penaliza variedades lingüísticas, estilos argumentativos, acentos, estructuras discursivas o formas culturalmente situadas de escritura, el daño no es solo estadístico, también afecta el reconocimiento académico de la voz estudiantil. La literatura psicométrica reciente propone marcos para evaluar la función algorítmica diferencial, análoga a preocupaciones existentes sobre sesgo en pruebas, pero adaptada a decisiones automatizadas.<sup>13</sup>

Ahora bien, un hallazgo transversal que se pudo identificar en esta revisión es la falta de evidencia robusta sobre grupos históricamente poco estudiados, en donde el género, la raza y la etnicidad concentran buena parte de la literatura, especialmente en Estados Unidos. En cambio, discapacidad, condición migrante, lengua indígena, ruralidad, conectividad, primera generación universitaria, maternidad o paternidad estudiantil, trabajo remunerado, edad, orientación sexual o interseccionalidad aparecen de manera dispersa. En América Latina, el vacío es más grave, los sistemas educativos tienen desigualdades profundas, pero existen pocos estudios que midan cómo los algoritmos educativos funcionan en esos contextos.

Es así como, los escenarios donde se pueden manifestar sesgos descritos anteriormente permiten plantear dos tipos de brechas, la primera es epistémica, derivada de la manera en que los sistemas se entrenan, evalúan o justifican con evidencia producida en otros contextos, se corre el riesgo de invisibilizar desigualdades locales. La segunda es de naturaleza regulatoria, sin evidencia situada, las políticas públicas pueden adoptar fórmulas generales de transparencia o protección de datos sin atender daños concretos de lengua, territorio, modalidad educativa o infraestructura, tal es el caso de regiones

<sup>12</sup>COGHLAN, SIMON; MILLER, TIM y PATERSON, JEANNIE, «Good Proctor or 'Big Brother'? Ethics of Online Exam Supervision Technologies», *Philosophy & Technology*, vol. 34, 2021, pp. 1581-1606.

<sup>13</sup>SUK, YOUNG y HAN, KYUNG T., «A Psychometric Framework for Evaluating Fairness in Algorithmic Decision Making: Differential Algorithmic Functioning», *Applied Measurement in Education*, vol. 49, núm. 2, 2024.

con alta desigualdad digital, donde un modelo que usa frecuencia de conexión como indicador de compromiso puede confundir precariedad material con desinterés académico.

También existe un vacío sobre impactos acumulativos, muchos estudios evalúan un modelo en un momento específico, pero las trayectorias educativas se componen de decisiones sucesivas; tales como, admisión, diagnóstico, nivelación, asignación de grupos, becas, evaluación, tutoría, alertas, permanencia y egreso. Pequeñas disparidades en cada etapa pueden acumularse hasta producir exclusiones significativas, por ello, la investigación futura debería analizar procesos de decisión y no solo herramientas aisladas. En síntesis, la IA educativa puede ayudar a detectar barreras, ofrecer retroalimentación o ampliar apoyos, pero solo bajo condiciones estrictas de diseño y gobernanza. El riesgo no es la automatización por sí misma, sino su uso acrítico, opaco y descontextualizado en instituciones que ya distribuyen oportunidades de manera desigual.

#### **4. IMPLICACIONES ÉTICAS Y JURÍDICAS**

Desde una perspectiva ética, el sesgo algorítmico en el ámbito educativo compromete dignidad, autonomía, justicia, transparencia, beneficencia, no maleficencia y responsabilidad institucional. Desde una perspectiva jurídica, se vincula con igualdad y no discriminación, derecho a la educación, protección de datos personales, debido proceso, transparencia, acceso a información significativa y disponibilidad de medios. La UNESCO y el Consejo de Europa colocan estos principios en el centro de la gobernanza de IA, porque los sistemas deben respetar derechos humanos, igualdad, privacidad, supervisión y rendición de cuentas durante todo su ciclo de vida.<sup>14,15</sup>

El reglamento europeo de inteligencia artificial ofrece un referente comparado relevante porque clasifica como de alto riesgo los sistemas de IA usados en educación cuando pueden determinar acceso, evaluación o trayectoria profesional, e impone obligaciones sobre gestión de riesgos, calidad de datos, documentación, información al usuario, supervisión humana, robustez y trazabilidad. Además, prohíbe ciertas prácticas como el reconocimiento de emociones en instituciones educativas. Aunque este marco no es directamente aplicable en América Latina, su propuesta para analizar el riesgo permite identificar qué decisiones educativas requieren mayor control jurídico.<sup>16</sup>

En el contexto mexicano, el análisis puede anclarse en el derecho a la educación, la igualdad sustantiva, la no discriminación y la protección de datos personales. La educación debe orientarse al respeto de la dignidad, los derechos humanos y la igualdad; por ello, una plataforma que distribuya apoyos o evalúe estudiantes mediante criterios opacos no puede tratarse únicamente como herramienta administrativa. Si procesa datos personales, su operación debe observar finalidad, proporcionalidad, información, seguridad, consentimiento cuando proceda, responsabilidad y derechos de acceso,

---

<sup>14</sup>UNESCO, Recommendation on the Ethics of Artificial Intelligence, París, UNESCO, 2022.

<sup>15</sup>COUNCIL OF EUROPE, Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, 2024.

<sup>16</sup>EUROPEAN COMMISSION, «AI Act», Shaping Europe's Digital Future, 2026. Disponible en: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>; Reglamento (UE) 2024/1689

rectificación, cancelación u oposición. Si actúa en instituciones públicas, el estándar de legalidad y rendición de cuentas se vuelve más importante.<sup>17</sup>

Una primera implicación jurídica es la necesidad de pasar de la transparencia formal a la explicabilidad significativa, entendida como el requerimiento por parte de las personas afectadas para conocer, en términos comprensibles, qué tipo de datos se usan, con qué finalidad, qué decisión se apoya o automatiza, qué margen de intervención humana existe, cómo puede controvertirse el resultado y qué consecuencias puede producir. La opacidad es especialmente problemática cuando el algoritmo influye en admisión, becas, permanencia, evaluación, disciplina o asignación de apoyos.<sup>18</sup>

Una segunda implicación es la rendición de cuentas, bajo el entendido de que las guías éticas de IA suelen coincidir en transparencia, justicia, no maleficencia, responsabilidad y privacidad, pero la literatura advierte que esos principios pueden volverse retórica si no se traducen en obligaciones, responsables, evidencia y remedios. En el contexto educativo, la responsabilidad no puede desplazarse al proveedor tecnológico ni diluirse en la complejidad del modelo, a razón de que la institución que adopta el sistema conserva deberes de selección diligente, evaluación de impacto, monitoreo, información a la comunidad y suspensión cuando el riesgo sea injustificado.<sup>19</sup>

De esta manera, la auditoría algorítmica interna y la documentación de modelos son instrumentos necesarios para cerrar la brecha de responsabilidad. Ante este escenario, registros de entrenamiento, reportes de desempeño por subgrupo, bitácoras de incidentes y evaluaciones de impacto permiten superar declaraciones genéricas de confianza y producir evidencia verificable. Estos documentos deben incluir finalidades pedagógicas, población prevista, variables sensibles, limitaciones, condiciones de uso, pruebas de equidad, protocolos de apelación y criterios para retiro del sistema.<sup>20,21</sup>

Una tercera implicación se ubica en la relación entre privacidad y equidad, ya que para medir sesgos suele ser necesario conocer variables sensibles o de contexto; pero recopilar esos datos puede aumentar riesgos de vigilancia, perfilamiento o uso secundario. Por esto, la respuesta no debe ser ignorar las diferencias, porque eso impide detectar discriminación, ni recolectar todo sin límites, porque eso vulnera privacidad. La salida jurídicamente razonable, de acuerdo con especialistas, consiste en la gobernanza de datos; mediante la finalidad explícita, minimización, separación de bases, controles

<sup>17</sup>CÁMARA DE DIPUTADOS DEL H. CONGRESO DE LA UNIÓN, Constitución Política de los Estados Unidos Mexicanos, texto vigente; Ley General de Educación; Ley Federal de Protección de Datos Personales en Posesión de los Particulares; Ley General de Protección de Datos Personales en Posesión de Sujetos Obligados.

<sup>18</sup>MITTELSTADT, BRENT DANIEL; ALLO, PATRICK; TADDEO, MARIAROSARIA; WACHTER, SANDRA y FLORIDI, LUCIANO, «The Ethics of Algorithms: Mapping the Debate», *Big Data & Society*, vol. 3, núm. 2, 2016.

<sup>19</sup>JOBIN, ANNA; IENCA, MARCELLO y VAYENA, EFFY, «The Global Landscape of AI Ethics Guidelines», *Nature Machine Intelligence*, vol. 1, 2019, pp. 389-399.

<sup>20</sup>RAJI, INIOLUWA DEBORAH; SMART, ANDREW; WHITE, REBECCA N.; MITCHELL, MARGARET; GEBRU, TIMNIT; HUTCHINSON, BEN; SMITH-LOUD, JAMILA; THERON, DANIEL y BARNES, PARKER, «Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing», *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020.

<sup>21</sup>MITCHELL, MARGARET; WU, SIMONE; ZALDIVAR, ANDREW; BARNES, PARKER; VASSERMAN, LUCY; HUTCHINSON, BEN; SPITZER, ELENA; RAJI, INIOLUWA DEBORAH y GEBRU, TIMNIT, «Model Cards for Model Reporting», *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 2019, pp. 220-229.

de acceso, anonimización o seudonimización cuando sea posible, evaluación de impacto y participación de grupos afectados.

Una cuarta implicación es el derecho a impugnar decisiones, dado que los sistemas educativos inteligentes no deberían producir decisiones finales e incontestables, especialmente cuando afectan trayectoria académica. La revisión humana debe ser significativa, es decir, no un trámite en que la persona revisora confirma automáticamente el resultado, sino una evaluación con autoridad real para corregir, contextualizar, suspender o descartar la recomendación algorítmica. Esto exige capacitación docente y administrativa, así como reglas claras sobre evidencia, carga argumentativa y tiempos de respuesta.

## **5. IMPLICACIONES PEDAGÓGICAS**

La discusión pedagógica en cuanto a los sesgos que pueden cometer los algoritmos no debe limitarse a preguntar si la IA únicamente mejora la eficiencia del proceso de aprendizaje. La cuestión de fondo es qué concepción de conocimiento, estudiante, docente y evaluación se manifiesta en el sistema. Un modelo que reduce aprendizaje a número de clics, permanencia en pantalla o patrones de respuesta puede ser útil para ciertas tareas, pero pobre para valorar comprensión, creatividad, colaboración, ética o pensamiento crítico. Por esto, la pedagogía aporta criterios para decidir qué puede automatizarse, qué debe permanecer bajo juicio docente y qué nunca debe delegarse sin garantías reforzadas.<sup>22</sup>

La analítica del aprendizaje suele justificarse como forma de prevención institucional, bajo el entendido de que conocer mejor al estudiante permitiría intervenir antes. Sin embargo, la misma práctica puede convertirse en vigilancia si no existe transparencia, consentimiento informado, proporcionalidad y participación. La frontera entre cuidado y control depende de quién define el problema, qué datos se observan, qué inferencias se hacen, cómo se comunican los resultados y qué oportunidades reales tiene el estudiante de comprender o cambiar su perfil.<sup>23</sup>

La ética en torno a los indicadores de aprendizaje propone principios de transparencia, control, seguridad, consentimiento, acceso y mejora educativa. Desde una mirada pedagógica, estas bases deben traducirse en prácticas concretas; por ejemplo, explicar a estudiantes cómo se usa su información, permitirles ver y corregir datos relevantes, evitar inferencias que excedan la finalidad educativa, diferenciar recomendaciones de decisiones obligatorias y asegurar que las intervenciones algorítmicas amplíen oportunidades en lugar de restringir trayectorias.<sup>24</sup>

Al mismo tiempo, las percepciones de los estudiantes deben tomarse en cuenta, más aún, porque la aceptación de este tipo de sistemas no puede suponerse por el hecho de que estén orientados a mejorar resultados. Los estudiantes pueden valorar

---

<sup>22</sup>AKGUN, SELIN y GREENHOW, CHRISTINE, «Artificial Intelligence in Education: Addressing Ethical Challenges in K-12 Settings», *AI and Ethics*, vol. 2, 2022, pp. 431-440.

<sup>23</sup>SLADE, SHARON y PRINSLOO, PAUL, «Learning Analytics: Ethical Issues and Dilemmas», *American Behavioral Scientist*, vol. 57, núm. 10, 2013, pp. 1510-1529.

<sup>24</sup>PARDO, ABELARDO y SIEMENS, GEORGE, «Ethical and Privacy Principles for Learning Analytics», *British Journal of Educational Technology*, vol. 45, núm. 3, 2014, pp. 438-450.

ciertos usos de datos y rechazar otros; pueden aceptar alertas académicas y oponerse al monitoreo continuo; pueden considerar legítimo recibir retroalimentación y, a la vez, demandar límites frente a inferencias sobre motivación, salud mental, honestidad o riesgo de abandono; de manera ideal, una gobernanza pedagógica debe incorporar estas diferencias y no tratar al estudiante como simple fuente de datos.<sup>25</sup>

Por otra parte, el papel docente también cambia, a partir de una reflexión, una plataforma puede producir recomendaciones, pero es el docente quien puede asumirlas como apoyo interpretativo o como autoridad superior. En el primer caso, el algoritmo enriquece la deliberación profesional; en el segundo, desplaza el juicio pedagógico. Por ello, la formación docente en IA debe incluir la lectura crítica de métricas, comprensión de errores, análisis de sesgos, diseño de intervenciones equitativas y criterios para explicar a estudiantes las razones de una decisión apoyada por tecnología.

Desde esta mirada, la alfabetización algorítmica se perfila como dimensión emergente del derecho educativo en la que no basta enseñar a usar herramientas; es necesario comprender cómo clasifican, qué datos requieren, qué sesgos pueden reproducir, qué límites tienen y cómo cuestionarlos. En una sociedad donde algoritmos influyen en admisión, crédito, empleo, salud, seguridad y educación, formar ciudadanía democrática exige que estudiantes puedan interpretar, auditar y disputar sistemas automatizados.

Sin embargo, los avances no se detienen, la irrupción de IA generativa intensifica estos dilemas, porque a diferencia de sistemas predictivos tradicionales, los modelos generativos producen texto, retroalimentación, explicaciones y materiales que pueden influir directamente en el aprendizaje. Sus sesgos pueden aparecer como estereotipos, omisiones, alucinaciones, traducciones deficientes, recomendaciones culturalmente inadecuadas o estilos de corrección que favorecen ciertos registros lingüísticos. Ante esto, la guía de la UNESCO insiste en una aproximación centrada en la persona, con regulación, privacidad, edad apropiada, validación ética y diseño pedagógico significativo.<sup>26</sup>

En términos generales, las implicaciones pedagógicas pueden sintetizarse en cuatro exigencias. Primera, no usar IA para decisiones educativas de alto impacto sin evidencia suficiente de validez, equidad y pertinencia contextual. Segunda, preservar espacios de interacción humana, especialmente en evaluación, orientación y apoyo socioemocional. Tercera, convertir la transparencia en práctica didáctica, no solo en aviso legal. Cuarta, diseñar intervenciones que reparen desigualdades, no que administren de manera más eficiente su reproducción.

## 6. HACIA UNA GOBERNANZA EDUCATIVA DE LA JUSTICIA ALGORÍTMICA

La gobernanza educativa basada en la justicia algorítmica debe partir de la siguiente premisa formal, los sistemas no se vuelven justos por cumplir una métrica aislada, porque las decisiones técnicas operan dentro de contextos institucionales, culturales y jurídicos. La literatura sobre sistemas sociotécnicos advierte contra las trampas de abstraer demasiado el problema, ignorar el contexto, formalizar de manera estrecha

<sup>25</sup>IFENTHALER, DIRK y SCHUMACHER, CLARA, «Student Perceptions of Privacy Principles for Learning Analytics», *Educational Technology Research and Development*, vol. 64, 2016, pp. 923-938.

<sup>26</sup>MIAO, Fengchun y Holmes, Wayne, *Guidance for Generative AI in Education and Research*, París, UNESCO, 2023.

la equidad o suponer que el diseño técnico puede reemplazar procesos deliberativos. En las instituciones educativas, estas trampas son especialmente visibles cuando se implementa una herramienta sin revisar currículo, población, infraestructura y cultura institucional.<sup>27</sup>

De tal manera que se plantean los siguientes criterios como base para una gobernanza en este contexto, el primero es la evaluación de impacto algorítmico antes de la adopción. Toda institución que pretenda usar IA para admisión, evaluación, permanencia, becas, disciplina, vigilancia, orientación o asignación de apoyos debería documentar finalidad, base jurídica, necesidad, proporcionalidad, alternativas menos invasivas, datos usados, riesgos para grupos vulnerables, pruebas de desempeño, mecanismos de supervisión y procedimientos de reclamación. Esta evaluación debe actualizarse durante el ciclo de vida del sistema, porque los modelos pueden degradarse o producir efectos no previstos.

El segundo criterio es la documentación verificable, esto comienza desde el proveedor, que debe entregar información técnica y funcional suficiente para que la institución comprenda límites, población objetivo, variables, entrenamiento, supuestos, errores conocidos y condiciones de uso. Si el secreto comercial impide auditar sesgos, la herramienta no debería usarse en decisiones de alto impacto. La contratación pública o institucional de tecnología educativa debe incluir cláusulas de acceso a evidencia, auditoría independiente, seguridad de datos, eliminación al terminar el servicio y prohibición de usos secundarios no autorizados.

El tercer criterio es la evaluación por subgrupos y por daños en la que no basta reportar precisión global; deben medirse falsos positivos, falsos negativos, sensibilidad, especificidad, calibración y consecuencias por grupos relevantes para el contexto. En México e Iberoamérica, esos grupos no pueden definirse copiando categorías de países desarrollados, estos deben dialogar con desigualdades locales, como lengua indígena, ruralidad, discapacidad, brecha digital, tipo de institución, modalidad educativa, trabajo remunerado, género, condición migrante y nivel socioeconómico.

El cuarto criterio es la participación de la comunidad, en donde estudiantes, docentes, familias, personal de orientación, áreas jurídicas, especialistas en datos y representantes de grupos potencialmente afectados deben intervenir en la definición de finalidades, riesgos aceptables, indicadores de equidad y procedimientos de revisión. La participación no es un mero trámite; mejora la calidad epistémica del sistema porque permite identificar daños que no aparecen en métricas técnicas.

El quinto criterio es el control humano significativo, comenzando por entender que no toda intervención humana es relevante; la supervisión exige capacitación, acceso a explicaciones, tiempo institucional, autoridad para corregir resultados y responsabilidad por decisiones finales. También requiere evitar automatización de facto, la cual puede darse cuando la carga de apartarse del proceso automatizado es demasiado alta y el resultado algorítmico termina gobernando aunque formalmente exista revisión humana.

---

<sup>27</sup>SELBST, ANDREW D.; BOYD, DANAH; FRIEDLER, SORELLE A.; VENKATASUBRAMANIAN, SURESH y VERTESI, JANET, «Fairness and Abstraction in Sociotechnical Systems», FAT\* 2019 - Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency, 2019, pp. 59-68.

El sexto criterio es la proporcionalidad pedagógica, es aquí donde vale la pena preguntarse si toda ventaja tecnológica justifica el tratamiento de datos o el riesgo de sesgo. Una plataforma de recomendación de lecturas no demanda el mismo escrutinio que un sistema de asignación de becas; un chatbot de apoyo general no plantea el mismo riesgo que un evaluador automático de exámenes finales. A partir de este principio, la intensidad de la regulación institucional debe aumentar con el impacto sobre derechos, trayectoria, evaluación y dignidad del estudiante.

Por último, el séptimo criterio es la reparación, si un sistema ha perjudicado a un grupo, la institución debe contar con medidas de corrección, entre las que se pueden encontrar, la revisión individual, rectificación de expedientes, reentrenamiento o retiro del modelo, comunicación a personas afectadas, ajuste de políticas, medidas compensatorias cuando procedan y aprendizaje institucional. Sin remedios, la auditoría se convierte en ejercicio descriptivo sin impacto real.

Estos criterios deben integrarse en políticas institucionales, no quedar como recomendaciones voluntarias de un comité técnico. Una universidad o sistema escolar que adopta IA debería contar con un inventario público de sistemas automatizados, responsables internos, clasificación de riesgo, periodicidad de auditoría, reglas de contratación, protocolos de incidentes, canales de queja y mecanismos de participación. La gobernanza algorítmica educativa requiere continuidad administrativa, no puede depender de la buena voluntad de un equipo o de la promesa comercial de un proveedor. En instituciones con capacidades técnicas limitadas, la cooperación interinstitucional puede ser decisiva, esto reduciría costos y evitaría que cada institución enfrente sola problemas que son estructurales.

## 7. CONCLUSIONES

El análisis de literatura permitió sostener que los sesgos algorítmicos en sistemas educativos que usan inteligencia artificial constituyen potenciales problemas de derechos y justicia en el contexto educativo. estas herramientas pueden apoyar la enseñanza, diagnóstico y gestión, pero también pueden reproducir desigualdades cuando aprenden de datos históricos, operan sin transparencia, carecen de evaluación por subgrupos o desplazan el juicio pedagógico hacia modelos opacos de trabajo.

La primera conclusión es de naturaleza conceptual, el sesgo algorítmico debe comprenderse como fenómeno social, el cual no nace exclusivamente en el código, sino en decisiones institucionales sobre qué medir, para qué predecir, qué consecuencias asociar a la predicción y quién puede cuestionarla. Por eso, la justicia algorítmica educativa requiere combinar métricas técnicas con principios jurídicos, deliberación pedagógica y participación de comunidades afectadas.

La segunda conclusión es jurídica, los sistemas que influyen en acceso, evaluación, permanencia, apoyos o vigilancia deben tratarse como tecnologías de alto impacto. Su adopción exige evaluación de necesidad y proporcionalidad, protección robusta de datos personales, información significativa, documentación, auditoría, revisión humana y remedios, en donde la opacidad contractual o técnica no puede justificar decisiones que afecten derechos educativos.

La tercera conclusión es pedagógica, porque la IA educativa debe subordinarse a fines formativos, si una herramienta reduce aprendizaje a datos conductuales, personaliza hacia trayectorias empobrecidas o convierte la intervención en vigilancia, su eficiencia puede ser pedagógicamente regresiva. El criterio de calidad no debe ser solo predicción correcta, sino ampliación efectiva de oportunidades de aprendizaje, reconocimiento de diversidad y fortalecimiento de autonomía estudiantil.

Finalmente, se afirma que existe un vacío importante de evidencias contextualizadas en este tema, debido a que la mayor parte de la literatura sobre sesgo algorítmico educativo proviene de países con categorías, datos e instituciones distintas. Esto obliga a impulsar estudios locales, especialmente sobre discapacidad, lengua, ruralidad, nivel socioeconómico, brecha digital e interseccionalidad. La ausencia de evidencia no debe interpretarse como ausencia de sesgo; al contrario, indica una zona de riesgo no observada.

El artículo propone, en suma, una base de criterios de gobernanza educativa para la justicia algorítmica que incluyen evaluación de impacto, documentación verificable, auditoría diferenciada por grupos, participación, control humano significativo, proporcionalidad pedagógica, protección de datos y reparación. Estos principios permiten alinear la innovación educativa con derechos humanos, no discriminación y Estado de derecho.

## **8. BIBLIOGRAFÍA**

- AKGUN, SELIN y GREENHOW, CHRISTINE, «Artificial Intelligence in Education: Addressing Ethical Challenges in K-12 Settings», *AI and Ethics*, vol. 2, 2022, pp. 431-440. En línea. [Disponible en: <https://doi.org/10.1007/s43681-021-00096-7>].
- BAKER, RYAN S. Y HAWN, AARON, «Algorithmic Bias in Education», *International Journal of Artificial Intelligence in Education*, vol. 32, 2022, pp. 1052-1092. En línea. [Disponible en: <https://doi.org/10.1007/s40593-021-00285-9>].
- BAYER, VACLAV; HLOSTA, MARTIN y FERNANDEZ, MIRIAM, «Learning Analytics and Fairness: Do Existing Algorithms Serve Everyone Equally?», en *Artificial Intelligence in Education. AIED 2021, Lecture Notes in Computer Science*, vol. 12749, Cham, Springer, 2021, pp. 71-75. En línea. [Disponible en: [https://doi.org/10.1007/978-3-030-78270-2\\_12](https://doi.org/10.1007/978-3-030-78270-2_12)].
- BLODGETT, SU LIN; BAROCAS, SOLON; DAUMÉ III, HAL y WALLACH, HANNA, «Language (Technology) is Power: A Critical Survey of 'Bias' in NLP», *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, pp. 5454-5476. En línea. [Disponible en: <https://doi.org/10.18653/v1/2020.acl-main.485>].
- CÁMARA DE DIPUTADOS DEL H. CONGRESO DE LA UNIÓN, *Leyes federales vigentes*, México. En línea. [Disponible en: <https://www.diputados.gob.mx/LeyesBiblio/index.htm>].
- COGHLAN, SIMON; MILLER, TIM y PATERSON, JEANNIE, «Good Proctor or 'Big Brother'? Ethics of Online Exam Supervision Technologies», *Philosophy & Technology*, vol. 34, 2021, pp. 1581-1606. En línea. [Disponible en: <https://doi.org/10.1007/s13347-021-00476-1>].
- COUNCIL OF EUROPE, *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, Estrasburgo, Council of Europe, 2024. En línea. [Disponible en: <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>].
- DWORK, CYNTHIA; HARDT, MORITZ; PITASSI, TONIANN; REINGOLD, OMER y ZEMEL, RICHARD, «Fairness through Awareness», *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, 2012, pp. 214-226. En línea. [Disponible en: <https://doi.org/10.1145/2090236.2090255>].

- EUROPEAN COMMISSION, «AI Act», Shaping Europe's Digital Future, Bruselas, 2026. En línea. [Disponible en: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>].
- IFENTHALER, DIRK y SCHUMACHER, CLARA, «Student Perceptions of Privacy Principles for Learning Analytics», *Educational Technology Research and Development*, vol. 64, 2016, pp. 923-938. En línea. [Disponible en: <https://doi.org/10.1007/s11423-016-9477-y>].
- JOBIN, ANNA; IENCA, MARCELLO y VAYENA, EFFY, «The Global Landscape of AI Ethics Guidelines», *Nature Machine Intelligence*, vol. 1, 2019, pp. 389-399. En línea. [Disponible en: <https://doi.org/10.1038/s42256-019-0088-2>].
- KIZILCEC, RENÉ F. Y LEE, HANSOL, «Algorithmic Fairness in Education», en Holmes, Wayne y Porayska-Pomsta, Kaska (eds.), *The Ethics of Artificial Intelligence in Education*, Londres, Routledge, 2022, pp. 174-202. En línea. [Disponible en: <https://doi.org/10.4324/9780429329067-10>].
- KLEINBERG, JON; MULLAINATHAN, SENDHIL y RAGHAVAN, MANISH, «Inherent Trade-Offs in the Fair Determination of Risk Scores», 8th Innovations in Theoretical Computer Science Conference, LIPIcs, vol. 67, 2017, art. 43. En línea. [Disponible en: <https://doi.org/10.4230/LIPIcs.ITCS.2017.43>].
- LEE, HANSOL y KIZILCEC, RENÉ F., «Evaluation of Fairness Trade-Offs in Predicting Student Success», arXiv, 2020. En línea. [Disponible en: <https://doi.org/10.48550/arXiv.2007.00088>].
- MIAO, FENGCHUN y HOLMES, WAYNE, *Guidance for Generative AI in Education and Research*, París, UNESCO, 2023. En línea. [Disponible en: <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>].
- MITCHELL, MARGARET; WU, SIMONE; ZALDIVAR, ANDREW; BARNES, PARKER; VASSERMAN, LUCY; HUTCHINSON, BEN; SPITZER, ELENA; RAJI, INIOLUWA DEBORAH y GEBRU, TIMNIT, «Model Cards for Model Reporting», *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 2019, pp. 220-229. En línea. [Disponible en: <https://doi.org/10.1145/3287560.3287596>].
- MITTELSTADT, BRENT DANIEL; ALLO, PATRICK; TADDEO, MARIAROSARIA; WACHTER, SANDRA y FLORIDI, LUCIANO, «The Ethics of Algorithms: Mapping the Debate», *Big Data & Society*, vol. 3, núm. 2, 2016. En línea. [Disponible en: <https://doi.org/10.1177/2053951716679679>].
- OECD, «Algorithmic Bias: The State of the Situation and Policy Recommendations», en *OECD Digital Education Outlook 2023*, París, OECD Publishing, 2023. En línea. [Disponible en: [https://www.oecd.org/en/publications/oecd-digital-education-outlook-2023\\_c74f03de-en/full-report/algorithmic-bias-the-state-of-the-situation-and-policy-recommendations\\_a0b7cec1.html](https://www.oecd.org/en/publications/oecd-digital-education-outlook-2023_c74f03de-en/full-report/algorithmic-bias-the-state-of-the-situation-and-policy-recommendations_a0b7cec1.html)].
- PARDO, ABELARDO y SIEMENS, GEORGE, «Ethical and Privacy Principles for Learning Analytics», *British Journal of Educational Technology*, vol. 45, núm. 3, 2014, pp. 438-450. En línea. [Disponible en: <https://doi.org/10.1111/bjet.12152>].
- RAJI, INIOLUWA DEBORAH; SMART, ANDREW; WHITE, REBECCA N.; MITCHELL, MARGARET; GEBRU, TIMNIT; HUTCHINSON, BEN; SMITH-LOUD, JAMILA; THERON, DANIEL y BARNES, PARKER, «Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing», *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pp. 33-44. En línea. [Disponible en: <https://doi.org/10.1145/3351095.3372873>].
- RIAZY, SHIRIN; SIMBECK, KATHARINA y SCHRECK, VANESSA, «Fairness in Learning Analytics: Student At-risk Prediction in Virtual Learning Environments», *Proceedings of the 12th International Conference on Computer Supported Education*, vol. 1, 2020, pp. 15-25. En línea. [Disponible en: <https://doi.org/10.5220/0009324100150025>].
- SELBST, ANDREW D.; BOYD, DANAH; FRIEDLER, SORELLE A.; VENKATASUBRAMANIAN, SURESH y VERTESI, JANET, «Fairness and Abstraction in Sociotechnical Systems», *FAT\* 2019 - Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency*, 2019, pp. 59-68. En línea. [Disponible en: <https://doi.org/10.1145/3287560.3287598>].
- SLADE, SHARON y PRINSLOO, PAUL, «Learning Analytics: Ethical Issues and Dilemmas», *American Behavioral Scientist*, vol. 57, núm. 10, 2013, pp. 1510-1529. En línea. [Disponible en: <https://doi.org/10.1177/0002764213479366>].

- SUK, YOUMI y HAN, KYUNG T., «A Psychometric Framework for Evaluating Fairness in Algorithmic Decision Making: Differential Algorithmic Functioning», *Applied Measurement in Education*, vol. 49, núm. 2, 2024. En línea. [Disponible en: <https://doi.org/10.3102/10769986231171711>].
- SURESH, HARINI y GUTTAG, JOHN, «A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle», *Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 2021. En línea. [Disponible en: <https://doi.org/10.1145/3465416.3483305>].
- UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, París, UNESCO, 2022. En línea. [Disponible en: <https://unesdoc.unesco.org/ark:/48223/pf0000380455>].
- ZAWACKI-RICHTER, OLAF; MARÍN, VICTORIA I.; BOND, MELISSA y GOUVERNEUR, FRANZISKA, «Systematic Review of Research on Artificial Intelligence Applications in Higher Education: Where Are the Educators?», *International Journal of Educational Technology in Higher Education*, vol. 16, art. 39, 2019. En línea. [Disponible en: <https://doi.org/10.1186/s41239-019-0171-0>].



# Desarrollo y regulación de la inteligencia artificial en México: Un análisis jurídico de sus avances y retos con base en los derechos humanos\*

## *Development and regulation of artificial intelligence in Mexico: A legal analysis of its advances and challenges based on human rights*

Marcos Cruz González\*\*

### RESUMEN

El presente trabajo aborda dos aspectos fundamentales que se ha planteado el gobierno de México, nos referimos al desarrollo y regulación de la Inteligencia Artificial (IA), ambos aspectos son de suma importancia dado que se encuentra cambiando a la humanidad en todos sus ámbitos. El Estado mexicano es consciente de esta situación, como también lo es de que no basta con que la sociedad mexicana se limite a adoptar a la cambiante IA, sino que debe ser abordada a través de los derechos públicos subjetivos en todas sus dimensiones, dada su inacabada y constante evolución, en este tenor, México como el resto de los países latinoamericanos, no pueden ni deben esperar que sean compartidos los avances en el tema por potencias tecnológicas, sino que se debe construir de manera soberana y de acuerdo con su propio contexto y necesidades su propia Inteligencia artificial y regulación.

### PALABRAS CLAVE

Inteligencia Artificial, Derechos Humanos, Desarrollo, Regulación.

### ABSTRACT

This paper addresses two fundamental issues that the Mexican government has raised: the development and regulation of Artificial Intelligence (AI). Both aspects are of paramount importance given that AI is transforming humanity in all its spheres. The Mexican State is aware of this situation, as well as the fact that it is not enough for Mexican society to simply adopt the ever-changing nature of AI. Rather, it must be addressed through subjective public rights in all their dimensions, given its ongoing and continuous evolution. In this regard, Mexico, like the rest of the Latin American countries, cannot and should not wait for technological advancements to be shared by leading powers. Instead, it must sovereignly develop its own artificial intelligence and regulations, tailored to its own context and needs.

### KEYWORDS

Artificial Intelligence, Human Rights, Development, Regulation.

---

\* Artículo de investigación

\*\* Docente e investigador en el Complejo Regional Mixteca de la Benemérita Universidad Autónoma de Puebla, y Miembro del Sistema Nacional de Investigadores Nivel 1 (marcos.cruzg@correo.buap.mx). <https://orcid.org/0009-0007-5933-475X>.

## SUMARIO

1. Introducción
2. Aspectos Generales
3. Definición y Clasificación de la Inteligencia Artificial
4. Los Derechos Humanos y su Relación con la Inteligencia Artificial
5. Desarrollo y Regulación de la IA México
6. Conclusiones
7. Referencias Bibliográficas

### 1. INTRODUCCIÓN

El artículo que se presenta a continuación tiene como objetivo general, describir y analizar los esfuerzos emprendidos en México, para regular y desarrollar a la Inteligencia Artificial (IA), con base en los derechos humanos, además, de determinar el lugar que ocupa nuestro país en Latinoamérica sobre estos y otros rubros. Así las cosas, el problema en cuestión que se aborda es el de la necesidad imperante que existe en el Estado mexicano, no solo por adoptar a la IA como parte de esta nueva realidad en la que está inmersa la humanidad, sino en afrontar el reto de regular a la misma a pesar de ser una tecnología inacabada y en constante evolución, y el de implementar una serie de políticas públicas para lograr su desarrollo de acuerdo al contexto y necesidades del país, siendo los derechos humanos la clave en todo el proceso, por ser estos los derechos básicos de la humanidad, incluso tratándose de su desarrollo.

El trabajo que presentamos se compone de un marco teórico conceptual en el que se aborda a la IA, su clasificación y la relación que tiene con los derechos humanos, para posteriormente tratar en el último apartado el desarrollo y regulación de nuestro objeto de estudio en México, e identificar el lugar que ocupa el País en el contexto latinoamericano, sobre estos y otros indicadores. La presente investigación es del tipo analítica, descriptiva, basada en técnicas documentales a través de la cual se examinan tanto aspectos cualitativos como cuantitativos, siendo los principales métodos utilizados el analítico, deductivo, inductivo y sistemático.

### 2. ASPECTOS GENERALES

Previo a abordar la presente investigación, resulta fundamental establecer algunos conceptos clave, junto con su respectiva clasificación y su relación con los derechos humanos, es decir, comenzaremos por conceptualizar y/o definir a nuestro objeto de estudio, la Inteligencia Artificial (IA), su clasificación, y su relación con los derechos humanos, todo lo anterior con el fin de estar en aptitud de poder abordar con mayor precisión el trabajo que ha llevado el Estado mexicano para regular y desarrollar a nuestro tema de estudio desde una postura acorde con los Derechos humanos.

### 3. DEFINICIÓN Y CLASIFICACIÓN DE LA INTELIGENCIA ARTIFICIAL

Como punto de partida para definir y/o conceptualizar a nuestro objeto de estudio, comenzaremos con la definición que nos brinda la Real Academia Española, la cual señala que la Inteligencia Artificial (IA) es una "...Disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico."<sup>1</sup>

Otro de los primeros acercamientos al tema que nos ocupa lo encontramos en el autor Rodríguez Riquelme al señalar que la Inteligencia Artificial (IA), es el "...conjunto de tecnologías que sirven para automatizar procesos y simular ciertos rasgos de la inteligencia humana mediante sistemas informáticos. Entre las capacidades de la IA se encuentra el reconocimiento de voz, texto e imágenes, la traducción de idiomas y la generación de textos."<sup>2</sup>

Una conceptualización más ad hoc al ámbito tecnológico nos la brinda Google Cloud al precisar que la IA "...es un conjunto de tecnologías que permite a las computadoras aprender, razonar y realizar una variedad de tareas avanzadas de formas que antes requerían inteligencia humana, como comprender el lenguaje, analizar datos y hasta proporcionar sugerencias útiles. Es una tecnología transformadora que puede generar un cambio positivo y significativo en las personas, las sociedades y el mundo."<sup>3</sup>

Mientras que para el Parlamento Europeo, la inteligencia artificial simplemente "...es la habilidad de una máquina de presentar las mismas capacidades que los seres humanos, como el razonamiento, el aprendizaje, la creatividad y la capacidad de planear."<sup>4</sup>

En este sentido, a través del Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, tenemos que la Inteligencia Artificial es concebida como "un sistema basado en una máquina que está diseñado para funcionar con distintos niveles de autonomía y que puede mostrar capacidad de adaptación tras el despliegue, y que, para objetivos explícitos o implícitos, infiere de la información de entrada que recibe la manera de generar resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que pueden influir en entornos físicos o virtuales;"<sup>5</sup>

Por último, tenemos que la Comisión Europea señala que la inteligencia artificial es "una combinación de tecnologías que agrupa datos, algoritmos y capacidad informática. Los avances en computación y la creciente disponibilidad de datos son, por tanto, un motor fundamental en el pronunciado crecimiento actual de la inteligencia artificial."<sup>6</sup>

Como se puede observar, la mayoría de las definiciones "...apuntan a que la Inteligencia Artificial pueda dotar a las máquinas de lo que llamamos inteligencia y

<sup>1</sup> REAL ACADEMIA ESPAÑOLA, «inteligencia artificial», Diccionario de la lengua española.

<sup>2</sup> RODRÍGUEZ, RIQUELME, "¿México está preparado para la Inteligencia Artificial?" en Periódico en línea el Economista, de fecha 4 de julio de 2024.

<sup>3</sup> GOOGLE CLOUD, «¿Qué es la inteligencia artificial (IA)?», Google Cloud.

<sup>4</sup> PARLAMENTO EUROPEO, «¿Qué es la inteligencia artificial y cómo se usa?», Temas, 8 de septiembre de 2020; actualización de 26 de marzo de 2021.

<sup>5</sup> REGLAMENTO (UE) 2024/1689 DEL PARLAMENTO EUROPEO Y DEL CONSEJO, de fecha 13 de junio de 2024, publicado en el diario oficial de la Unión Europea, núm. 1689, de fecha 12 de julio de 2024, artículo 3, inciso 1), p.46.

<sup>6</sup> COMISIÓN EUROPEA, Libro Blanco, sobre la Inteligencia- un enfoque Europeo orientado a la Excelencia y la confianza, 19 de febrero, Bruselas, 2020, p.2.

dependiendo de la naturaleza de cada una de las aproximaciones que se tengan, se podría entender como un proceso centrado en: a) lo que realiza el ser humano y por lo tanto antropocéntrico o b) en la racionalidad del proceso."<sup>7</sup>

Ahora bien, respecto a la clasificación de la IA, tenemos que existen muchos criterios para dicho fin, sin embargo, uno de los que consideramos más completos toma en cuenta las capacidades y funcionalidad, el procesamiento del lenguaje natural, la visión por computadora, la robótica y los sistemas expertos, para entonces determinar que existe la IA Reactiva; IA de Memoria Limitada; Teoría de la mente IA; e IA autoconsciente.<sup>8</sup>

A continuación, describimos brevemente cada una:

1. IA Reactiva.- Los sistemas de IA reactiva, como su nombre lo indica reaccionan ante situaciones específicas, pero no son capaces de aprender de experiencias pasadas, por lo tanto, funcionan únicamente con la información que se les proporciona en tiempo real, como por ejemplo la computadora que juega al ajedrez y los vehículos autónomos.<sup>9</sup>
2. IA de memoria limitada.- "...esta forma de IA puede recordar eventos y resultados pasados y monitorizar objetos o situaciones específicas a lo largo del tiempo. La IA de memoria limitada puede utilizar datos del momento pasado y presente para decidir el curso de acción más probable para ayudar a lograr el resultado deseado."<sup>10</sup> Estos sistemas utilizan datos históricos para mejorar su rendimiento a lo largo del tiempo, como por ejemplo los asistentes personales como Siri y Alexa.<sup>11</sup>
3. Teoría de la mente IA.- Estos sistemas, están diseñados para comprender las emociones y los estados mentales como los seres humanos; estos sistemas por lo tanto son capaces de sentir empatía, aunque no emociones, este tipo de IA ha avanzado en etapas de desarrollo.<sup>12</sup>
4. IA Autoconsciente.- "La IA autoconsciente es un tipo de clase de IA funcional para aplicaciones que poseerían capacidades de súper IA. Al igual que la teoría de la IA mental, la IA autoconsciente es estrictamente teórica. Si alguna vez se logra, tendría la capacidad de comprender sus propias condiciones y rasgos internos junto con las emociones y pensamientos humanos. También tendría su propio conjunto de emociones, necesidades y creencias."<sup>13</sup>

#### 4. LOS DERECHOS HUMANOS Y SU RELACIÓN CON LA INTELIGENCIA ARTIFICIAL

Previo a abordar a la Inteligencia artificial y su relación con los derechos humanos, es preciso recordar brevemente qué son los derechos humanos y su clasificación, motivo

---

<sup>7</sup>BARRERA ARRESTEGUI, Luis, "Fundamentos Históricos y Filosóficos de la Inteligencia Artificial UCV-Hacer, en Revista de Investigación y Cultura, Vol. 1, Núm. 1, julio-diciembre, 2012, Perú, p. 88 (pp. 87-92).

<sup>8</sup>Cfr. MIT SLOAN MANAGEMENT REVIEW MÉXICO, «¿Cuáles son los 4 tipos de Inteligencia Artificial que existen?», MIT Sloan Management Review México, 29 de marzo de 2026.

<sup>9</sup>Cfr. Ídem.

<sup>10</sup>IBM DATA AND AI TEAM, «Comprender los diferentes tipos de inteligencia artificial», IBM Think.

<sup>11</sup>Cfr. Ídem.

<sup>12</sup>Cfr. MIT Sloan Management Review México, op. cit., nota 8.

<sup>13</sup>IBM, comprender los diferentes tipos de inteligencia artificial, op. cit., nota 10.

por el cual comenzaremos por señalar que, de conformidad con la Comisión Nacional de Derechos Humanos, estos pueden ser definidos como el "...conjunto de prerrogativas sustentadas en la dignidad humana, cuya realización efectiva resulta indispensable para el desarrollo integral de la persona."<sup>14</sup>

Además, estas prerrogativas son inherentes a todos los seres humanos, sin distinción alguna de nacionalidad, sexo, origen, religión, lengua, o cualquier otro aspecto, por lo que, en resumen, todos tenemos los mismos derechos humanos sin discriminación alguna, lo que evidencia que dentro de sus características o atributos se encuentran la de ser universales, interrelacionados, interdependientes e indivisibles.<sup>15</sup>

Mientras que para Luigi Ferrajoli los derechos humanos: "son los derechos primarios de las personas y conciernen indistintamente a todos los seres humanos, como por ejemplo... el derecho a la vida y a la integridad de la persona, la libertad personal, la libertad de conciencia y de manifestación del pensamiento, el derecho a la salud y a la educación y las garantías penales y procesales..."<sup>16</sup>

En el entendido que los derechos humanos obedecen a un tiempo y contexto, y por ello se les ha clasificado en generaciones de derechos humanos, dentro de las cuales podemos citar los siguientes: Derechos de Primera Generación (Civiles y Políticos); Derechos de Segunda Generación (Económicos, Sociales y Culturales); y Derechos de Tercera Generación (Colectivos, de la Solidaridad o de la Fraternidad.); y en un contexto relativamente reciente se han agregado otras generaciones, uno de estos criterios señala una cuarta generación de derechos humanos, que surge a partir de las nuevas formas que se presentan los derechos de primera, segunda y tercera generación en el entorno del ciberespacio, dentro de los cuales podemos encontrar el internet, las Tecnologías de Información y comunicación, y la IA, entre otros.<sup>17</sup>

Ahora bien, respecto a la relación entre nuestro objeto de investigación y los derechos humanos, cabe aclarar que dicha relación es multifacética, e incluso compleja, ya que por principio de cuentas, la IA puede tener un impacto positivo en todos los derechos humanos, dado el impacto y trascendencia que tiene en la humanidad, así por ejemplo podemos observar como a través de nuestro objeto de estudio se ha potenciado y ampliado el derecho a la salud, educación, trabajo, acceso a la información, entre muchos otros derechos, en beneficio de la propia humanidad, pero al mismo tiempo hemos constatado como ante la ausencia de una normativa o directriz sustentada en los mismos derechos humanos, la IA puede plantear serios riesgos en estos mismos derechos básicos y esenciales, por lo tanto consideramos que tanto en la investigación, desarrollo, aplicación y regulación de nuestro objeto de estudio, los derechos humanos deben ser el

---

<sup>14</sup>COMISIÓN NACIONAL DE LOS DERECHOS HUMANOS (CNDH), «¿Qué son los derechos humanos?», México.

<sup>15</sup>Cfr. Ídem.

<sup>16</sup>FERRAJOLI, Luigi, *Derechos y Garantías, la Ley del más Débil*, 4ª ed., Ed. Trotta, Madrid España. 2004, p.40.

<sup>17</sup>Cfr. MOLINARES HASSAN, Viridiana, *Notas sobre Constitución, Organización del Estado y Derechos Humanos*, 2ª ed., Ed. Universidad del Norte, Barranquilla, Colombia, 2011, p. 77; así como: GUERRERO MARTÍNEZ, Rodolfo, "Derechos humanos de cuarta generación y las tecnologías de la información y de la comunicación." En *Revista Derechos Fundamentales a Debate*, No. 12, Enero-Abril, Guadalajara, Jalisco, Año 2020, p.137.

punto de partida en cualquiera de los ámbitos previamente señalados, todo lo anterior con el fin de potenciar sus beneficios y reducir sus riesgos.<sup>18</sup>

En otras palabras, consideramos que dado que en algunos casos la IA ha vulnerado aquellos derechos básicos que nos conciernen indistintamente a todos los seres humanos, y que existe el riesgo de que con su evolución se sigan vulnerando los mismos, debe existir una regulación fuerte basada en los propios derechos Humanos, además aseveramos que necesitamos un punto de partida común entre los Estados-Nación para regular en el contexto global a nuestro objeto de estudio en todas sus facetas, incluida la investigación, desarrollo y aplicación, por lo que consideramos que los derechos públicos subjetivos son el medio adecuado y práctico para dicho fin.<sup>19</sup>

Además de lo anterior, es preciso resaltar que la relación entre la IA y los derechos humanos no solo se limita a los esfuerzos por regular y desarrollar una tecnología inacabada y en constante evolución, ya que incluso cuando se trata del derecho que le asiste a la humanidad a acceder a esta nueva tecnología, nuevamente encontramos en los derechos públicos subjetivos el fundamento adecuado para exigir la inclusión de todo individuo en los avances científicos, incluida desde nuestra perspectiva la tecnología (IA) y todos sus beneficios que brinda a la sociedad, tal y como se desprende del artículo 27 de la Declaración de los derechos humanos de 1948 que a la letra reza:

"Toda persona tiene derecho a tomar parte libremente en la vida cultural de la comunidad, a gozar de las artes y a participar en el progreso científico y en los beneficios que de él resulten,"<sup>20</sup>

De lo que se desprende entonces, que la Inteligencia Artificial puede y debe ser una herramienta tecnológica accesible para toda la humanidad, a través de la cual se combatan las diferentes brechas en las que se desenvuelve el ser humano, sin embargo, la realidad de las cosas es que tal ideal y derecho humano, se encuentra amenazado como muchos otros por los intereses de quienes de manera egoísta buscan mantener su hegemonía, tal y como se advierte de la carrera tecnológica por el desarrollo de la IA entre Estados Unidos, China y la Unión Europea, por lo que la forma de evitar ser excluido de los beneficios tecnológicos que implica el desarrollo y aplicación de nuestro tema de investigación, es sumándose los diversos países, incluido México al campo de su estudio, investigación y desarrollo, lo cual nos lleva a un modelo económico basado en el conocimiento.<sup>21</sup>

## 5. DESARROLLO Y REGULACIÓN DE LA IA MÉXICO

Es del conocimiento general, que en los últimos años la inteligencia artificial ha experimentado un crecimiento significativo a nivel global, y México entiende la importancia de sumarse al reto de desarrollar y regular la IA, sin embargo, comenzó por centrar su

---

<sup>18</sup>Cfr. CRUZ GONZÁLEZ, Marcos, "La inteligencia artificial y los derechos humanos. Una carrera tecnológica entre Estados Unidos, China y la Unión Europea", en Revista IUS, Vol. 20, No. 56. Enero-Agosto, 2025, pp. 39, 40 y 41.

<sup>19</sup>Cfr. Ibidem, pp. 40 y 41.

<sup>20</sup>ORGANIZACIÓN DE LAS NACIONES UNIDAS (ONU), Declaración Universal de Derechos Humanos, 10 de diciembre de 1948, art. 27.1.

<sup>21</sup>\*Opinión del autor con base en el material bibliográfico y hemerográfico consultado hasta el momento.

atención en el aspecto regulatorio, antes de encausar sus esfuerzos en la investigación y desarrollo de nuestro objeto de estudio, tal y como se puede observar a través de las diversas iniciativas de ley presentadas en el Congreso de la Unión, así como de la Declaración de ética y buenas prácticas para el uso y desarrollo de IA, y las recientes políticas públicas llevadas a cabo por el gobierno de México, de las cuales seleccionaremos y describiremos algunas de ellas de manera concisa; estableciendo al final del presente trabajo el lugar que ocupa México en América Latina y el Caribe en relación al estado de avance, adopción, gobernanza y madurez de la inteligencia artificial (IA).<sup>22</sup>

### 5.1 Regulación de la IA en México

Por lo que respecta a las diversas iniciativas de ley presentadas en el Congreso de la Unión, tenemos que estas están principalmente enfocadas a reformar nuestra Constitución Política Mexicana, mientras que otras se enfocan al ámbito académico y/o científico, por ser este un importante impulsor del tema en cuestión; incluso en fechas recientes contamos con una Declaración de ética y buenas prácticas para el uso y desarrollo de IA que si bien no es vinculante si constituye un marco orientador en el diseño, implementación y evaluación de sistemas de inteligencia artificial.<sup>23</sup>

A continuación, seleccionamos aquellas que consideramos importantes:

#### 5.1.1 Iniciativas de Reforma a la Constitución Política Mexicana en Materia de Inteligencia Artificial

I.- Iniciativa que reforma la fracción XVII al artículo 73 de la Constitución Política de los Estados Unidos Mexicanos, para facultar al Congreso de la Unión para emitir las normas necesarias para regular la investigación, desarrollo y aplicaciones de la inteligencia artificial, de fecha 05 de septiembre de 2023.<sup>24</sup>

II.- Iniciativa que Reforma la Fracción XVII al artículo 73 de la Constitución Política de los Estados Unidos Mexicanos, en materia de Inteligencia Artificial, Ciberseguridad y Neuroderechos de fecha 20 de septiembre de 2023.<sup>25</sup>

III.- Iniciativa que Adiciona el artículo 73 de la Constitución Política de los Estados Unidos Mexicanos, en materia de inteligencia artificial de fecha 05 de marzo de 2024.<sup>26</sup>

IV.- iniciativa de Reforma a la Constitución Política, Presentada por el Diputado Eruviel Ávila Villegas, con el objetivo de plasmar en la Carta Magna la ética y perspectiva

---

<sup>22</sup>Cfr. AYALA, Juan, "Actualidades de la inteligencia artificial en México", en el Heraldo de México, de fecha 14 de mayo de 2024.

<sup>23</sup>Cfr. MEZA RUIZ, Ivan Vladimir, El camino hacia una regulación de la IA en México, de fecha 7 de Julio de 2024 Así como: SECIHTI, Ciencia y Tecnología, Presentan Declaración de ética y buenas prácticas para el uso y desarrollo de la IA en México: SECIHTI y ATDT, de fecha 29 de enero de 2026.

<sup>24</sup>Cfr. Iniciativa con proyecto de decreto por el que se reforma la fracción XVII al artículo 73 de la Constitución Política de los Estados Unidos Mexicanos, de fecha 5 de septiembre de 2023, presentada en la Ciudad de México, en la Gaceta Parlamentaria, de fecha 5 de septiembre de 2023.

<sup>25</sup>Cfr. Iniciativa con Proyecto de Decreto por el que se Reforma la Fracción XVII, al artículo 73 de la Constitución Política de los Estados Unidos Mexicanos, en Materia de Inteligencia Artificial, Ciberseguridad y Neuroderechos, de fecha 20 de septiembre de 2023, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 20 de septiembre de 2023.

<sup>26</sup>Cfr. Iniciativa que Adiciona el artículo 73 de la Constitución Política de los Estados Unidos Mexicanos, en Materia de Inteligencia Artificial, de fecha 5 de marzo de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria, de fecha 5 de marzo de 2024.

de derechos humanos como ejes rectores del uso y desarrollo de la Inteligencia Artificial (IA).<sup>27</sup>

(PEVM) Cabe destacar que respecto de las tres iniciativas de reforma antes señaladas, versan sobre otorgarle facultades al Congreso de la Unión para legislar en materia de Inteligencia artificial y que una de las ultimas iniciativas de las que tenemos conocimiento es la presentada por el diputado, con el objetivo de plasmar en la Carta Magna la ética y perspectiva de derechos humanos como ejes rectores del uso y desarrollo de la Inteligencia Artificial (IA).

### **5.1.2 Iniciativas de Ley que Pretenden Regular e Impulsar la IA desde el Ámbito Científico y/o Académico**

Respecto a las Iniciativas de Ley tendientes a regular el uso de la Inteligencia Artificial en el ámbito Académico y/o científico de manera directa o Indirecta podemos destacar los siguientes Proyectos:

I.- Iniciativa que Reforma y Adiciona Diversas Disposiciones de La Ley de Ciencia y Tecnología y de la Ley General de Salud, de fecha 28 de abril de 2020.<sup>28</sup>

II.- Iniciativa que expide la Ley que crea la Universidad de Tecnologías de la Información, Comunicaciones e Innovación, de fecha 12 de agosto de 2020.<sup>29</sup>

III.- Iniciativa que Reforma y Adiciona Diversas Disposiciones de la Ley de Ciencia y Tecnología, de fecha 07 de enero de 2021.<sup>30</sup>

IV.- Iniciativa del Congreso de Jalisco, con Proyecto de Decreto por el que se Reforma y Adiciona el Artículo 30 de la Ley General de Educación, de fecha 22 de febrero de 2021.<sup>31</sup>

V.- Iniciativa que Reforma los Artículos 30 y 70 de la Ley General de Educación y 7° de la Ley General de Educación Superior de fecha 04 de octubre de 2023.<sup>32</sup>

VI.- Iniciativa que reforma la fracción IV del artículo 30 de la Ley General de Educación en Materia de Inteligencia Artificial como parte de los planes y Programas de Estudio, de fecha 22 de septiembre de 2023.<sup>33</sup>

---

<sup>27</sup>Cfr. Cámara de Diputados LXVI Legislatura, Nota No. 2321, Presenta Eruviel Ávila iniciativa de reforma constitucional en materia de Inteligencia Artificial y Tecnologías Emergentes.

<sup>28</sup>Cfr. Iniciativa que Reforma y Adiciona Diversas Disposiciones de la Ley de Ciencia y Tecnología y de la Ley General de Salud, de fecha 28 de abril de 2020, presentada en la Ciudad de México, en la Gaceta Parlamentaria, de fecha 28 de abril de 2020.

<sup>29</sup>Cfr. Iniciativa que expide la Ley que crea la Universidad de Tecnologías de la Información, Comunicaciones e Innovación, de fecha 12 de agosto de 2020, presentada en la Ciudad de México, en la Gaceta Parlamentaria, de fecha 12 de agosto de 2020.

<sup>30</sup>Cfr. Iniciativa que Reforma y Adiciona Diversas Disposiciones de la Ley de Ciencia y Tecnología, de fecha 7 de enero de 2021, Presentada en la Ciudad de México, en la Gaceta Parlamentaria, de fecha 7 de enero de 2021.

<sup>31</sup>Cfr. Iniciativa del Congreso de Jalisco, con Proyecto de Decreto por el que se Reforma y Adiciona el Artículo 30 de la Ley General de Educación, de fecha 22 de febrero de 2021, Presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 18 de marzo de 2021.

<sup>32</sup>Cfr. Iniciativa que Reforma los Artículos 30 Y 70 de la Ley General de Educación y 7° de la Ley General de Educación Superior, de fecha 4 de octubre de 2023, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 4 de octubre de 2023.

<sup>33</sup>Cfr. Iniciativa que reforma la fracción IV del artículo 30 de la Ley General de Educación en Materia de Inteligencia Artificial como parte de los planes y Programas de Estudio, de fecha 22 de septiembre de 2023, presentada en la Ciudad de México, en la Gaceta parlamentaria, de fecha 26 de septiembre de 2023.

VII.- Iniciativa que reforma diversas disposiciones de la Ley General de Educación y de la Ley General de Educación Superior, en materia de inteligencia artificial de fecha 05 de diciembre de 2023.<sup>34</sup>

Viii.- Iniciativa que adiciona diversas disposiciones de la Ley General en Materia de Humanidades, Ciencias, Tecnologías e Innovación, en materia de ética y derechos humanos, de fecha 17 de enero de 2024.<sup>35</sup>

IX.- Iniciativa que reforma y adiciona diversas disposiciones de la Ley General en materia de Humanidades, Ciencias, Tecnologías e Innovación, de fecha 14 de febrero de 2024.<sup>36</sup>

X.- Iniciativa que adiciona una fracción IV Bis al artículo 30 y adiciona un artículo 86 de la Ley General de Educación de fecha 13 de febrero de 2024.<sup>37</sup>

XI.- Iniciativa que adiciona una nueva fracción VIII Bis al artículo 11 de la Ley General en Materia de Humanidades, Ciencias, Tecnologías e Innovación de fecha 05 de marzo de 2024.<sup>38</sup>

XII.- Iniciativa que adiciona diversas disposiciones de la Ley General en materia de Humanidades, Ciencias, Tecnologías e Innovación, en materia de lineamientos de políticas públicas para el fomento y promoción del uso de Inteligencia Artificial, de fecha 14 de febrero de 2024.<sup>39</sup>

Xiii.- Iniciativa que reforma y adiciona diversas disposiciones de la Ley General en materia de Humanidades, Ciencias, Tecnologías e Innovación, relativa a la inteligencia artificial de fecha 01 de abril de 2024.<sup>40</sup>

### 5.1.3 Declaración de ética y buenas prácticas para el uso y desarrollo de IA

Además de lo anterior, el gobierno de México también ha trabajado en un marco orientador para el diseño, implementación y evaluación de sistemas de inteligencia

---

<sup>34</sup>Cfr. Iniciativa que reforma diversas disposiciones de la Ley General de Educación y de la Ley General de Educación Superior, en materia de Inteligencia Artificial, de fecha 5 de diciembre de 2023, presentada en la Ciudad de México, en la Gaceta Parlamentaria, de fecha 5 de diciembre de 2023.

<sup>35</sup>Cfr. Iniciativa que adiciona diversas disposiciones de la Ley General en Materia de Humanidades, Ciencias, Tecnologías e Innovación, en materia de ética y derechos humanos, de fecha 17 de enero de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 17 de enero de 2024.

<sup>36</sup>Cfr. Iniciativa que reforma y adiciona diversas disposiciones de la Ley General en Materia de Humanidades, Ciencias, Tecnologías e Innovación, de fecha 14 de febrero de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 14 de febrero de 2024.

<sup>37</sup>Cfr. Iniciativa que adiciona una fracción IV Bis al artículo 30 y adiciona un artículo 86 de la Ley General de Educación de fecha 13 de febrero de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 28 de febrero de 2024.

<sup>38</sup>Cfr. Iniciativa que adiciona una nueva fracción VIII Bis al artículo 11 de la Ley General en Materia de Humanidades, Ciencias, Tecnologías e Innovación de fecha 5 de marzo de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 6 de marzo de 2024.

<sup>39</sup>Cfr. Iniciativa que adiciona diversas disposiciones de la Ley General en materia de Humanidades, Ciencias, Tecnologías e Innovación, en materia de lineamientos de políticas públicas para el fomento y promoción del uso de inteligencia artificial, de fecha 14 de febrero de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 14 de febrero de 2024.

<sup>40</sup>Cfr. Iniciativa que reforma y adiciona diversas disposiciones de la Ley General en Materia de Humanidades, Ciencias, Tecnologías e Innovación, Relativa a la Inteligencia Artificial de fecha 1 de abril de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 2 de marzo de 2024.

artificial (IA) en el país, el cual fue desarrollado y presentado por la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (Secihti) junto con la Agencia de Transformación Digital y Telecomunicaciones (ATDT); en palabras de la titular de la Secihti, Rosaura Ruiz Gutiérrez, tal Declaración constituye una hoja de ruta o guía para orientar decisiones de política pública sobre el desarrollo y uso de la IA, promoviendo una innovación responsable que contribuya al bienestar social, la inclusión, la soberanía tecnológica del país, y protección de los derechos humanos, es decir, tiene como objetivo que sean los derechos humanos la pieza clave a través de la cual sean adoptadas toda clase medidas, acciones y estrategias que garanticen un uso ético y responsable de la IA.<sup>41</sup>

Sin más preámbulos, la Declaración de ética y buenas prácticas para el uso y desarrollo de IA también conocido como los principios de Chapultepec, establecen lo siguiente:

I.- La IA debe ampliar derechos, nunca reducirlos: Implica que los sistemas de IA deben evitar tomar decisiones automatizadas desvinculadas de los derechos humanos, ya que de lo contrario nuestro objeto de estudio podría reproducir formas de exclusión, prejuicios sociales o violentar el derecho a la información, por lo tanto, deberán revisarse de manera continua los datos y criterios que utiliza un sistema de este tipo atenta contra los derechos humanos.<sup>42</sup>

Esto significa que el uso de la IA debe tener una visión pública, de protección de derechos humanos y sociales, y garantizar un marco regulatorio responsable, ya que existe el riesgo de que la IA profundice brechas de desigualdad, genere discriminación y sesgos en sus contenidos, como por ejemplo el derecho a la información.<sup>43</sup>

II. Toda decisión apoyada por IA debe tener responsables humanos: En relación a este principio ético presentado por la Secihti no pasa desapercibido para nosotros que el mismo guarda gran parecido con la recomendación sobre la ética de la Inteligencia Artificial proporcionada por la Unesco durante su 41ª reunión durante el periodo del 9 al 24 de noviembre de 2021, a través de la cual se indica que la responsabilidad ética y la obligación de rendir cuentas de las decisiones y las acciones basadas de alguna manera en un sistema de IA siempre deberían ser atribuibles, en última instancia, a los actores de la IA conforme a la función que tengan en el ciclo de vida del sistema de IA.<sup>44</sup>

---

<sup>41</sup>Cfr. SECIHTI, ciencia y tecnología, Presentan Declaración de ética y buenas prácticas para el uso y desarrollo de la IA en México: SECIHTI y ATDT, op. cit., nota 22. Así como: PANTOJA, Sara, "Presentan declaración de ética y buenas prácticas de Inteligencia Artificial; este es el decálogo" en revista proceso, de fecha 29 de enero de 2026.

<sup>42</sup>Cfr. GOBIERNO DE MÉXICO, Principios de Chapultepec. Declaración de ética y buenas prácticas para el uso y desarrollo de la inteligencia artificial.

<sup>43</sup>Cfr. POY SOLANA, Laura, "Presenta Secretaría de Ciencia decálogo de principios de ética y buen gobierno de la IA en México" en periódico la jornada, de fecha 29 de enero de 2026.

<sup>44</sup>Cfr. UNESCO, Recomendación sobre la Ética de la Inteligencia Artificial, de fecha 23 de noviembre de 2021, p.74.

A decir de la propia Secihti, La inteligencia artificial no puede operar sin responsables, es decir, su uso en el ámbito público debe contar y sustentarse en marcos institucionales a través de los cuales queden precisados los responsables que intervienen en el diseño, la toma de decisiones, la supervisión, y a quien se le debe atribuir en consecuencia la responsabilidad por sus efectos.<sup>45</sup>

III.- Si una decisión no puede explicarse, no debe automatizarse: Por principio de cuentas, hay que precisar que este principio ético se centra en la IA explicable, la cual en síntesis se trata de un campo dentro de la inteligencia artificial que busca que los sistemas basados en IA, sobre todo los más complejos, puedan explicar cómo y por qué toman ciertas decisiones, lo cual es de suma importancia cuando los modelos de IA operan en áreas complejas y sensibles dentro nuestra sociedad, como por ejemplo en la salud, cuando se emite un diagnóstico médico, o en el campo de la justicia, cuando es utilizada para las evaluaciones de riesgos o sentencias, o incluso en el área de las finanzas, donde puede utilizarse para evaluar aprobaciones de créditos, por lo tanto, los algoritmos deben ser comprensibles tanto para los usuarios técnicos como para los usuarios finales, ya que de no existir esa transparencia, la IA puede reforzar sesgos y reproducir desigualdades sin que nadie lo advierta.<sup>46</sup>

Aclarado lo anterior, podemos entender la relevancia del tercer principio ético que nos ocupa, y por qué se recomienda que no debe automatizarse una decisión a la que se haya llegado a través de la IA, si la misma no puede explicarse en términos claros y sencillos.

IV.- La IA se gobierna mejor cuando se decide en colectivo: De acuerdo a los propios principios de Chapultepec, esto implica que las decisiones sobre el uso de la IA deben ser tomadas en grupo o por el mayor número de miembros de la sociedad y no de manera individual o a través de una sola perspectiva o área; ya que involucrar diferentes perspectivas y enfoques ayuda a fortalecer el criterio, reducir riesgos y asegurar que la tecnología se utilice con responsabilidad y transparencia, ya que como se indica en el propio principio ético, este principio busca asegurar que las decisiones sobre el desarrollo y uso de la IA no se tomen de manera aislada ni exclusivamente técnica, sino que se decidan en un contexto de diálogo y participación, lo cual asegura que se responda a necesidades reales, se consideren distintos puntos de vista y se oriente al interés público, lo que se traduce en un fortalecimiento de su legitimidad y desarrollo en el país.<sup>47</sup>

V.- La IA sólo es valiosa si genera bienestar para las personas: En relación con este principio ético, cabe señalar que existen principalmente dos posturas sobre la IA que resultan contradictorias; la primera considera a nuestro objeto de estudio como una herramienta fundamental para enfrentar los grandes desafíos globales como la pobreza, desigualdad y el cambio climático, los cuales son pilares de los Objetivos de Desarrollo

---

<sup>45</sup>Cfr. GOBIERNO DE MÉXICO, Principios de Chapultepec, Declaración de ética y buenas prácticas para el uso y desarrollo de la Inteligencia Artificial, op. cit., nota 40.

<sup>46</sup>Cfr. IA INSTITUTO DE INNOVACIÓN DIGITAL, «Redes neuronales explicables: el desafío de entender las decisiones de la IA», Noticias, 13 de marzo de 2025. A mayor abundamiento, véase IBM, «¿Qué es la interpretabilidad de la IA?», IBM Think.

<sup>47</sup>Cfr. Gobierno de México, Principios de Chapultepec, Declaración de ética y buenas prácticas para el uso y desarrollo de la Inteligencia Artificial, op. cit., nota 40.

Sostenible (ODS) de la Organización de las Naciones Unidas (ONU), ya que a través de los mismos se pretende el bienestar social, económico y ambiental.<sup>48</sup>

Sin embargo, la segunda postura en torno a la IA nos hace reflexionar en el hecho de que detrás de este avance vertiginoso y estos fines loables debe tenerse en cuenta que nuestro tema de investigación no puede considerarse neutral, ya que tanto su desarrollo y aplicación también pueden acentuar desigualdades sociales existentes, generando nuevas formas de exclusión, explotación laboral y desequilibrio ambiental. Así, mientras algunos sectores se benefician del poder transformador de estas tecnologías, otros enfrentan los costos invisibles de su funcionamiento. La pregunta fundamental sería: ¿cómo construir una inteligencia artificial equitativa y al servicio de todos?<sup>49</sup>

VI.- Antes de automatizar, hay que comprender a quién y qué afecta: Para entender mejor este principio ético, hay que tener en cuenta, qué debemos entender por automatización en el contexto de nuestro tema de estudio, en el entendido que se refiere a un proceso de sustitución de trabajo humano por máquinas, lo anterior con el fin de hacer más eficientes determinadas tareas de los puestos de trabajo, lo cual ha derivado en que muchos empleos, tanto de baja como de alta calificación, y muchos trabajadores se hayan visto amenazados de ser reemplazados, y han perdido la carrera con la tecnología.<sup>50</sup>

En relación con la segunda parte del principio que nos ocupa, y que hace énfasis en ¿a quién? y ¿qué afecta?, tenemos que de una revisión de los efectos, llevada a cabo en relación con las habilidades, los empleos y los trabajadores, la automatización todavía está centrada en reemplazar tareas y empleos rutinarios y de baja cualificación, no obstante, la automatización y la IA son capaces de impulsar la eficiencia hasta el punto de que se creen más y mejores puestos de trabajo, y en tal sentido si los trabajadores adquieren las habilidades para interactuar con la IA, también es posible un futuro de trabajo de entendimiento entre personas y máquinas.<sup>51</sup>

VII. La Tecnología Estratégica debe responder a las necesidades del país: De acuerdo con la Declaración de ética y buenas prácticas para el uso y desarrollo de la Inteligencia Artificial, presentada por la Secihti, México debe proponerse desarrollar sus tecnologías estratégicas (la inteligencia artificial), con base en un enfoque autónomo y encauzado a fortalecer las capacidades del país, para lo cual debe tener en cuenta las necesidades nacionales, así como una visión a largo plazo.<sup>52</sup>

Este principio es de suma importancia, si tenemos en cuenta que México refleja un rezago sobre en el desarrollo, adopción y regulación de nuestro tema de estudio, lo que hace evidente que se requieren acciones decididas para atender dicha problemática,

---

<sup>48</sup>Cfr. DÍAZ, Viridiana, "La IA como motor clave para un país sostenible e inclusivo", en *El Economista*, de fecha 14 de julio de 2025.

<sup>49</sup>Cfr. UNAM GLOBAL REVISTA, ¿La inteligencia artificial reduce desigualdades o las profundiza aún más?, de fecha 10 de junio de 2025.

<sup>50</sup>Cfr. FICAPAL-CUSÍ, Pilar, "IA, automatización y trabajo humano: de la carrera al entendimiento" en *Oikonomics Revista de economía, empresa y sociedad*, Núm. 23, noviembre, Catalunya, España, 2024, p. 1.

<sup>51</sup>Cfr. Ídem.

<sup>52</sup>Cfr. Gobierno de México, Principios de Chapultepec, Declaración de ética y buenas prácticas para el uso y desarrollo de la Inteligencia Artificial, op. cit., nota 40.

sobre todo si tenemos en cuenta que los países que hoy inviertan en infraestructura, talento y gobernanza serán los países que definan las reglas del juego para la próxima década en torno a la IA. En este mismo orden de ideas, los principios de Chapultepec, como el Laboratorio Nacional de Inteligencia Artificial, o el Centro Público de Formación en IA, son una muestra de los esfuerzos que existen por crear una ruta de desarrollo de acuerdo con las necesidades del país, pero tales esfuerzos aún distan de ser una Estrategia Nacional de Inteligencia Artificial debidamente definida y operativa.<sup>53</sup>

VIII.- El Desarrollo de la IA requiere fortalecer la Educación y el Conocimiento en el País: Este principio ético de Chapultepec, por principio de cuentas cabe señalar que guarda estrecha relación con el modelo económico del conocimiento como base de la economía, ya que de acuerdo con el mismo el conocimiento es el activo más importante que los bienes de capital y mano de obra, y donde la cantidad y sofisticación de este permea en las actividades económicas y sociales, por lo que el conocimiento está en el centro de dicha estrategia, esto finalmente implica que la educación y capacitación nacional son la piedra angular.<sup>54</sup>

A mayor abundamiento, para impulsar el desarrollo de la IA en el país, el Estado mexicano y las Universidades del país se han dado a la tarea de fomentar la educación y la investigación sobre nuestro tema de estudio, principalmente a través de la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (Secihti), así como, a través de otras instituciones educativas de renombre en el país, como el Centro de Ciencias Matemáticas (CCM) por parte de la UNAM; el Centro de Investigación y Estudios Avanzados (CINVESTAV) en el Departamento de Ciencias de la Computación en su laboratorio de sistemas inteligentes, por parte del Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE); el Instituto Politécnico Nacional (IPN) integra su investigación en IA en la Escuela de Computación (ESCOM), la Buap se concentra en la IA a través del Laboratorio de Sistemas Robóticos (SIRO) y más reciente mediante la Agencia de Inteligencia Artificial y Ética Algorítmica (AURA); entre muchas otras Instituciones en la República Mexicana.<sup>55</sup>

IX.- La IA no puede ser ajena a la diversidad cultural y lingüística del País: En México se estima que se hablan 68 diferentes lenguas, por lo que el país se encuentra dentro de las diez naciones con mayor diversidad lingüística del mundo, y en segundo lugar en América Latina, en la inteligencia, que una lengua es considerada como un sistema completo de comunicación con su propia gramática, sintaxis, fonología y léxico.<sup>56</sup>

En relación con lo anterior, deben llevarse a cabo esfuerzos por preservar las lenguas indígenas en México, no solo por la riqueza cultural que representa para nuestro País,

---

<sup>53</sup>Cfr. CENTRO MÉXICO DIGITAL, *Inteligencia artificial en México: de la promesa al impacto económico*, Adopción empresarial e implicaciones económicas, Ciudad de México, 2026, p.8, y 49.

<sup>54</sup>Cfr. CARLOS SÁNCHEZ y HUMBERTO RÍOS, "La economía del conocimiento como base del crecimiento económico en México" en *Revista Venezolana de Información, Tecnología y Conocimiento*, vol. 8, núm. 2, mayo-agosto, 2011, p. 46.

<sup>55</sup>Cfr. BAÑOS-GONZÁLEZ, "La inteligencia artificial, estudio de su evolución y aplicación en México", en *Pädi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI*, Vol. 12 No. Especial 4, noviembre, Hidalgo, Mexico, 2024, p.258 Así como: AGENCIA DE INTELIGENCIA ARTIFICIAL Y ÉTICA ALGORÍTMICA (AURA) BUAP.

<sup>56</sup>Cfr. UDLAP, «¿Cómo usar la IA para preservar y difundir la diversidad lingüística?», IA UDLAP, 27 de febrero de 2026; así como SECRETARÍA DE CULTURA, «¿Sabías que en México hay 68 lenguas indígenas, además del español?», Gobierno de México, 21 de febrero de 2018.

sino también porque se trata de un asunto de derechos humanos, ya que cuando las personas se ven obligadas a abandonar su lengua nativa, lo hacen con el objetivo de evitar la exclusión social, motivo por el cual existe una vulneración del derecho de igualdad, identidad cultural y no discriminación.<sup>57</sup>

La Constitución Política Mexicana no establece al español como idioma oficial, y, por otra parte, reconoce plenamente a las lenguas indígenas, sin embargo no es un secreto que las personas que hablan una lengua distinta al español en México han enfrentado discriminación, lo que ha llevado a un menor uso de las mismas, por lo tanto, la IA puede ser una herramienta poderosa para visibilizar a los pueblos indígenas, y rescatar sus lenguas originarias, sin embargo también conlleva riesgos significativos, como la perpetuación de sesgos culturales, la apropiación indebida de conocimientos ancestrales y la exclusión informativa.<sup>58</sup>

En síntesis, si bien México necesita desarrollar su propio enfoque sobre la inteligencia artificial, a través del cual se refleje la realidad pluricultural presente en el país, no menos cierto es que previo a hablar de inteligencia artificial, resulta fundamental resolver algunos problemas estructurales, como por ejemplo la brecha digital que aún persiste, la falta de infraestructura, y el acceso desigual a la tecnología en toda la sociedad.<sup>59</sup>

X.- Los datos son un bien público que debe cuidarse con responsabilidad:.

La IA se construye a partir de datos, estos datos pueden ser imágenes, documentos, información personal, etcétera, esto significa que dicho elemento reviste una doble importancia, en primer lugar, porque sin datos de calidad, actualizados y bien estructurados, los modelos no pueden aprender, tomar decisiones fiables ni generar valor práctico, en otras palabras, los datos son el combustible que permite que los algoritmos funcionen, y que la IA se convierta en una herramienta útil para empresas, gobiernos y sociedad.<sup>60</sup>

Por otra parte, este elemento también resulta importante dado que los datos y su uso son un tema recurrente de preocupación en la medida en que esté en riesgo la privacidad de los mismos, sobre todo porque muchas personas que usan a diario a nuestro objeto de estudio están alimentando a los sistemas con datos sensibles, y por consiguiente erosionando su propia privacidad, por eso el Estado Mexicano reconoce la importancia de este recurso y el riesgo que persiste ante la ausencia de controles de acceso adecuados, o de gobierno de la IA.<sup>61</sup>

## 5.2 Desarrollo de la Inteligencia Artificial en México

Dentro de las acciones que han tenido lugar para emprender el desarrollo de la IA en nuestro País, podemos destacar tres, dentro de la cuales se puede observar la partici-

---

<sup>57</sup>Cfr. Ídem.

<sup>58</sup>Cfr. ZARATE Itzel, "La IA y los pueblos indígenas", en Unam Gacetas Políticas, 10 de octubre, Ciudad de México, 2025. Así como: Constitución Política de los Estados Unidos Mexicanos, Última reforma publicada en el Diario Oficial de la Federación el 3 de marzo de 2026, artículo 2 fracción V, p.3.

<sup>59</sup>Cfr. Ídem.

<sup>60</sup>Cfr. DATOS.GOB.ES, «¿Cómo se construye un modelo de inteligencia artificial?», Gobierno de España, 4 de febrero de 2025.

<sup>61</sup>Cfr. IBM, «Guía sobre la protección de los datos de la IA y el machine learning», IBM Think, 16 de diciembre de 2025.

pación del Congreso de la Unión; el sector privado, las Universidades y el Gobierno de México, entre otros muchos actores, nos referimos específicamente a las siguientes:

1.- La iniciativa de Creación de la Agenda Nacional de IA (2024-2030): esta Iniciativa fue presentada por la senadora Alejandra Lagunes Soto Ruíz en el año 2024, se trata de una agenda que busca consolidar un marco de trabajo conjunto entre el gobierno, la academia y el sector privado, en la inteligencia que para su elaboración se realizaron 18 mesas de trabajo y 90 reuniones para mapear el ecosistema de IA en México, y si bien hasta el momento no ha sido aprobado por el Congreso de la Unión, si demuestra el interés que existe en México por generar ciencia y tecnología en el país.<sup>62</sup>

La citada agenda es importante, ya que a través de la misma se fijan una serie puntos clave sobre la IA en México, uno de esos puntos tiene que ver con su aspecto regulatorio, pero además de este aspecto se aborda la Gobernanza, la Creación de infraestructura, y la creación de un ecosistema para dicha tecnología, fomentándose en consecuencia la inversión en investigación y desarrollo; además de que se precisa la necesidad de construir las capacidades humanas que preparen a nuestra sociedad para la transformación que implica el desarrollo y aplicación de la Inteligencia artificial en territorio Mexicano.<sup>63</sup>

En resumen, esta iniciativa busca establecer una hoja de ruta para regular, fomentar y diagnosticar el uso ético y desarrollo de la IA en el país, con el objetivo de integrar estas estrategias en el Plan Nacional de Desarrollo del Gobierno de Mexico.<sup>64</sup>

2.- El Desarrollo de IA Soberana a través de la Iniciativa «MÉXICO IA+ Inversión Acelerada (2025): A través del sector privado se lanzó la presente iniciativa, específicamente, se trata de una plataforma que articula inteligencia artificial, infraestructura digital e inversión con enfoque de soberanía tecnológica y desarrollo económico regional, cuya primera edición o foro tuvo lugar del 12 al 13 de noviembre del 2025, a través de la misma se busca convertir al país en un centro para las inversiones en infraestructura tecnológica e inteligencia artificial (IA), es decir, se pretende posicionar al país como un hub regional de IA, con un enfoque en desarrollar un lenguaje de inteligencia artificial propio y adaptado a las necesidades nacionales.<sup>65</sup>

A mayor abundamiento, México IA+ Inversión Acelerada fue un evento clave organizado por el Consejo Coordinador Empresarial (CCE) y Nvidia, contando con la participación de la Secretaría de Economía que "...busca articular a gobierno, industria, academia y sociedad civil en torno a un objetivo: que el país desarrolle y controle su

---

<sup>62</sup>Cfr. SENADO DE LA REPÚBLICA, LXVI Legislatura, «Presentan en el Senado Agenda Nacional de la Inteligencia Artificial para México 2024-2030», 15 de mayo de 2024.

<sup>63</sup>Cfr. LAGUNES SOTO, Alejandra, MARTÍNEZ MANCILLA, Yolanda, et al., "Propuesta de Agenda Nacional de la Inteligencia Artificial para México (2024 - 2030)", en Alianza Nacional de Inteligencia Artificial (ANIA), Mayo 2024, México, pp.16, 17, 18, 19, 20, 21 y 22.

<sup>64</sup>Cfr. SENADO DE LA REPÚBLICA, LXVI Legislatura, Presentan en el Senado Agenda Nacional de la Inteligencia Artificial para México 2024-2030, op. cit., nota 61.

<sup>65</sup>Cfr. OLIVIA ANDREA, "México IA+ / Inversión Acelerada: el evento que consolidará a México como hub regional de innovación y tecnología", en Periódico el Universal, de fecha 9 de septiembre de 2026 Por otra parte, el termino ingles hub, puede traducirse al español como: "...intercambiador, centro logístico o punto de conexión, concentrador o nodo, según los casos." Véase: fundéu RAE, hub, alternativas en español.

propia infraestructura y capacidades de inteligencia artificial, sin depender de tecnologías externas. ".<sup>66</sup>

Se puede asegurar entonces que dicho evento atrajo inversión estructural y posicionó al país en la ruta hacia un mercado de IA de alto impacto, y que dicho encuentro fue el inicio de una arquitectura nacional y regional de inteligencia artificial, que busca consolidar a México como líder en soberanía digital y desarrollo tecnológico estratégico hacia 2030.<sup>67</sup>

3.- Formación de Talento Especializado (2026): Dentro de las acciones emprendidas por el Gobierno de México para desarrollar la Inteligencia artificial en México, podemos destacar aquella llevada a cabo con fecha 26 de enero de 2026 por la Agencia de Transformación Digital y Telecomunicaciones (ATDT), a través de la cual se puso en marcha el Centro Público de Formación en Inteligencia Artificial, mismo que es considerado el más grande de Latinoamérica por su alcance Nacional, y su enfoque estratégico en el desarrollo de talento tecnológico; así las cosas, en la fecha antes mencionada 10 mil estudiantes iniciaron actividades para certificarse de manera gratuita, sin embargo el objetivo real es alcanzar la cifra de 25 mil personas al finalizar el año 2026 en áreas estratégicas de alta demanda laboral.<sup>68</sup>

La iniciativa pretende responder a la creciente demanda de talento en inteligencia artificial y habilidades digitales especializadas a nivel nacional e internacional, colocando al alcance de los ciudadanos mexicanos nuevas y mejores oportunidades laborales para las y los profesionales del sector tecnológico, en el entendido que dicha formación es gratuita y permite dobles certificaciones avaladas por instituciones públicas y privadas.<sup>69</sup>

A mayor abundamiento, el citado programa ofrece un modelo de entrenamiento acelerado sobre la Inteligencia Artificial, así como el análisis de datos, ciberseguridad, computación en la nube, programación en Java, y el desarrollo de habilidades psicosociales e inglés conversacional, todo lo anterior con el fin de llevar a cabo la transformación digital de México, y mejorar la empleabilidad de quienes deseen integrarse al mercado laboral, tecnológico, tanto en México como a nivel internacional.<sup>70</sup>

---

<sup>66</sup>Cfr. RIQUELME, Rodrigo, "México IA+ busca impulsar la soberanía tecnológica de México" en el Economista, de fecha 3 de octubre, México, 2025.

<sup>67</sup>Cfr. Ídem.

<sup>68</sup>Cfr. CIREL Alexa, "México pone en marcha el mayor Centro Público de Formación en Inteligencia Artificial de Latinoamérica", en diario en línea Infobae, de fecha 27 de enero de 2026.

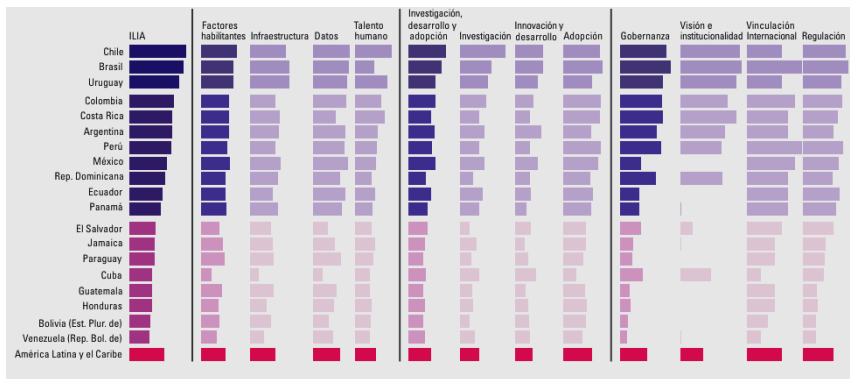
<sup>69</sup>Cfr. GOBIERNO DE MÉXICO, Convocatoria. Centro Público de Formación en Inteligencia Artificial, noviembre de 2025.

<sup>70</sup>CIREL Alexa, op. cit., nota 67.

En este sentido, el arranque de la primera generación se realizó de forma simultánea en 10 ciudades del país: Mérida, Tijuana, Morelos, Veracruz, Puebla, Morelia, Oaxaca, Tamaulipas, Nayarit y la Ciudad de México, consolidando un proyecto de alcance nacional, siendo imperante destacar que el programa se desarrolla en coordinación con la ATDT, INFOTEC, TecNM y la Asociación Mexicana de la Industria de Tecnologías de Información (AMITI), además de contar con el respaldo de empresas líderes del sector tecnológico, que facilitarán la vinculación laboral de los estudiantes, como por ejemplo Microsoft, Google, Amazon, IBM, entre otras.<sup>71</sup>

Finalmente, toda vez que se han abordado las principales acciones emprendidas por el gobierno de México, con el fin de alcanzar a corto y largo plazo el desarrollo y regulación de la Inteligencia Artificial en el País. A continuación, presentaremos a través de un gráfico proporcionado por la Comisión Económica para América Latina y el Caribe (CEPAL), los avances generales de la IA en Latinoamérica:

Gráfico 1.<sup>72</sup>



De lo antes expuesto, podemos apreciar que en el año 2025, México ocupó el octavo puesto con base en tres indicadores que son: 1.- Factores habitantes (infraestructura, datos y talento humano; 2.- Investigación, desarrollo y adopción (Investigación, innovación y desarrollo, y adopción); y 3.- Gobernanza (visión e institucionalidad, vinculación internacional, y regulación), en la inteligencia que Chile ocupa el primer puesto, Brasil el segundo, y Uruguay el tercero, resultando además que dentro de las fortalezas de México se encuentran los indicadores de infraestructura, datos, adopción, vinculación internacional, y regulación, más no así en talento humano, innovación y desarrollo, y visión e institucionalidad.<sup>73</sup>

Estas deficiencias actualmente son abordadas principalmente a través de las 3 acciones señaladas al inicio del presente apartado, ya que, a través de las mismas, el Estado mexicano se esfuerza por reposicionar al país e incluso convertirlo en un hub

<sup>71</sup> Cfr. Ídem.

<sup>72</sup> SOTO, Álvaro, DURÁN, Rodrigo, et al., (Coords.), Índice Latinoamericano del Centro Nacional de Inteligencia Artificial (ILIA), Ed. Publicación de Naciones Unidas, Chile, 2025, p.17.

<sup>73</sup> Cfr. Ídem.

regional de inteligencia artificial, con alianzas estratégicas entre gobierno, y empresas globales.<sup>74</sup>

## 6. CONCLUSIONES

Primero: La Inteligencia Artificial puede y debe ser una herramienta tecnológica accesible para toda la humanidad, a través de la cual se combatan las diferentes brechas en las que se desenvuelve el ser humano, tal y como lo estipula el artículo 27.1 de la Declaración Universal de los derechos humanos de 1948, sin embargo, tal ideal y derecho se encuentra amenazado por aquellos países que solo pretenden mantener su influencia y poder sobre el resto del mundo, por lo que la forma más eficaz de no ser excluido de los beneficios tecnológicos que implica el desarrollo y aplicación de la IA, es sumándonos a su campo de estudio, investigación, desarrollo y regulación de conformidad con los derechos públicos subjetivos, y de acuerdo con la realidad y contexto social, político y económico de México, adoptando una economía basada en el conocimiento como una medida ideal para diversificar e impulsar el crecimiento económico del país.

Segundo: De acuerdo con los datos de 2025 que arroja la Cepal, México presenta un rezago significativo en torno a la IA sobre los indicadores de talento humano, innovación y desarrollo, así como en visión e institucionalidad, que colocaron a México en el octavo lugar de los países Latinoamericanos, no obstante, el Gobierno Mexicano ha asumido el reto de atender cada uno de estos indicadores a través de una serie de acciones, dentro de las cuales podemos destacar La iniciativa de Creación de la Agenda Nacional de IA (2024-2030); El Desarrollo de IA Soberana a través de la Iniciativa «MÉXICO IA+ Inversión Acelerada (2025); y Formación de Talento Especializado (2026), entre otras acciones, es decir, se pretende reposicionar al país como un hub regional mediante la inversión en infraestructura, talento y gobernanza, no obstante es a través de la última acción señalada (Formación de Talento Especializado (2026), la que consideramos más significativa, dado que es a través de la Educación que podemos comenzar a construir las bases para lograr un verdadero desarrollo de la IA de acuerdo con nuestros intereses nacionales.

Tercero: Como se ha planteado a lo largo de la presente investigación, los derechos humanos son la pieza clave a través de la cual el Estado mexicano debe adoptar toda clase medidas, acciones y estrategias que garanticen el desarrollo y regulación ético y responsable de la IA, lo anterior en virtud, de que nuestro objeto de estudio está en constante evolución y desarrollo, pero además porque requerimos una serie de parámetros básicos y flexibles que por una parte no asfixien a nuestro tema de estudio, pero por otra parte tampoco se vulneren los derechos básicos de los seres humanos para la realización de cualquier plan de vida.

## 7. REFERENCIAS BIBLIOGRÁFICAS

Hemerográficas:

---

<sup>74</sup>Cfr. CONSEJO COORDINADOR EMPRESARIAL, Se presenta MÉXICO IA+ | Inversión Acelerada: Hub regional de inteligencia artificial soberana, inversión estructural y desarrollo con propósito, de fecha 29 de julio de 2025.

- ABELIUK, Andrés y GUTIÉRREZ Claudio, "Historia y Evolución de la Inteligencia Artificial", en Revista bits de ciencia del departamento de Ciencias de la computación de la facultad de Ciencias físicas y Matemáticas de la Universidad de Chile, Núm. 21, Chile, 2021.
- ARANA, Laura, "InfoCDMX presenta iniciativa para regular el uso de inteligencia artificial y protección de datos personales" en Periódico el Universal, de fecha 24 de octubre de 2024. [Disponible en: <https://shorturl.at/oexhA>].
- AYALA, Juan, "Actualidades de la inteligencia artificial en México", en el Heraldo de México, de fecha 14 de mayo de 2024. [Disponible en: <https://shorturl.at/wO8vS>].
- BERNARD, Marr, "Estos son los 15 mayores riesgos de la IA", en Revista Forbes, de fecha 28 de julio, España, 2023,. [Disponible en: <https://tinyurl.com/5azdzkde>].
- CANCELA, Outeda, Celso, "La carrera de la inteligencia artificial y el nuevo orden mundial", Revista CIDOB d'Afers Internacionals, núm. 131, Mayo-Septiembre, 2022, Barcelona, España. [Disponible en: <https://shorturl.at/UeR6Z>].
- CORONA, Nakamura, Luis Antonio y González Madrigal, Joseline Adriana, "La perspectiva ética y jurídica de la Inteligencia Artificial en México", en Revista Misión Jurídica, Vol. 16, Núm. 25, Julio - Diciembre de 2023, México. [Disponible en: <https://tinyurl.com/5ae2tybw>].
- DÍAZ ARMENDÁRIZ, Antonio, "Legislaciones internacionales que regulan la inteligencia artificial", en Reporte Cesop, Inteligencia Artificial, Número 152, Agosto, México, 2023. [Disponible en: <https://shorturl.at/Uc0tq>].
- LOMBANA BERMÚDEZ, Andrés, "La evolución de las brechas digitales y el auge de la Inteligencia Artificial (IA)" en Revista mexicana de bachillerato a distancia, Vol. 10, Núm. 20, Agosto, México, 2018. [Disponible en: <https://tinyurl.com/yn4kjakp>].
- O'BRIEN, MATT, "Trump revoca orden de Biden sobre seguridad en IA para diferenciarse de su antecesor" en Los Ángeles Times, de fecha 22 de enero de 2025. [Disponible en: <https://shorturl.at/XmShc>].
- RÍOS, Sebastián, "Los riesgos de la Inteligencia Artificial", en Revista Mensaje, Mayo, Chile, 2023. [Disponible en: <https://tinyurl.com/47prh927>].
- RIQUELME, "Rodrigo, La Ciudad de México se enfrenta al desafío de regular la Inteligencia Artificial", en Diario el Economista, de fecha 9 de noviembre de 2024. [Disponible en: <https://shorturl.at/hblBg>].
- VEGA, Paloma, "Sin Protecciones en Ciberseguridad, la IA es una Apuesta que no Podemos Permitirnos" en Diario el universal, de fecha 20 de diciembre, México, 2024. [Disponible en: <https://tinyurl.com/hhr3v9dz>].

#### Legisgráficas:

- Iniciativa con proyecto de decreto por el que se reforma la fracción XVII al artículo 73 de la Constitución Política de los Estados Unidos Mexicanos, de fecha 5 de septiembre de 2023, presentada en la Ciudad de México, en la Gaceta Parlamentaria, de fecha 5 de septiembre de 2023. [Disponible en: <https://shorturl.at/9H3mW>].
- Iniciativa con Proyecto de Decreto por el que se Reforma la Fracción XVII, al artículo 73 de la Constitución Política de los Estados Unidos Mexicanos, en Materia de Inteligencia Artificial, Ciberseguridad y Neuroderechos, de fecha 20 de septiembre de 2023, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 20 de septiembre de 2023. [Disponible en: <https://shorturl.at/Lfrei>].
- Iniciativa del Congreso de Jalisco, con Proyecto de Decreto por el que se Reforma y Adiciona el Artículo 30 de la Ley General de Educación, de fecha 22 de febrero de 2021, Presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 18 de marzo de 2021. [Disponible en: <https://shorturl.at/4x9Vv>].
- Iniciativa que adiciona diversas disposiciones de la Ley General en Materia de Humanidades, Ciencias, Tecnologías e Innovación, en materia de ética y derechos humanos, de fecha 17 de enero de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 17 de enero de 2024, [Disponible en: <https://shorturl.at/NVfvR>].

- Iniciativa que adiciona diversas disposiciones de la Ley General en materia de Humanidades, Ciencias, Tecnologías e Innovación, en materia de lineamientos de políticas públicas para el fomento y promoción del uso de inteligencia artificial, de fecha 14 de febrero de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 14 de febrero de 2024. [Disponible en: <https://shorturl.at/OjIKv>].
- Iniciativa que adiciona una fracción IV Bis al artículo 30 y adiciona un artículo 86 de la Ley General de Educación de fecha 13 de febrero de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 28 de febrero de 2024. [Disponible en: <https://shorturl.at/hGP0v>].
- Iniciativa que adiciona una nueva fracción VIII Bis al artículo 11 de la Ley General en Materia de Humanidades, Ciencias, Tecnologías e Innovación de fecha 5 de marzo de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 6 de marzo de 2024. [Disponible en: <https://shorturl.at/vZCDK>].
- Iniciativa que expide la Ley que crea la Universidad de Tecnologías de la Información, Comunicaciones e Innovación, de fecha 12 de agosto de 2020, presentada en la Ciudad de México, en la Gaceta Parlamentaria, de fecha 12 de agosto de 2020. [Disponible en: <https://shorturl.at/uQbK8>].
- Iniciativa que reforma diversas disposiciones de la Ley General de Educación y de la Ley General de Educación Superior, en materia de Inteligencia Artificial, de fecha 5 de diciembre de 2023, presentada en la Ciudad de México, en la Gaceta Parlamentaria, de fecha 5 de diciembre de 2023. [Disponible en: <https://shorturl.at/zGRQN>].
- Iniciativa que reforma la fracción IV del artículo 30 de la Ley General de Educación en Materia de Inteligencia Artificial como parte de los planes y Programas de Estudio, de fecha 22 de septiembre de 2023, presentada en la Ciudad de México, en la Gaceta parlamentaria, de fecha 26 de septiembre de 2023. [Disponible en: <https://shorturl.at/z5jqV>].
- Iniciativa que Reforma los Artículos 30 Y 70 de la Ley General de Educación y 7° de la Ley General de Educación Superior, de fecha 4 de octubre de 2023, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 4 de octubre de 2023. [Disponible en: <https://shorturl.at/mew3H>].
- Iniciativa que Reforma y Adiciona Diversas Disposiciones de la Ley de Ciencia y Tecnología y de la Ley General de Salud, de fecha 28 de abril de 2020, presentada en la Ciudad de México, en la Gaceta Parlamentaria, de fecha 28 de abril de 2020. [Disponible en: <https://shorturl.at/m0Gt2>].
- Iniciativa que Reforma y Adiciona Diversas Disposiciones de la Ley de Ciencia y Tecnología, de fecha 7 de enero de 2021, Presentada en la Ciudad de México, en la Gaceta Parlamentaria, de fecha 7 de enero de 2021. [Disponible en: <https://shorturl.at/hhZD7>].
- Iniciativa que reforma y adiciona diversas disposiciones de la Ley General en Materia de Humanidades, Ciencias, Tecnologías e Innovación, de fecha 14 de febrero de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 14 de febrero de 2024. [Disponible en: <https://shorturl.at/7Ehtc>].
- Iniciativa que reforma y adiciona diversas disposiciones de la Ley General en Materia de Humanidades, Ciencias, Tecnologías e Innovación, Relativa a la Inteligencia Artificial de fecha 1 de abril de 2024, presentada en la Ciudad de México, en la Gaceta Parlamentaria de fecha 2 de marzo de 2024. [Disponible en: <https://shorturl.at/FzNFX>].
- Mesográficas:  
Gaceta Parlamentaria, de fecha 5 de marzo de 2024. [Disponible en: <https://shorturl.at/LyCwB>].  
GARCIA, Laura, Ciencia Unam, Inteligencia Artificial, la herencia de Alan Turing. [Disponible en: <https://shorturl.at/aNIFx>].
- Grupo Ático 34, ¿Qué derechos puede violar la inteligencia artificial? [Disponible en: <https://tinyurl.com/57b298kh>].
- ISO, ¿Qué es la inteligencia artificial? [Disponible en: <https://shorturl.at/sedDu>], página consultada el 9 de diciembre de 2024.

- MEZA RUIZ, Ivan Vladimir, Meza Ruiz, El camino hacia una regulación de la IA en México, de fecha 7 de Julio de 2024. [Disponible en: <https://shorturl.at/ljErD>].
- Real Academia Española, Inteligencia Artificial. [Disponible en: <https://shorturl.at/xfFJk>].
- Temas, Parlamento Europeo, ¿Qué es la inteligencia artificial y cómo se usa? [Disponible en: <https://shorturl.at/jXhsT>]
- CNDH, ¿Qué son los Derechos Humanos? [Consultado en: 3 de abril de 2026]. [Disponible en: <https://shorturl.at/wcF0g>].
- COMISIÓN EUROPEA, Libro Blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza, Bruselas, 19 de febrero de 2020, p. 2. [Consultado en: 12 de febrero de 2026]. [Disponible en: <https://shorturl.at/qqFjy>].
- FERRAJOLI, Luigi, Derechos y garantías. La ley del más débil, 4.ª ed., Madrid, Trotta, 2004.
- GUERRERO MARTÍNEZ, Rodolfo, "Derechos humanos de cuarta generación y las tecnologías de la información y de la comunicación", Derechos Fundamentales a Debate, núm. 12, enero-abril de 2020, Guadalajara, Jalisco, p. 137. [Consultado en: 4 de abril de 2026]. [Disponible en: <https://historico.cedhj.org.mx/revista%20DF%20Debate/revista%20pdf/ADEBATE%2012-2020.pdf>].
- IBM, "Comprender los diferentes tipos de inteligencia artificial". [Consultado en: 1 de abril de 2026]. [Disponible en: <https://shorturl.at/lOsT0>].
- IBM, "¿Qué es la interpretabilidad de la IA?". [Consultado en: 14 de marzo de 2026]. [Disponible en: <https://www.ibm.com/es-es/think/topics/interpretability>].
- MOLINARES HASSAN, Viridiana, Notas sobre Constitución, Organización del Estado y Derechos Humanos, 2.ª ed., Barranquilla, Colombia, Universidad del Norte, 2011.
- UNESCO, Recomendación sobre la Ética de la Inteligencia Artificial, 23 de noviembre de 2021. [Consultado en: 10 de marzo de 2026]. [Disponible en: <https://shorturl.at/cGhU6>].
- GOBIERNO DE MÉXICO, "¿Sabías que en México hay 68 lenguas indígenas, además del español?". [Consultado en: 27 de marzo de 2026]. [Disponible en: <https://shorturl.at/54lpe>].
- GOBIERNO DE ESPAÑA, "¿Cómo se construye un modelo de inteligencia artificial?". [Consultado en: 2 de mayo de 2026]. [Disponible en: <https://shorturl.at/GkXRd>].
- SENADO DE LA REPÚBLICA, LXVI Legislatura, "Presentan en el Senado Agenda Nacional de la Inteligencia Artificial para México 2024-2030", 15 de mayo de 2024. [Consultado en: 8 de abril de 2026]. [Disponible en: <https://comunicacionsocial.senado.gob.mx/informacion/comunicados/8982-presentan-en-el-senado-agenda-nacional-de-la-inteligencia-artificial-para-mexico-2024-2030>].
- CEPAL y CENTRO NACIONAL DE INTELIGENCIA ARTIFICIAL, Índice Latinoamericano de Inteligencia Artificial (ILIA) 2025, Documentos de Proyectos, LC/TS.2025/68, 2025.
- "Cabe mencionar que la mayoría de los grupos y centros de investigación de IA mexicanos cuentan con activa participación tanto en proyectos nacionales como internacionales, en colaboración cercana tanto con la industria privada como el gobierno para contribuir al avance de la IA en el país."<sup>75</sup>

---

<sup>75</sup>Cfr. BAÑOS-GONZÁLEZ, op. cit., nota 53, p.258.



# Los derechos humanos de la nación mexicana: Rompiendo paradigmas teóricos para el pleno ejercicio de derechos mediante la inteligencia artificial\*

*The human rights of the mexican nation: Breaking theoretical paradigms for the full exercise of rights through artificial intelligence*

David Espinosa – González\*\* Alejandra Cedallin Martínez – Sánchez\*\*\* Gabriel Pérez – Galmiche\*\*\*\*

## RESUMEN

La teoría de los derechos humanos ha evolucionado constantemente ante las necesidades de la población, como la búsqueda de reconocimiento de los Derechos Humanos como "nación mexicana", y, considerar a los avances tecnológicos, como la IA, para contribuir con el ejercicio de derechos de primera y segunda generación. La presente investigación documental, emplea el método analítico-sintético para lograr cumplir el objetivo general "contribuir a la teoría de los derechos humanos mediante el desarrollo de un nuevo paradigma 'los derechos humanos de una Nación' y vincularlo con la IA como medio para contribuir al ejercicio pleno de los derechos humanos de primera y segunda generación". Como principales resultados y aportaciones, se ofrece el desarrollo teórico de la propuesta de reconocimiento de los derechos humanos de la nación mexicana y una propuesta para el ejercicio y goce de derechos humanos de primera y segunda generación.

## PALABRAS CLAVE

derechos humanos; inteligencia artificial (IA); Nación; persona moral.

## ABSTRACT

The theory of Human Rights has constantly evolved in response to the needs of the population, such as the search for the recognition of Human Rights of the "Mexican Nation," and the consideration of technological advances, such as Artificial Intelligence (AI), as tools to contribute to the exercise of first- and second-generation rights. This documentary research employs the analytic-synthetic method in order to fulfill the general objective: "To contribute to the theory of Human Rights through the development of a new paradigm, 'the Human Rights of a Nation,' and to link it with AI as a means of contributing to the full exercise of first- and second-generation Human Rights." As its main results and contributions, the study offers the theoretical development of the proposal for the recognition of the Human Rights of the Mexican Nation, as well as a proposal aimed at facilitating the exercise and enjoyment of first- and second-generation Human Rights.

## KEYWORDS

Human Rights; Artificial Intelligence (AI); Nation; Legal Entity.

\* Artículo de Investigación

\*\* Doctor en Derecho y Educación, maestro en derecho, y abogado, notario y actuario por la Benemérita Universidad Autónoma de Puebla. (david.espinosa@correo.buap.mx). <https://orcid.org/0000-0002-6993-3349>.

\*\*\* Doctora en Educación y Profesora investigadora en la Benemérita Universidad Autónoma de Puebla (alejandra.martinezsa@correo.buap.mx). <https://orcid.org/0000-0002-7798-7670>.

\*\*\*\* Doctor en derecho y profesor investigador de la Benemérita Universidad Autónoma de Puebla (gabriel.galmiche@correo.buap.mx). <https://orcid.org/0000-0003-0843-1887>.

## SUMARIO

1. Introducción
2. Prolegómenos de la investigación: De los derechos humanos: una perspectiva innovadora
3. De lo teórico a lo jurídico
4. De la importancia de la IA para el ejercicio y desarrollo de los derechos humanos de la Nación
5. Conclusiones
6. Referencias bibliográficas

### 1. INTRODUCCIÓN

Los derechos humanos han sido a lo largo de la historia del ser humano un reto para los dirigentes de territorios para lograr su reconocimiento, respeto y establecer los medios óptimos para su pleno goce y ejercicio, en el mejor de los casos dentro de un plano ético, político y gubernamental.

Por lo que, desde el plano de los integrantes del pueblo, la nación, es decir, los ciudadanos y residentes de determinado país, han buscado desde diversas perspectivas buscar el respeto y garantía de ejercicio de los derechos humanos, emprendiendo diversos desarrollos teóricos, no únicamente jurídicos, para establecer medios efectivos para encontrar certeza en el ejercicio de las prerrogativas y gozar de estas como los textos constitucionales lo señalan.

Continuando con este orden de ideas, la presente investigación busca a través de una estructura de tres partes contribuir a la solución de las diversas problemáticas de violación de los derechos humanos en México. En una primera parte, se ofrece una propuesta teórica jurídica para considerar la existencia de los derechos humanos de la nación mexicana como una persona moral, como estrategia. Posteriormente, se realiza el desarrollo jurídico del argumento principal, denominado: De lo teórico a lo jurídico.

Y, en tercer punto, se considera a la IA como medio viable para buscar el desarrollo y ejercicios de los derechos civiles, Políticos, Económicos, Sociales y Culturales (Derechos de Primera y segunda generación).

De igual forma, el artículo busca brindar respuesta a la pregunta de investigación ¿Cómo la reconsideración de la teoría de los derechos humanos, al contemplar a la Nación como titular de estas prerrogativas podría contribuir a buscar una mayor protección apoyándose en el desarrollo tecnológico como la IA para el ejercicio pleno y goce de derechos (de primera y segunda generación)? Considerando que la reconfiguración de la teoría clásica sobre los titulares de estos derechos abre un mayor campo de opciones para brindar soluciones a las problemáticas de violaciones ante las acciones y omisiones de la autoridad, en este caso, mexicana. Estableciendo como medio oportuno para algunos derechos humanos, el empleo de la IA como medio para entender, comprender y aprender información que contribuya al goce de derechos de primera y segunda generación.

El lector tiene en sus manos una propuesta teórica y debatible sobre una aportación teórica en materia de derechos humanos y la implementación de IA, para el ejercicio

y goce de derechos de primera y segunda generación, sin duda alguna, será pieza de consulta, discusión y debate.

## **2. PROLEGÓMENOS DE LA INVESTIGACIÓN: DE LOS DERECHOS HUMANOS: UNA PERSPECTIVA INNOVADORA**

La teoría clásica de los derechos humanos atiende a la necesidad de abordar un tema inacabado de constantes aportaciones de diversas plumas y tinteros, que buscan abonar al compromiso de garantía de ejercicio y disfrute de estas prerrogativas humanas.

Desde los anales de la historia se han identificado personajes que han buscado contribuir a la protección de la naturaleza humana e, incluso, aquellos valores y necesidades que la propia filosofía del derecho los ha denominado "principios infrahumanos".

En este orden de ideas, la corriente ideológica que ubica a los derechos humanos como las prerrogativas propias y únicamente reconocidas y dirigidas al Ser humano, se ha contemplado en diversos momentos por los autores del presente artículo, considerados como: "principios humanos con fundamento internacional ubicados en diversos instrumentos u objetos, que representan la responsabilidad de protección de la existencia del ser y sus necesidades básicas de existencia. Adecuados al tiempo y espacio"<sup>1</sup>, está siendo ubicada una vez más en la mesa de discusión ante la búsqueda de protección en el mayor sentido.

Por lo que es indispensable, en pro de ejercitar y exigir a "Todas las autoridades, en el ámbito de sus competencias, tienen la obligación de promover, respetar, proteger y garantizar los derechos humanos"<sup>2</sup>, conocer estrategias con mayor peso ante la existencia de violaciones de estas prerrogativas humanas a grupos o entidades como la Nación.

Exponiendo que la Nación es una persona moral reconocida por el Código Civil Federal y que, con base en los instrumentos jurisprudenciales, votos razonados y posiciones – recomendaciones de las autoridades judiciales, caso concreto, la Suprema Corte de Justicia de la Nación, es posible establecer que la nación mexicana como persona moral cuenta con derechos humanos exigibles ante la autoridad.

Bajo este razonamiento, es pertinente definir los elementos que conformarán la base argumentativa y teórica de esta investigación. Principalmente, el término "nación" tiene su origen en el latín "natio – ōnis" que refiere al "lugar de nacimiento" e incluso "pueblo" o "tribu"<sup>3</sup> considerado, en una primera acepción como el "conjunto de los habitantes de un país regido por el mismo Gobierno"<sup>4</sup> y, en segundo plano, como el "conjunto de personas de un mismo origen y que generalmente hablan un mismo idioma y tienen una tradición común"<sup>5</sup>, considerando este último, más descriptivo y viable para el término.

---

<sup>1</sup>PÉREZ-GALMICHE, Gabriel; ESPINOSA-GONZÁLEZ, David, *La crisis de los derechos humanos: estudio comparativo México-Italia*, Ciudad de México, Montiel Et Soriano Editores y BUAP, 2021, p. 257.

<sup>2</sup>MÉXICO, Constitución Política de los Estados Unidos Mexicanos, art. 1.º, Cámara de Diputados del H. Congreso de la Unión, p. 1.

<sup>3</sup>REAL ACADEMIA ESPAÑOLA, *Diccionario de la lengua española*, voz «nación», 2024.

<sup>4</sup>Ibidem.

<sup>5</sup>Ibidem.

Respecto a la conceptualización de "nación", se considera pertinente retomar a Benedict Anderson con su fundamento de nación, el cual, se identifica como "una comunidad política imaginada como inherentemente limitada y soberana"<sup>6</sup>.

Para Max Weber, la figura de "Nación" posee un enfoque sociológico dirigido al poder y prestigio: "La nación es una comunidad de sentimiento que se manifiesta adecuadamente en un Estado propio"<sup>7</sup> no obstante, se requiere atribuirle voluntad colectiva que se origina de la individualidad de sus integrantes, es decir y como lo sustenta Ernest Renan "una nación es un plebiscito cotidiano"<sup>8</sup>, en otras palabras, día a día se refuerza su cohesión.

Desde otro punto de vista, más contemporáneo, Smith comenta que "las naciones modernas tienen raíces en comunidades étnicas preexistentes"<sup>9</sup>, por lo que es posible identificar una característica de la nación, es decir, los rasgos étnicos preexistentes.

Continuando, el concepto de nación posee algunas características únicas que son abordadas por diversos autores, sin embargo, una de las perspectivas viables para el presente estudio es la postura de Renan, al expresar que se requieren dos elementos esenciales para que exista una nación ya que se basa en un proceso de memoria compartida y la intención y voluntad de convivir vinculados, sin considerar prioritario la raza y / o el idioma o lengua.<sup>10</sup>

Por su parte, Anthony Smith estructura la base de la Nación en mitos, símbolos, territorio y memoria histórica.<sup>11</sup>

Como es posible observar el nutrido término de Nación es amplio y complejo que es fundamento de la presente investigación, no obstante, es preciso definir dos elementos más, el primero el de derechos humanos que, siendo un elemento internacionalmente abordado, se ofrecerán algunos puntos concretos.

En primer acercamiento, las Naciones Unidas los consideran como:

Los derechos humanos son derechos inherentes a todos los seres humanos, sin distinción alguna de raza, sexo, nacionalidad, origen étnico, lengua, religión o cualquier otra condición. Entre los derechos humanos se incluyen el derecho a la vida y a la libertad; a no estar sometido ni a esclavitud ni a torturas; a la libertad de opinión y de expresión; a la educación y al trabajo, entre otros muchos. Estos derechos corresponden a todas las personas, sin discriminación alguna.<sup>12</sup>

Continuando con la Comisión Nacional de derechos humanos:

Los derechos humanos son el conjunto de prerrogativas sustentadas en la dignidad humana, cuya realización efectiva resulta indispensable para el desarrollo integral de la persona. Este conjunto de prerrogativas se encuentra establecido dentro del orden jurídico nacional, en nuestra Constitución Política, tratados internacionales y las leyes.<sup>13</sup>

<sup>6</sup>ANDERSON, Benedict, Comunidades imaginadas: reflexiones sobre el origen y la difusión del nacionalismo, México, Fondo de Cultura Económica, 1993, p. 23.

<sup>7</sup>WEBER, Max, Economy and Society, California, University of California Press, 1978, p. 922.

<sup>8</sup>RENAN, Ernest, ¿Qué es una nación?, Madrid, Alianza, 2010, p. 53.

<sup>9</sup>SMITH, Anthony, The Ethnic Origins of Nations, Oxford y Nueva York, Blackwell, 1986, p. 57.

<sup>10</sup>Véase RENAN, Ernest, op. cit., nota 8, pp. 53–55.

<sup>11</sup>Véase SMITH, Anthony, op. cit., nota 9, pp. 14–16 y 66–75.

<sup>12</sup>ORGANIZACIÓN DE LAS NACIONES UNIDAS, «Desafíos globales. Derechos humanos: ¿qué son los derechos humanos?», 2025.

<sup>13</sup>COMISIÓN NACIONAL DE LOS DERECHOS HUMANOS, «Derechos humanos: ¿qué son los derechos humanos?».

Concluyendo que los derechos humanos "Es el vínculo existente entre la vida humana - existencia- y el poder soberano, sujetos ante la intención y obligación de la regulación para su tutela mas no reconocimiento"<sup>14</sup>.

De igual forma, la figura de "persona moral" es abordada por diversos diccionarios y enciclopedias, por lo que es posible retomarla como sinónimo de persona jurídica como lo considera el Diccionario Panhispánico del español jurídico que, continuando con esa sinonimia, se expone que puede definirse como "Institución dotada de personalidad propia e independiente y plena capacidad para el cumplimiento de sus fines, creada por las leyes o conforme a lo establecido en las mismas"<sup>15</sup>. Finalmente, es indispensable abordar la figura de persona moral desde la perspectiva del Quinto Tribunal Colegiado del Primer Circuito Judicial de los denominados Tribunales de Circuito que a la letra expresa:

Las personas morales son entes jurídicos que, en su origen, fueron creados y reconocidos en el derecho con el propósito de regular y fomentar actividades útiles a sus integrantes y a la sociedad, esto es, se crearon y regularon para fines lícitos.<sup>16</sup>

Como es posible identificar esta persona dotada por diversas facultades y características, entre ellas la capacidad y personalidad jurídica, son pieza fundamental del desenvolvimiento jurídico, económico, comercial e incluso político en México.

### **3. DE LO TEÓRICO A LO JURÍDICO**

La investigación se orienta, en un primer plano, a conocer los derechos humanos atribuidos a la nación mexicana, por lo que es prioritario presentar un desarrollo jurídico, partiendo desde la Constitución Política de los Estados Unidos Mexicanos para culminar en la jurisprudencia, ya que serán los instrumentos de mayor valía en el presente artículo científico.

En un primer sentido, la Constitución Política de los Estados Unidos Mexicanos, es considerada en el artículo 2.º constitucional donde caracteriza a la nación mexicana de manera amplia desde la reforma constitucional de 30 de septiembre de 2024:

Artículo 2o. La nación mexicana es única e indivisible, basada en la grandeza de sus pueblos y culturas.

La Nación tiene una composición pluricultural y multiétnica sustentada originalmente en sus pueblos indígenas, que son aquellas colectividades con una continuidad histórica de las sociedades precoloniales establecidas en el territorio nacional; y que conservan, desarrollan y transmiten sus instituciones sociales, normativas, económicas, culturales y políticas, o parte de ellas.<sup>17</sup>

De igual forma, un elemento esencial es la soberanía que emana y reside en el Pueblo, que en sentido amplio es la nación mexicana, por lo que el artículo 39 constitucional, y presentando el espíritu de la formulación original de 1917, a la letra expone:

<sup>14</sup>PÉREZ-GALMICHE, Gabriel; ESPINOSA-GONZÁLEZ, David, op. cit., nota 1, p. 257.

<sup>15</sup>REAL ACADEMIA ESPAÑOLA, Diccionario panhispánico del español jurídico, voces «persona moral» y «persona jurídica».

<sup>16</sup>MÉXICO, tesis I.5o.C.64 C (10a.), «Personas morales. Su protección legal depende de la buena fe con la que se conduzcan», Semanario Judicial de la Federación y su Gaceta, libro XXIII, t. 3, agosto de 2013, p. 1693.

<sup>17</sup>MÉXICO, Constitución Política de los Estados Unidos Mexicanos, art. 2.º, Cámara de Diputados del H. Congreso de la Unión, versión citada de 3 de marzo de 2026, p. 2.

Artículo 39. La soberanía nacional reside esencial y originariamente en el pueblo. Todo poder público dimana del pueblo y se instituye para beneficio de éste. El pueblo tiene en todo tiempo el inalienable derecho de alterar o modificar la forma de su gobierno.<sup>18</sup>

Identificando dos elementos esenciales, primero, el sujeto, si se puede determinar de esta forma, el origen y residencia de la soberanía: el pueblo. Y, en segundo punto, se percibe un vínculo entre el Pueblo y el Poder Público, ya que se establece que el pueblo requiere de los Poderes de la Unión para ejercer su soberanía:

Artículo 41. El pueblo ejerce su soberanía por medio de los Poderes de la Unión, en los casos de la competencia de éstos, y por los de los Estados y la Ciudad de México, en lo que toca a sus regímenes interiores, en los términos respectivamente establecidos por la presente Constitución Federal y las particulares de cada Estado y de la Ciudad de México, las que en ningún caso podrán contravenir las estipulaciones del Pacto Federal.<sup>19</sup>

Identificando un nexo legislativo constitucional entre la nación, el pueblo, la soberanía y ejercicio del Poder emanado de la nación.

Continuando con el hilo de lo razonado, la Nación, es considerada por el Código Civil Federal como una persona moral, como a la letra lo expone el instrumento civil en su artículo 25: "son personas morales: la Nación, los Estados y los Municipios"<sup>20</sup>.

Por lo que, legalmente, "las personas morales pueden ejercitar todos los derechos que sean necesarios para realizar el objeto de su institución"<sup>21</sup>, en este caso, el ejercicio de sus derechos y facultades para que sus integrantes puedan contar con un desarrollo y ejercicio pleno de sus derechos civiles, políticos, económicos, sociales, culturales e, incluso, ambientales.

De igual forma, se debe tener presente que las "personas morales obran y se obligan por medio de los órganos que las representan sea por disposición de la ley o conforme a las disposiciones relativas de sus escrituras constitutivas y de sus estatutos"<sup>22</sup>, en otras palabras, la Nación se hará representar por los poderes públicos, sin embargo, cabe matizar esta postura lógica jurídica, ya que con base en el párrafo 2.º del artículo 2.º constitucional, se tendría que visualizar otras figuras de representación de la Nación, no solo encasillar su vinculación con el Estado.

En otras palabras, si se considera a la Nación como un sujeto histórico y eminentemente social, se tendría que establecer que el Estado únicamente existe para generar una posible organización política que permea a otros ámbitos, no obstante, no abarcaría el todo de la Nación.

Este orden de ideas, orilla a considerar primordial la existencia del Derecho Consuetudinario (usos y costumbres), donde las poblaciones indígenas no son representadas por el Estado e, incluso, no lo reconocen como autoridad organizativa. Por lo que establecen y cimentan su propio ordenamiento social, político e incluso normativo. Por lo que sus representantes, obedecen a un aspecto tradicional.

---

<sup>18</sup>Ibidem, p. 51.

<sup>19</sup>Ibidem, p. 52.

<sup>20</sup>MÉXICO, Código Civil Federal, Cámara de Diputados, versión citada de 14 de noviembre de 2025, p. 4.

<sup>21</sup>Ibidem, art. 26, p. 4.

<sup>22</sup>Ibidem, art. 27, p. 4.

Este posicionamiento es interesante, ya que si se establece un nexo directo con la jurisprudencia con rubro Pueblos y barrios originarios de la Ciudad de México. Se asimilan al concepto de pueblos tribales establecido en el Convenio 169 sobre pueblos indígenas y tribales en países independientes, de la Organización Internacional del Trabajo (OIT), para efecto de cumplir con el deber de consulta previa, se reconoce, incluso, con base al instrumento internacional mencionado, los pueblos y barrios originarios de esta ciudad:

La OIT ha determinado una serie de elementos subjetivos para considerar a un grupo como tribal: a) estilos tradicionales de vida; b) una cultura, organización social, condiciones económicas y formas de vida distintas a las de otros segmentos de la población nacional; y c) tradiciones y costumbres propias, leyes tradicionales propias y/o un reconocimiento jurídico especial. Los pueblos y barrios originarios de la Ciudad de México cumplen estas características, toda vez que se les reconoce como integrantes de esa entidad federativa desde tiempos de la colonización y son colectivos que tienen estilos tradicionales de vida, una cultura y organización social y económica propia que los distingue de otros segmentos de la población nacional y local y tienen sistemas normativos propios reconocidos por la Ciudad de México.<sup>23</sup>

Ergo, aplicando el método analógico, esta tesis se puede aplicar en sentido amplio a todos los barrios y pueblos de las Ciudades que integran a la República Mexicana.

Como es posible observar, no es un tema sencillo la aseveración de que la Nación, como persona moral, posee derechos humanos y que la empodera mediante sus representantes para hacer exigir los medios suficientes para el ejercicio de sus derechos empleando como instrumento a la IA, no obstante, es un derecho que poseen los integrantes de la nación y, no únicamente el Estado está obligado a emprender estrategias suficientes para el acceso y accesibilidad de este tipo de adelanto tecnológico.

En un siguiente instrumento jurídico y fuente del derecho, se ubica a la jurisprudencia que, sin introducirse en un debate de la actuación de los actuales miembros de la Suprema Corte de Justicia de la Nación o de los Tribunales colegiados, se expone y retoma una tesis con panorama para identificar el reconocimiento de los derechos humanos de la nación mexicana y lograr acceder a la segunda parte de la investigación.

La principal tesis jurisprudencial, es la titulada Personas morales o jurídicas. Deben gozar no sólo de los derechos fundamentales constituidos por los derechos humanos reconocidos en la constitución y en los tratados internacionales, y de las garantías para su protección, siempre y cuando estén encaminados a proteger su objeto social, sino también de aquellos que aparezcan como medio o instrumento necesario para la consecución de la finalidad que persiguen, que en su primera parte, aborda la justificación sobre la protección de las personas morales, formulando un argumento apegado al artículo 1.º constitucional, donde se establece que "deben de gozar" de los derechos fundamentales constituidos por los derechos humanos:

Las personas morales o jurídicas son sujetos protegidos por el artículo 1o. de la Constitución Política de los Estados Unidos Mexicanos, por lo que deben gozar de los derechos fundamentales constituidos por los derechos humanos reconocidos en la propia Constitución y en los

<sup>23</sup>MÉXICO, tesis I.19o.A.1 A (11a.), "Pueblos y barrios originarios de la Ciudad de México. Se asimilan al concepto de pueblos tribales establecido en el Convenio 169 sobre pueblos indígenas y tribales en países independientes, de la Organización Internacional del Trabajo (OIT), para efecto de cumplir con el deber de consulta previa", México, Gaceta del Semanario Judicial de la Federación, libro 47, tomo I, volumen 1, marzo de 2025, p. 981.

tratados internacionales de los que el Estado Mexicano sea parte, así como de las garantías para su protección, siempre y cuando sean acordes con la finalidad que persiguen, por estar encaminados a la protección de su objeto social, así como de aquellos que aparezcan como medio o instrumento necesario para la consecución de la finalidad referida.<sup>24</sup>

En la segunda parte, se observa un análisis lingüístico jurídico sobre el término "persona" a la luz del mismo artículo constitucional:

Lo anterior es así, porque en la palabra «personas», para efectos del artículo indicado, no sólo se incluye a la persona física, o ser humano, sino también a la moral o jurídica, quien es la organización creada a partir de la agrupación voluntaria de una pluralidad de personas físicas, con una finalidad común y una identidad propia y diferenciada que trasciende la de los individuos que la integran, dotada de órganos que expresan su voluntad independiente de la de sus miembros y de un patrimonio propio, separado del de sus integrantes, a la que el ordenamiento jurídico atribuye personalidad y, consecuentemente, reconoce capacidad para actuar en el tráfico jurídico, como sujeto independiente de derechos y obligaciones, acorde al título segundo del libro primero del Código Civil Federal, al artículo 9o. de la Carta Magna y conforme a la interpretación de protección más amplia que, en materia de derechos humanos se autoriza en el párrafo segundo del artículo 1o. constitucional.<sup>25</sup>

Finalmente, ubica claramente la importancia del Derecho Interno y el interés de buscar una protección y ejercicio pleno de los derechos humanos reconocidos por la naturaleza de las personas morales:

Sin que sea obstáculo que los derechos fundamentales, en el sistema interamericano de derechos humanos, sean de los seres humanos, pues tal sistema no sustituye a las jurisdicciones nacionales, sino que otorga una protección coadyuvante o complementaria de la que ofrece el derecho interno de los Estados Americanos, por lo que una vez arraigados los derechos humanos en el derecho constitucional propio y singular del Estado Mexicano, éstos se han constituido en fundamentales, y su goce, así como el de las garantías para su protección, ha sido establecido por el propio derecho constitucional a favor de las personas y no sólo del ser humano.<sup>26</sup>

Con este punto jurisprudencial se da por concluida la exposición y argumentación sobre la propuesta de reconocimiento de los derechos humanos de la nación mexicana y, consecutivamente, se expondrá el importante vínculo de la inteligencia artificial para el ejercicio pleno de las prerrogativas humanas de carácter civil, político, económico, cultural y social.

#### **4. DE LA IMPORTANCIA DE LA IA PARA EL EJERCICIO Y DESARROLLO DE LOS DERECHOS HUMANOS DE LA NACIÓN**

Posterior a la Segunda Guerra Mundial, la creación de las Naciones Unidas, los juicios de Nuremberg y la celebración de la Declaración Universal de derechos humanos, se brindó especial importancia a los Derechos en materia Civil y política y, por otra parte, a los derechos económicos, sociales y culturales.

---

<sup>24</sup>MÉXICO, tesis XXVI.5o. (V Región) 2 K (10a.), "Personas morales o jurídicas. Deben gozar no sólo de los derechos fundamentales constituidos por los derechos humanos reconocidos en la constitución y en los tratados internacionales, y de las garantías para su protección, siempre y cuando estén encaminados a proteger su objeto social, sino también de aquellos que aparezcan como medio o instrumento necesario para la consecución de la finalidad que persiguen", México, Semanario Judicial de la Federación y su Gaceta, libro XI, tomo 2, agosto de 2012, p. 1876.

<sup>25</sup>Ibidem.

<sup>26</sup>Ibidem.

Para su mejor protección y compromiso, se adoptaron los pactos internacionales en materia de derechos civiles y Políticos (2200 A [XXI] – Parte A), Derechos Económicos, sociales y culturales (2200 A [XXI] – Parte B) y el Primer Protocolo Facultativo del Pacto de derechos civiles y Políticos (2200 A [XXI] – Parte C), durante la Asamblea General de las Naciones Unidas el 16 de diciembre de 1966 (Nueva York, EE. UU.) con resoluciones distintas; entrando en vigor el 23 de marzo de 1976.

De igual forma, existió una propuesta académica para el estudio de los derechos humanos, emitida por el Dr. Karel Vašák que buscaba establecer un estudio a partir del establecimiento de las diversas etapas en las que fueron reconocidos estos derechos, denominándolas como Generaciones.

En donde los derechos civiles y Políticos serían considerados como de primera generación, y los Económicos, Sociales y Culturales, se ubicaron dentro de la segunda generación. Considerándose viable esta clasificación aceptada para el estudio cronológico, más no como una forma de división.

Si bien es cierto que, a lo largo de la historia de los derechos humanos, se trataron de manera separada, la realidad es que en la actualidad y bajo el principio de indivisibilidad e interdependencia, se han vinculado. Justamente lo expresan las propias Naciones Unidas ante la duda "Conceptos clave sobre los DESC – ¿Son los derechos económicos, sociales y culturales fundamentalmente diferentes de los derechos civiles y políticos?", por lo que se da a la tarea de explicar una respuesta conciliadora y vinculante:

No. Años atrás se solía considerar que los derechos económicos, sociales y culturales eran totalmente diferentes de los derechos civiles y políticos. Si bien la Declaración Universal de derechos humanos no hacía ninguna distinción entre derechos, ésta apareció en el contexto de la intensificación de las tensiones de la guerra fría entre el Este y el Oeste. (...) Eso dio lugar a la negociación y aprobación de los pactos: uno de derechos civiles y políticos y otro de derechos económicos, sociales y culturales.<sup>27</sup>

No obstante, se debe tener presente que esta separación se ha hecho a un lado, es decir, se ha generado una innovadora estructura de derechos humanos donde no existe separación de hecho o derecho:

Sin embargo, esa rígida separación se ha abandonado y se ha producido un restablecimiento de la arquitectura original de la Declaración Universal. En los últimos decenios, los tratados de derechos humanos, como la Convención sobre los Derechos del Niño o la Convención sobre los derechos de las personas con discapacidad, han integrado todos los derechos.<sup>28</sup>

En este sentido, es posible observar que el ser humano ha tenido un mayor acceso a la tecnología que al ejercicio de derechos civiles, participación política, desenvolvimiento económico, vinculación social y desarrollo cultural, empero, es visible que el acceso a los teléfonos celulares ha sido exponencial en el último sexenio de la segunda década del siglo XXI, ya que 96.6 millones de personas poseen un celular, de estos, el 96.6% emplean un Smartphone y 3.2% emplea un dispositivo celular básico (mayo de 2025)<sup>29</sup>, aunado al empleo de internet: "en 2024, México alcanzó por primera vez los 100.2 millones de

<sup>27</sup> NACIONES UNIDAS, «Conceptos clave sobre los DESC: ¿son los derechos económicos, sociales y culturales fundamentalmente diferentes de los derechos civiles y políticos?», Oficina del Alto Comisionado, 2026.

<sup>28</sup> Ibidem.

<sup>29</sup> TÉLLEZ, Rodrigo, «¿Cuántos celulares hay en México? Esto dice el INEGI», Líder Empresarial, 13 de enero de 2026.

personas usuarias de internet, lo que representa el 83.1% de la población de seis años o más”<sup>30</sup>.

En este orden de ideas, el acceso a las plataformas de inteligencia artificial con base y empleo de redes de internet de pago o gratuitos, son de constante consulta. Por lo que los autores de esta investigación “le preguntaron” a la IA ChatGPT si considera que, si el acceso total e ilimitado a la inteligencia artificial ayudaría al ejercicio y disfrute de los Derechos civiles, políticos, económicos, sociales y culturales, a lo que respondió:

Sí, el acceso amplio —e incluso cercano a universal— a la inteligencia artificial podría contribuir significativamente al ejercicio y disfrute de los derechos civiles, políticos, económicos, sociales y culturales en México, siempre que existan regulación, alfabetización digital y acceso equitativo. La IA puede funcionar como una herramienta facilitadora de derechos, pero no sustituye las obligaciones del Estado.<sup>31</sup>

La IA tiene presente, algorítmicamente, de emitir una postura clara. Donde la contribución sería significativa y de alto impacto en el desarrollo y ejercicio de Derechos de esta índole, empero, considera pertinente contemplar medios de regulación y no busca suplantar las obligaciones del Estado.

Sus respuestas se dirigieron a cada categoría de derechos; en primer lugar, al valorar su contribución a los Derechos civiles, formula que puede ayudar a:

- Acceso a información jurídica para cualquier persona.
- Orientación legal básica automatizada para personas que no pueden pagar abogados.
- Traducción automática a lenguas indígenas, facilitando el acceso a la justicia.
- Detección de discriminación o abuso de autoridad mediante análisis de datos.<sup>32</sup>

Aún y cuando se habían posicionado los desarrolladores de IA, al no contemplar más las asesorías jurídicas, esta plataforma de Open AI (OA), establece oportuno y viable mantener este tipo de asesorías en materia jurídica.

Respecto a los Derechos políticos, plantea tres elementos interesantes que podrían impulsar al ejercicio de estos, sin embargo, el elemento valioso es que el ser humano comprendería las bases de la participación ciudadana y el ejercicio democrático más allá de un simple voto:

- Facilitar información sobre procesos electorales y participación ciudadana.
- Analizar políticas públicas y programas gubernamentales.
- Permitir a los ciudadanos comprender iniciativas legislativas complejas.<sup>33</sup>

El resultado de este entendimiento resultará en un mayor compromiso sobre la actividad Estatal y, por ende, un mayor proceso de transparencia al inmiscuirse en las actividades de las autoridades.

Por otra parte, al abordar a los Derechos económicos con base en los millones de datos de acceso, considera que podría contribuir mediante:

<sup>30</sup> RODRÍGUEZ CATARINO, Giovanni, «¿Cuántos mexicanos usan internet? INEGI revela que ya son más de 100 millones», Diario Récord, 7 de mayo de 2025.

<sup>31</sup> OPENAI, «¿Consideras que el acceso total e ilimitado a la inteligencia artificial ayudaría al ejercicio y disfrute de los derechos civiles, políticos, económicos, sociales y culturales en México?», ChatGPT, modelo identificado por los autores como «versión 4», 18 de abril de 2026.

<sup>32</sup> Ibidem.

<sup>33</sup> Ibidem.

- Productividad empresarial.
- Creación de nuevos empleos y sectores económicos.
- Acceso a educación financiera y asesoría para emprendedores.<sup>34</sup>

Al representar un recurso valioso y gratuito para las estrategias contables, análisis de mercados e incluso marketing, las mipymes que son el motor de México, se impulsarían contribuyendo a crecimiento y desarrollo económico del país, significando, en conjunto, un bloque competitivo en contra de las grandes empresas.

Desde el paradigma amplio de los Derechos sociales, como la salud, educación y bienestar social, la Inteligencia artificial podría impulsar:

- Apoyar el diagnóstico médico temprano.
- Mejorar la educación personalizada.
- Facilitar servicios públicos digitales.<sup>35</sup>

Si bien es cierto, todos los Derechos previamente abordados son los que son base del bienestar del ser humano.

En el sentido cultural, las plataformas de inteligencia artificial, podría contribuir al disfrute y gozo de los derechos culturales sin necesidad de desplazarse:

- Preservar lenguas indígenas mediante traducción y documentación.
- Difundir patrimonio cultural mexicano.
- Facilitar el acceso a bibliotecas, archivos y conocimiento científico.<sup>36</sup>

Finalmente, no es posible dejar de lado la existencia o generación de riesgos ante el empleo ilimitado y gratuito de la IA, tal como:

- Desinformación
- Sesgos algorítmicos
- Vigilancia masiva
- Brecha digital.<sup>37</sup>

Como es posible vislumbrar, es un panorama prometedor, y de manera sincera es una experiencia valiosa "dialogar" con la inteligencia artificial por parte de estos autores.

Concluyentemente, es un imponente y valioso recurso digital, empero, nunca superará el razonamiento, la innovación y el intelecto humano, por lo que priorizar la ética tecnológica y la regulación oportuna, jurídicamente hablando, garantizará un mayor porcentaje de empleo seguro para los integrantes de la nación mexicana.

## 5. CONCLUSIONES

Derivado de la presente investigación, argumentos y discurso principal, es posible confirmar que se cumplieron los objetivos de investigación, se brindó respuesta a la pregunta de investigación y se dio cumplimiento a la postura hipotética principal. De igual forma, se pueden exponer las siguientes conclusiones:

Primera. - La Nación mexicana cuenta con reconocimiento legal para ser constituida como una persona moral y, por ende, ser considerada como titular de derechos humanos

---

<sup>34</sup>Ibidem.

<sup>35</sup>Ibidem.

<sup>36</sup>Ibidem.

<sup>37</sup>Ibidem.

que surjan y/o coincidan con su naturaleza reconocida por la Constitución Política de los Estados, el Código Civil y la jurisprudencia.

Segunda. – La inteligencia artificial de libre acceso ofrecida por el Estado, será de gran utilidad para el desarrollo de los derechos civiles, Políticos, Económicos, Sociales y Culturales, sin una mayor inversión económica o debatirse entre los conceptos de acceso y accesibilidad.

Tercera. – Los derechos humanos que históricamente se han separado, es decir, los derechos civiles y Políticos, y los Económicos, Sociales y Culturales, en la actualidad son unidos y estudiados conforme a los principios de interdependencia e indivisibilidad, por lo que, en la actualidad, pueden ser englobados dentro de la misma importancia.

Cuarta. – La búsqueda por el desarrollo y ejercicio pleno de los derechos humanos, retomados para la presente investigación, ha dirigido la mira hacia la inteligencia artificial. Siendo desde esta misma, la mención de la necesidad de su empleo ético y jurídicamente, viable.

Quinta. – La propuesta de un desarrollo en modalidad de Plataforma de inteligencia artificial diseñada por el Estado Mexicano, contribuiría al ejercicio de los Derechos Humanos de la nación mexicana.

Sexta. – Esta propuesta no suplantarán las obligaciones del Estado ni los medios y estrategias gubernamentales para lograr contribuir al ejercicio pleno de los Derechos de primera y segunda generación, considerados así únicamente ante la posibilidad de estudiarlos al establecer un orden cronológico, ya que con base a los principios de estas prerrogativas humanas, no es posible dividirlos o separarlos.

## 6. REFERENCIAS BIBLIOGRÁFICAS

ANDERSON, Benedict. Comunidades imaginadas: reflexiones sobre el origen y la difusión del nacionalismo. México: Fondo de Cultura Económica, 1993.

COMISIÓN NACIONAL DE LOS DERECHOS HUMANOS. «Derechos humanos: ¿qué son los derechos humanos?». [Disponible en: <https://bsu.buap.mx/f3b>].

MÉXICO. Constitución Política de los Estados Unidos Mexicanos. Cámara de Diputados del H. Congreso de la Unión. [Disponible en: <https://bsu.buap.mx/f24>].

MÉXICO. Código Civil Federal. Cámara de Diputados, versión citada de 14 de noviembre de 2025. [Disponible en: <https://bsu.buap.mx/gcS>].

MÉXICO. Tesis I.19o.A.1 A (11a.), "Pueblos y barrios originarios de la Ciudad de México. Se asimilan al concepto de pueblos tribales establecido en el Convenio 169 sobre pueblos indígenas y tribales en países independientes, de la Organización Internacional del Trabajo (OIT), para efecto de cumplir con el deber de consulta previa", México, Gaceta del Semanario Judicial de la Federación, libro 47, tomo I, volumen 1, marzo de 2025, p. 981. [Disponible en: <https://bsu.buap.mx/gcT>] (Fecha de consulta: 18 de abril de 2026).

MÉXICO. Tesis XXVI.5o. (V Región) 2 K (10a.), "Personas morales o jurídicas. Deben gozar no sólo de los derechos fundamentales constituidos por los derechos humanos reconocidos en la constitución y en los tratados internacionales, y de las garantías para su protección, siempre y cuando estén encaminados a proteger su objeto social, sino también de aquellos que aparezcan como medio o instrumento necesario para la consecución de la finalidad que persiguen", México, Semanario Judicial de la Federación y su Gaceta, libro XI, tomo 2, agosto de 2012, p. 1876. [Disponible en: <https://bsu.buap.mx/gcU>] (Fecha de consulta: 18 de abril de 2026).

- MÉXICO. Tesis I.5o.C.64 C (10a.), "Personas morales. Su protección legal depende de la buena fe con la que se conduzcan", México, Semanario Judicial de la Federación y su Gaceta, libro XXIII, tomo 3, agosto de 2013, p. 1693. [Disponible en: <https://bsu.buap.mx/ga8>] (Fecha de consulta: 20 de marzo de 2026).
- NACIONES UNIDAS. «Conceptos clave sobre los DESC: ¿son los derechos económicos, sociales y culturales fundamentalmente diferentes de los derechos civiles y políticos?». Oficina del Alto Comisionado, 2026. [Disponible en: <https://bsu.buap.mx/gcV>].
- OPENAI. «¿Consideras que el acceso total e ilimitado a la inteligencia artificial ayudaría al ejercicio y disfrute de los derechos civiles, políticos, económicos, sociales y culturales en México?». ChatGPT, versión 4 según los autores, 18 de abril de 2026. [Disponible en: <https://bsu.buap.mx/gcY>].
- ORGANIZACIÓN DE LAS NACIONES UNIDAS. «Desafíos globales. Derechos humanos». 2025. [Disponible en: <https://bsu.buap.mx/f3a>].
- PÉREZ-GALMICHE, Gabriel; ESPINOSA-GONZÁLEZ, David. La crisis de los derechos humanos: estudio comparativo México-Italia. Ciudad de México: Montiel Et Soriano Editores y BUAP, 2021.
- REAL ACADEMIA ESPAÑOLA. Diccionario de la lengua española, voz «nación». 2024. [Disponible en: <https://bsu.buap.mx/f26>].
- REAL ACADEMIA ESPAÑOLA. Diccionario panhispánico del español jurídico, voces «persona moral» y «persona jurídica». [Disponible en: <https://bsu.buap.mx/f3X>].
- RENAN, Ernest. ¿Qué es una nación? Madrid: Alianza, 2010.
- RODRÍGUEZ CATARINO, Geovanni. «¿Cuántos mexicanos usan internet? INEGI revela que ya son más de 100 millones». Diario Récord, 7 de mayo de 2025. [Disponible en: <https://bsu.buap.mx/gcX>].
- SMITH, Anthony. The Ethnic Origins of Nations. Oxford y Nueva York: Blackwell, 1986.
- TÉLLEZ, Rodrigo. «¿Cuántos celulares hay en México? Esto dice el INEGI». Líder Empresarial, 13 de enero de 2026. [Disponible en: <https://bsu.buap.mx/gcW>].
- WEBER, Max. Economy and Society. California: University of California Press, 1978.



# Predecir no es juzgar: Sistemas algorítmicos de riesgo penal y presunción de inocencia en América Latina\*

## *Predicting is not judging: Algorithmic systems of criminal risk and presumption of innocence in Latin America*

Claudia Lorena Morffi Collado\*\*

### RESUMEN

En el año 2023, un juez penal de Bogotá incorporó en una decisión cautelar el resultado arrojado por PRISMA, sistema algorítmico de evaluación de riesgo de la fiscalía general de la Nación, sin que la defensa pudiera conocer las variables utilizadas ni su ponderación. Este hecho, documentado en la Sentencia T-529/2023 de la Corte Constitucional colombiana, constituye el punto de partida de este artículo. El objetivo es demostrar que la incorporación de sistemas algorítmicos de evaluación de riesgo en la justicia penal latinoamericana no ha supuesto una modernización garantista, sino la reproducción automatizada de desigualdades preexistentes. La investigación se desarrolló mediante un enfoque cualitativo, de carácter jurídico-doctrinal y comparado. A partir del examen de los casos de Argentina, Colombia, Brasil y Chile, se identifican como principal resultado las funciones atribuidas a los sistemas algorítmicos de evaluación de riesgo, sus niveles de transparencia y las tensiones que generan frente a las garantías del debido proceso.

### PALABRAS CLAVE

inteligencia artificial judicial; presunción de inocencia; opacidad algorítmica; evaluación de riesgo procesal; positivismo tecnificado.

### ABSTRACT

In 2023, a criminal court judge in Bogotá relied, in a precautionary ruling, on the output produced by PRISMA, a risk assessment algorithm used by the Office of the Attorney General, without the defense being able to examine either the variables considered or the weight assigned to them. This situation, recorded in Judgment T-529/2023 of the Colombian Constitutional Court, serves as the starting point for this article. The paper argues that the incorporation of algorithmic risk assessment tools into Latin American criminal justice has not led to a rights-based modernization, but rather to the automated reproduction of pre-existing inequalities. The study follows a qualitative, doctrinal, and comparative approach, grounded in the analysis of case law, statutory sources, and specialized literature. Based on a review of Argentina, Colombia, Brazil, and Chile, the findings identify the functions assigned to these systems, their levels of transparency, and the tensions they generate in relation to due process guarantees.

### KEYWORDS

judicial artificial intelligence; presumption of innocence; algorithmic opacity; pretrial risk assessment; technicised positivism.

\* Artículo de Investigación

\*\* Doctora en Ciencias Jurídicas y Políticas por la Universidad Pablo de Olavide, Sevilla, España y Máster en Ciencias de la Educación Superior (lorenamorffi91@gmail.com). <https://orcid.org/0000-0002-0556-6578>.

## SUMARIO

1. Introducción
2. Fundamentos y tensiones del riesgo algorítmico en el proceso penal
3. Circulación y adopción de tecnologías de riesgos en América Latina
4. Sesgos algorítmicos y reproducción de desigualdades estructurales
5. Garantías procesales bajo tensión
6. El positivismo tecnificado como marco explicativo
7. Hacia una regulación de la transparencia algorítmica judicial: referentes y garantías procesales
8. Conclusiones
9. Referencias bibliográficas

### 1. INTRODUCCIÓN

En el Derecho procesal penal, la restricción de la libertad antes de una condena firme constituye la intervención estatal más grave sobre la esfera individual. Su legitimidad descansa sobre una doble exigencia: que se apoye en hechos concretos y probados, y que el razonamiento que la sustenta pueda ser examinado, contradicho y recurrido. Los sistemas algorítmicos de evaluación de riesgo cuestionan esa doble exigencia en su raíz.

El episodio de Bogotá no fue una anomalía. Fue una consecuencia lógica de una tendencia regional: la adopción de herramientas de predicción actuarial en decisiones penales sin el soporte normativo ni institucional que esa adopción requiere. En Argentina, el sistema Prometea opera en la Fiscalía de la Ciudad de Buenos Aires automatizando actos procesales desde 2017 sin que exista un marco legal que defina con precisión qué puede hacer y qué no. En Brasil, el Supremo Tribunal Federal utiliza Victor para la gestión de recursos y varios tribunales han desplegado modelos predictivos a través de la plataforma Sinapses, en un ecosistema tecnológico que ninguna norma sustantiva disciplina de manera integral. En Colombia, PRISMA estima probabilidades de reincidencia que informan solicitudes de medidas cautelares. En Chile, la Defensoría Penal Pública ha documentado la incorporación de informes de riesgo no auditables en audiencias de prisión preventiva<sup>1</sup>.

La pregunta que organiza este trabajo es: ¿en qué condiciones el uso de sistemas algorítmicos de evaluación de riesgo procesal en América Latina es compatible con la presunción de inocencia y el debido proceso, considerando la opacidad estructural de estos sistemas y el vacío regulatorio regional? La respuesta que se defiende aquí es que esa compatibilidad no existe en las condiciones actuales de implementación, y que construirla requiere algo más que regulación técnica: requiere repensar qué tipo de racionalidad puede legítimamente fundar una decisión penal.

La literatura sobre algoritmos y justicia penal se ha concentrado, con razón, en el caso estadounidense: COMPAS, *State v. Loomis*, el análisis de ProPublica sobre tasas de error

---

<sup>1</sup>BARONA VILAR, Silvia, "Dataización de la justicia (algoritmos, inteligencia artificial y justicia, ¿el comienzo de una gran amistad?)", *Revista Boliviana de Derecho*, n.º 36, 2023, pp. 14-45; BARROS, Filipe et al., "Machine learning algorithms applied on classification of processes for conciliation on Brazilian labour judiciary", en *Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*, 20.º, Belo Horizonte, Sociedade Brasileira de Computação, 2023, pp. 389-402.

diferencial por raza<sup>2</sup>. Ese debate es imprescindible como referencia comparada, pero no sustituye el análisis de lo que ocurre en sistemas jurídicos con tradición continental, estructuras institucionales distintas y déficits regulatorios más pronunciados.

La contribución específica del trabajo es doble. Por un lado, ofrece el primer análisis sistemático comparado de cuatro jurisdicciones latinoamericanas (Argentina, Colombia, Brasil y Chile) que evalúa los sistemas implementados no desde su eficiencia operativa, sino desde su compatibilidad con garantías procesales concretas: motivación, contradicción y presunción de inocencia. Por otro, desarrolla el concepto de positivismo tecnificado como categoría analítica propia para describir el fenómeno por el cual la autoridad del dato estadístico desplaza la argumentación jurídica en la decisión judicial, con consecuencias que trascienden el error técnico y comprometen la legitimidad del acto jurisdiccional.

La investigación parte de un estudio de Derecho comparado con enfoque analítico-crítico. No produce datos empíricos propios: trabaja sobre fuentes documentales, jurisprudenciales y técnicas, y su contribución es interpretativa. El análisis cubre cuatro jurisdicciones: Argentina, Colombia, Brasil y Chile. Los tres primeros porque han institucionalizado sistemas con nombre propio, función declarada y documentación pública suficiente: Prometea en Argentina, PRISMA y PretorIA en Colombia, Victor y Sinapses en Brasil. Chile se incorpora porque ofrece un caso analíticamente productivo y diferente: veinte años de reforma procesal penal acusatoria en funcionamiento, un debate documentado sobre herramientas de apoyo a decisiones de prisión preventiva, y una resistencia institucional activa protagonizada por la Defensoría Penal Pública que no tiene equivalente visible en los otros casos. Esa resistencia es en sí misma un dato relevante.

México se consideró en etapas previas de investigación, pero fue excluido del análisis comparado sistemático por ausencia de documentación pública suficiente sobre sistemas algorítmicos penales concretos. Se incluye una referencia descriptiva al contexto mexicano en el apartado 4.3 por su importancia doctrinal regional, pero sin pretensión de comparabilidad sistemática con los otros cuatro casos.

Se han consultado tres tipos de fuentes. Fuentes primarias: decisiones judiciales *State v. Loomis* (Wisconsin, 2016), Sentencia T-529/2023 de la Corte Constitucional colombiana, jurisprudencia del TJUE en *SCHUFA Holding* (C-634/21, 2023) y *Dun & Bradstreet* (C-203/22, 2024), textos normativos Reglamento (UE) 2024/1689 sobre Inteligencia Artificial, RGPD y documentación técnica de los sistemas analizados cuando es públicamente accesible. Fuentes secundarias: literatura académica en derecho procesal, filosofía del derecho y ciencias de la computación<sup>3</sup>. Fuentes técnicas: el análisis de

<sup>2</sup>ANGWIN, Julia, et al., "Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks", en *ProPública*.

<sup>3</sup>BARONA VILAR, Silvia, "Algoritmización de la prueba y de la decisión judicial en el proceso penal: ¿utopía o distopía?", en Pillado González, E. (ed.), *El proceso penal ante una nueva realidad tecnológica europea*, Pamplona, Aranzadi, 2023; BELLOSO MARTÍN, Nuria, "Algoritmos predictivos al servicio de la justicia: ¿una nueva forma de minimizar el riesgo y la incertidumbre?", *Revista da Faculdade Mineira de Direito*, vol. 22, n.º 43, 2019, pp. 2-31; CORVALÁN, Juan Gustavo, "Inteligencia artificial: retos, desafíos y oportunidades. Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la justicia", *Revista de Investigações Constitucionais*, vol. 5, n.º 1, 2018, pp. 55-89; CRUZ GONZÁLEZ, Marcos, "La inteligencia artificial y los derechos humanos. Una carrera tecnológica entre Estados Unidos, China y la Unión Europea", *Revista IUS*, vol. 20, N.º 56, 2025, pp. 31-51.

ProPublica sobre COMPAS, los estudios de DRESSEL y FARID sobre precisión predictiva<sup>4</sup>, el análisis de CHOULDECHOVA sobre incompatibilidad matemática entre métricas de equidad, e informes institucionales de los propios sistemas<sup>5</sup>.

Se adopta un método funcional comparado: los sistemas se analizan en función de lo que hacen, no de cómo se denominan institucionalmente, y se evalúan según su compatibilidad con estándares procesales identificados a partir del derecho constitucional comparado y los instrumentos internacionales de derechos humanos aplicables. El análisis no es neutral respecto de sus conclusiones: parte de la premisa de que las garantías procesales son condiciones de legitimidad del sistema penal, no variables de ajuste en función de la eficiencia operativa.

Por último, el artículo no evalúa el rendimiento técnico de los sistemas analizados ni realiza análisis estadístico de sus tasas de error, tampoco examina sistemas en fase puramente experimental o sin documentación pública suficiente. Las conclusiones tienen alcance regional y no pretenden generalizarse a contextos con estructuras institucionales radicalmente distintas.

## 2. FUNDAMENTOS Y TENSIONES DEL RIESGO ALGORÍTMICO EN EL PROCESO PENAL

Un sistema de evaluación de riesgo algorítmico es una arquitectura computacional que transforma datos en decisiones probabilísticas. Su resultado es una puntuación o score que pretende anticipar comportamientos humanos futuros a partir de regularidades observadas en datos históricos. En el ámbito penal, esa puntuación se usa para informar decisiones sobre medidas cautelares, libertad condicional o determinación de la pena.

No se trata simplemente de un software auxiliar, su rasgo distintivo es que traduce trayectorias sociales y datos históricos en una puntuación de riesgo que pretende anticipar conductas. Para ello, el sistema procesa variables como antecedentes penales, edad, entorno socioeconómico o historial judicial, y las combina según patrones extraídos de bases de datos previas<sup>6</sup>. El resultado suele expresarse en una categoría (alto, medio, bajo riesgo) o en un valor numérico que se presenta como objetivable. Su lógica es actuarial, pues no examina el hecho concreto en términos de culpabilidad, sino que sitúa al individuo dentro de un grupo estadístico y le asigna el comportamiento típico de ese grupo. En ese desplazamiento (del acto al perfil) radica su principal implicación jurídica.

Además, estos sistemas no operan en el vacío, dependen de datos producidos por instituciones que ya seleccionan, clasifican y filtran la realidad (policía, fiscalía, tribunales). Por eso, más que descubrir riesgo, lo que hacen es reorganizar y proyectar hacia el futuro patrones contenidos en esos datos.

El funcionamiento del sistema opera en tres fases. Primero, la agregación: grandes volúmenes de información antecedentes, variables socioeconómicas, entorno territorial, comportamiento institucional previo se recopilan y codifican. Segundo, el entrenamiento: esos datos se procesan mediante modelos estadísticos o de aprendizaje automático

---

<sup>4</sup>DRESSEL, Julia; FARID, Hany, "The accuracy, fairness, and limits of predicting recidivism", *Science Advances*, vol. 4, n.º 1, 2018, pp. 1-56.

<sup>5</sup>CHOULDECHOVA, Alexandra, "Fair prediction with disparate impact: a study of bias in recidivism prediction instruments", *Big Data*, vol. 5, n.º 2, 2017, pp. 153-163.

<sup>6</sup>PARRA SEPÚLVEDA, Darío y CONCHA MACHUCA, Ricardo, "Inteligencia artificial y derecho. Problemas, desafíos y oportunidades", en *Revista de Derecho*, vol. 70, 2021, pp. 1-25.

que identifican correlaciones entre variables de entrada y resultados históricos, generalmente decisiones judiciales pasadas. Tercero, la predicción: el sistema aplica los patrones aprendidos al caso concreto y produce una puntuación que, lejos de ser neutral, condensa decisiones previas sobre qué medir, cómo codificarlo y qué objetivo optimizar<sup>7</sup>.

La distinción analíticamente central para este trabajo no es entre modelos más o menos precisos, sino entre modelos de caja negra y modelos de caja blanca. Los primeros generan resultados cuya lógica interna no puede ser reconstruida, ya sea por complejidad técnica (arquitecturas de aprendizaje profundo con millones de parámetros) o por restricciones comerciales, invocación de secreto industrial para bloquear el acceso al código. Los segundos están diseñados para ser interpretables: permiten identificar qué variables incidieron en el resultado y en qué medida.

Esta distinción no es técnica en sentido estricto: es constitucional. Un sistema de caja negra en una decisión penal crea un problema de motivación irresoluble<sup>8</sup>. El Reglamento (UE) 2024/1689 clasifica como sistemas de alto riesgo, entre otros, los empleados en la administración de justicia y los destinados a la predicción de reincidencia delictiva. Esa clasificación activa obligaciones reforzadas de transparencia, supervisión humana, auditabilidad que en América Latina no tienen equivalente normativo.

La opacidad algorítmica no es un defecto técnico, sino una fricción directa con las garantías básicas del Estado de Derecho. Aparece cuando ni el contenido ni la lógica interna de un sistema pueden examinarse de modo que la decisión final quede desacoplada de cualquier posibilidad real de comprensión, impugnación o control.

La opacidad tiene dos fuentes distintas que a veces se confunden<sup>9</sup>. La primera es técnica: los modelos de aprendizaje profundo operan mediante arquitecturas de alta complejidad donde la trazabilidad de una decisión concreta se diluye en millones de parámetros ajustados iterativamente. La segunda es jurídica: las empresas desarrolladoras invocan secreto comercial para impedir el acceso a la lógica del sistema<sup>10</sup>. Cuando eso ocurre, la opacidad deja de ser una limitación técnica y se convierte en una decisión institucionalizada. Las consecuencias son cualitativamente distintas: la opacidad técnica puede mitigarse mediante herramientas de inteligencia artificial explicable (XAI); la opacidad jurídica requiere una respuesta normativa que establezca que la protección de la propiedad intelectual no puede operar como límite absoluto frente al derecho de defensa.

La falta de explicabilidad impacta en tres dimensiones procesales concretas. Primera, el derecho de defensa: sin acceso a los criterios que guiaron la evaluación, ni el juez ni las partes pueden someter el resultado a contradicción real porque no hay base sobre la que discutir su validez. Segunda, la motivación de las resoluciones: si el órgano

---

<sup>7</sup>BARONA VILAR, Silvia, "Justicia penal líquida (desde la mirada de Bauman)", *Teoría y Derecho, Revista de pensamiento crítico*, n.º 22, 2017, p. 44.

<sup>8</sup>CONCHA CAMACHO, William, "El derecho a la transparencia algorítmica como estándar ético en una democracia liberal", *Revista Jurídica Austral*, vol. 5, n.º 1, 2024, pp. 529-550.

<sup>9</sup>PASQUALE, Frank, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge, Harvard University Press, 2015.

<sup>10</sup>MEDINA GUERRERO, Manuel, "El derecho a conocer los algoritmos utilizados en la toma de decisiones: aproximación desde la perspectiva del derecho fundamental a la protección de datos", *Teoría y Realidad Constitucional*, n.º 49, 2022, pp. 55-74.

judicial se limita a asumir la salida del sistema, fenómeno documentado en la psicología cognitiva como automation bias, la decisión pierde anclaje en razones verificables. Tercera, el control del sesgo: la opacidad actúa como pantalla para desigualdades presentes en los datos de entrenamiento que se reeditan a escala masiva revestidas de apariencia técnica<sup>11</sup>.

El desplazamiento desde el hecho probado hacia la probabilidad estimada no es un cambio instrumental: es un cambio de paradigma. La decisión deja de apoyarse en lo que ocurrió y empieza a apoyarse en lo que el modelo anticipa que podría ocurrir. Ese cambio erosiona la presunción de inocencia con independencia de la precisión técnica del sistema, porque el problema no es si el modelo predice bien, sino si predecir es una forma legítima de fundar una decisión penal. La respuesta del Estado de Derecho a esa pregunta ha sido históricamente negativa: la peligrosidad futura no puede fundar por sí sola una restricción de libertad. Los algoritmos no cambian esa respuesta; la eluden.

### **3. CIRCULACIÓN Y ADOPCIÓN DE TECNOLOGÍAS DE RIESGOS EN AMÉRICA LATINA**

#### **3.1 Panorama regional: adopción sin regulación**

América Latina presenta un rasgo que la distingue del modelo europeo: la tecnología avanza con rapidez en la práctica institucional mientras el correlato normativo permanece fragmentario. No existen leyes integrales que disciplinen el uso de sistemas algorítmicos en la toma de decisiones públicas. Predomina un enfoque de soft law basado en lineamientos de política pública y principios generales que no resuelven las cuestiones concretas: qué variables puede usar un sistema predictivo penal, quién lo audita, cómo puede impugnarlo la defensa, qué responsabilidad tienen los desarrolladores.

El resultado es una fase de hibridación en la que la toma de decisiones combina intervención humana y asistencia algorítmica sin condiciones claras que definan los límites de cada una<sup>12</sup>. Ante ese vacío, algunos tribunales han comenzado a asumir un papel correctivo casuístico: limitar o suspender el uso de tecnologías cuando no se acreditan garantías suficientes. Esa respuesta es necesaria pero insuficiente: llega siempre a posteriori, caso a caso, sin efecto estructural sobre el ecosistema tecnológico.

La pregunta no es ya si la región debe adoptar estas herramientas porque ya las adoptó, sino qué condiciones mínimas deben cumplirse para que esa adopción sea compatible con las garantías que los sistemas jurídicos latinoamericanos reconocen formalmente.

#### **3.2 Análisis comparado: Argentina, Colombia, Brasil y Chile**

En Argentina, Prometea aparece configurado principalmente como una herramienta de apoyo para la automatización de tareas repetitivas y la elaboración de borradores o

---

<sup>11</sup>RUIZ RAJO, Carlos Enrique, "Entre algoritmos y garantías: una revisión crítica sobre la inteligencia artificial en el proceso penal", *Revista de Investigación Científica Erga Omnes*, vol. 5, n.º 2, 2025, pp. 1-8; GARCÍA, Manuel Lorenzo, "La inteligencia artificial predictiva al servicio de la prevención e investigación del delito y del proceso penal", *Ciencia Policial*, vol. 183, n.º 1, 2024, pp. 91-132.

<sup>12</sup>DE GREGORIO, Giovanni, "From constitutional freedoms to the power of the platforms: protecting fundamental rights online in the algorithmic society", *European Journal of Legal Studies*, vol. 11, n.º 2, 2019, pp. 65-104; GIMÉNEZ, María, "Inteligencia artificial y derecho penal", *Actualidad Jurídica Iberoamericana*, n.º 10, 2019, pp. 792-843.

dictámenes a partir de precedentes, lo que justifica su clasificación como un sistema asistencial y no como un modelo de predicción penal en sentido estricto. Las fuentes institucionales y técnicas disponibles destacan su aporte a la celeridad y eficiencia del trabajo judicial y fiscal, pero no permiten reconstruir públicamente, con el mismo nivel de detalle, la lógica interna completa del sistema ni verificar la existencia de auditorías externas abiertas, razón por la cual su transparencia puede calificarse como parcial.

En Colombia, la experiencia es funcionalmente más compleja. PretorlA fue desarrollada por la Corte Constitucional para apoyar el proceso de selección y clasificación de acciones de tutela, de modo que su naturaleza es eminentemente organizativa y auxiliar, aun cuando utilice técnicas avanzadas de procesamiento de información<sup>13</sup>. En cambio, PRISMA se proyecta sobre un terreno más sensible, pues fue concebida para estimar riesgo de reincidencia en el contexto de solicitudes de medidas de aseguramiento. Esa diferencia explica que, dentro de un mismo país, convivan herramientas de ordenación judicial con otras de perfil claramente predictivo. No obstante, mientras PretorlA cuenta con reconocimiento institucional público sobre su finalidad y funcionamiento general, en PRISMA la información accesible externamente no permite un escrutinio suficiente de variables, ponderaciones o auditorías independientes, lo que justifica hablar de una transparencia reducida.

En Brasil, el panorama revela un mayor grado de institucionalización. Victor, en el Supremo Tribunal Federal, fue presentado como una herramienta destinada a apoyar la clasificación y triagem de recursos, especialmente en materia de repercusión general, mientras que Sinapses fue formalizada por el Conselho Nacional de Justiça como una plataforma nacional de almacenamiento, entrenamiento y compartición de modelos de inteligencia artificial para el Poder Judicial. Más recientemente, la regulación brasileña ha reforzado de manera expresa exigencias de gobernanza, auditoría, monitoreo, prevención de sesgos y transparencia, lo que sitúa al caso brasileño como uno de los más estructurados de la región. Con todo, esa mayor institucionalización no elimina la heterogeneidad: el nivel de documentación pública y de trazabilidad sigue siendo desigual según la herramienta concreta, por lo que la transparencia debe valorarse como variable y no como homogénea.

En Chile, a diferencia de los casos anteriores, no se advierte en las fuentes primarias revisadas la existencia de un sistema judicial nacional consolidado comparable a Prometea, PretorlA o Sinapses. Lo que sí aparece con claridad es un escenario de discusión doctrinal y cautela institucional, especialmente en materias sensibles como la prisión preventiva, junto con referencias aisladas a exploraciones o usos localizados de herramientas de análisis de riesgo en ámbitos del Ministerio Público. Por ello, calificar a Chile como un caso experimental o en discusión constituye una inferencia prudente a partir del estado actual de la evidencia disponible, más que una afirmación sobre un despliegue nacional consolidado.

Desde esta perspectiva, el valor del cuadro no radica solo en identificar tecnologías, sino en mostrar que el verdadero contraste comparado se produce entre modelos de apoyo administrativo, sistemas de clasificación judicial y herramientas predictivas

---

<sup>13</sup> GUÍO, Armando, Marco ético para la inteligencia artificial en Colombia, Colombia, Consejería Presidencial para Asuntos Económicos y Transformación Digital, 2020.

con incidencia penal. Esa diferencia es decisiva, porque los riesgos jurídicos no son los mismos cuando la inteligencia artificial agiliza trámites, cuando ayuda a priorizar expedientes o cuando introduce inferencias sobre conductas futuras. En consecuencia, la comparación regional confirma que el examen crítico de estos sistemas debe centrarse en su finalidad, en la posibilidad de control humano efectivo y en las garantías de transparencia, explicabilidad y revisión.

### **Argentina: Prometea y el diseño asistencial**

Prometea es el caso más documentado y, en cierta medida, el más cauteloso en su diseño<sup>14</sup>. Surgió con un propósito acotado, agilizar tareas repetitivas y detectar expedientes prioritarios en áreas como salud o vivienda y su evolución posterior ha incorporado funciones de búsqueda de precedentes y automatización de actos procesales. Lo relevante es la lógica bajo la cual se integra: como herramienta de asistencia, con insistencia explícita en preservar la intervención humana como instancia decisoria final. Esa cautela de diseño es real y distingue a Prometea de sistemas predictivos penales como PRISMA. Sin embargo, la ausencia de marco normativo específico significa que esa cautela depende de decisiones institucionales que pueden cambiar sin necesidad de modificar ninguna norma.

### **Colombia: PRISMA y el salto a la predicción penal**

Colombia es el caso más problemático del análisis. PRISMA no gestiona expedientes ni clasifica tutelas: estima la probabilidad de reincidencia de una persona para informar decisiones sobre medidas de aseguramiento<sup>15</sup>. Ese es un uso cualitativamente distinto al de los otros sistemas, porque introduce la lógica actuarial en la fase más sensible del proceso penal: aquella en que se decide si una persona permanece en libertad mientras se sustancia la causa.

Las críticas documentadas son convergentes: falta de transparencia en la ponderación de variables, dependencia de datos históricos que reflejan prácticas policiales y judiciales selectivas, imposibilidad práctica de impugnación técnica por parte de la defensa. La Sentencia T-529/2023 de la Corte Constitucional comenzó a delimitar el terreno, pero lo hizo en un caso concreto y sin establecer estándares generales vinculantes para todos los usos algorítmicos en decisiones públicas.

PretorIA, por contraste, opera en la jurisdicción constitucional para gestionar el volumen masivo de acciones de tutela. Su función es organizativa, no predictiva en sentido penal. La distinción importa: las consecuencias de un error en la clasificación de una tutela y las de un error en la predicción de reincidencia no son equivalentes.

### **Brasil: escala sin marco**

Brasil destaca por la escala de su despliegue y por la existencia de una estrategia institucional de coordinación que no tiene equivalente en la región. Victor, Sinapses y GAIA no son proyectos aislados: forman parte de un ecosistema donde distintos tribunales

---

<sup>14</sup>ROA AVELLA, Marcela; SANABRIA-MOYANO, Jesús; DINAS-HURTADO, Katherin, "Uso del algoritmo COMPAS en el proceso penal y los riesgos a los derechos humanos", *Revista Brasileira de Direito Processual Penal*, vol. 8, n.º 1, 2022.

<sup>15</sup>VALENCIA GALLO, Lina María, et al., "Implicaciones éticas de la utilización de la inteligencia artificial (IA) en la administración de justicia de Colombia", *Revista IUS*, vol. 20, n.º 56, 2025, pp. 47-87.

comparten y entrenan modelos, con un nivel de integración tecnológica avanzado<sup>16</sup>. Eso es exactamente lo que hace más urgente la ausencia de regulación: cuando el despliegue es masivo y coordinado, los errores sistémicos, incluyendo los sesgos, también lo son.

La amplitud del despliegue brasileño no ha ido acompañada de mecanismos equivalentes de transparencia y rendición de cuentas. La documentación técnica pública de Victor es mayor que la de otros sistemas, pero la trazabilidad de las decisiones que informan y el control judicial sobre sus resultados permanecen sin respuesta normativa clara.

### **Chile: resistencia institucional como dato analítico**

Chile es el único caso de los cuatro en que no existe un sistema de evaluación de riesgo penal desplegado a escala nacional. Lo que existe es algo más instructivo para el análisis comparado: un debate activo sobre si debe existir, con posiciones institucionales encontradas y con resistencia documentada desde la Defensoría Penal Pública.

La Defensoría ha sostenido con consistencia que el estándar del artículo 140 del Código Procesal Penal que exige fundar la prisión preventiva en antecedentes concretos que justifiquen la necesidad de cautela es incompatible con predicciones probabilísticas basadas en perfiles. Esa posición no es retórica: ha sido sostenida en audiencias concretas y ha generado pronunciamientos de tribunales de garantía que rechazaron argumentos fiscales basados en informes de riesgo no auditables. El caso chileno muestra que la resistencia institucional es posible, funciona, y requiere capacidad técnica para materializarse. También muestra que esa resistencia no puede depender indefinidamente de la voluntad institucional de la defensa: necesita anclaje normativo.

### **3.3 Contexto mexicano: influencia doctrinal sin despliegue documentado**

Hoy por hoy, México no dispone de una regulación jurídica integral que discipline el uso de la inteligencia artificial dentro del sistema de justicia; el país opera, más bien, en un terreno fragmentario donde predominan principios constitucionales como la igualdad, no discriminación, debido proceso y esquemas de soft law que intentan contener riesgos sin ofrecer reglas técnicas precisas.

Esta ausencia de diseño normativo detallado no es excepcional en la región: como ocurre en otros países latinoamericanos, la producción legislativa avanza con rezago frente al desarrollo tecnológico, lo que deja sin respuesta cuestiones clave como la validez, control y límites del uso de algoritmos en decisiones judiciales. En contraste, México sí ocupa un lugar protagónico en el plano académico, donde universidades y centros de investigación han construido marcos teóricos influyentes sobre argumentación jurídica, transformación digital y sesgo algorítmico, con especial atención a sus efectos en contextos sensibles como la violencia de género<sup>17</sup>.

Ante el vacío normativo interno, el referente práctico tiende a desplazarse hacia estándares internacionales, especialmente el modelo europeo de regulación de IA, que clasifica los sistemas aplicados a la justicia como de alto riesgo y exige condiciones

---

<sup>16</sup>SILVA VIANNA, Ludmila da, "Entre algoritmos e sentenças: o papel da inteligência artificial na administração da justiça", *Inovações jurídicas e direito contemporâneo: reflexões sobre sociedade, economia e tecnologia*. [S. l.]: Editorial Científica Digital, 2025, pp. 59-71.

<sup>17</sup>DÍEZ SPELZ, Juan Francisco, "¿Robots con derechos? La frontera entre lo humano y lo no humano. Reflexiones desde la teoría de los derechos humanos", *Revista IUS*, vol. 15, n.º 48, 2021, pp. 259-287.

estrictas de transparencia, explicabilidad y supervisión humana. Sin embargo, la falta de reglas propias no es neutra: abre espacios a la opacidad algorítmica y a la protección de secretos comerciales que pueden bloquear el escrutinio público, debilitando la posibilidad de impugnar decisiones automatizadas y, en consecuencia, erosionando garantías procesales básicas.

De ahí que la discusión doctrinal insista en la urgencia de introducir salvaguardas concretas auditorías independientes, trazabilidad de datos, derecho a explicaciones comprensibles como parte de una reformulación del debido proceso en clave tecnológica. En este escenario, México se sostiene en un equilibrio inestable: sin ley específica.

Ante el vacío normativo interno, el referente práctico se desplaza hacia estándares internacionales, especialmente el modelo europeo de regulación de IA, que clasifica los sistemas aplicados a la justicia como de alto riesgo y exige condiciones estrictas de transparencia, explicabilidad y supervisión humana. La producción académica mexicana funciona por ahora como el principal mecanismo de orientación y crítica frente a una incorporación de IA que, cuando se produzca a escala, encontrará el mismo vacío normativo que en el resto de la región.

### 3.4 La influencia de COMPAS: trasplante sin controles

Compas (Correctional Offender Management Profiling for Alternative Sanctions) se ha convertido en el arquetipo de los sistemas de evaluación de riesgo penal. Su centralidad no se explica por su precisión técnica sino por su validación institucional en el sistema estadounidense y por su capacidad para fijar un estándar replicable: una herramienta que traduce información biográfica y delictiva en una puntuación usable en decisiones judiciales<sup>18</sup>.

State v. Loomis (Wisconsin Supreme Court, 2016) es el punto de inflexión<sup>19</sup>. El tribunal avaló el uso de COMPAS como herramienta de apoyo con tres límites declarados: naturaleza meramente orientativa del score, obligación de advertir al juzgador sobre las limitaciones técnicas, y preservación de la autonomía judicial. Esos límites son razonables en abstracto y vacíos en la práctica si no existen mecanismos institucionales que los hagan efectivos.

Existen estudios que han documentado que COMPAS clasificaba desproporcionadamente a personas afrodescendientes como de alto riesgo en comparación con personas blancas con trayectorias similares. Los autores DRESSEL y FARID demostraron que la precisión predictiva del sistema no superaba la de personas sin formación especializada<sup>20</sup>. Chouldechova mostró que las distintas métricas de equidad son matemáticamente incompatibles entre sí: un sistema no puede ser simultáneamente equitativo en todas las dimensiones relevantes<sup>21</sup>.

Lo que el análisis técnico de COMPAS deja al descubierto no es un error de calibración corregible<sup>22</sup>. Es el resultado lógico de entrenar un modelo con datos históricos produ-

---

<sup>18</sup>ANGWIN, Julia, et al., *Machine Bias...*, cit., p. 6.

<sup>19</sup>TRIBUNAL SUPREMO DE WISCONSIN. *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016).

<sup>20</sup>DRESSEL, Julia; FARID, Hany, *The accuracy, fairness, and limits of predicting...* cit., p. 45.

<sup>21</sup>CHOUDECHOVA, Alexandra, *Fair prediction with disparate impact: a study of bias in recidivism...*, cit., p. 153.

<sup>22</sup>RUIZ PRADILLO, Juan Carlos, *La investigación del delito en la era digital. Los derechos fundamentales frente a las nuevas medidas tecnológicas de investigación*, Madrid, Fundación Alternativas, 2013.

cidos por un sistema de justicia penal con sesgos estructurales documentados. El sistema aprendió exactamente lo que le enseñaron.

El trasplante de esa lógica a Colombia materializado en PRISMA introduce una capa adicional de riesgo. Los sesgos estructurales de los sistemas judiciales latinoamericanos por clase, etnia, territorio son al menos tan pronunciados como los del sistema estadounidense, y los marcos de control son más débiles. Importar el modelo sin importar también los mecanismos de auditoría y control que incluso en Estados Unidos resultaron insuficientes es una decisión institucional que requiere justificación.

#### 4. SESGOS ALGORÍTMICOS Y REPRODUCCIÓN DE DESIGUALDADES ESTRUCTURALES

##### 4.1 Herencia estructural del sesgo: bias in, bias out

Los modelos de aprendizaje automático se entrenan con datos históricos y, en ese proceso, incorporan las regularidades que esos datos contienen: tanto decisiones correctas como prácticas problemáticas. La inteligencia artificial no actúa como corrector externo del sistema. Actúa como su heredero. Hereda sus sesgos con la misma fidelidad con que hereda sus aciertos, y los reproduce con una consistencia estadística que los hace más difíciles de cuestionar<sup>23</sup>.

El mecanismo tiene nombre en la literatura especializada: biased training data. Cuando los datos de entrenamiento están marcados por prácticas policiales selectivas, decisiones judiciales inconsistentes, o subregistro estructural de ciertos delitos, el sistema aprende exactamente esas mismas relaciones. No distingue entre correlación estadística y justicia sustantiva. Decisiones históricamente arbitrarias o discriminatorias se convierten en insumos legítimos para predicciones futuras. La desigualdad se estabiliza bajo apariencia de coherencia estadística.

En el ámbito latinoamericano esto tiene una dimensión específica. Los sistemas de justicia penal en la región presentan sesgos documentados por clase social, etnia y territorio que son anteriores y más profundos que los que COMPAS reprodujo en el contexto estadounidense. Entrenar un modelo predictivo con datos producidos por esos sistemas no neutraliza los sesgos: los codifica. La eficiencia del algoritmo se convierte en eficiencia para reproducir la desigualdad preexistente.

##### 4.2 Discriminación estructural por clase, etnia y territorio

El sesgo en los modelos algorítmicos no requiere incluir explícitamente variables como raza o clase para producir efectos discriminatorios. Les basta con incluir variables que se correlacionen con ellas: el entorno residencial, la trayectoria laboral, el nivel educativo, la estabilidad familiar. Esas variables funcionan como proxies. El resultado es discriminación que no se declara pero que se materializa en sus efectos. Chouldechova formalizó matemáticamente que las distintas métricas de equidad son incompatibles entre sí optimizar una empeora necesariamente otra<sup>24</sup>.

El sesgo racial constituye el caso empírico más robusto en la literatura sobre evaluación algorítmica del riesgo, siendo el sistema COMPAS su exponente paradigmático:

<sup>23</sup>FURINGO, Carla Irene, "La inteligencia artificial: ¿innovación o amenaza para las garantías procesales en la justicia penal?", *Revista de Derecho de la UNED*, n.º 36, 2026, pp. 101-132.

<sup>24</sup>CHOULDCHOVA, Alexandra, Fair prediction with disparate impact: a study of bias in recidivism..., cit., p. 153.

evidencia independiente mostró que, ante historiales comparables, clasifica a personas afroamericanas como de alto riesgo con una tasa de falsos positivos sustancialmente superior a la de personas blancas, mientras que invierte el patrón en los falsos negativos, favoreciendo a estas últimas al infraestimar su probabilidad de reincidencia<sup>25</sup>. Este comportamiento no se explica por la inclusión explícita de la variable racial formalmente ausente, sino por la explotación de variables proxy (residencia, entorno familiar, trayectorias institucionales) que permiten reconstruir correlaciones étnicas y reproducir, en clave estadística, desigualdades estructurales preexistentes.

En paralelo, los modelos incorporan un sesgo de clase al traducir indicadores de vulnerabilidad desempleo, baja escolarización, inestabilidad laboral o precariedad financiera en señales de peligrosidad o incumplimiento, de modo que la pobreza deviene un factor de riesgo autónomo: no se penaliza el hecho, sino la posición socioeconómica del sujeto, con efectos que se extienden más allá del ámbito penal hacia sistemas de scoring crediticio que restringen el acceso a bienes básicos.

A ello se suma un sesgo territorial que convierte el espacio en variable incriminatoria; el uso del código postal o de zonas catalogadas como de alta criminalidad introduce una imputación colectiva que afecta incluso a individuos sin antecedentes, y en el caso de la vigilancia predictiva genera bucles de retroalimentación: la intensificación policial en determinados barrios produce más registros y detenciones, reforzando los datos que el propio sistema interpreta como validación de su hipótesis inicial.

Este problema se agrava cuando modelos entrenados en contextos específicos se transfieren sin adaptación a otros entornos socioculturales, exportando así patrones discriminatorios desanclados de la realidad local. El resultado, en términos jurídicos, es un desplazamiento desde un derecho penal centrado en el acto hacia uno orientado al autor estadístico, donde la imputación descansa en perfiles probabilísticos más que en conductas probadas, erosionando principios nucleares como la presunción de inocencia y la igualdad ante la ley; de ahí que la respuesta doctrinal insista en auditorías independientes, trazabilidad de los modelos y mecanismos efectivos de supervisión humana capaces de intervenir sobre decisiones automatizadas con potencial discriminatorio<sup>26</sup>.

La dimensión territorial añade la dinámica más problemática: los bucles de retroalimentación. La vigilancia predictiva identifica determinadas zonas como de alto riesgo a partir de datos previos; eso orienta la asignación de recursos policiales hacia esos espacios; eso aumenta la probabilidad de registrar incidentes en ellos; eso refuerza la hipótesis inicial del algoritmo. Con el tiempo, el modelo construye una imagen de la realidad que es, en buena medida, producto de su propia intervención. Eso no es predicción: es profecía autocumplida.

El problema tiene un efecto adicional que la literatura señala con menos frecuencia: desplaza el eje del juicio penal. Cuando la decisión se apoya en perfiles contruidos a partir del pasado, se diluye la centralidad del hecho concreto y reaparece, por vía técnica, una lógica próxima al derecho penal de autor: se valora al individuo por su pertenencia o su entorno más que por su conducta específica.

<sup>25</sup>SILVA VIANNA, Ludmila da, Entre algoritmos e sentenças: o papel da inteligência artificial...cit., p. 63.

<sup>26</sup>Saavedra Vera, et al., "La incidencia del sesgo algorítmico en la justicia predictiva del sistema judicial", Tzhoeoen, vol. 15, n.º 2, 2023, pp. 79-97; REBOLLO DELGADO, Lucrecio, Inteligencia artificial y derechos fundamentales, Madrid, Dykinson, 2023.

### 4.3 El círculo vicioso institucionalizador

El mecanismo completo funciona así: datos históricos con sesgos alimentan el modelo; el modelo produce predicciones que informan decisiones; esas decisiones generan nuevos datos que confirman las predicciones iniciales; esos datos se incorporan al siguiente ciclo de entrenamiento. La discriminación sistémica previa se convierte en patrón estadístico; el patrón estadístico se convierte en criterio de decisión; el criterio de decisión produce resultados que parecen confirmar que el patrón era correcto.

Lo que hace especialmente difícil interrumpir este círculo es que sus resultados tienen apariencia empírica<sup>27</sup>. No son opiniones ni prejuicios: son correlaciones estadísticas respaldadas por datos reales. Una discriminación que se presenta como opinión puede impugnarse argumentativamente. Una discriminación que se presenta como estadística requiere una refutación técnica que la mayoría de los defensores no está en condiciones de producir, especialmente cuando el sistema es opaco.

## 5. GARANTÍAS PROCESALES BAJO TENSIÓN

### 5.1 El derecho a la motivación: delegación encubierta

La motivación de las resoluciones judiciales no es una formalidad externa: es el mecanismo por el cual el acto jurisdiccional se distingue del acto de poder. Una resolución motivada puede ser revisada, impugnada, corregida. Una resolución que delega su fundamento en un sistema opaco no puede ser ninguna de esas cosas.

Este deber, positivizado en preceptos como el artículo 120.3 de la Constitución española y el artículo 93, IX de la Constitución brasileña, impone al órgano jurisdiccional la carga de hacer explícita la arquitectura racional de su decisión, esto es, de exponer de manera verificable la conexión entre los hechos acreditados, las normas aplicadas y las inferencias que conducen al fallo, de modo que la resolución no se agote en una conclusión autoritativa sino que revele los criterios jurídicos que la hacen inteligible, controlable y susceptible de contradicción.

La incorporación de sistemas algorítmicos de caja negra produce lo que puede denominarse motivación delegada: la resolución menciona el resultado del sistema como si fuera un dato y construye sobre él sin poder explicar cómo se produjo. El juez no puede hacer explícito un itinerario lógico que desconoce. La consecuencia no es una motivación defectuosa: es la ausencia de motivación disfrazada de tecnología.

La doctrina ha propuesto una motivación reforzada cuando se incorporan herramientas algorítmicas: el juez debe explicitar la incidencia del resultado en su convencimiento<sup>28</sup>. Esa propuesta es correcta pero insuficiente si el sistema es opaco. No se puede motivar lo que no se puede conocer. La solución no está solo en las exigencias que se imponen al juez, sino en las exigencias que se imponen al sistema antes de que el juez lo reciba. La motivación reforzada solo funciona si el sistema que informa la decisión puede explicarse.

<sup>27</sup>MORÁN ESPINOSA, Alejandra, "Responsabilidad penal de la inteligencia artificial (IA). ¿La próxima frontera?", Revista IUS, vol. 15, n.º 48, 2021, pp. 290-323; MARTÍN FABA, José María, "La información sobre la lógica aplicada en el credit scoring automatizado: fácil, concisa y sin tecnicismos", Revista CESCO de Derecho de Consumo, n.º 54, 2025, pp. 15-30.

<sup>28</sup>COTINO HUESO, Lorenzo, "La primera sentencia del Tribunal de Justicia de la Unión Europea sobre decisiones automatizadas y sus implicaciones para la protección de datos y el Reglamento de inteligencia artificial", Diario La Ley, n.º 80, 2024, pp. 5-38.

## 5.2 La contradicción ante una puntuación inimpugnable

El principio de contradicción exige que las partes puedan acceder, comprender y cuestionar todos los elementos que sustentan la decisión judicial. Cuando uno de esos elementos es una puntuación producida por un sistema propietario cuya lógica está protegida como secreto comercial, esa exigencia se vuelve inoperante en la práctica.

La defensa se encuentra ante una asimetría epistémica de primer orden: sabe que existe una puntuación, conoce su resultado, y no puede hacer nada más. No puede reconstruir el razonamiento que la produjo, ni identificar qué variables influyeron en ella, ni siquiera verificar si los datos de entrada eran correctos. Impugnar esa puntuación requeriría acceso a información que el secreto comercial bloquea.

State v. Loomis ilustra el problema. El tribunal admitió el uso de COMPAS sin revelar íntegramente los factores que sustentaban la puntuación. La defensa pudo conocer el resultado, pero no impugnar su fundamento. Eso no es contradicción: es acceso a la conclusión sin acceso a las premisas. En términos procesales, equivale a presentar una pericia cuyo método no puede ser examinado.

La jurisprudencia del TJUE en SCHUFA Holding (C-634/21, 2023) y Dun & Bradstreet (C-203/22, 2024) establece que el interesado tiene derecho a conocer la lógica aplicada en los sistemas de puntuación automatizada, y que ese derecho no cede frente a la protección de secretos comerciales cuando están en juego garantías fundamentales. Ese estándar no tiene equivalente jurisprudencial consolidado en América Latina.

## 5.3 La presunción de inocencia y el etiquetado predictivo

La presunción de inocencia impone que toda persona sea tratada como inocente hasta condena firme. Los sistemas de etiquetado predictivo de riesgo no condenan: clasifican. Esa distinción formal no resuelve el problema sustantivo. Una clasificación de alto riesgo en el contexto de una decisión cautelar produce efectos similares a los de una presunción de culpabilidad: aumenta la probabilidad de que se imponga prisión preventiva, lo que en la práctica latinoamericana puede significar meses o años de privación de libertad antes de que exista siquiera una sentencia.

El problema no es que los sistemas predigan mal, aunque en muchos casos lo hagan. El problema es que predecir no es una forma legítima de fundar una restricción de libertad en un Estado de Derecho. El proceso penal de acto exige que la decisión sobre la libertad de una persona se apoye en hechos concretos probados, no en inferencias estadísticas sobre lo que podría hacer en el futuro. Ese principio no admite excepción tecnológica.

Las variables que estos sistemas utilizan (nivel educativo, estabilidad laboral, entorno territorial) no acreditan riesgo concreto: reproducen desigualdad estructural. Una persona pobre, con empleo precario y residente en una zona estigmatizada obtendrá una puntuación de riesgo más alta que alguien con características individuales similares pero distinta posición socioeconómica. Eso no es predicción: es penalización automatizada de la vulnerabilidad.

## 6. EL POSITIVISMO TECNIFICADO COMO MARCO EXPLICATIVO COMO MARCO EXPLICATIVO

### 6.1 Genealogía del positivismo tecnificado

La literatura sobre algoritmos y derecho ha descrito con precisión los mecanismos técnicos del sesgo, la opacidad y la discriminación automatizada. Lo que ha abordado con menor profundidad es el fenómeno epistemológico que los sostiene: la sustitución de la autoridad argumentativa por la autoridad del dato. A ese fenómeno lo denominamos aquí positivismo tecnificado<sup>29</sup>.

La denominación no es arbitraria. El positivismo jurídico clásico planteó que el derecho es lo que las normas dicen, con independencia de su contenido moral. Kelsen separó la validez de la justicia; Hart separó la existencia del derecho de su mérito. El positivismo tecnificado da un paso adicional: la decisión es lo que los datos indican, con independencia de la argumentación jurídica. La diferencia no es filosófica en sentido abstracto: tiene consecuencias institucionales directas.

La distinción entre esta categoría y conceptos próximos disponibles en la literatura merece precisión. La tecnocracia judicial describe una orientación institucional hacia la gestión técnica de la justicia, pero no captura el desplazamiento epistemológico específico que aquí se analiza. Delegación algorítmica describe un fenómeno de transferencia de decisiones, pero permanece en el plano descriptivo. Automatización decisional captura el aspecto procedimental pero no el ontológico: la transformación de qué cuenta como razón para decidir. El positivismo tecnificado designa precisamente ese desplazamiento: la reconfiguración de la racionalidad judicial de modo que la autoridad del dato estadístico sustituye a la justificación argumentativa como fundamento de legitimidad.

### 6.2 El mecanismo institucional

En el contexto judicial, el positivismo tecnificado se manifiesta cuando el juez acepta el resultado del sistema algorítmico no porque haya examinado y validado su razonamiento, sino porque confía en la autoridad del dato cuantificado. El automation bias, documentado en la psicología cognitiva y en la literatura sobre sistemas de apoyo a decisiones, describe exactamente esa dinámica: la tendencia a aceptar acríticamente los outputs de sistemas automatizados, especialmente en contextos de sobrecarga de trabajo o presión institucional. En jurisdicciones latinoamericanas con altas tasas de litigiosidad y déficits estructurales de recursos, esa presión es la norma.

El positivismo tecnificado no implica mala fe del juez. Implica algo más preocupante: una transformación silenciosa de la racionalidad judicial que no requiere ninguna decisión explícita para producirse. El juez sigue siendo formalmente el decisor; materialmente, el centro de gravedad de la decisión se ha desplazado hacia el algoritmo. La legitimidad del acto jurisdiccional descansa entonces en la autoridad del sistema, no en el razonamiento del juez. Y la autoridad del sistema descansa, a su vez, en datos que nadie examina, producidos por instituciones cuyos sesgos nadie ha corregido.

Romper esa lógica requiere algo más que regulación técnica. Requiere reafirmar que la decisión judicial es un acto de argumentación en el sentido de MacCormick, de

<sup>29</sup>ESPÍN ALBA, Isabel, "Protección de datos, inteligencia artificial, score crediticio: últimos movimientos judiciales y normativos", *Revista Electrónica de Direito*, vol. 35, n.º 3, 2024, pp. 2-6; BORBÓN, Diego, "Neuroética y neuroderechos ante las propuestas de neurointervenciones coercitivas con fines de mejora moral en ámbitos penales", *Cuadernos Electrónicos de Filosofía del Derecho*, n.º 53, 2025, pp. 56-80.

Alexy, de Atienza, no de cálculo; que su legitimidad no deriva de su eficiencia sino de su justificabilidad pública; y que ningún sistema algorítmico, por preciso que sea, puede sustituir la responsabilidad del juez de razonar sobre por qué una persona merece ser privada de su libertad<sup>30</sup>.

El caso *State v. Loomis* opera como punto de inflexión en la incorporación de sistemas algorítmicos al proceso penal: no inaugura solo una técnica, sino una forma de racionalidad decisional que tensiona los límites clásicos del debido proceso. En ese expediente, Eric Loomis fue condenado en Wisconsin con apoyo de un informe generado por COMPAS, herramienta que, a partir de un cuestionario extenso y datos históricos, produce una puntuación de riesgo de reincidencia; la defensa objetó su uso por la opacidad del modelo (blindado por secreto comercial), su potencial discriminatorio y su incompatibilidad con una sentencia verdaderamente individualizada.

La Wisconsin Supreme Court desestimó la impugnación, pero lo hizo bajo un esquema condicional: el output algorítmico no puede erigirse en fundamento exclusivo, debe ser tratado como insumo auxiliar, el juzgador ha de ser advertido de sus limitaciones empíricas y sesgos potenciales, y, sobre todo, debe preservar una capacidad crítica autónoma frente al resultado automatizado. Sin embargo, esta validación pragmática no neutraliza los déficits estructurales que el propio caso visibiliza: la opacidad impide un contradictorio efectivo al vedar el examen de la lógica interna del sistema; la evidencia empírica posterior, como la difundida por ProPublica, mostró patrones sistemáticos de sesgo racial en la clasificación del riesgo; y, en un plano más profundo, el modelo desplaza el eje de imputación desde el hecho probado hacia la pertenencia estadística, erosionando la individualización de la pena y aproximando el sistema a una lógica actuarial. Loomis, así, no legitima sin más la justicia algorítmica: la expone, en sus propias condiciones de posibilidad, como un dispositivo útil pero normativamente inestable, cuya adopción sin estándares de explicabilidad y control público amenaza con sustituir la deliberación judicial por una autoridad técnica difícilmente escrutable.

## **7. HACIA UNA REGULACIÓN DE LA TRANSPARENCIA ALGORÍTMICA JUDICIAL: REFERENTES Y GARANTÍAS PROCESALES**

El Reglamento (UE) 2024/1689 ofrece la referencia normativa más desarrollada disponible. Clasifica como sistemas de alto riesgo los empleados en la administración de justicia y en la predicción de reincidencia, y les impone obligaciones reforzadas: transparencia funcional (artículo 13), supervisión humana significativa (artículo 14) y documentación técnica suficiente para auditoría<sup>31</sup>. La jurisprudencia del TJUE en *SCHUFA* y *Dun & Bradstreet* establece que el derecho a conocer la lógica aplicada en sistemas de puntuación automatizada no cede frente al secreto comercial cuando están en juego garantías procesales fundamentales.

---

<sup>30</sup>ATIENZA, Manuel, *Las razones del Derecho. Teorías de la argumentación jurídica*, México, Universidad Nacional Autónoma de México, 2005.

<sup>31</sup>PÉREZ-UGENA, María, "Análisis comparado de los distintos enfoques regulatorios de la inteligencia artificial en la Unión Europea, EE. UU., China e Iberoamérica", *Anuario Iberoamericano de Justicia Constitucional*, vol. 28, n.º 1, 2024, pp. 129-156; DÍAZ SENÉS, Juan Pedro, "Panorámica internacional de la regulación de la inteligencia artificial: Reglamento europeo de inteligencia artificial (AI Act)", *Revista CESCO de Derecho de Consumo*, n.º 49, 2024, pp. 117-130.

Ese estándar europeo es útil como referencia para América Latina, pero no es directamente trasplantable. La región necesita un marco adaptado a sus condiciones institucionales: sistemas de justicia con mayor carga litigiosa, déficits estructurales de recursos, menor capacidad técnica instalada en los operadores jurídicos, y presencia dominante de proveedores privados con escasa disposición a la transparencia. Un marco regulatorio que copie las obligaciones europeas sin construir las instituciones que las hagan efectivas desplaza el problema sin resolverlo.

Independientemente del modelo regulatorio que cada jurisdicción adopte, hay tres estándares sin los cuales ningún sistema algorítmico debería utilizarse en una decisión penal.

Primero, explicabilidad. Todo sistema que informe una decisión penal debe poder explicar a la defensa, en términos comprensibles y verificables, qué variables consideró, cómo las ponderó y por qué llegó al resultado que llegó. La protección de la propiedad intelectual no puede operar como límite absoluto frente al derecho de defensa. Cuando esa información sea necesaria para impugnar una decisión, debe ser accesible. El mecanismo puede ser la mediación de una autoridad judicial o de control como el modelo europeo establece, pero la información debe llegar.

Segundo, auditabilidad. Todo sistema de evaluación de riesgo penal debe ser objeto de auditoría independiente periódica, con acceso al código fuente, a los datos de entrenamiento y a los criterios de ponderación. Esa auditoría debe incluir evaluación de sesgos y debe ser pública en sus conclusiones, aunque no necesariamente en sus detalles técnicos protegibles. Los resultados de la auditoría deben estar disponibles para la defensa en cualquier proceso en que el sistema haya informado una decisión.

Tercero, supervisión humana efectiva. El juez no puede ser un validador formal de resultados algorítmicos. Debe poder confirmar, matizar o rechazar motivadamente la recomendación del sistema, y esa capacidad requiere que el sistema sea comprensible y que el juez tenga formación suficiente para ejercer ese control de manera sustantiva. Cuando el juez incorpore una recomendación algorítmica a su decisión, debe motivar específicamente por qué la considera fiable y pertinente en el caso concreto. La referencia genérica al resultado del sistema no es motivación suficiente.

La auditoría algorítmica no es un mecanismo de mejora técnica: es una condición de legitimidad procesal. Su función en el proceso judicial es análoga a la cadena de custodia de la prueba material: garantiza que lo que llega al juez es lo que dice ser y que ha llegado de manera verificable.

La auditoría debe ser continua e integrada en todo el ciclo de vida del sistema. La auditoría *ex ante* evalúa el diseño antes del despliegue. La auditoría *ex post* periódica detecta desviaciones durante la operación. La auditoría extraordinaria, activable a instancia de parte cuando existan indicios fundados de fallos o sesgos, es la que más directamente protege el derecho de defensa.

Ninguna de estas auditorías es efectiva sin acceso a los datos de entrenamiento, al código y a los criterios de ponderación. Eso entra en tensión con los intereses comerciales de los proveedores. Esa tensión debe resolverse en favor del derecho de defensa, no del secreto empresarial, cuando los sistemas se usan en decisiones que afectan la libertad de personas. Los sistemas que no puedan ser auditados en esas condiciones no deberían utilizarse en el ámbito penal.

## 8. CONCLUSIONES

Primera: Lo que este trabajo ha documentado no es un problema técnico que la ingeniería pueda resolver ni un problema normativo que la regulación pueda contener por sí sola. Es un problema de racionalidad: la racionalidad actuarial y la racionalidad del proceso penal responden a lógicas diferentes, y esa diferencia no se resuelve diseñando mejores algoritmos.

Segunda: Los sistemas de evaluación de riesgo penal en América Latina comparten un conjunto de rasgos que los hacen estructuralmente incompatibles con las garantías que los propios ordenamientos de la región reconocen. Operan con opacidad suficiente para impedir la motivación real y la contradicción efectiva. Se entrenan con datos que reflejan desigualdades históricas y las reproducen con apariencia estadística de objetividad. Se despliegan en ausencia de marcos normativos que establezcan condiciones mínimas de legitimidad. Y funcionan en contextos institucionales donde la sobrecarga judicial hace probable el automation bias.

Tercera: El concepto de positivismo tecnificado describe el fenómeno resultante: la sustitución de la argumentación jurídica por la autoridad del dato, que transforma silenciosamente la racionalidad de la decisión judicial sin necesidad de ninguna decisión explícita. Ese desplazamiento se produce como resultado lógico de la interacción entre sistemas opacos, operadores sin formación técnica suficiente para cuestionarlos, y condiciones institucionales que incentivan la eficiencia por encima de la deliberación.

Cuarta: Las conclusiones del análisis comparado son coherentes con esa diagnosis. Argentina ha construido el caso más cauteloso, con un diseño asistencial que preserva formalmente la intervención humana, pero sin marco normativo que garantice que esa cautela no cambie. Colombia ha desplegado el caso más preocupante PRISMA con opacidad documentada, sin auditabilidad pública y sin respuesta regulatoria estructural tras la Sentencia T-529/2023. Brasil ha apostado por la escala sin construir los mecanismos de control equivalentes. Chile demuestra que la resistencia institucional es posible y tiene efectos reales, pero no es sustituto de regulación.

Quinta: Las propuestas de este trabajo son las condiciones mínimas sin las cuales la incorporación de sistemas algorítmicos en la justicia penal no puede considerarse compatible con los principios que esos sistemas dicen proteger. Explicabilidad, auditabilidad y supervisión humana efectiva no son aspiraciones técnicas avanzadas: son traducciones al lenguaje del diseño tecnológico de exigencias que el derecho constitucional ya establece. La cuestión de fondo no es tecnológica: es si los sistemas de justicia latinoamericanos van a utilizar la tecnología para corregir sus inercias o para darles una apariencia de objetividad que las haga más difíciles de ver y más costosas de combatir. La respuesta a esa pregunta no está en los algoritmos: está en las decisiones políticas e institucionales que determinan cómo, con qué garantías y con qué rendición de cuentas esos algoritmos se usan.

## 9. REFERENCIAS BIBLIOGRÁFICAS

ANGWIN, Julia, et al., "Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks", ProPública, [Consultada 23 de mayo de 2016]. [Disponible en: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>].

- ATIENZA, Manuel, *Las razones del Derecho. Teorías de la argumentación jurídica*, México, Universidad Nacional Autónoma de México, 2005.
- BARONA VILAR, Silvia, "Algoritmización de la prueba y de la decisión judicial en el proceso penal: ¿utopía o distopía?" En Pillado González, E., ed. *El proceso penal ante una nueva realidad tecnológica europea*, Pamplona, Aranzadi, 2023. [Consultada 23 de abril de 2026]. [Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=8857602>].
- BARONA VILAR, Silvia, "Dataización de la justicia (algoritmos, inteligencia artificial y justicia, ¿el comienzo de una gran amistad?)", *Revista Boliviana de Derecho*, n.º 36, 2023, pp. 14-45. [Consultada 20 de abril de 2026]. [Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=9043836>].
- BARONA VILAR, Silvia, "Justicia penal líquida (desde la mirada de Bauman)", *Teoría y Derecho. Revista de pensamiento crítico*, n.º 22, 2017. [Consultada 20 de marzo de 2026]. [Disponible en: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3510764](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3510764)].
- BARROS, Filipe, et al., "Machine learning algorithms applied on classification of processes for conciliation on Brazilian labour judiciary", *Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*, 20.º, Belo Horizonte. Belo Horizonte: Sociedade Brasileira de Computação, 2023, pp. 389-402. [Consultada 19 de abril de 2026]. [Disponible en: <https://sol.sbc.org.br/index.php/eniac/article/view/25716>].
- BELLOSO MARTÍN, Nuria, "Algoritmos predictivos al servicio de la justicia: ¿una nueva forma de minimizar el riesgo y la incertidumbre?", *Revista da Faculdade Mineira de Direito*, vol. 22, n.º 43, 2019, pp. 2-31. [Consultada 2 de marzo de 2026]. [Disponible en: <https://periodicos.pucminas.br/Direito/article/view/20780>].
- BORBÓN, Diego, "Neuroética y neuroderechos ante las propuestas de neurointervenciones coercitivas con fines de mejora moral en ámbitos penales", *Cuadernos Electrónicos de Filosofía del Derecho*, n.º 53, 2025. [Consultada 20 de marzo de 2026]. [Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=10214876>].
- CHOULDCHOVA, Alexandra, "Fair prediction with disparate impact: a study of bias in recidivism prediction instruments", *Big Data*, vol. 5, n.º 2, 2017, pp. 153-163. [Consultada 14 de febrero de 2026]. [Disponible en: <https://n9.cl/kl1ln0>].
- CONCHA CAMACHO, William, "El derecho a la transparencia algorítmica como estándar ético en una democracia liberal", *Revista Jurídica Austral*, vol. 5, n.º 1, 2024, pp. 529-550. [Consultada 25 de abril de 2026]. [Disponible en: <https://ojs.austral.edu.ar/juridicaaustral/article/view/1239>].
- CORVALÁN, Juan Gustavo, "Inteligencia artificial: retos, desafíos y oportunidades. Prometea: la primera inteligencia artificial de Latinoamérica al servicio de la justicia", *Revista de Investigações Constitucionais*, vol. 5, n.º 1, 2018, pp. 55-89. [Consultada 15 de abril de 2026]. [Disponible en: <https://n9.cl/ua5s13>].
- COTINO HUESO, Lorenzo, "La primera sentencia del Tribunal de Justicia de la Unión Europea sobre decisiones automatizadas y sus implicaciones para la protección de datos y el Reglamento de inteligencia artificial", *Diario La Ley*, 2024, n.º 80. [Consultada 20 de marzo de 2026]. [Disponible en: <https://www.cotino.es/el-primer-estudio-sobre-la-primer-sentencia-del-tjue-sobre-decisiones-automatizadas/>].
- CRUZ GONZÁLEZ, Marcos, "La inteligencia artificial y los derechos humanos. Una carrera tecnológica entre Estados Unidos, China y la Unión Europea", *Revista IUS*, vol. 20, n.º 56, 2025, pp. 31-51. [Consultada 27 de marzo de 2026]. [Disponible en: <https://www.revistaius.com/index.php/ius/article/view/1059>].
- DE GREGORIO, Giovanni, "From constitutional freedoms to the power of the platforms: protecting fundamental rights online in the algorithmic society", *European Journal of Legal Studies*, vol. 11, n.º 2, 2019, pp. 65-104. [Consultada 5 de marzo de 2026]. [Disponible en: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3365106](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3365106)].
- DÍAZ SENÉS, Juan Pedro, "Panorámica internacional de la regulación de la inteligencia artificial: Reglamento europeo de inteligencia artificial (AI Act)", *Revista CESCO de Derecho de Consumo*, n.º 49,

- 2024, pp. 117-130. [Consultada 25 de abril de 2026]. [Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=9462903>].
- DÍEZ SPELZ, Juan Francisco, "¿Robots con derechos? La frontera entre lo humano y lo no humano. Reflexiones desde la teoría de los derechos humanos", *Revista IUS*, vol. 15, n.º 48, 2021, pp. 259-287. [Consultada 25 de abril de 2026]. [Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=8953275>].
- DRESSEL, Julia; FARID, Hany, "The accuracy, fairness, and limits of predicting recidivism", *Science Advances*, vol. 4, n.º 1, 2018, pp. 1-56. [Consultada 27 de marzo de 2026]. [Disponible en: <https://pubmed.ncbi.nlm.nih.gov/29376122>].
- ESPÍN ALBA, Isabel, "Protección de datos, inteligencia artificial, score crediticio: últimos movimientos judiciales y normativos", *Revista Electrónica de Direito*, vol. 35, n.º 3, 2024, pp. 2-6. [Consultada 26 de abril de 2026]. [Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=9773689>].
- FURINGO, Carla Irene, "La inteligencia artificial: ¿innovación o amenaza para las garantías procesales en la justicia penal?", *Revista de Derecho de la UNED*, n.º 36, 2026, pp. 101-132. [Consultada 23 de abril de 2026]. [Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=10611408>].
- GARCÍA, Manuel Lorenzo, "La inteligencia artificial predictiva al servicio de la prevención e investigación del delito y del proceso penal", *Ciencia Policial*, vol. 183, 2024, n.º 1, pp. 91-132. [Consultada 20 de marzo de 2026]. [Disponible en: <https://n9.cl/g2x93>].
- GIMÉNEZ, María, "Inteligencia artificial y derecho penal", *Actualidad Jurídica Iberoamericana*, n.º 10, 2019, pp. 792-843. [Consultada 6 de marzo de 2026]. [Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=6978830>].
- GUÍO, Armando, Marco ético para la inteligencia artificial en Colombia, Colombia, Consejería Presidencial para Asuntos Económicos y Transformación Digital, 2020. [Consultada 25 de marzo de 2026]. [Disponible en: <https://www.usergioarboleda.edu.co/wp-content/uploads/2021/11/Marco-etico-para-la-inteligencia-artificial-en-Colombia-Maestria-en-Inteligencia-artificial.pdf>].
- MARTÍN FABA, José María, "La información sobre la lógica aplicada en el credit scoring automatizado: fácil, concisa y sin tecnicismos", *Revista CESCO de Derecho de Consumo*, n.º 54, 2025, pp. 15-30. [Consultada 20 de marzo de 2026]. [Disponible en: <https://n9.cl/jn1917>].
- MEDINA GUERRERO, Manuel, "El derecho a conocer los algoritmos utilizados en la toma de decisiones: aproximación desde la perspectiva del derecho fundamental a la protección de datos", *Teoría y Realidad Constitucional*, n.º 49, 2022, pp. 55-74. [Consultada 20 de marzo de 2026]. [Disponible en: <https://revistas.uned.es/index.php/TRC/article/view/33847>].
- MORÁN ESPINOSA, Alejandra, "Responsabilidad penal de la inteligencia artificial (IA). ¿La próxima frontera?", *Revista IUS*, vol. 15, n.º 48, 2021, pp. 290-323. [Consultada 20 de marzo de 2026]. [Disponible en: <https://revistaius.com/index.php/ius/article/view/706>].
- PARRA SEPÚLVEDA, Darío y CONCHA MACHUCA, Ricardo, "Inteligencia artificial y derecho. Problemas, desafíos y oportunidades", *Revista de Derecho*, vol. 70, 2021, pp. 1-25. [Consultada 25 de abril de 2026]. [Disponible en: <https://revistas.javeriana.edu.co/index.php/vnijuri/article/view/35364>].
- PASQUALE, Frank, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge, Harvard University Press, 2015. [Consultada 20 de marzo de 2026]. [Disponible en: <https://n9.cl/47qzjg>].
- PÉREZ-UGENA, María, "Análisis comparado de los distintos enfoques regulatorios de la inteligencia artificial en la Unión Europea, EE. UU., China e Iberoamérica". *Anuario Iberoamericano de Justicia Constitucional*, vol. 28, n.º 1, 2024, pp. 129-156. [Consultada 20 de marzo de 2026]. [Disponible en: <https://recyt.fecyt.es/index.php/AJJC/article/view/108247>].
- REBOLLO DELGADO, Lucrecia, *Inteligencia artificial y derechos fundamentales*, Madrid, Dykinson, 2023.
- ROA AVELLA, Marcela; SANABRIA-MOYANO, Jesús; DINAS-HURTADO, Katherin, "Uso del algoritmo COMPAS en el proceso penal y los riesgos a los derechos humanos", *Revista Brasileira de Direito Processual Penal*, vol. 8, n.º 1, 2022. [Consultada 26 de marzo de 2026]. [Disponible en: <https://revista.ibraspp.com.br/RBDPP/article/view/615>].

- RUIZ PRADILLO, Juan Carlos, *La investigación del delito en la era digital. Los derechos fundamentales frente a las nuevas medidas tecnológicas de investigación*. Madrid, Fundación Alternativas, 2013. [Consultada 20 de marzo de 2026]. [Disponible en: <https://n9.cl/wo2vj>].
- RUIZ RAJO, Carlos Enrique, "Entre algoritmos y garantías: una revisión crítica sobre la inteligencia artificial en el proceso penal", *Revista de Investigación Científica Erga Omnes*, vol. 5, n.º 2, 2025, pp. 1-8. [Consultada 20 de marzo de 2026]. [Disponible en: <https://revistas.unh.edu.pe/index.php/rceo/article/view/763>].
- SAAVEDRA VERA, César Orlando, et al., "La incidencia del sesgo algorítmico en la justicia predictiva del sistema judicial". *Tzhoeoen*, vol. 15, n.º 2, 2023, pp. 79-97. [Consultada 25 de marzo de 2026]. [Disponible en: <https://revistas.uss.edu.pe/index.php/tzh/article/view/2592>].
- SILVA VIANNA, Ludmila da, "Entre algoritmos e sentenças: o papel da inteligência artificial na administração da justiça", *Inovações jurídicas e direito contemporâneo: reflexões sobre sociedade, economia e tecnologia*. [S. l.]: Editorial Científica Digital, 2025, pp. 59-71. [Consultada 13 de marzo de 2026]. [Disponible en: <https://periodicos.unesc.net/ojs/index.php/AnaisDirH/article/view/10501>].
- VALENCIA GALLO, Lina María et al., "Implicaciones éticas de la utilización de la inteligencia artificial (IA) en la administración de justicia de Colombia". *Revista IUS*, vol. 20, n.º 56, 2025, pp. 47-87. [Consultada 20 de marzo de 2026]. [Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=10453773>].

### **Fuentes jurisprudenciales**

- CORTE CONSTITUCIONAL. Sentencia T-529/2023, Bogotá, Corte Constitucional, 2023. [Consultada 20 de marzo de 2026]. [Disponible en: <https://www.corteconstitucional.gov.co/relatoria/2023/t-529-23.htm>].
- TRIBUNAL DE JUSTICIA DE LA UNIÓN EUROPEA. Asunto C-203/22. Dun & Bradstreet Austria GmbH. Sentencia de 27 de febrero de 2025. [Consultada 8 de marzo de 2026]. [Disponible en: <https://n9.cl/lrt5n2>].
- TRIBUNAL DE JUSTICIA DE LA UNIÓN EUROPEA. Asunto C-634/21. SCHUFA Holding AG. Sentencia de 7 de diciembre de 2023. [Consultada 8 de marzo de 2026]. [Disponible en: <https://n9.cl/bmjomq>].
- TRIBUNAL SUPREMO DE WISCONSIN. *State v. Loomis*, 881 N.W.2d 749 (Wis. 2016). [Consultada 25 de marzo de 2026]. [Disponible en: <https://courts.ca.gov/sites/default/files/courts/default/2024-12/btb24-21-3.pdf>].

### **Fuentes legales**

- 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. *Diario Oficial de la Unión Europea*, 2016, L 119. [Consultada 20 de abril de 2026]. [Disponible en: <https://n9.cl/is1sch>].
- UNIÓN EUROPEA. 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 (Reglamento de Inteligencia Artificial). *Diario Oficial de la Unión Europea*, 2024, L. [Consultada 21 de abril de 2026]. [Disponible en: <https://n9.cl/uwg0c>].



## Traducción normativa en sistemas algorítmicos: Insuficiencias del paradigma de derechos humanos por diseño y bases para una teoría fuerte\*

*Normative translation in algorithmic systems: Structural shortcomings of the human rights by design paradigm and foundations for a strong theory*

Johanna Mildred Pinto García\*\*

### RESUMEN

El paradigma de los derechos humanos por diseño presenta deficiencias teóricas al carecer de un modelo eficaz para la traducción normativa del derecho hacia la arquitectura computacional, operando mediante concreciones fragmentarias de los derechos fundamentales. Ante esto, se propone una teoría fuerte de traducción normativa estructurada en cuatro niveles: principios jurídicos, normas operativas, especificaciones técnicas e implementación sistémica. Desde una perspectiva del Sur Global, se concluye que la constitucionalización de la inteligencia artificial exige mecanismos de mediación entre derecho y código sensibles al contexto geopolítico y a las asimetrías epistémicas. La investigación subraya la necesidad de superar las reducciones semánticas actuales para asegurar una gobernanza algorítmica alineada con los derechos humanos.

### PALABRAS CLAVE

Inteligencia artificial; Gobernanza de la inteligencia artificial; Derechos humanos; Sistemas algorítmicos; Traducción normativa.

### ABSTRACT

The human rights by design paradigm faces theoretical flaws due to the lack of an effective model for translating legal norms into computational architecture, operating through fragmentary concretizations of fundamental rights. In response, this paper proposes a robust theory of normative translation structured across four levels: legal principles, operational rules, technical specifications, and systemic implementation. From a Global South perspective, it concludes that the constitutionalization of artificial intelligence demands law-to-code mediation mechanisms sensitive to geopolitical contexts and epistemic asymmetries. This research underscores the necessity of overcoming current semantic reductions to ensure an algorithmic governance strictly aligned with human rights.

### KEYWORDS

Artificial intelligence; Artificial intelligence governance; Human rights; Algorithmic systems; Normative translation.

\* Artículo de investigación

\*\* Licenciada en Derecho por la Universidad Militar Nueva Granada y especialista en Derecho Público por la Universidad Externado de Colombia. Actualmente cursa la Maestría en Derecho, con orientación en Filosofía del Derecho, en la Universidad Nacional Autónoma de México (johanna.pinto@comunidad.unam.mx). <https://orcid.org/0009-0007-2850-5747>.

## SUMARIO

1. Introducción
2. El paradigma «derechos humanos por diseño»: consolidación y alcances normativos
3. Crítica estructural: los déficits del enfoque «por diseño»
4. El problema de la traducción normativa: derecho y código como lenguajes inconmensurables
5. Una perspectiva desde el Sur Global: colonialidad algorítmica y asimetría epistémica
6. Teoría de la traducción normativa en sentido fuerte
7. Operativización a lo largo del ciclo de vida de la inteligencia artificial
8. Implicaciones constitucionales y dimensión institucional
9. Conclusiones
10. Referencias bibliográficas

### 1. INTRODUCCIÓN

En los últimos años, la gobernanza de la inteligencia artificial ha experimentado un desplazamiento significativo: de respuestas regulatorias reactivas hacia enfoques preventivos anclados en el propio diseño tecnológico. Este giro se manifiesta con nitidez en el paradigma de los «derechos humanos por diseño», que propone incorporar principios iusfundamentales desde las fases iniciales del desarrollo de sistemas algorítmicos. Lejos de ser una tendencia marginal, ha sido adoptado por organismos internacionales de primer orden —UNESCO, OCDE, Consejo de Europa— como eje articulador de la regulación emergente, y ha dado lugar a instrumentos de primera generación como el Reglamento (UE) 2024/1689 de Inteligencia Artificial y el Convenio Marco del Consejo de Europa sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de Derecho.

Ese desplazamiento plantea, sin embargo, una pregunta de teoría jurídica que los marcos regulatorios no han respondido satisfactoriamente: ¿es posible traducir derechos humanos en sistemas técnicos que operan mediante formalización, cuantificación y ejecución automatizada? La bibliografía especializada ha producido contribuciones valiosas en equidad algorítmica, explicabilidad y responsabilidad, pero de forma sectorial y sin articular un modelo teórico integrador. A esta limitación interna se suma un déficit de perspectiva geopolítica: la mayor parte de la producción académica y regulatoria sobre derechos humanos e inteligencia artificial proviene del Norte Global y asume, sin cuestionarlo, la validez universal de sus categorías jurídicas. Las voces críticas surgidas desde América Latina, África y Asia alertan sobre un proceso de colonialidad algorítmica mediante el cual se exportan marcos normativos elaborados en contextos de alta capacidad institucional hacia países con realidades radicalmente distintas.

El presente trabajo parte de una premisa diferente a la que domina el debate: el problema no radica en la ausencia de un enfoque basado en derechos humanos, sino en la falta de una teoría de traducción normativa capaz de articular el lenguaje del derecho con el del código, articulación atravesada por asimetrías epistémicas y de poder que condicionan tanto la producción de los sistemas algorítmicos como la de los marcos que buscan regularlos, el trabajo reconstruye el paradigma de los «derechos humanos por diseño» y sus fundamentos recientes; formula una crítica estructural de sus deficiencias;

plantea la traducción normativa como tensión entre lenguajes incommensurables desde la perspectiva del Sur Global; desarrolla sobre esa base una teoría de la traducción normativa en sentido fuerte, operativizada a lo largo del ciclo de vida de la inteligencia artificial; y examina sus implicaciones constitucionales e institucionales.

Desde el punto de vista metodológico, este trabajo se inscribe en la investigación jurídica de carácter teórico-dogmático y cualitativo, con un enfoque crítico-propositivo. No constituye un estudio empírico ni estadístico, sino una investigación documental orientada a la construcción teórica: a partir de la revisión sistemática del estado del arte se emplea el método analítico-conceptual para descomponer las categorías del paradigma dominante y evaluar su consistencia interna; un método comparativo, que contrasta los marcos del Norte Global con los desarrollos latinoamericanos y africanos; y un método crítico de raíz decolonial, que interroga el lugar de enunciación desde el cual se produce el conocimiento sobre gobernanza algorítmica. Sobre esa base, la investigación asume una orientación reconstructiva y propositiva, cuyo estatuto epistemológico se precisa en el apartado VI.

## 2. EL PARADIGMA «DERECHOS HUMANOS POR DISEÑO»: CONSOLIDACIÓN Y ALCANCES NORMATIVOS

La pregunta de cómo incorporar valores normativos en los artefactos técnicos precede, con mucho, a la inteligencia artificial, por lo que conviene rastrear tres tradiciones intelectuales que confluyeron, en las últimas tres décadas, en el enfoque «derechos humanos por diseño». La primera arranca de la privacidad por diseño, formulada por Ann Cavoukian, entonces Comisionada de Información y Privacidad de Ontario, sobre una premisa de hondas implicaciones prácticas: la protección de la privacidad no puede garantizarse solo mediante el cumplimiento posterior al daño, sino que debe inscribirse en la arquitectura de los sistemas desde su concepción.<sup>1</sup> Esa orientación se articuló en siete principios fundacionales reconocidos internacionalmente en 2010 y que la Unión Europea incorporó al Reglamento General de Protección de Datos de 2016.<sup>2</sup> Su relevancia radica en haber demostrado que era técnicamente posible —y jurídicamente exigible— trasladar un derecho fundamental al interior de la infraestructura computacional, sin reducirlo a una declaración de intenciones.

La segunda es el diseño sensible a los valores (Value Sensitive Design o VSD), inaugurado por Batya Friedman a mediados de los años noventa.<sup>3</sup> y desarrollado junto con Peter Kahn, Alan Borning y David Hendry.<sup>4</sup> El VSD sostiene que los valores humanos con trascendencia ética —privacidad, justicia, autonomía, bienestar— quedan necesariamente implicados en cada decisión de diseño, de forma consciente o inadvertida; de ahí

<sup>1</sup>CAVOUKIAN, Ann, 2009. *Privacy by Design: The 7 Foundational Principles*. Toronto: Information and Privacy Commissioner of Ontario.

<sup>2</sup>UNIÓN EUROPEA, 2016. 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. Diario Oficial de la Unión Europea, L 119, pp. 1-88; sobre la incorporación de la privacidad por diseño, véase art. 25 del citado Reglamento.

<sup>3</sup>FRIEDMAN, Batya; NISSENBAUM, Helen, 1996. «Bias in Computer Systems». *ACM Transactions on Information Systems* [en línea], vol. 14, núm. 3, pp. 330-347. DOI: <https://doi.org/10.1145/230538.230561>.

<sup>4</sup>FRIEDMAN, Batya; HENDRY, David G., 2019. *Value Sensitive Design: Shaping Technology with Moral Imagination*. Cambridge (MA): MIT Press.

que la pregunta pertinente no sea si los sistemas incorporan valores, sino qué valores incorporan, para quién y con qué consecuencias.<sup>5</sup> Su metodología ofrece un precedente procedimental directo para el modelo que se desarrolla más adelante. La tercera es la gobernanza de la arquitectura tecnológica, cuya formulación más influyente se debe a Lawrence Lessig, quien sostuvo que el código informático, al determinar qué conductas son posibles y cuáles están técnicamente excluidas, opera como una modalidad de regulación tan efectiva como el derecho positivo, las normas sociales o el mercado.<sup>6</sup> La tesis, condensada en la fórmula «el código es ley», no afirma la identidad entre código y derecho, sino que las decisiones de arquitectura tienen consecuencias normativas de primer orden y que quienes escriben el código asumen un rol análogo al del legislador.<sup>7</sup>

Las tres comparten la tesis de que los sistemas tecnológicos no son neutros y de que sus consecuencias normativas deben gestionarse desde las fases más tempranas del desarrollo; su extensión al campo de la inteligencia artificial da lugar al paradigma «derechos humanos por diseño». Entre 2023 y 2025, el paradigma alcanzó una consolidación institucional sin precedentes. El Reglamento (UE) 2024/1689 constituye el primer marco jurídico horizontal y vinculante sobre inteligencia artificial: adopta un enfoque basado en el riesgo y clasifica los sistemas en cuatro categorías según su peligro para la salud, la seguridad y los derechos fundamentales; los de alto riesgo quedan sujetos a evaluación de conformidad, registro, transparencia y supervisión humana, mientras que el artículo 9 exige un sistema de gestión de riesgos a lo largo del ciclo de vida y el artículo 10 impone requisitos de calidad de los datos para reducir sesgos discriminatorios. El Convenio Marco del Consejo de Europa, adoptado en mayo de 2024 y abierto a la firma en septiembre, es el primer tratado internacional vinculante sobre la materia y, a diferencia del Reglamento, ancla la gobernanza en los estándares consolidados del derecho internacional de los derechos humanos: legalidad, necesidad, proporcionalidad, transparencia, responsabilidad y no discriminación. La Resolución A/RES/78/265 de la Asamblea General de Naciones Unidas, aprobada por consenso en marzo de 2024, reafirma que los derechos humanos deben respetarse a lo largo de todo el ciclo de vida de los sistemas e insta a los Estados a abstenerse de utilizar aquellos incompatibles con el derecho internacional de los derechos humanos.

El impulso internacional ha encontrado eco en América Latina, aunque de forma heterogénea y sin que ningún país cuente todavía con una ley general vigente. Brasil registra el avance más significativo con el Proyecto de Ley 2.338/2023, aprobado por su Senado en diciembre de 2024, que clasifica los sistemas por niveles de riesgo, exige evaluaciones de impacto algorítmico e impone obligaciones de transparencia.<sup>8</sup> Colombia protagoniza el proceso más dinámico con el Proyecto de Ley 043/2025, que adopta

<sup>5</sup>FRIEDMAN, Batya; KAHN, Peter H.; BORNING, Alan, 2006. «Value Sensitive Design and Information Systems». En: ZHANG, P.; GALLETTA, D. (eds.). *Human-Computer Interaction in Management Information Systems: Foundations*. Nueva York: M.E. Sharpe, pp. 348-372.

<sup>6</sup>LESSIG, Lawrence, 1999. *Code and Other Laws of Cyberspace*. Nueva York: Basic Books.

<sup>7</sup>LESSIG, Lawrence, 1999. «The Law of the Horse: What Cyberlaw Might Teach». *Harvard Law Review* [en línea], vol. 113, núm. 2, pp. 501-546.

<sup>8</sup>BRASIL, SENADO FEDERAL, 2024. Projeto de Lei núm. 2.338/2023: Marco Legal da Inteligência Artificial. Aprobado por el Senado Federal el 10 de diciembre de 2024; en trámite ante la Cámara de Diputados. Véase también: DERECHOS DIGITALES, 2024. «Regulación de la IA en Brasil: media sanción para un proyecto de ley que seguirá en discusión».

un enfoque basado en riesgos, designa a MinCiencias como Autoridad Nacional, prevé evaluaciones de impacto en derechos fundamentales y reconoce como referentes la UNESCO, la OCDE y el Reglamento europeo.<sup>9</sup> México presenta la situación más compleja: desde 2020 se han presentado más de cincuenta iniciativas, ninguna convertida en ley general; la más sustantiva es la Iniciativa de Ley Federal para el Desarrollo Ético, Soberano e Inclusivo de la Inteligencia Artificial (abril de 2025), que sitúa la dignidad humana como eje del marco, crea un Registro y un Consejo Nacional e incorpora protecciones para menores, aún en comisiones.<sup>10</sup>

El patrón regional ilustra la limitación estructural del paradigma dominante: todos estos instrumentos adoptan la taxonomía de riesgos y los principios del enfoque «derechos humanos por diseño», pero ninguno especifica los mecanismos mediante los cuales esos principios se traducen en obligaciones computacionalmente operativas. La recepción del modelo europeo en contextos de menor capacidad institucional agrava, además, el riesgo de colonialidad normativa que se examina en la sección V: se importan las categorías sin importar los presupuestos técnicos, institucionales y epistémicos que las hacen funcionar. En el plano académico, el paradigma ha cristalizado en tres líneas convergentes, valiosas pero insuficientes. La desarrollada por Alessandro Mantelero y colaboradores se centra en los marcos de evaluación del impacto en derechos humanos (HRIA) y aporta una metodología basada en evidencia empírica, así como contribuciones al diseño de la plantilla FRIA del Reglamento europeo.<sup>11</sup> La impulsada por Yeung, Howes y Pogrebna propone una gobernanza por diseño deliberativo que reivindica el derecho internacional de los derechos humanos frente al ethics washing, pero permanece en un nivel de abstracción que no desciende a la operacionalización computacional.<sup>12</sup> La proveniente de Kate Jones examina la relación entre ética de la IA y derechos humanos en la política global, abogando por una gobernanza internacional coherente, sin proporcionar una teoría de cómo los principios jurídicos se incorporan en la arquitectura algorítmica.<sup>13</sup> Las tres carecen de una teoría de la traducción normativa que especifique cómo los derechos fundamentales pueden convertirse en restricciones computacionales sin pérdida sustancial de contenido.

<sup>9</sup>COLOMBIA, MINISTERIO DE CIENCIA, TECNOLOGÍA E INNOVACIÓN, 2025. Proyecto de Ley núm. 043/2025 Senado: «Por medio de la cual se regula la Inteligencia Artificial en Colombia para garantizar su desarrollo ético, responsable, competitivo e innovador, y se dictan otras disposiciones». Radicado el 28 de julio de 2025; mensaje de urgencia el 8 de septiembre de 2025.

<sup>10</sup>MÉXICO, CÁMARA DE DIPUTADOS, 2025. Iniciativa de Ley Federal para el Desarrollo Ético, Soberano e Inclusivo de la Inteligencia Artificial. 22 de abril de 2025; en trámite en comisiones. México D.F.: Palacio Legislativo de San Lázaro.

<sup>11</sup>MANTELERO, Alessandro; ESPOSITO, Maria Samantha, 2024. «An evidence-based methodology for human rights impact assessment (HRIA) in the development of AI data-intensive systems». *Computer Law & Security Review* [en línea], vol. 41, art. 105561.

<sup>12</sup>YEUNG, Karen; HOWES, Andrew; POGREBNA, Ganna, 2020. «AI Governance by Human Rights-Centered Design, Deliberation, and Oversight». En: FLORIDI, L. et al. (eds.). *The Oxford Handbook of AI Ethics*. Oxford: Oxford University Press. ; YEUNG, Karen, 2025. *AI Equality by Design, Deliberation and Oversight*. RAI UK.

<sup>13</sup>JONES, Kate, 2023. *AI Governance and Human Rights: Resetting the Relationship*. Londres: Chatham House; CHATHAM HOUSE, 2024. *Artificial Intelligence and the Challenge for Global Governance* [en línea].

### 3. CRÍTICA ESTRUCTURAL: LOS DÉFICITS DEL ENFOQUE «POR DISEÑO»

El paradigma «derechos humanos por diseño», pese a su consolidación, padece déficits estructurales que limitan su capacidad de producir efectos jurídicamente vinculantes y computacionalmente operativos. El primer déficit es la fragmentación conceptual. La bibliografía y los marcos regulatorios abordan de manera desarticulada dimensiones intrínsecamente interdependientes —equidad algorítmica, explicabilidad, responsabilidad, privacidad— sin un marco integrador de sus relaciones lógicas y normativas. Las consecuencias son de primera magnitud: como ha demostrado la investigación en aprendizaje automático, los objetivos de equidad y de exactitud estadística suelen estar en tensión, y las definiciones técnicas de equidad son con frecuencia mutuamente incompatibles<sup>14</sup> un sistema puede satisfacer la paridad de tasas de falsos positivos entre grupos y vulnerar a la vez la paridad predictiva. La elección entre métricas no es técnicamente neutral, sino una decisión de fondo que exige fundamentación jurídica<sup>15</sup> pese a ello, los marcos actuales la delegan en los equipos de desarrollo sin suministrar los criterios normativos necesarios.

El segundo déficit, de mayor calado, es la reducción semántica: los conceptos nucleares del derecho internacional de los derechos humanos son sometidos a operacionalizaciones que los reducen a magnitudes computacionales, con pérdida sustancial de contenido normativo. El caso más llamativo es la igualdad. En el derecho comprende dimensiones formales y materiales, abarca la prohibición de la discriminación directa y la protección frente a la indirecta, exige el análisis de los efectos estructurales de las prácticas institucionales y admite diferenciaciones justificadas por proporcionalidad. En la bibliografía sobre equidad algorítmica, en cambio, se traduce en una de varias decenas de métricas, cada una de las cuales captura solo una dimensión parcial del concepto jurídico. Algo análogo ocurre con el debido proceso, reducido en el plano computacional a «explicabilidad»: la capacidad de describir los factores que influyeron en una decisión automatizada. Los sistemas algorítmicos asumen un nivel de formalización estructuralmente incompatible con la naturaleza interpretativa del derecho. La explicabilidad técnica no equivale a la justificación jurídica, que requiere no solo descripción, sino razonamiento normativo, consideración del contexto y apertura a la impugnación. En este sentido, Niamh Kinchin ha identificado una «brecha procedimental» en la justicia algorítmica: mientras las métricas de equidad responden a la justicia distributiva, no abordan la dimensión procedimental, v. gr. consentimiento informado, derecho a ser escuchado antes de que la decisión produzca efectos.

El tercer déficit es la ausencia de un modelo sistemático de traducción normativa. Lessig estableció que «el código es ley» en el sentido de que la arquitectura regula conductas de manera análoga al derecho positivo<sup>16</sup> Balkin extendió el análisis a la infraestructura de la información y Hildebrandt ha explorado la noción de derecho como tecnología para examinar las relaciones entre formalización jurídica y computacional.

<sup>14</sup>CHOULDCHOVA, Alexandra, 2017. «Fair prediction with disparate impact: A study of bias in recidivism prediction instruments». *Big Data* [en línea], vol. 5, núm. 2, pp. 153-163.

<sup>15</sup>GIANNOPOULOS, Giorgos et al., 2024. «Fairness in AI: challenges in bridging the gap between algorithms and law». En: 2024 IEEE 40th International Conference on Data Engineering Workshops (ICDEW). IEEE, pp. 217-225.

<sup>16</sup>LESSIG, L., 1999a, op. cit. Véase también BALKIN, Jack M., 2018. «Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation». *UC Davis Law Review*, vol. 51, núm. 3, pp. 1149-1210.

Ninguno de estos enfoques ofrece, sin embargo, un modelo que especifique los mecanismos, instancias y procedimientos mediante los cuales una norma jurídica puede convertirse en restricción computacional sin perder su contenido esencial. El trabajo reciente sobre «proceso debido por diseño» reconoce la posibilidad de alinear la tecnología con marcos socio-normativos<sup>17</sup>, pero permanece en la afirmación programática sin desarrollar la teoría que la haría posible. El aporte de este artículo consiste en colmar ese vacío.

#### 4. EL PROBLEMA DE LA TRADUCCIÓN NORMATIVA: DERECHO Y CÓDIGO COMO LENGUAJES INCONMENSURABLES

Para comprender la raíz del déficit estructural conviene introducir una distinción analítica que la filosofía del derecho computacional ha comenzado a perfilar: la diferencia entre inconmensurabilidad semántica e inconmensurabilidad operativa.<sup>18</sup> La inconmensurabilidad semántica alude a la diferencia entre los criterios de validez, justificación y aplicación del discurso del derecho y del discurso del código. El derecho, como práctica normativa, opera mediante conceptos deliberadamente abiertos (igualdad, proporcionalidad, debido proceso) cuyo contenido se determina a través de procesos interpretativos contextuales y dialógicos. Esta apertura semántica constitutiva —lo que Hart denominó la textura abierta del derecho—<sup>19</sup> no es un defecto que deba corregirse, sino una condición funcional indispensable: esa vaguedad permite al derecho adaptarse a situaciones no previstas por el legislador, incorporar el desarrollo moral de la comunidad y responder a la singularidad de cada caso, y adquiere relevancia decisiva cuando el derecho debe articularse con la formalización computacional.<sup>20</sup> Los sistemas computacionales, en cambio, se rigen por la formalización, la cuantificación y la ejecución determinista. Para que una instrucción sea ejecutable es necesario traducirla a un lenguaje formal suficientemente preciso: un umbral numérico, una función de pérdida, una restricción de optimización. Incluso en el aprendizaje automático, cuyo comportamiento no está explícitamente programado, la lógica subyacente permanece anclada en la minimización de funciones de error formalizadas. Como han señalado Cooper, Frankle y De Sa, el propio no determinismo del aprendizaje automático escapa al marco tradicional del «código como ley» basado en reglas deterministas,<sup>21</sup> lo que plantea desafíos inéditos a la regulación.

La divergencia entre derecho y código puede iluminarse mediante una analogía que aquí se propone con la tesis de la traducción radical de Willard Van Orman Quine.<sup>22</sup>

<sup>17</sup>FISCHMAN-AFORI, Orit, 2025. Due process by design: Enhancing fairness and trust in AI decision-making [en línea]. SSRN.

<sup>18</sup>FLÓREZ, Diego Tarapués, 2020. «¿Hay inconmensurabilidad tecnológica?». Cuadernos de Filosofía [en línea], núm. 38, pp. 119-141.

<sup>19</sup>Sobre la textura abierta (open texture) del lenguaje jurídico como propiedad estructural y no como defecto corregible, véase HART, Herbert Lionel Adolphus, 1963. El concepto de derecho. Traducido por Genaro R. Carrió. Buenos Aires: Abeledo-Perrot, cap. VII, §1, pp. 155-169.

<sup>20</sup>HILDEBRANDT, Mireille, 2020. «A philosophy of technology for computational law». En: Is Law Computable? Critical Perspectives on Law and Artificial Intelligence. Oxford: Hart Publishing, pp. 23-46.

<sup>21</sup>COOPER, A. Feder; FRANKLE, Jonathan; DE SA, Christopher, 2022. «Non-Determinism and the Lawlessness of Machine Learning Code». En: Proceedings of the 2022 Symposium on Computer Science and Law. Nueva York: ACM, pp. 1-12.

<sup>22</sup>QUINE, Willard Van Orman, 1968. Palabra y objeto. Traducido por Manuel Sacristán. Barcelona: Labor, cap. 2, §7

Conviene precisar que Quine no se ocupó del problema jurídico-computacional; su tesis se refiere a la indeterminación de la traducción entre lenguas naturales, y es su estructura argumental —no una afirmación suya sobre la inteligencia artificial— la que se traslada aquí. Quine mostró que, ante una lengua desconocida, un traductor podría elaborar múltiples manuales compatibles con toda la evidencia observacional pero radicalmente distintos entre sí: la expresión «gavagai», proferida ante un conejo, podría significar «conejo», «parte no separable de conejo», «estadio temporal de conejo» o «manifestación de la conejidad», sin que la observación empírica permita decidir la interpretación correcta.<sup>23</sup>

«De manera análoga, ante el mandato de «no discriminación» un equipo de desarrollo encuentra una pluralidad de métricas de equidad —paridad demográfica, igualdad de oportunidades, calibración— que son otras tantas «traducciones» plausibles del principio, a menudo mutuamente excluyentes; elegir una no es un acto técnico neutral, sino de interpretación normativa, hoy realizado sin fundamentación jurídica explícita ni control democrático. Si el diagnóstico de Quine parece abocarnos a la imposibilidad de una traducción perfecta, la filosofía de Paul Ricœur ofrece una vía constructiva para gestionar esa tensión sin negarla. Ricœur propone un «paradigma de la traducción» que renuncia al ideal inalcanzable de la equivalencia exacta y asume la traducción como trabajo continuo de mediación y «hospitalidad lingüística»<sup>24</sup>. Sobre esa base, Cassin precisa que traducir no es hallar un doble perfecto del original, sino construir un equivalente funcional que, aun imperfecto y sujeto a pérdida de significado, hace posible la comunicación entre universos simbólicos distintos.<sup>25</sup> La perspectiva es decisiva: la traducción de un derecho humano a código siempre implicará cierta reducción semántica, pero esa reducción puede gestionarse, documentarse y someterse a escrutinio mediante procedimientos adecuados.

Se revela así la verdadera carencia del paradigma: no la inconmensurabilidad semántica —inevitable—, sino la operativa, esto es, la ausencia de un proceso institucionalizado, trazable y auditable que guíe la mediación entre derecho y código. En la práctica, ello se traduce en decisiones de diseño opacas, imposibles de rastrear hasta su fundamento normativo y rígidas frente al caso concreto. La teoría de la traducción normativa en sentido fuerte que aquí se propone no aspira, pues, a una conversión algorítmica perfecta de los derechos humanos, sino a subsanar la inconmensurabilidad operativa mediante un modelo procedimental que, al modo ricœuriano, asuma la pérdida inevitable de contenido semántico pero la gestione con trazabilidad, fundamentación de cada paso y apertura a la revisión. Se abandona así la equivalencia perfecta por una exigencia más rigurosa y democrática: la de una traducción responsable y justificable.

<sup>23</sup>Ibidem, cap. 2, §12.

<sup>24</sup>RICŒUR, Paul, 2001. «El paradigma de la traducción». *Revista de Occidente* [en línea], núm. 240, pp. 160-177.

<sup>25</sup>CASSIN, Barbara. La traducción como paradigma para las humanidades. Traducido por Ana Olivares Muñoz-Ledo. *Interpretatio. Revista de hermenéutica* [en línea]. 2022, vol. 7, n.º 2, pp. 167-197.

## 5. UNA PERSPECTIVA DESDE EL SUR GLOBAL: COLONIALIDAD ALGORÍTMICA Y ASIMETRÍA EPISTÉMICA

La crítica de las secciones anteriores es interna al paradigma: identifica sus déficits lógicos y teóricos sin cuestionar el lugar de enunciación desde el que fue construido. Existe una crítica de mayor alcance, articulada desde el Sur Global, según la cual el propio paradigma constituye una forma de ejercicio del poder epistémico. En el debate contemporáneo ha tomado fuerza el concepto de colonialismo digital<sup>26</sup> para describir las relaciones de extracción, dependencia y subordinación epistémica que caracterizan el despliegue global de la inteligencia artificial. La descolonización de la gobernanza algorítmica exige tanto dismantelar los marcos cognitivos eurocéntricos como reconocer las disparidades de acceso y representación en los sistemas de datos.<sup>27</sup> El concepto de colonialidad del poder, formulado por Aníbal Quijano y aplicado a los sistemas de inteligencia artificial por Mohamed, Png e Isaac<sup>28</sup> permite comprender cómo el desarrollo de la IA reproduce una cadena de suministro colonial sostenida en una división internacional injusta del trabajo digital y en la subordinación histórica de los países del Sur Global.

Este dominio tiene consecuencias directas sobre la traducción normativa: cuando un sistema entrenado mayoritariamente con datos de Estados Unidos o Europa occidental se despliega en contextos latinoamericanos, africanos o asiáticos, no solo produce resultados estadísticamente inadecuados para esas poblaciones, sino que impone categorías jurídicas y valores que no son los de las comunidades afectadas. La investigación empírica en India y Brasil muestra que los sistemas de traducción, evaluación y recomendación rinden sistemáticamente peor con lenguas regionales, discursos culturalmente situados y tradiciones retóricas no occidentales.<sup>29</sup>

El marco regulatorio latinoamericano avanza de manera desigual, con Brasil, Chile y Colombia a la vanguardia, mientras la mayoría de los marcos nacionales reproducen acríticamente los modelos europeos o angloamericanos sin adaptarlos a las realidades locales.<sup>30</sup> En ciertos ámbitos, los Estados de la región convergen con los estándares internacionales sin disponer de los mecanismos institucionales para implementarlos con eficacia.<sup>31</sup> Organizaciones como TEDIC (Paraguay) y Derechos Digitales han articulado propuestas ante la Comisión Interamericana de Derechos Humanos para que los instrumentos del sistema interamericano incorporen una perspectiva de ciclo de vida de las

<sup>26</sup>El concepto de colonialismo digital fue formulado por COULDRY, Nick; MEJIAS, Ulises A., 2019. *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. Stanford: Stanford University Press; y KWET, Michael, 2019. «Digital Colonialism: US Empire and the New Imperialism in the Global South». *Race & Class* [en línea], vol. 60, núm. 4, pp. 3-26. DOI: <https://doi.org/10.1177/0306396818823172>.

<sup>27</sup>AYANA, G. et al., 2024. «Decolonizing Global AI Governance: Assessment of the State of Decolonized AI Governance in Sub-Saharan Africa». *Royal Society Open Science* [en línea], vol. 11, núm. 8, art. 231994.

<sup>28</sup>QUIJANO, Anibal, 2000. «Colonialidad del poder y clasificación social». *Journal of World-Systems Research*, vol. 6, núm. 2, pp. 342-386. Su aplicación a los sistemas de inteligencia artificial se desarrolla en MOHAMED, Shakir; PNG, Marie-Therese; ISAAC, William, 2020. «Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence». *Philosophy & Technology*, vol. 33, núm. 4, pp. 659-684.

<sup>29</sup>AHMED, Samar A., 2026. «Algorithmic Dependence and Digital Colonialism: A Conceptual Framework for AI in Education and Knowledge Systems of the Global South». *Frontiers in Education* [en línea].

<sup>30</sup>ULLOA, Mónica, 2024. *Avances en la regulación de la Inteligencia Artificial en América Latina* [en línea]. ORCG. Disponible en: <https://orcg.info/articulos/avances-en-la-regulacin-de-la-inteligencia-artificial-en-amrica-latina>

<sup>31</sup>CONTRERAS, Pablo, 2024. «Convergencia internacional y caminos propios: regulación de la inteligencia artificial en América Latina». *Actualidad Jurídica Iberoamericana*, agosto 2024, núm. 21, pp. 468-493.

políticas tecnológicas, atendiendo a los efectos sistémicos del racismo algorítmico sobre poblaciones históricamente marginadas.<sup>32</sup> La Iniciativa Latinoamericana de Justicia de Datos ha demandado el reconocimiento de la titularidad colectiva sobre los datos y la soberanía territorial sobre la información como contrapeso a los modelos individualistas —de cuño anglosajón— que dominan los marcos de privacidad.<sup>33</sup>

Desde el continente africano, la Estrategia Continental de Inteligencia Artificial de la Unión Africana (julio de 2024) propone un marco de desarrollo inclusivo y centrado en las personas, articulado con la Agenda 2063, que resiste la importación acrítica de estándares OCDE o UE.<sup>34</sup> El índice global de «Inteligencia Artificial Responsable», con liderazgo del Global Center on AI Governance, define la IA responsable como aquella que respeta los derechos humanos y los principios éticos en cada etapa del ciclo de vida y de la cadena de valor, subrayando la necesidad de marcos que reflejen la diversidad sociolingüística y cultural del continente.<sup>35</sup> En el plano filosófico, la «colonialidad algorítmica» señala que los sistemas de IA no son herramientas neutras, sino artefactos que materializan jerarquías epistémicas: priorizan determinadas formas de conocer, sujetos de preocupación y valores políticos, todos ellos elaborados en el Norte Global.<sup>36</sup> La obra de Boaventura de Sousa Santos sobre las epistemologías del Sur permite problematizar el paradigma desde sus fundamentos: la epistemología occidental dominante produce lo que denomina «epistemicidio», la supresión de formas de conocimiento alternativas mediante jerarquías cognitivas que presentan el saber occidental como universal y objetivo.<sup>37</sup> Su «ecología de saberes» exige, en el plano de la traducción normativa, que las normas operativas del modelo no sean meras transcripciones de principios elaborados en otro contexto, sino resultado de procesos deliberativos que incluyan las perspectivas de las comunidades afectadas.<sup>38</sup> La implicación metodológica es concreta: la identificación de los principios jurídicos aplicables no puede ser una operación técnica cerrada, sino un proceso abierto a la participación de las comunidades afectadas, en particular las indígenas, afrodescendientes y lingüísticamente marginadas, que son las más vulnerables a los efectos discriminatorios de los sistemas algorítmicos.

## 6. TEORÍA DE LA TRADUCCIÓN NORMATIVA EN SENTIDO FUERTE

Frente a los déficits identificados, y con la perspectiva crítica de la sección anterior, se propone la teoría de la traducción normativa en sentido fuerte: un modelo que articula la especificación progresiva del contenido normativo de los derechos fundamentales a través de cuatro niveles conectados mediante procedimientos explícitos de derivación

<sup>32</sup>TEDIC, 2025. Inteligencia artificial y derechos humanos: aportes a la 192.ª sesión de la CIDH [en línea].

<sup>33</sup>RAHA, Saptaparni; PATRA, Joydeb, 2024. «Data Colonialism, Algorithmic Justice, and the Law: A Global South Perspective». *Anusandhanvallari*, vol. 1, núm. 1, pp. 105-113.

<sup>34</sup>UNIÓN AFRICANA, 2024. Continental Artificial Intelligence Strategy [en línea]. Addis Abeba.Unión Africana.

<sup>35</sup>GARCÍA, Juan Manuel, 2024. ¿Qué sabemos realmente sobre el uso de IA en América Latina? Revisión de informes, índices y estudios (2019-2023) [en línea]. Santiago de Chile: Derechos Digitales.

<sup>36</sup>BIRHANE, Abeba, 2020. «Algorithmic Colonization of Africa». *SCRIPTed*, vol. 17, núm. 2, pp. 389-409;

<sup>37</sup>DE SOUSA SANTOS, Boaventura, et al. *Epistemologías del Sur - Epistemologias do Sul*. Coordinación de Maria Paula Meneses y Karina Andrea Bidaseca. 1.ª ed. Ciudad Autónoma de Buenos Aires: CLACSO; Coímbra: Centro de Estudos Sociais - CES, 2018. [en línea].

<sup>38</sup>Ibidem, pp. 21-65

y trazabilidad. El modelo es deliberadamente contextual: no ofrece un algoritmo de decisión independiente del contexto jurídico, cultural y geopolítico, sino un marco procedimental que estructura la mediación entre derecho y código. La «traducción fuerte» se distingue de las aproximaciones débiles predominantes. La traducción débil formula principios generales de derechos humanos como criterios de diseño sin especificar los mecanismos de conversión en requisitos técnicos concretos. La fuerte, en cambio, exige una cadena de derivación explícita y auditable desde el principio jurídico hasta la implementación sistémica, con justificación en cada paso.

El primer nivel comprende los derechos fundamentales y principios jurídicos aplicables al sistema, identificados mediante análisis jurídico contextual. No puede predeterminarse de manera genérica: la selección depende del ámbito de aplicación, la naturaleza de los efectos sobre las personas y el contexto regulatorio. Requiere al menos tres operaciones: determinar el ámbito material de los derechos en juego; analizar las obligaciones positivas y negativas que generan para el Estado y los actores privados; y evaluar los conflictos entre derechos y los criterios de ponderación aplicables. En el Reglamento (UE) 2024/1689 esta operación ha sido parcialmente institucionalizada mediante la evaluación del impacto en los derechos fundamentales para los sistemas de alto riesgo.<sup>39</sup> Desde la perspectiva del Sur Global, el nivel debe incorporar una auditoría epistémica: ¿los principios seleccionados capturan las concepciones de dignidad, igualdad y justicia vigentes en las comunidades afectadas, o reproducen sin más el canon liberal-occidental?

Las normas operativas son formulaciones derivadas de los principios seleccionados que especifican sus exigencias en el contexto particular del sistema, con concreción suficiente para orientar decisiones de diseño sin ser aún especificaciones técnicas. Por ejemplo, el principio de no discriminación puede derivar, para un sistema de selección de personal en un país latinoamericano con historia de discriminación racial y de género, en normas como «el sistema no puede producir tasas de selección estadísticamente inferiores para personas afrodescendientes o indígenas sin justificación objetiva» o «el sistema debe justificar cada decisión individual en términos comprensibles para la persona afectada, en su lengua materna si así lo solicita». Tales normas mantienen la carga normativa del principio y orientan la selección de métricas en el nivel siguiente. El tercer nivel es la formalización computacional: elección de métricas de equidad, umbrales de aceptabilidad, mecanismos de explicabilidad y restricciones de optimización. La contribución del modelo es doble: la selección de especificaciones técnicas no puede ser una decisión puramente técnica, sino que debe fundarse en las normas operativas del Nivel II y, por tanto, en los principios del Nivel I; y las especificaciones deben ser trazables en sentido ascendente, de modo que, dado un comportamiento del sistema, sea posible rastrear las especificaciones que lo generaron, las normas que las fundamentan y los principios que las justifican. El cuarto nivel comprende la implementación efectiva —código, arquitectura del modelo, entrenamiento, evaluación, supervisión humana y auditoría—, donde las especificaciones se vuelven comportamiento observable. La traducción fuerte exige aquí trazabilidad completa con los niveles superiores: documentar cada decisión de implementación y su fundamento; verificar mediante evaluación el

<sup>39</sup>UNIÓN EUROPEA, 2024. Reglamento (UE) 2024/1689, op. cit., art. 27.

cumplimiento de las especificaciones; revisar periódicamente para actualizar el sistema cuando cambien las normas operativas o se detecten incumplimientos; y preservar una supervisión humana capaz de intervenir cuando los resultados resulten incompatibles con los principios jurídicos.

Conviene precisar el estatuto epistemológico de la propuesta, que se adscribe a las teorías normativo-procedimentales de carácter reconstructivo. Es normativa: no constata un estado de cosas, sino que prescribe los procedimientos que deberían seguirse para que la mediación entre lo jurídico y lo computacional sea legítima, auditable y respetuosa del contenido esencial de los derechos; opera como un ideal regulativo, análogo al de los estándares de derechos humanos en el derecho internacional público, que orienta el diseño institucional y técnico hacia un horizonte de legitimidad aun cuando su realización plena sea asintótica.<sup>40</sup> Es procedimental: no predetermina resultados sustantivos —no fija ex ante qué métrica de equidad ni qué umbral—, sino que estructura los procedimientos para que esas decisiones estén fundamentadas, sean trazables y queden sujetas a escrutinio público, en línea con la metodología HUDERIA del Consejo de Europa, que ofrece un proceso estructurado para identificar y mitigar impactos adversos sin predeterminar el contenido de las medidas.<sup>41</sup> Y es reconstructiva: sistematiza y dota de coherencia interna a prácticas ya existentes de forma fragmentaria —la evaluación de impacto del artículo 27 del Reglamento (UE) 2024/1689, las metodologías de HRIA basadas en evidencia<sup>42</sup>, los marcos de auditoría algorítmica participativa del Algorithmic Impact Methods Lab<sup>43</sup> y los esfuerzos por construir estándares de evaluación de modelos de lenguaje desde los derechos humanos<sup>44</sup>—, elevándolas a un nivel de sistematicidad del que hoy carecen.

Que la teoría sea normativo-procedimental y no directamente falsable por los procedimientos de las ciencias naturales no la exime de abrir vías de validación operativa. Futuras líneas de investigación deberían aplicar la matriz a sistemas concretos desplegados en contextos sensibles a los derechos fundamentales —selección de personal en el sector público latinoamericano, asignación de beneficios sociales—. Asimismo, ejercicios de trazabilidad inspirados en el marco HumRights-Bench, que evalúa la capacidad de los sistemas para identificar violaciones de derechos humanos, recordar marcos legales aplicables y aplicar disposiciones apropiadas,<sup>45</sup> podrían adaptarse para verificar si es posible rastrear las decisiones de diseño hasta principios jurídicos fundamentados. La colaboración con organismos que desarrollen evaluaciones de impacto algorítmico, según el modelo de los programas piloto de HUDERIA,<sup>46</sup> es una vía particularmente prometedora.

<sup>40</sup>LÓPEZ OLVERA, Carmen Patricia, 2023. «Epistemología jurídica aplicada». En: *Epistemología jurídica aplicada* [en línea]. México: Universidad Nacional Autónoma de México, Instituto de Investigaciones Jurídicas, pp. 7-42.

<sup>41</sup>SPINDLOW, Sam, 2024. «Council of Europe adopts Turing-developed human rights risk and impact assessment for AI systems». The Alan Turing Institute [en línea], 5 de diciembre de 2024.

<sup>42</sup>MANTELERO, Alessandro; ESPOSITO, Maria Samantha, 2024, op. cit.

<sup>43</sup>DATA & SOCIETY, 2024. Algorithmic Impact Methods Lab [en línea]. Data & Society.

<sup>44</sup>THAIS, Savannah et al., 2025. HumRights-Bench: The First Human Rights Standard for AI Evaluation [en línea]. Women at the Table.

<sup>45</sup>Idem.

<sup>46</sup>SPINDLOW, Sam, 2024, op. cit.

## 7. OPERATIVIZACIÓN A LO LARGO DEL CICLO DE VIDA DE LA INTELIGENCIA ARTIFICIAL

La traducción normativa en sentido fuerte no puede aplicarse como un control puntual previo al despliegue: debe integrarse en cada fase del ciclo de vida. El Reglamento (UE) 2024/1689 adopta un enfoque de ciclo de vida compatible con el modelo, aunque sin la profundidad teórica que exige una traducción genuinamente fuerte. La aprobación de la metodología HUDERIA por el Comité de Ministros del Consejo de Europa (febrero de 2025), junto con su modelo complementario COBRA para el análisis de riesgos basado en el contexto, ofrece una guía estructurada para evaluar riesgos e impactos a lo largo de todo el ciclo de vida<sup>47</sup> aun así, estos instrumentos requieren complementarse con una teoría explícita de la traducción que garantice la trazabilidad entre principios jurídicos y decisiones técnicas.

En la fase de diseño, la traducción fuerte exige una evaluación previa del impacto en los derechos humanos que identifique los principios aplicables y derive las normas operativas correspondientes. Esta exigencia encuentra respaldo en instrumentos metodológicos recientes: el Toolkit de Evaluación de Impacto en Derechos Humanos para la IA del PNUD ofrece una metodología estructurada del diseño al despliegue.<sup>48</sup> y la guía FRIA del European Center for Not-for-Profit Law (ECNL) y el Instituto Danés de Derechos Humanos subraya que una FRIA adecuada no puede reducirse a una lista de verificación, sino que debe constituir un mecanismo de gobernanza que facilite una discusión estructurada sobre los impactos adversos.<sup>49</sup> La fase de diseño debe incluir además el análisis de sesgos en los datos de entrenamiento y la verificación de la representatividad de los conjuntos respecto de los grupos protegidos: la investigación reciente ha demostrado que los sesgos codificados en los datos son uno de los principales impulsores de la discriminación algorítmica, y que la subrepresentación de poblaciones vulnerables puede generar efectos discriminatorios que no siempre se ajustan a las intuiciones convencionales sobre la relación entre representatividad y equidad.<sup>50</sup>

Desde la perspectiva del Sur Global, esta fase debe incorporar procesos de participación comunitaria que permitan a las poblaciones afectadas contribuir a la definición de los principios y normas operativas. La literatura crítica latinoamericana ha puesto de relieve que la gobernanza participativa, basada en la soberanía de datos y la justicia cognitiva, puede orientar el uso de la IA hacia el bien común y contrarrestar las lógicas coloniales que reproducen los modelos algorítmicos desplegados sin mediación comunitaria.<sup>51</sup> proyectos como Favel IA, que investiga cómo las prácticas de IA en las

<sup>47</sup>CONSEJO DE EUROPA, 2025. HUDERIA – risk and impact assessment of AI systems [en línea]. Estrasburgo: Consejo de Europa, 26 de febrero de 2025.

<sup>48</sup>PROGRAMA DE NACIONES UNIDAS PARA EL DESARROLLO (PNUD), 2025. Human Rights Impact of AI Assessment Toolkit [en línea]. PNUD, 19 de diciembre de 2025.

<sup>49</sup>EUROPEAN CENTER FOR NOT-FOR-PROFIT LAW (ECNL); DANISH INSTITUTE FOR HUMAN RIGHTS, 2025. A Guide to Fundamental Rights Impact Assessments (FRIA) [en línea]. ECNL, 12 de diciembre de 2025.

<sup>50</sup>BAROCAS, Solon; SELBST, Andrew D., 2016. «Big Data's Disparate Impact». California Law Review [en línea], vol. 104, núm. 3, pp. 671-732. DOI: <https://doi.org/10.15779/Z38BG31>. Para una sistematización de las fuentes de sesgo, véase MEHRABI, Ninareh; MORSTATTER, Fred; SAXENA, Nripsuta; LERMAN, Kristina; GALSTYAN, Aram, 2021. «A Survey on Bias and Fairness in Machine Learning». ACM Computing Surveys [en línea], vol. 54, núm. 6, art. 115, pp. 1-35. DOI: <https://doi.org/10.1145/3457607>.

<sup>51</sup>Véase, en este sentido, DERECHOS DIGITALES, 2024. «Regulación de la IA en Brasil: media sanción para un proyecto de ley que seguirá en discusión» [en línea]; y TEDIC, 2025. Inteligencia artificial y derechos humanos: aportes a la 192.<sup>a</sup>

favelas brasileñas pueden informar una gobernanza más equitativa desde los márgenes, ilustran la necesidad de repensar la IA desde las experiencias de quienes más sufren las desigualdades sistémicas.<sup>52</sup>

En la fase de desarrollo, la traducción opera mediante la selección de métricas de equidad, la definición de funciones de pérdida y la implementación de restricciones de optimización, derivadas explícitamente de las normas operativas de la fase de diseño. La derivación no es trivial: las obligaciones jurídicas no pueden operacionalizarse directamente en el código fuente, sino que requieren una operacionalización indirecta que admite múltiples interpretaciones válidas con distintos grados de adecuación jurídica.<sup>53</sup> Para abordar esta complejidad se han propuesto marcos interdisciplinarios que integran el análisis jurídico y el técnico durante el desarrollo de modelos de aprendizaje automático,<sup>54</sup> lo que exige estructuras organizativas capaces de traducir el razonamiento jurídico en elecciones de operacionalización y métricas de evaluación, garantizando a la vez la viabilidad computacional.<sup>55</sup> Para salvar la distancia entre los requisitos de alto nivel del Reglamento y las actividades de verificación implementables se han propuesto mapeos estructurados que descomponen los requisitos legales en sub-requisitos operacionales vinculados a actividades de verificación a lo largo del ciclo de vida.<sup>56</sup> En la fase de supervisión continua, la traducción exige mecanismos de control humano que preserven la capacidad de intervención cuando el comportamiento del sistema se aparte de los principios que deben orientarlo. La metodología HUDERIA incorpora una revisión iterativa que exige reevaluación periódica a medida que evolucionan las tecnologías y los contextos sociales, reconociendo que la evaluación de riesgos es un proceso dinámico y no un ejercicio estático.<sup>57</sup> La Coalición por la Libertad en Línea, en su Declaración Conjunta sobre Inteligencia Artificial y Derechos Humanos de 2025, ha subrayado la necesidad de incorporar trazabilidad y explicabilidad que permitan acceder a información sobre el razonamiento, las capacidades y las limitaciones del sistema.<sup>58</sup> Esta exigencia no es meramente instrumental: es condición de posibilidad del control democrático y de la rendición de cuentas ante las comunidades afectadas. La supervisión continua no es, por tanto, un mero mecanismo técnico de monitoreo, sino un espacio institucional de deliberación e impugnación de las decisiones algorítmicas que afectan derechos fundamentales.

---

sesión de la CIDH [en línea]. Sobre soberanía de datos y justicia cognitiva como ejes de una gobernanza participativa, RAHA, Saptaparni; PATRA, Joydeb, 2024. «Data Colonialism, Algorithmic Justice, and the Law: A Global South Perspective». *Anusandhanvallari*, vol. 1, núm. 1, pp. 105-113.

<sup>52</sup> DATA & SOCIETY, 2025. Favel IA: AI Governance and Literacy from the Margins in the Global South [en línea]. Data & Society.

<sup>53</sup> HANSON, Mathias; LEWKOWICZ, Gregory; VERBOVEN, Sam, 2025. Engineering the Law-Machine Learning Translation Problem: Developing Legally Aligned Models [en línea]. FARI Institute, junio de 2025.

<sup>54</sup> Ídem.

<sup>55</sup> Ídem.

<sup>56</sup> BUSCEMI, Alessio; DECKENBRUNNEN, Tom; KABIR, Fahria; MISHCHENKO, Kateryna; MOWLA, Nishat, 2025. «Assessing High-Risk AI Systems under the EU AI Act: From Legal Requirements to Technical Verification». *arXiv* [en línea], arXiv:2512.13907v2, 22 de diciembre de 2025.

<sup>57</sup> CONSEJO DE EUROPA, 2025. HUDERIA, op. cit.

<sup>58</sup> FREEDOM ONLINE COALITION, 2025. Joint Statement on Artificial Intelligence and Human Rights [en línea]. Freedom Online Coalition, junio de 2025.

## 8. IMPLICACIONES CONSTITUCIONALES Y DIMENSIÓN INSTITUCIONAL

El recorrido argumental conduce a una conclusión ineludible: la traducción normativa en sentido fuerte no es un mero problema técnico de operacionalización, sino una cuestión constitucional de primer orden. La irrupción de los sistemas algorítmicos en decisiones que afectan derechos individuales transforma la idea de los derechos fundamentales como límites externos al poder. Como advierte la doctrina constitucionalista reciente, la inteligencia artificial generativa está reconfigurando las categorías tradicionales del derecho constitucional en la imputación de responsabilidad, la motivación de las decisiones y la protección efectiva de las garantías.<sup>59</sup> Los derechos deben, por tanto, convertirse en componentes constitutivos del diseño de los sistemas. Este tránsito del paradigma del límite al del diseño constitutivo es la transformación más profunda que la IA impone a la teoría constitucional contemporánea: si la arquitectura computacional es, como sostenía Lessig, una modalidad de regulación tan efectiva como el derecho positivo, su producción es un acto normativo sometido a controles de legitimidad equivalentes a los de la producción legislativa. La verdadera cuestión constitucional de nuestro tiempo no es ya si el código es ley, sino cómo garantizar que la producción de esa arquitectura sea constitucionalmente legítima.<sup>60</sup>

La separación de poderes se ve doblemente interpelada: el ejecutivo puede eludir controles escudándose en la supuesta neutralidad técnica de los algoritmos, y el judicial afronta decisiones cuya motivación resulta inescrutable. De ahí que la traducción fuerte exija, junto a la transparencia en el diseño y la participación de las comunidades afectadas, mecanismos de revisión judicial de las arquitecturas algorítmicas, para los que la trazabilidad ascendente y descendente del modelo de cuatro niveles ofrece el andamiaje procedimental.

El escenario latinoamericano merece consideración específica. A inicios de 2026, ningún país cuenta aún con una ley general de inteligencia artificial plenamente vigente, y la mayoría de las iniciativas permanecen en trámite, a la expectativa de los desarrollos externos<sup>61</sup>, dependencia que reproduce en el plano normativo las asimetrías epistémicas identificadas en la sección V. No obstante, el constitucionalismo latinoamericano ofrece recursos propios. La doctrina del *Ius Constitutionale Commune* en América Latina (ICCAL) integra los órdenes nacionales, transnacionales e internacionales al servicio de los derechos humanos, la democracia y el Estado de derecho<sup>62</sup>, proporcionando un marco fértil para abordar la gobernanza algorítmica sin limitarse a la recepción pasiva de estándares externos. La Corte Constitucional de Colombia, en su Sentencia T-067 de 2025, estableció que la transparencia algorítmica es una garantía fundamental cuando el Estado emplea sistemas de decisión automatizada y que los algoritmos públicos deben

<sup>59</sup> BALAGUER CALLEJÓN, Francisco, 2025. La constitución del algoritmo. Zaragoza: Fundación Manuel Giménez Abad.

<sup>60</sup> *Idem*.

<sup>61</sup> COMISIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE (CEPAL), 2025. Índice Latinoamericano de Inteligencia Artificial (ILIA) 2025 [en línea]. CEPAL.; DEL POZO, Claudia. América Latina busca su propio camino en la regulación de la IA. *Wired* [en línea]. 20 de enero de 2026.

<sup>62</sup> BOGDANDY, Armin von; MORALES ANTONIAZZI, Mariela; FERRER MAC-GREGOR, Eduardo (eds.), 2017. *Ius Constitutionale Commune en América Latina*. México: UNAM.

ser auditables<sup>63</sup> en la Sentencia T-008 de 2026, pese a declarar improcedentes las tutelas por falta de subsidiariedad, reconoció que el uso de IA generativa en actos administrativos plantea retos inéditos al control judicial.<sup>64</sup> Estos pronunciamientos evidencian una incipiente pero significativa conciencia constitucional sobre la necesidad de someter los sistemas algorítmicos al escrutinio de los derechos fundamentales.

En el Sur Global, la construcción de instituciones supervisoras enfrenta obstáculos añadidos —recursos limitados, escasez de personal con formación híbrida, dependencia de infraestructuras controladas por empresas transnacionales y subordinación a marcos diseñados sin su participación—, a los que se suma la soberanía digital: ¿cómo auditar un sistema cuya infraestructura escapa a la jurisdicción nacional? Una teoría con pretensiones universales debe reconocer estas asimetrías y promover mecanismos de cooperación internacional que refuercen —no sustituyan— la capacidad institucional local: transferencia de conocimientos sin condicionar la autonomía regulatoria, infraestructuras digitales públicas y deliberación regional como la del sistema interamericano de derechos humanos.

## 9. CONCLUSIONES

El paradigma «derechos humanos por diseño» representa un avance normativo significativo, y la proliferación de instrumentos internacionales en 2023-2025 refleja un consenso emergente sobre la necesidad de someter los sistemas algorítmicos a los estándares del derecho internacional de los derechos humanos. No obstante, este artículo ha mostrado que, en su formulación actual, el paradigma carece del componente teórico más decisivo: una teoría de la traducción normativa que especifique cómo los principios jurídicos pueden convertirse en restricciones computacionalmente operativas sin perder su contenido esencial. Los déficits identificados no son problemas de implementación, sino deficiencias estructurales que afectan la coherencia interna del paradigma. A esa limitación se suma el déficit de perspectiva geopolítica: el paradigma dominante reproduce, sin cuestionarla, la pretensión de universalidad de los marcos elaborados en el Norte Global, ignorando tanto las asimetrías de poder epistémico como las realidades jurídico-sociales del Sur Global. La teoría de la traducción normativa en sentido fuerte ofrece un modelo de cuatro niveles —principios jurídicos, normas operativas, especificaciones técnicas e implementación sistémica— que procura articular la brecha entre derecho y código de manera sistemática, auditable y contextualmente sensible, incorporando la perspectiva del Sur Global no como complemento periférico, sino como condición de posibilidad de una teoría genuinamente universal.

Las limitaciones del modelo son reales y deben reconocerse con honestidad intelectual. La indeterminación jurídica inherente a los principios de derechos humanos no puede eliminarse en el Nivel II, sino solo reducirse mediante procesos deliberativos. Los conflictos entre métricas técnicas en el Nivel III reflejan conflictos normativos sustantivos que la teoría no resuelve por sí sola, sino que hace visibles y somete al escrutinio

---

<sup>63</sup>COLOMBIA, CORTE CONSTITUCIONAL, 2025. Sentencia T-067 de 2025. Magistrada Ponente: Natalia Angel Cabo. 14 de marzo de 2025.

<sup>64</sup>COLOMBIA, CORTE CONSTITUCIONAL, 2026. Sentencia T-008 de 2026. Magistrado Ponente: Juan Carlos Cortés González. 22 de enero de 2026.

democrático. Y el riesgo de un nuevo formalismo —que aparente cumplir la trazabilidad sin garantizar la protección efectiva de los derechos— es real y debe gestionarse mediante mecanismos institucionales sólidos. En definitiva, la constitucionalización de la inteligencia artificial no se completa diseñando mejores sistemas: requiere construir instituciones jurídicas y técnicas capaces de supervisar, auditar y corregir las arquitecturas algorítmicas de manera continua, en colaboración con las comunidades que soportan sus efectos y con plena conciencia de las asimetrías de poder que estructuran la gobernanza global de la tecnología.

## 10. REFERENCIAS BIBLIOGRÁFICAS

- ASAMBLEA GENERAL DE NACIONES UNIDAS, 2024. Resolución A/RES/78/265: Aprovechamiento de las oportunidades que ofrecen los sistemas de inteligencia artificial seguros, protegidos y fiables para el desarrollo sostenible. 21 de marzo de 2024. Nueva York: Naciones Unidas. [Disponible en: <https://undocs.org/A/RES/78/265>].
- AYANA, G. et al., 2024. «Decolonizing Global AI Governance: Assessment of the State of Decolonized AI Governance in Sub-Saharan Africa». *Royal Society Open Science*, vol. 11, núm. 8, art. 231994. DOI: <https://doi.org/10.1098/rsos.231994>.
- AHMED, Samar A., 2026. «Algorithmic Dependence and Digital Colonialism: A Conceptual Framework for AI in Education and Knowledge Systems of the Global South». *Frontiers in Education*. DOI: <https://doi.org/10.3389/feduc.2026.1720563>.
- BALAGUER CALLEJÓN, Francisco, 2025. *La constitución del algoritmo*. Zaragoza: Fundación Manuel Giménez Abad.
- BALKIN, Jack M., 2018. «Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation». *UC Davis Law Review*, vol. 51, núm. 3, pp. 1149-1210.
- BIRHANE, Abeba, 2020. «Algorithmic Colonization of Africa». *SCRIPTed*, vol. 17, núm. 2, DOI: <https://doi.org/10.2966/scrip.170220.389>.
- BOGDANDY, Armin von; MORALES ANTONIAZZI, Mariela; FERRER MAC-GREGOR, Eduardo (eds.), 2017. *lus Constitutionale Commune en América Latina*. México: UNAM.
- BRASIL, SENADO FEDERAL, 2024. Projeto de Lei núm. 2.338/2023: Marco Legal da Inteligência Artificial. Aprobado por el Senado Federal el 10 de diciembre de 2024; en trámite ante la Cámara de Diputados. Brasília: Senado Federal.
- BUSCEMI, Alessio; DECKENBRUNNEN, Tom; KABIR, Fahria; MISHCHENKO, Kateryna; MOWLA, Nishat, 2025. «Assessing High-Risk AI Systems under the EU AI Act: From Legal Requirements to Technical Verification». *arXiv*, arXiv:2512.13907v2, 22 de diciembre de 2025. DOI: <https://doi.org/10.48550/arXiv.2512.13907>.
- CASSIN, Barbara. La traducción como paradigma para las humanidades. Traducido por Ana Olivares Muñoz-Ledo. *Interpretatio. Revista de hermenéutica*. 2022, vol. 7, n.º 2, pp. 167-197. [Consultado en: 16 de abril de 2026]. ISSN 2448-864X. [Disponible en: [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S2448-864X2022000200167&lng=es&nrm=iso](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2448-864X2022000200167&lng=es&nrm=iso)].
- CAVOUKIAN, Ann, 2009. *Privacy by Design: The 7 Foundational Principles*. Toronto: Information and Privacy Commissioner of Ontario. [Disponible en: [https://student.cs.uwaterloo.ca/~cs492/papers/7foundationalprinciples\\_longer.pdf](https://student.cs.uwaterloo.ca/~cs492/papers/7foundationalprinciples_longer.pdf)].
- CHATHAM HOUSE, 2024. *Artificial Intelligence and the Challenge for Global Governance*. Londres: Chatham House. [Disponible en: <https://www.chathamhouse.org/2024/06/artificial-intelligence-and-challenge-global-governance>].
- CHOULDECHOVA, Alexandra, 2017. «Fair prediction with disparate impact: A study of bias in recidivism prediction instruments». *Big Data*, vol. 5, núm. 2, <https://pubmed.ncbi.nlm.nih.gov/28632438/>.

- COLOMBIA, CORTE CONSTITUCIONAL, 2025. Sentencia T-067 de 2025. Magistrada Ponente: Natalia Angel Cabo. 14 de marzo de 2025. [Disponible en: <https://www.corteconstitucional.gov.co/relatoria/2025/T-067-25.htm>].
- COLOMBIA, CORTE CONSTITUCIONAL, 2026. Sentencia T-008 de 2026. Magistrado Ponente: Juan Carlos Cortés González. 22 de enero de 2026. [Disponible en: <https://www.corteconstitucional.gov.co/relatoria/2026/T-008-26.htm>].
- COLOMBIA, MINISTERIO DE CIENCIA, TECNOLOGÍA E INNOVACIÓN, 2025. Proyecto de Ley núm. 043/2025 Senado: «Por medio de la cual se regula la Inteligencia Artificial en Colombia para garantizar su desarrollo ético, responsable, competitivo e innovador, y se dictan otras disposiciones». Radicado el 28 de julio de 2025; mensaje de urgencia el 8 de septiembre de 2025. Bogotá D.C.: MinCiencias. [Disponible en: [https://minciencias.gov.co/sites/default/files/upload/noticias/pl\\_ia\\_finalizado.pdf](https://minciencias.gov.co/sites/default/files/upload/noticias/pl_ia_finalizado.pdf)].
- COMISIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE (CEPAL), 2025. Índice Latinoamericano de Inteligencia Artificial (ILIA) 2025. CEPAL. [Disponible en: <https://indicadores.cepal.org/ilia/>].
- CONSEJO DE EUROPA, 2024. Convenio Marco del Consejo de Europa sobre la Inteligencia Artificial y los Derechos Humanos, la Democracia y el Estado de Derecho. Vilna, 5 de septiembre de 2024. Council of Europe Treaty Series, núm. 225. Estrasburgo: Consejo de Europa. [Disponible en: <https://rm.coe.int/1680afae3c>].
- CONSEJO DE EUROPA, 2025. HUDERIA – risk and impact assessment of AI systems. Estrasburgo: Consejo de Europa, 26 de febrero de 2025. [Disponible en: <https://www.coe.int/en/web/artificial-intelligence/huderia-risk-and-impact-assessment-of-ai-systems>].
- CONTRERAS, Pablo, 2024. «Convergencia internacional y caminos propios: regulación de la inteligencia artificial en América Latina». Actualidad Jurídica Iberoamericana, núm. 21, pp. 468-493. ISSN 2386-4567. [Disponible en: [https://revista-aji.com/wp-content/uploads/2024/07/AJ121\\_Art17.pdf](https://revista-aji.com/wp-content/uploads/2024/07/AJ121_Art17.pdf)].
- COOPER, A. Feder; FRANKLE, Jonathan; DE SA, Christopher, 2022. «Non-Determinism and the Lawlessness of Machine Learning Code». En: Proceedings of the 2022 Symposium on Computer Science and Law. Nueva York: ACM, DOI: <https://doi.org/10.1145/3511265.3550446>.
- DATA & SOCIETY, 2024. Algorithmic Impact Methods Lab. Data & Society. [Disponible en: <https://datasociety.net/research/algorithmic-impact-methods-lab/>].
- DATA & SOCIETY, 2025. Favel IA: AI Governance and Literacy from the Margins in the Global South. Data & Society. [Disponible en: <https://datasociety.net/research/algorithmic-impact-methods-lab/>].
- DERECHOS DIGITALES, 2024. «Regulación de la IA en Brasil: media sanción para un proyecto de ley que seguirá en discusión». [Disponible en: <https://www.derechosdigitales.org/recursos/regulacion-de-la-ia-en-brasil-media-sancion-para-un-proyecto-de-lei-que-seguira-en-discusion/>].
- DEL POZO, Claudia. América Latina busca su propio camino en la regulación de la IA. Wired. 20 de enero de 2026. [Consultado en: 16 de abril de 2026]. [Disponible en: <https://es.wired.com/articulos/america-latina-busca-su-propio-camino-en-la-regulacion-de-la-ia>].
- DE SOUSA SANTOS, Boaventura, et al. Epistemologías del Sur – Epistemologias do Sul. Coordinación de Maria Paula Meneses y Karina Andrea Bidaseca. 1.ª ed. Ciudad Autónoma de Buenos Aires: CLACSO; Coímbra: Centro de Estudos Sociais – CES, 2018. Libro digital, PDF. [Consultado en: 16 de abril de 2026]. [Disponible en: <https://biblioteca-repositorio.clacso.edu.ar/bitstream/handle/11096/9385/Epistemologias-del-Sur.pdf>].
- ERICKSON, J.; GREGORY, M., 2025. «Against Algorithmic Clarity: Law Beyond Specification». International Journal for the Semiotics of Law. DOI: 10.1007/s11196-025-10379-5.
- EUROPEAN CENTER FOR NOT-FOR-PROFIT LAW (ECNL); DANISH INSTITUTE FOR HUMAN RIGHTS, 2025. A Guide to Fundamental Rights Impact Assessments (FRIA). ECNL, 12 de diciembre de 2025. [Disponible en: <https://ecn1.org/publications/guide-fundamental-rights-impact-assessments-fria>].
- FISCHMAN-AFORI, Orit, 2025. Due process by design: Enhancing fairness and trust in AI decision-making. SSRN. <https://dx.doi.org/10.2139/ssrn.5884162>.

- FLÓREZ, Diego Tarapués, 2020. «¿Hay inconmensurabilidad tecnológica?». Cuadernos de Filosofía, núm. 38, pp. 119-141. ISSN 0716-9884. [Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=7606370>].
- FREEDOM ONLINE COALITION, 2025. Joint Statement on Artificial Intelligence and Human Rights. Freedom Online Coalition, junio de 2025. [Disponible en: <https://freedomonlinecoalition.com/joint-statement-on-ai-and-human-rights-2025/>].
- FRIEDMAN, Batya; HENDRY, David G., 2019. Value Sensitive Design: Shaping Technology with Moral Imagination. Cambridge (MA): MIT Press. ISBN 978-0-262-03953-6.
- FRIEDMAN, Batya; KAHN, Peter H.; BORNING, Alan, 2006. «Value Sensitive Design and Information Systems». En: ZHANG, P.; GALLETTA, D. (eds.). Human-Computer Interaction in Management Information Systems: Foundations. Nueva York: M.E. Sharpe.
- GARCÍA, Juan Manuel, 2024. ¿Qué sabemos realmente sobre el uso de IA en América Latina? Revisión de informes, índices y estudios (2019-2023). Santiago de Chile: Derechos Digitales. [Disponible en: <https://www.derechosdigitales.org/recursos/informes-indices-y-estudios-que-sabemos-realmente-sobre-el-uso-de-ia-en-america-latina/>].
- GIANNOPOULOS, Giorgos et al., 2024. «Fairness in AI: challenges in bridging the gap between algorithms and law». En: 2024 IEEE 40th International Conference on Data Engineering Workshops (ICDEW). IEEE.
- HANSON, Mathias; LEWKOWICZ, Gregory; VERBOVEN, Sam, 2025. Engineering the Law-Machine Learning Translation Problem: Developing Legally Aligned Models. FARI Institute, junio de 2025. [Disponible en: <https://www.fari.brussels/research-and-innovation/publication/machine-learning-law>].
- HILDEBRANDT, Mireille, 2020. «A philosophy of technology for computational law». En: Is Law Computable? Critical Perspectives on Law and Artificial Intelligence. Oxford: Hart Publishing.
- JONES, Kate, 2023. AI Governance and Human Rights: Resetting the Relationship. Londres: Chatham House.
- KINCHIN, Niamh, 2024. «Voiceless»: The Procedural Gap in Algorithmic Justice». International Journal of Law and Information Technology, vol. 32, eaae024. DOI: <https://doi.org/10.1093/ijlit/eaae024>.
- LESSIG, Lawrence, 1999a. Code and Other Laws of Cyberspace. Nueva York: Basic Books.
- LESSIG, Lawrence, 1999b. «The Law of the Horse: What Cyberlaw Might Teach». Harvard Law Review, vol. 113, núm. 2.
- LÓPEZ OLVERA, Carmen Patricia, 2023. «Epistemología jurídica aplicada». En: Epistemología jurídica aplicada. México: Universidad Nacional Autónoma de México, Instituto de Investigaciones Jurídicas. [Disponible en: <https://archivos.juridicas.unam.mx/www/bjv/libros/15/7099/5.pdf>].
- MANTELERO, Alessandro; ESPOSITO, Maria Samantha, 2024. «An evidence-based methodology for human rights impact assessment (HRIA) in the development of AI data-intensive systems». Computer Law & Security Review, vol. 41, art. 105561. DOI: <https://doi.org/10.1016/j.clsr.2021.105561>.
- MÉXICO, CÁMARA DE DIPUTADOS, 2025. Iniciativa de Ley Federal para el Desarrollo Ético, Soberano e Inclusivo de la Inteligencia Artificial. Suscrita el 22 de abril de 2025; en trámite en comisiones. México D.F.: Palacio Legislativo de San Lázaro. [Disponible en: [https://sil.gobernacion.gob.mx/Archivos/Documentos/2025/04/asun\\_4896460\\_20250430\\_1747757009.pdf](https://sil.gobernacion.gob.mx/Archivos/Documentos/2025/04/asun_4896460_20250430_1747757009.pdf)].
- MOHAMED, Shakir; PNG, Marie-Therese; ISAAC, William, 2020. «Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence». Philosophy & Technology, vol. 33, núm. 4.
- PROGRAMA DE NACIONES UNIDAS PARA EL DESARROLLO (PNUD), 2025. Human Rights Impact of AI Assessment Toolkit. PNUD, 19 de diciembre de 2025. [Disponible en: <https://www.undp.org/eurasia/publications/human-rights-impact-ai-assessment-toolkit>].
- QUIJANO, Anibal, 2000. «Colonialidad del poder y clasificación social». Journal of World-Systems Research, vol. 6, núm. 2.

- QUINE, Willard Van Orman, 1968. Palabra y objeto. Traducido por Manuel Sacristán. Barcelona: Labor. ISBN 84-335-2915-6.
- RAHA, Saptarni; PATRA, Joydeb, 2024. «Data Colonialism, Algorithmic Justice, and the Law: A Global South Perspective». *Anusandhanvallari*, vol. 1, núm. 1, ISSN 2229-3388. [Disponible en: <https://psvmkendra.com/index.php/journal/article/download/167/113>].
- RICŒUR, Paul, 2001. «El paradigma de la traducción». *Revista de Occidente*, núm. 240. ISSN 0034-8635.
- SPINDLOW, Sam, 2024. «Council of Europe adopts Turing-developed human rights risk and impact assessment for AI systems». The Alan Turing Institute, 5 de diciembre de 2024. [Disponible en: <https://www.turing.ac.uk/news/council-europe-adopts-turing-developed-human-rights-risk-and-impact-assessment-ai-systems>].
- TEDIC, 2025. Inteligencia artificial y derechos humanos: aportes a la 192.<sup>a</sup> sesión de la CIDH. TEDIC. [Disponible en: <https://www.tedic.org/inteligencia-artificial-cidh/>].
- THAIS, Savannah et al., 2025. HumRights-Bench: The First Human Rights Standard for AI Evaluation. Women at the Table. [Disponible en: <https://www.womenatthetable.net/project/humrights-bench/>].
- ULLOA, Mónica, 2024. Avances en la regulación de la Inteligencia Artificial en América Latina. ORCG. [Disponible en: <https://orcg.info/articulos/avances-en-la-regulacin-de-la-inteligencia-artificial-en-amrica-latina>].
- UNIÓN AFRICANA, 2024. Continental Artificial Intelligence Strategy. Addis Abeba: Unión Africana. [Disponible en: <https://au.int/en/documents/20240712/continental-artificial-intelligence-strategy>].
- UNIÓN EUROPEA, 2016. 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. Diario Oficial de la Unión Europea. 4 de mayo de 2016, L 119.
- UNIÓN EUROPEA, 2024. Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas sobre inteligencia artificial (Reglamento de Inteligencia Artificial). Diario Oficial de la Unión Europea. 12 de julio de 2024, L 2024/1689. [Disponible en: <https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX:32024R1689>].
- YEUNG, Karen, 2025. AI Equality by Design, Deliberation and Oversight. RAI UK. [Disponible en: [https://rai.ac.uk/new\\_projects/ai-equality-by-design-deliberation-and-oversight/](https://rai.ac.uk/new_projects/ai-equality-by-design-deliberation-and-oversight/)].
- YEUNG, Karen; HOWES, Andrew; POGREBNA, Ganna, 2020. «AI Governance by Human Rights-Centered Design, Deliberation, and Oversight». En: FLORIDI, L. et al. (eds.). *The Oxford Handbook of AI Ethics*. Oxford: Oxford University Press. [Disponible en: <https://ssrn.com/abstract=3435011>].

# Inteligencia artificial y la protección del patrimonio neuro-genómico: Hacia una regulación ética frente a los nuevos desafíos del estado de derecho\*

*Artificial intelligence and the protection of neuro-genomic heritage: Towards ethical regulation in the face of the new challenges to the rule of law*

Omar Fernando Becerra Partida\*\* María de Jesús Méndez Verduzco\*\*\*

## RESUMEN

El presente artículo científico analiza la convergencia disruptiva entre la Inteligencia Artificial (IA) y la biojurídica contemporánea, centrando su atención en la urgente necesidad de proteger el patrimonio neuro-genómico del individuo. A través de un método analítico-deductivo y una revisión sistemática de la doctrina internacional y el bloque de constitucionalidad mexicano, se examina cómo el procesamiento de datos biológicos masivos por algoritmos de aprendizaje profundo vulnera los pilares tradicionales del Estado de Derecho, tales como la autonomía de la voluntad y el libre desarrollo de la personalidad. Los resultados de la investigación demuestran que el actual marco normativo de protección de datos personales es insuficiente ante la capacidad de reidentificación y predictibilidad de la IA. Se concluye con la propuesta de elevar los neuroderechos al rango de garantías fundamentales, asegurando una gobernanza ética que impida el determinismo biológico-digital y salvaguarde la dignidad humana como eje rector de la innovación tecnológica.

## PALABRAS CLAVE

Inteligencia Artificial, Neuroderechos, Derecho Genómico, Bioética, Estado de Derecho, Patrimonio Genético, Privacidad Mental.

## ABSTRACT

This scientific article analyzes the disruptive convergence between Artificial Intelligence (AI) and contemporary biojurisprudence, focusing on the urgent need to protect the individual's neuro-genomic heritage. Through an analytical-deductive method and a systematic review of international scholarship and Mexico's constitutional framework, it examines how the processing of massive biological data by deep-learning algorithms undermines traditional pillars of the rule of law, such as autonomy of the will and the free development of personality. The research findings show that the current personal-data protection framework is insufficient in light of AI's capacity for re-identification and prediction. It concludes by proposing that neurorights be elevated to the status of fundamental guarantees, ensuring ethical governance that prevents biological-digital determinism and safeguards human dignity as the guiding principle of technological innovation.

## KEYWORDS

Artificial Intelligence, Neurorights, Genomic Law, Bioethics, Rule of Law, Genetic Heritage, Mental Privacy.

\* Artículo de investigación

\*\* Doctor en Bioética y Derechos Humanos por el Colegio de Bioética y Terapia de Jalisco A. C.; maestro en Derecho Corporativo, Ética y Bioética y Derecho y abogado por la Universidad de Guadalajara con especialidad en Bioética (omar.bpartida@academicos.udg.mx). <https://orcid.org/0000-0002-8142-3333>.

\*\*\* Maestra en Ética y Doctora en Bioética y Derechos Humanos por el Instituto de Ética y Terapia de Jalisco; Licenciada en Nutrición por la Universidad del Valle de Atemajac y Profesora adscrita al Departamento de Morfología del Centro Universitario de Ciencias de la Salud de la Universidad de Guadalajara (maria.mendez@academicos.udg.mx). <https://orcid.org/0009-0007-8542-9453>.

## SUMARIO

1. Introducción
2. El derecho genómico ante el procesamiento de datos masivos (big data)
3. Neuroderechos e inteligencia artificial: la frontera de la privacidad mental
4. Hacia una gobernanza biojurídica en México: una propuesta de reforma integral
5. Conclusiones y prospectiva: el destino del Estado de Derecho en la era biodigital
6. Referencias Bibliográficas

### 1. INTRODUCCIÓN

La evolución tecnológica del siglo XXI ha desplazado la frontera del derecho desde la protección de la integridad física y el patrimonio material hacia la salvaguarda de la integridad biológica y mental. Nos encontramos en un punto de inflexión donde la tecnociencia no solo rodea al sujeto, sino que pretende decodificar su esencia interna. La emergencia de la Inteligencia Artificial (IA), potenciada por el procesamiento de volúmenes masivos de datos (Big Data), ha dado lugar a una capacidad de análisis sobre el genoma humano y la actividad neuronal que plantea dilemas éticos y jurídicos sin precedentes en la historia del pensamiento jurídico. Como señalan Ienca y Andorno, la sofisticación de las interfaces cerebro-computadora y los algoritmos predictivos permiten hoy el acceso a niveles de la psique y la biología humana que anteriormente eran considerados el «sagrario» inviolable de la individualidad.<sup>1</sup>

Este fenómeno se inserta en lo que Schwabdenomina la «Cuarta Revolución Industrial», una era caracterizada por la fusión de tecnologías que desdibujan las líneas entre lo físico, lo biológico y lo digital. Sin embargo, desde la perspectiva del Estado de Derecho, esta fusión no es neutral. El derecho, tradicionalmente reactivo y basado en la regulación de conductas tangibles, se enfrenta a una crisis de efectividad ante la ubicuidad de los algoritmos de aprendizaje profundo (<sup>2</sup> Deep Learning). Estos sistemas poseen la capacidad de inferir, con una precisión estadística inquietante, predisposiciones biológicas, estados emocionales y patrones conductuales que escapan al control del propio sujeto. Estamos ante el surgimiento de lo que Zuboffdenomina «capitalismo de vigilancia», donde la bioinformática se convierte en la nueva materia prima de extracción para la predicción y modificación de la conducta humana.<sup>3</sup>

El concepto de «patrimonio neuro-genómico» que se propone en esta investigación surge como una categoría de resistencia jurídica frente a esta intrusión. Si bien la Declaración Universal sobre el Genoma Humano y los Derechos Humanos de la UNESCO estableció que el genoma es «patrimonio de la humanidad», esta declaración debe ser actualizada a la luz de la inteligencia artificial. El genoma ya no es solo un mapa biológico; es una base de datos dinámica. Por otro lado, la actividad neuronal, procesada a través de algoritmos, se ha convertido en el último reducto de la privacidad. La

---

<sup>1</sup> IENCA, Marcello; ANDORNO, Roberto, «Towards new human rights in the age of neuroscience and neurotechnology», *Life Sciences, Society and Policy*, 2017, vol. 13, art. 5.

<sup>2</sup> SCHWAB, Klaus, *The Fourth Industrial Revolution*, Geneva, World Economic Forum, 2016.

<sup>3</sup> ZUBOFF, Shoshana, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York, PublicAffairs, 2019.

convergencia de ambos —el código genético y el flujo neurofisiológico— conforma el patrimonio más íntimo del ser humano: su identidad biometabólica y cognitiva.<sup>4</sup>

La problemática central radica en que el libre albedrío, pilar fundamental sobre el cual se construye la responsabilidad jurídica y la autonomía privada, se ve amenazado por la posibilidad de una manipulación neurodigital que opera por debajo del umbral de la conciencia. Si un algoritmo puede predecir nuestra susceptibilidad a ciertas decisiones o patologías basándose en nuestro patrimonio neuro-genómico, ¿hasta qué punto podemos hablar de un consentimiento verdaderamente libre? El Estado de Derecho, para seguir siendo tal, debe garantizar que la tecnología no se convierta en un instrumento de eugenesia digital o de determinismo biológico.<sup>5</sup>

En consecuencia, esta introducción plantea que la regulación ética de la IA no es un complemento opcional, sino una necesidad existencial para la democracia. La asimetría de información entre las corporaciones biotecnológicas y el ciudadano crea un estado de indefensión estructural. El presente artículo se propone, por tanto, realizar un análisis exhaustivo de la insuficiencia de los mecanismos legales actuales, especialmente en el contexto mexicano, y proponer una ruta crítica hacia la integración de los neuroderechos en el bloque de convencionalidad. Solo mediante una gobernanza bioética robusta será posible asegurar que el avance científico permanezca al servicio de la persona y no como una herramienta de deshumanización algorítmica.

### **1.1 Metodología de la investigación**

Para respaldar el alcance del estudio, se llevó a cabo una investigación cualitativo-documental de corte dogmático-jurídico y bioético. El análisis se estructuró a partir de cuatro categorías analíticas fundamentales:

- a) Gobernanza y protección de datos biológico-genéticos**
- b) Principio de no discriminación algorítmica e integridad mental**
- c) Trazabilidad y reidentificación del dato genómico**
- d) Soberanía cognitiva frente a la inteligencia artificial**

La revisión sistemática de la literatura doctrinal, normativa y jurisprudencial abarcó el periodo 2010–2025, seleccionando fuentes indexadas en bases de datos académicas (Scopus, Web of Science, SciELO y Redalyc), así como resoluciones emitidas por la Suprema Corte de Justicia de la Nación (SCJN) y la Corte Interamericana de Derechos Humanos (Corte IDH). El contexto del estudio se sitúa en la urgencia de reconfigurar el Estado de Derecho mexicano ante los avances de los modelos de aprendizaje profundo y la biometría neuronal en el ámbito corporativo y público.

## **2. EL DERECHO GENÓMICO ANTE EL PROCESAMIENTO DE DATOS MASIVOS (BIG DATA)**

El patrimonio genético constituye la base biológica de la identidad humana y la continuidad de la especie. El Derecho Genómico, como rama especializada del ordenamiento jurídico, se ha estructurado tradicionalmente sobre tres pilares fundamentales:

---

<sup>4</sup>UNESCO, Declaración Universal sobre el Genoma Humano y los Derechos Humanos, París, UNESCO, 1997.

<sup>5</sup>YUSTE, Rafael; GOERING, Sara; AGÜERA Y ARCAS, Blaise; et al., «Four ethical priorities for neurotechnologies and AI», *Nature*, 2017, vol. 551, núm. 7679, pp. 159–163.

el consentimiento informado, la confidencialidad absoluta y la prohibición estricta de la discriminación. No obstante, la irrupción de la Inteligencia Artificial (IA) ha provocado una metamorfosis en la naturaleza jurídica del dato genético. Ya no estamos ante una secuencia de nucleótidos con fines puramente diagnósticos; nos encontramos ante un activo predictivo de escala.<sup>6</sup> macroscópica que alimenta modelos de Big Data con fines que trascienden la salud.<sup>7</sup>

## **2.1 El bloque de legalidad en México y la insuficiencia normativa**

El bloque de legalidad en el Estado mexicano, robustecido tras la reforma constitucional de junio de 2011, coloca a los derechos humanos de fuente internacional en el mismo nivel jerárquico que la Constitución Política de los Estados Unidos Mexicanos. Sin embargo, existe una brecha significativa entre la norma y la realidad algorítmica.<sup>8</sup>

### **2.1.1 Análisis del artículo 1.º constitucional y la discriminación indirecta**

El artículo 1.º de la CPEUM establece una prohibición absoluta contra toda forma de discriminación motivada por «condiciones de salud». En una interpretación evolutiva, la Suprema Corte de Justicia de la Nación (SCJN) ha señalado que el derecho al libre desarrollo de la personalidad protege el «núcleo de identidad» del individuo. No obstante, el riesgo de la IA trasciende la discriminación directa. Nos enfrentamos a la discriminación algorítmica indirecta, donde los sistemas de aprendizaje profundo toman decisiones basadas en correlaciones genómicas opacas. Estas decisiones son difícilmente impugnables mediante los recursos de amparo tradicionales, ya que el ciudadano no puede demostrar el sesgo biológico oculto en la «caja negra» del software.<sup>9</sup>

### **2.1.2 La LFPDPPP y el agotamiento del consentimiento informado**

La Ley Federal de Protección de Datos Personales en Posesión de los Particulares (LFPDPPP) clasifica a los datos genéticos como «sensibles». El modelo de consentimiento que propone es análogo y estático, asumiendo que el titular comprende para qué se usará su información. Sin embargo, la IA opera bajo una lógica de «aprendizaje continuo». Un dato aportado para una investigación clínica puede ser reutilizado para inferir riesgos financieros o rasgos de personalidad mediante correlaciones que el titular nunca pudo prever. La opacidad algorítmica rompe el nexo de causalidad entre la voluntad y el uso del dato, lo que exige transitar hacia una gobernanza proactiva.<sup>10</sup>

### **2.1.3 La Ley General de Salud y el Título Quinto Bis**

El Título Quinto Bis de la Ley General de Salud (LGS) declara que el genoma humano es «patrimonio básico de la humanidad». No obstante, su enfoque es eminentemente clínico. La ley es silenciosa frente al uso de la biometría genómica con fines de seguridad o perfiles comerciales. El Estado mexicano carece de una normativa que regule la

<sup>6</sup>SAN VICENTE PARADA, Aida del Carmen, Temas selectos de bioética y derecho, Valencia, Tirant lo Blanch, 2025.

<sup>7</sup>VAYENA, Effy; BLASIMME, Alessandro; COHEN, I. Glenn, «Machine learning in medicine: Addressing ethical challenges», PLOS Medicine, 2018, vol. 15, núm. 11, e1002689.

<sup>8</sup>MÉXICO, Constitución Política de los Estados Unidos Mexicanos, Diario Oficial de la Federación, 1917.

<sup>9</sup>SUPREMA CORTE DE JUSTICIA DE LA NACIÓN (MÉXICO), «Derecho al libre desarrollo de la personalidad. Su dimensión externa e interna», tesis 1a. J. 4/2019 (10a.), Semanario Judicial de la Federación, 2019.

<sup>10</sup>MÉXICO, Ley Federal de Protección de Datos Personales en Posesión de los Particulares, Diario Oficial de la Federación, 2010, art. 3, fracc. VI.

«trazabilidad del dato biológico» en entornos digitales, dejando el patrimonio genético expuesto a un escrutinio algorítmico permanente.<sup>11</sup>

## **2.2 Biometría genómica y el fenómeno de la reidentificación**

Uno de los riesgos más críticos es la reidentificación genética. El Derecho Genómico tradicional se basa en la premisa de que la anonimización es suficiente. Sin embargo, la IA ha demostrado que el ADN es la biometría definitiva y es imposible de anonimizar totalmente.

Investigaciones como las de Gymrek et al. han evidenciado que es posible identificar a individuos y sus familias a partir de bases de datos genéticas supuestamente anónimas, cruzándolas con registros de genealogía pública. En México, no existe una salvaguarda legal contra la reidentificación genética. Sin una regulación que contemple la trazabilidad, el patrimonio genético de los ciudadanos queda vulnerable. Es imperativo que la legislación evolucione para proteger no solo el dato, sino la relación biológica que este representa.<sup>12</sup>

## **2.3 Determinismo algorítmico y la nueva estratificación biológica**

La integración de la IA en la gestión de datos genómicos da paso al «determinismo algorítmico». Si instituciones financieras o aseguradoras obtienen acceso a perfiles genómicos procesados por IA, surge el peligro de crear una nueva estratificación social basada en la «calidad biológica».

Este escenario vulnera el principio de igualdad. Como advierte O'Neil, los algoritmos pueden actuar como «armas de destrucción matemática» que perpetúan sesgos. En un Estado de Derecho, el ADN no debe convertirse en un destino jurídico. La protección del patrimonio genético frente a la IA exige reconocer<sup>13</sup> el «derecho a no ser perfilado» y asegurar que la dignidad humana permanezca por encima de la eficiencia predictiva.

# **3. NEURODERECHOS E INTELIGENCIA ARTIFICIAL: LA FRONTERA DE LA PRIVACIDAD MENTAL**

La neurotecnología, potenciada por la Inteligencia Artificial (IA), ha generado una disrupción sin precedentes en la concepción clásica de los derechos humanos. Si el Derecho Genómico protege la información biológica hereditaria, los neuroderechos surgen para salvaguardar el flujo dinámico de la conciencia y la identidad. Esta sección analiza la urgencia de establecer un marco protector frente a la capacidad de los algoritmos para decodificar e influir en la actividad cerebral humana.

## **3.1 La decodificación algorítmica y el fin de la libertad de pensamiento:**

El asalto al *forum internum*

Históricamente, el derecho ha operado bajo la premisa de que el pensamiento es libre mientras no se manifieste externamente. Esta distinción entre el *forum internum* (inviolable) y el *forum externum* (regulable) constituye la base de la libertad de conciencia en las democracias liberales. Sin embargo, la integración de redes neuronales artificiales

<sup>11</sup>MÉXICO, Ley General de Salud, Diario Oficial de la Federación, 1984, art. 103 Bis.

<sup>12</sup>GYMREK, Melissa; MCGUIRE, Amy L.; GOLAN, David; HALPERIN, Eran; ERLICH, Yaniv, «Identifying personal genomes by surname inference», *Science*, 2013, vol. 339, núm. 6117, pp. 321–324.

<sup>13</sup>O'NEIL, Cathy, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York, Crown Publishing, 2016.

con sistemas de electroencefalografía (EEG) de alta resolución y resonancia magnética funcional (fMRI) está permitiendo lo que la doctrina denomina «lectura mental algorítmica».

Como advierten Ienca y Andorno, la capacidad de la IA para interpretar estados cognitivos, intenciones y emociones a partir de señales neuronales crudas erosiona el último reducto de la privacidad: la intimidad mental. Ya no se trata solo de proteger lo que el individuo dice o hace, sino lo que el individuo<sup>14</sup> procesa antes de actuar. La IA no solo observa; decodifica. Esta «transparencia cerebral» forzada por la tecnología elimina la asimetría necesaria entre el pensamiento privado y la interacción social, dejando al sujeto en un estado de vulnerabilidad absoluta frente a quienes poseen la tecnología de decodificación.

El riesgo se agrava con la neuromodulación asistida por IA, donde el flujo de información es bidireccional. Sistemas de estimulación cerebral profunda o interfaces no invasivas pueden ser utilizados no solo para extraer datos, sino para inyectar estímulos que alteren la toma de decisiones. En este punto, el libre albedrío deja de ser una categoría filosófica para convertirse en un bien jurídico en peligro. Si un algoritmo puede inducir una preferencia o mitigar una inhibición por debajo del umbral de la conciencia consciente, la estructura misma de la responsabilidad penal y civil —basada en la imputabilidad y la voluntad— colapsa.

### **3.2 Derecho comparado: El precedente constitucional de Chile y la vanguardia regulatoria**

Ante este vacío, el análisis comparado resulta imperativo. La República de Chile se ha consolidado como el referente global al promulgar en 2021 la primera reforma constitucional en la historia dedicada a proteger la integridad mental. Esta reforma no es un simple ajuste técnico; es el reconocimiento de que la actividad cerebral es la base ontológica del ser humano y, por tanto, merece una protección equiparable a la integridad física.<sup>15</sup>

El modelo chileno, influenciado por el grupo Morningside y liderado por Rafael Yuste, establece que la ley debe regular los requisitos para que el desarrollo científico y tecnológico esté al servicio de las personas, protegiendo especialmente la integridad psíquica. Este precedente obliga a México a replantear su bloque de convencionalidad. La propuesta de los cinco neuroderechos fundamentales —privacidad mental, identidad personal, libre albedrío, acceso equitativo y protección contra sesgos— debe ser integrada no como una ley secundaria, sino como una garantía constitucional de cuarta generación. Sin esta elevación de rango, la protección frente a las empresas tecnológicas que operan fuera de la jurisdicción nacional será nula.<sup>16</sup>

### **3.3 Situación en México: Del libre desarrollo de la personalidad a la soberanía cognitiva**

En México, la protección de la esfera mental carece de una autonomía normativa específica. Actualmente, la defensa de la mente se intenta articular a través del derecho

<sup>14</sup> IENCA, Marcello; ANDORNO, Roberto, «Hacia nuevos derechos humanos en la era de la neurociencia y la neurotecnología», *Análisis Filosófico*, 2021, vol. 41, núm. 1, pp. 141–185.

<sup>15</sup> CHILE, Reforma constitucional. Modificación al artículo 19, n.º 1, de la Carta Fundamental para establecer que el desarrollo científico y tecnológico estará al servicio de las personas, *Diario Oficial de la República de Chile*, 25 de octubre de 2021.

<sup>16</sup> YUSTE, Rafael; et al., op. cit., nota 5.

a la salud y el derecho a la privacidad. No obstante, estas categorías son insuficientes. La salud mental se aborda desde una perspectiva patológica, mientras que el riesgo de la IA es funcional y existencial.<sup>17</sup>

La Suprema Corte de Justicia de la Nación (SCJN) ha desarrollado la doctrina del Libre Desarrollo de la Personalidad, la cual constituye el asidero jurídico más sólido para la construcción de los neuroderechos en el país. Si este derecho fundamental protege la capacidad de elegir el propio plan de vida sin interferencias, es evidente que la manipulación neurodigital representa su violación más radical. La «soberanía cognitiva» debe ser entendida como el derecho del ciudadano mexicano a que su actividad neuronal no sea tratada como un<sup>18</sup> commodity o un activo de datos explotable por terceros.

El Estado mexicano enfrenta el desafío de reformar la Ley General de Salud y la Ley Federal de Protección de Datos Personales para incluir el concepto de «neurodatos» como una categoría de datos personales ultrasensibles. A diferencia de un correo electrónico o una dirección, un neurodato es la representación física de un pensamiento o un impulso. Su tratamiento exige una regulación de «consentimiento reforzado» y una prohibición estricta de su uso para fines de mercadotecnia, perfilamiento político o discriminación laboral, asegurando que el cráneo humano siga siendo el límite infranqueable de la vigilancia estatal y corporativa.

#### **4. HACIA UNA GOBERNANZA BIOJURÍDICA EN MÉXICO: UNA PROPUESTA DE REFORMA INTEGRAL**

La transición de un Estado de Derecho analógico a uno digital y neuroprotegido no puede postergarse más. La arquitectura jurídica mexicana, aunque robusta en su protección a la dignidad humana, presenta grietas estructurales frente a la potencia extractiva de la Inteligencia Artificial (IA) sobre el patrimonio neuro-genómico. Esta sección articula una propuesta de gobernanza tridimensional: constitucional, legislativa y administrativa.

##### **4.1 La paradoja del artículo 19: De la Libertad Física a la Soberanía Cognitiva**

Resulta paradójico que el Artículo 19 de la Constitución Política de los Estados Unidos Mexicanos (CPEUM) se haya erigido históricamente como el baluarte contra las limitaciones a la libertad física —centrándose en el auto de vinculación a proceso y la prisión preventiva—, pero ignore sistemáticamente las limitaciones a la libertad cognitiva. Mientras el Estado gasta recursos en vigilar que el cuerpo del ciudadano no sea recluido arbitrariamente, deja su psique —el núcleo de su voluntad— desprotegida frente a las intrusiones neurotecnológicas asistidas por IA.

##### **4.1.1 El fenómeno de la «Prisión Mental» y el Derecho al Habeas Mens**

Una reforma integral debe garantizar que el desarrollo tecnológico esté supeditado a la dignidad humana, estableciendo el derecho al «Habeas Mens». Si el Artículo 19 prohíbe la incomunicación y la tortura física, con mayor razón debe prohibir la intervención cognitiva no consentida. La IA actual permite la extracción de testimonios o «recuerdos»

<sup>17</sup>MÉXICO, Constitución Política de los Estados Unidos Mexicanos, op. cit., art. 4.º.

<sup>18</sup>SUPREMA CORTE DE JUSTICIA DE LA NACIÓN (MÉXICO), «Libre desarrollo de la personalidad», tesis 1a. J. 4/2019 (10a.), Semanario Judicial de la Federación, 2019.

mediante la detección de patrones neuronales (como el potencial P300), lo que constituye una forma de autoincriminación tecnológica.

Caso de Estudio: El precedente de *State of Maharashtra v. Aditi Sharma*. En este proceso, se utilizó el escaneo cerebral para determinar la culpabilidad de un acusado. En México, la ausencia de una prohibición explícita en el artículo 19 deja la puerta abierta a que algoritmos de IA «validen» testimonios, erosionando la presunción de inocencia y convirtiendo al proceso penal en un ejercicio de biometría punitiva.<sup>19</sup>

#### **4.2 La Creación del «Neurodato» como Categoría de Protección Ultrasensible**

La gobernanza biojurídica exige una redefinición técnica en la Ley Federal de Protección de Datos Personales en Posesión de los Particulares (LFPDPPP). La categoría actual de «datos sensibles» es insuficiente para contener la información que la IA extrae del cerebro y del genoma.

##### **4.2.1 El Neurodato y la Trazabilidad Biológica**

Se propone la creación jurídica de la categoría de «Neurodatos». A diferencia de un correo electrónico o una huella dactilar, el neurodato es la representación física de un impulso, un pensamiento o una predisposición biológica. La reforma debe prohibir explícitamente el tratamiento de neurodatos para fines de mercadotecnia, perfilamiento político o discriminación laboral (neurotargeting).

Caso de Estudio: Del escándalo de Cambridge Analytica al neuroprofiling. La evolución de la recolección de datos ha pasado de las preferencias en redes sociales al monitoreo en tiempo real de estados emocionales mediante dispositivos wearables. Sin una regulación que declare la inviolabilidad de la actividad cerebral, la IA puede ser utilizada para suprimir el pensamiento crítico del ciudadano antes de que este emita un voto o realice una compra, constituyendo una «detención de la voluntad» por medios algorítmicos.

#### **4.3 El Título Quinto Bis de la Ley General de Salud: Trazabilidad y Comités de Bioética**

Finalmente, la gobernanza debe aterrizar en la operatividad de la Ley General de Salud (LGS). El actual Título Quinto Bis, aunque declara al genoma como patrimonio básico de la humanidad, carece de dientes frente a la IA comercial.<sup>20</sup>

##### **4.3.1 Supeditación Terapéutica y Human in the Loop**

La propuesta legislativa exige que toda técnica de IA que pretenda manipular la actividad cerebral o procesar el genoma en territorio nacional deba:

1. Contar con una autorización expresa de un Comité de Bioética Autónomo, prohibiendo fines punitivos o comerciales puros.
2. Garantizar el principio de «Human in the Loop» (Humano en el ciclo): ninguna decisión judicial, médica o laboral basada en el patrimonio neuro-genómico puede ser tomada exclusivamente por un algoritmo. El ciudadano tiene el derecho irrenunciable a una explicación humana y comprensible sobre el sesgo de la máquina.

#### **4.4 Prospectiva Jurídica:**

La Reinención del Pacto Social en la Era de la Singularidad Biotecnológica.

<sup>19</sup>INDIA, *State of Maharashtra v. Sharma*, C.C. No. 508/07, Court of Sessions, Pune, 12 de junio de 2008.

<sup>20</sup>MÉXICO, Ley General de Salud, op. cit., título quinto bis, art. 103 Bis.

La conclusión de este estudio no debe entenderse como un punto final, sino como el umbral de una nueva ontología jurídica. El patrimonio neuro-genómico no es simplemente un conjunto de datos biológicos; representa la última frontera de la libertad humana y el reducto final de la individualidad. Si el Estado de Derecho no evoluciona con la celeridad que la convergencia entre la Inteligencia Artificial (IA) y la neurotecnología exige, corremos el riesgo inminente de transitar hacia una distopía de estratificación biológica y vigilancia mental permanente, donde la esencia misma del ser humano sea reducida a un código explotable por algoritmos de mercado o de control estatal.

#### **4.4.1 El fin del «Sujeto de Derecho» analógico y el nacimiento del Ciudadano Neuro-Protegido**

La dogmática jurídica tradicional se construyó sobre la base de un sujeto racional, libre y con una voluntad impenetrable en su *forum internum*. La prospectiva jurídica nos indica que este modelo ha caducado. La IA, al ser capaz de predecir comportamientos mediante el análisis del genoma y de decodificar estados mentales a través de interfaces cerebro-computadora, disuelve la barrera de la privacidad clásica.

En la próxima década, el Derecho no se limitará a regular acciones externas, sino que deberá arbitrar sobre la integridad del flujo de la conciencia. El concepto de «persona» deberá expandirse constitucionalmente para incluir la protección de los sustratos biológicos y neuronales que permiten la existencia de la voluntad. Sin esta protección, el ciudadano del futuro será un sujeto «transparente», cuya predictibilidad algorítmica anulará de facto la presunción de inocencia y el libre albedrío.

#### **4.4.2 El Riesgo de la Nueva Eugenesia Digital y la Estratificación Social**

Uno de los puntos más críticos de la prospectiva es la aparición de una brecha neuro-cognitiva. Si el acceso a la mejora cognitiva asistida por IA o a la edición genética de precisión queda supeditado al mercado, el Derecho enfrentará el surgimiento de una nueva clase social: una élite con capacidades cognitivas aumentadas frente a una mayoría «biológicamente rezagada».

El escudo biojurídico que aquí se propone debe contemplar el principio de justicia distributiva biotecnológica. La Constitución mexicana debe prever que la manipulación del patrimonio neuro-genómico no se convierta en una herramienta de castas. El riesgo no es solo la pérdida de privacidad, sino la cristalización de la desigualdad en el ADN y en las conexiones sinápticas, lo que haría que la movilidad social fuera biológicamente imposible. La ley debe intervenir para evitar que los seguros, los empleadores o el Estado utilicen el perfilamiento neuro-genómico como un mecanismo de selección eugénica «suave» o encubierta.

#### **4.4.3 El Desafío Punitivo: Hacia un Derecho Penal Preventivo y la Crisis de la Culpabilidad**

La prospectiva en el ámbito penal es disruptiva. Si la IA puede detectar una predisposición neurológica a la violencia o al incumplimiento normativo, ¿podrá el Estado sancionar antes de que el delito se cometa? Este escenario de «justicia predictiva» choca frontalmente con el artículos 19 y 20 de la CPEUM.

El llamado que este estudio hace al Poder Judicial es para resistir la tentación del determinismo neuro-tecnológico. La culpabilidad debe seguir siendo un juicio reprochable sobre la conducta, no sobre el escaneo cerebral. La construcción de este escudo

biojurídico implica blindar el proceso penal contra el uso de evidencias neurocientíficas que pretendan sustituir la voluntad del juzgador o violentar la dignidad del imputado mediante la extracción forzada de neurodatos.

#### **4.4.4 Llamado a la Acción: Legislador, Academia y Judicatura**

Este estudio es un exhorto a los tres pilares de la República para construir una Constitución Neuro-protégida.

1. Al Legislador: Para que abandone la reactividad y legisle con una visión de futuro, elevando los neuroderechos al rango de garantías individuales inalienables.
2. A la Academia: Para que profundice en la intersección entre el Derecho Genómico y la Neuroética, alejándose de los formalismos que ignoran la realidad tecnológica.
3. Al Poder Judicial: Para que sea el garante último de la dignidad humana, ejerciendo un control de convencionalidad que sitúe al ser humano por encima de la eficiencia del algoritmo.

#### **4.4.5 Conclusión Prospectiva: La Soberanía sobre el propio Destino**

En última instancia, la lucha por la regulación del patrimonio neuro-genómico es la lucha por la supervivencia de la condición humana. La Inteligencia Artificial debe ser un instrumento de dignificación que potencie nuestras capacidades, no un mecanismo de deshumanización que nos convierta en activos de datos. Solo a través de un marco biojurídico sólido, ético y trazable, podremos garantizar que la humanidad en el siglo XXI no sea un simple espectador de su propia obsolescencia, sino la dueña absoluta de su destino biológico y mental. La Constitución del futuro debe ser el último refugio de la libertad, asegurando que el pensamiento siga siendo, para siempre, el territorio más privado e inviolable de la persona.

En conclusión, la reforma al artículo 19 y la actualización de las leyes secundarias deben establecer que el cráneo humano es el límite infranqueable de la vigilancia estatal y corporativa. Solo así, el Estado de Derecho mexicano podrá transitar hacia una soberanía cognitiva real, donde la tecnología sea un instrumento de progreso y no un mecanismo de eugenesia digital.

### **5. CONCLUSIONES Y PROSPECTIVA: EL DESTINO DEL ESTADO DE DERECHO EN LA ERA BIODIGITAL**

La investigación exhaustiva presentada en este artículo permite concluir que nos encontramos ante la crisis más profunda del constitucionalismo desde la posguerra. La Inteligencia Artificial y la neurotecnología no son solo herramientas; son fuerzas disruptivas que redefinen la esencia misma del ser humano. A continuación, se presentan las conclusiones cardinales derivadas del análisis dogmático y legal realizado:

#### **5.1 La obsolescencia de la protección analógica y la emergencia del patrimonio neuro-genómico**

Se ha demostrado que el actual bloque de legalidad en México, centrado en la Ley Federal de Protección de Datos Personales (LFPDPPP), es insuficiente para contener los riesgos de la era genómica. El consentimiento informado, pilar de la bioética clásica, ha colapsado frente a la opacidad de los algoritmos de «caja negra». La conclusión es clara: el dato biológico es inalienable y su protección debe ser proactiva. El concepto de «patrimonio

neuro-genómico» debe ser reconocido como el bien jurídico supremo, pues de él emanan la identidad, el libre desarrollo de la personalidad y la soberanía del individuo sobre su propia biología.

### **5.2 El Imperativo Categórico de la Reforma al Artículo 19 Constitucional**

La paradoja detectada en el artículo 19 constitucional exige una respuesta inmediata del legislador. No podemos permitir un sistema que vigila con celo la libertad física mientras ignora la libertad cognitiva. La inclusión del derecho al «Habeas Mens» en la Constitución no es una opción teórica, sino una necesidad de supervivencia democrática. Como se analizó a través de los casos de manipulación algorítmica y «huellas de memoria», el Estado de Derecho carece de sentido si el ciudadano no tiene la garantía de que su proceso de pensamiento es libre de interferencias externas. El cráneo humano debe ser declarado el límite sagrado de la vigilancia estatal y corporativa.

### **5.3 Hacia una Soberanía Cognitiva y la Desconexión Biológica**

La gobernanza biojurídica propuesta en la Sección 4 —integrando el concepto de «Neurodato» y el fortalecimiento de los Comités de Bioética— traza la ruta hacia lo que denominamos Soberanía Cognitiva. El Estado mexicano debe dejar de ser un receptor pasivo de tecnología para convertirse en un regulador ético que garantice el «Derecho a la Desconexión Biológica». La dignidad humana debe prevalecer sobre la eficiencia técnica; ningún algoritmo, por sofisticado que sea, puede tener la autoridad legal para dictar el destino de un ser humano basado en su código genético o su actividad neuronal.

### **5.4 El Rol de México en la Vanguardia Latinoamericana**

Siguiendo el precedente de Chile, México tiene la oportunidad histórica de liderar la regulación de los neuroderechos en el bloque regional. La implementación de un régimen de responsabilidad objetiva para los desarrolladores de IA biotecnológica y la exigencia de la «Intervención Humana» en las decisiones algorítmicas sentarán un precedente necesario para frenar la eugenesia digital. El futuro del Derecho no está en la regulación de las máquinas, sino en la defensa irreductible del ser humano frente a la instrumentalización tecnológica.

### **5.5 Limitaciones del estudio**

El presente trabajo se delimita al análisis dogmático y regulatorio del marco normativo mexicano e interamericano respecto a la IA y el patrimonio neuro-genómico. No incluye una evaluación empírica cuantitativa sobre el nivel de implementación actual de estas tecnologías en centros sanitarios o empresas biotecnológicas en México, ni aborda las especificaciones técnicas de la arquitectura computacional de las redes neuronales utilizadas en la decodificación cerebral.

## **6. REFERENCIAS BIBLIOGRÁFICAS**

- CHILE. Reforma constitucional. Modificación al artículo 19, n.º 1, de la Carta Fundamental para establecer que el desarrollo científico y tecnológico estará al servicio de las personas. Diario Oficial de la República de Chile, 25 de octubre de 2021.
- GYMREK, Melissa; MCGUIRE, Amy L.; GOLAN, David; HALPERIN, Eran; ERLICH, Yaniv. «Identifying personal genomes by surname inference». *Science*, 2013, vol. 339, núm. 6117, pp. 321–324. DOI: 10.1126/science.1229566.

- IENCA, Marcello; ANDORNO, Roberto. «Hacia nuevos derechos humanos en la era de la neurociencia y la neurotecnología». *Análisis Filosófico*, 2021, vol. 41, núm. 1, pp. 141–185. DOI: 10.36446/AF.2021.386.
- IENCA, Marcello; ANDORNO, Roberto. «Towards new human rights in the age of neuroscience and neurotechnology». *Life Sciences, Society and Policy*, 2017, vol. 13, art. 5. DOI: 10.1186/s40504-017-0050-1.
- INDIA. *State of Maharashtra v. Sharma*, C.C. No. 508/07, Court of Sessions, Pune, 12 de junio de 2008.
- MÉXICO. *Constitución Política de los Estados Unidos Mexicanos*. Diario Oficial de la Federación, 1917.
- MÉXICO. *Ley Federal de Protección de Datos Personales en Posesión de los Particulares*. Diario Oficial de la Federación, 2010. Artículo 3, fracción VI.
- MÉXICO. *Ley General de Salud*. Diario Oficial de la Federación, 1984. Artículo 103 Bis.
- O'NEIL, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Publishing, 2016. ISBN 9780553418811.
- SAN VICENTE PARADA, Aida del Carmen. *Temas selectos de bioética y derecho*. Valencia: Tirant lo Blanch, 2025. ISBN 9788410562721.
- SCHWAB, Klaus. *The Fourth Industrial Revolution*. Geneva: World Economic Forum, 2016. ISBN 9781944835002.
- SUPREMA CORTE DE JUSTICIA DE LA NACIÓN (MÉXICO). «Derecho al libre desarrollo de la personalidad. Su dimensión externa e interna». Tesis 1a. J. 4/2019 (10a.). *Semanario Judicial de la Federación*, 2019.
- SUPREMA CORTE DE JUSTICIA DE LA NACIÓN (MÉXICO). «Libre desarrollo de la personalidad». Tesis 1a. J. 4/2019 (10a.). *Semanario Judicial de la Federación*, 2019.
- UNESCO. *Declaración Universal sobre el Genoma Humano y los Derechos Humanos*. París: UNESCO, 1997.
- VAYENA, Effy; BLASIMME, Alessandro; COHEN, I. Glenn. «Machine learning in medicine: Addressing ethical challenges». *PLOS Medicine*, 2018, vol. 15, núm. 11, e1002689. DOI: 10.1371/journal.pmed.1002689.
- YUSTE, Rafael; GOERING, Sara; AGÜERA Y ARCAS, Blaise; et al. «Four ethical priorities for neurotechnologies and AI». *Nature*, 2017, vol. 551, núm. 7679, pp. 159–163. DOI: 10.1038/551159a.
- ZUBOFF, Shoshana. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs, 2019. ISBN 9781610395694.

# Identidad narrativa y derechos humanos: Una aproximación hermenéutica a la migración forzada y la inteligencia artificial\*

## *Narrative identity and human rights: A hermeneutical approach to forced migration and artificial intelligence*

Francisco Raúl Matus Ramírez\*\*

### RESUMEN

Este artículo argumenta que los procesos algorítmicos de inteligencia artificial en contextos migratorios, particularmente ImmigrationOS, suprimen las condiciones de posibilidad desde las cuales los derechos humanos se constituyen en la identidad narrativa del sí mismo como otro. Se propone una comprensión hermenéutica de los derechos humanos, en contraposición a las teorías que buscan fundamentarlos, que tome en cuenta la historia particular desde la cual se experimentan y narran. Para comprender los derechos humanos, se analiza la identidad narrativa (*idem e ipse*) a partir de la propuesta filosófica de Paul Ricoeur, en la que el sujeto se reconoce como hombre capaz de actuar, narrar, decir e imputarse sus propios actos. Se concluye que la migración forzada interrumpe las condiciones para la constitución narrativa del migrante, y que el uso de la inteligencia artificial suprime su particularidad narrativa, anula su capacidad de imputación y niega las condiciones relacionales e institucionales que hacen posible su reconocimiento.

### PALABRAS CLAVE

fundamentación de derechos humanos, hermenéutica, identidad narrativa, migración forzada, inteligencia artificial.

### ABSTRACT

This article argues that algorithmic processes of artificial intelligence in migration contexts, particularly ImmigrationOS, suppress the conditions of possibility under which human rights are constituted within the narrative identity of the self as the other. This paper proposes a hermeneutic understanding of human rights, as opposed to theories that seek to establish their foundations, that takes into account the specific historical contexts in which they are experienced and narrated. To understand human rights, this paper analyzes narrative identity (*idem and ipse*) based on Paul Ricoeur's philosophical framework, in which the subject recognizes himself as a person capable of acting, narrating, speaking, and taking responsibility for his own actions. It is concluded that forced migration interrupts the conditions for the migrant's narrative construction, and that the use of artificial intelligence suppresses their narrative distinctiveness, nullifies their capacity for imputation, and denies the relational and institutional conditions that make their recognition possible.

### KEYWORDS

foundationalism of human rights, hermeneutics, narrative identity, forced migration, artificial intelligence.

\* Artículo de investigación

\*\* Profesor de la Facultad de Derecho y Ciencias Sociales de la UMSNH, y de la División de Estudios de Posgrado de la misma Facultad, México (raul.matus@umich.mx). <https://orcid.org/0009-0006-7869-8762>.

## SUMARIO

1. Introducción
2. Datos preliminares
3. Los derechos humanos y su fundamentación
4. Identidad narrativa
5. Experiencia migratoria
6. Inteligencia artificial e identidad narrativa
7. Conclusiones
8. Referencias bibliográficas

### 1. INTRODUCCIÓN

En la actualidad, se han desarrollado procesos algorítmicos de inteligencia artificial con la finalidad de optimizar el control migratorio. En Estados Unidos, por ejemplo, la implementación en 2025 de ImmigrationOS,<sup>1</sup> desarrollado por Palantir, ha representado una innovación técnica en la gestión fronteriza. Sin embargo, el uso de la inteligencia artificial en contextos que involucran la praxis como modo de acción humana suele ser mucho más complejo que la mera optimización técnica de un sistema.

Los procesos migratorios llevan implícitas personas con una identidad narrativa concreta en la que se encuentran constituidos derechos humanos. Afirmación que tiene implicaciones teóricas que no siempre son compatibles con las posturas clásicas sobre su fundamentación (enfoques positivistas, iusnaturalistas, realistas o dualistas, por nombrar solo algunos), pues dichas perspectivas tienden a presuponer un sujeto solipsista y ahistórico, abstraído de su experiencia concreta, al que se le otorgan derechos ajenos a él, lo que provoca contradicciones al momento de articular las dimensiones ética, moral y jurídica de los derechos.

Comprender los derechos humanos como parte constitutiva de la identidad narrativa del sí mismo como otro, perspectiva que se puede inferir de la filosofía de Paul Ricoeur,<sup>2</sup> permite articular sus distintas dimensiones. Esto es especialmente relevante en casos como los procesos algorítmicos relacionados con la migración, los cuales tienen implicaciones no solo jurídicas, sino que impactan directamente la identidad narrativa de las personas a quienes se aplican desde distintas dimensiones.

Mark Coeckelbergh expone la forma en que la inteligencia artificial reduce a las personas a datos y a clasificaciones estadísticas con la finalidad de hacer predicciones,<sup>3</sup> desplazando a la persona de ser autor del propio relato a convertirse en el objeto de un análisis estadístico. Sin embargo, su análisis se sitúa en el plano hermenéutico-existencial sin desarrollar las implicaciones de esa anulación sobre la constitución de

---

<sup>1</sup>ImmigrationOS es un acrónimo de Immigration Lifecycle Operating System (Sistema Operativo del Ciclo de Vida Migratorio). Es un conjunto de capacidades de inteligencia artificial incorporadas mediante la modificación P00009 al contrato federal número 70CTD022FR0000170, suscrito entre el Servicio de Inmigración y Control de Aduanas de Estados Unidos y Palantir Technologies Inc., con fecha 11 de abril de 2025. U.S. IMMIGRATION AND CUSTOMS ENFORCEMENT, Contract modification P00009, No. 70CTD022FR0000170, U.S. Department of Homeland Security, 2025.

<sup>2</sup>RICOEUR, Paul, *Sí mismo como otro*, México, Siglo Veintiuno Editores, 1996.

<sup>3</sup>Cfr. COECKELBERGH, Mark, "Narrative responsibility and artificial intelligence. How AI challenges human responsibility and sense-making", *AI & Society*, vol. 38, 2023, pp. 2437-2450.

los derechos humanos en la identidad narrativa de la persona concreta. El uso de la inteligencia artificial en el contexto migratorio no solo plantea el problema de la responsabilidad narrativa, sino que suprime las condiciones de posibilidad desde las cuales los derechos humanos se constituyen en el sí mismo como otro.

Por su parte, Hille Haker analiza el contexto migratorio desde el enfoque ético de Paul Ricoeur y Hannah Arendt, mostrando que el régimen internacional de derechos humanos se ha reducido en la práctica a una mera referencia retórica,<sup>4</sup> ya que no cuentan con la autoridad política para su protección efectiva.

Para Haker los derechos humanos existen como expresión del agente con capacidad de respuesta y responsabilidad que brotan del derecho a tener derechos, articulándolos con la ética ricoeuriana de la responsabilidad y el reconocimiento.<sup>5</sup>

Sin embargo, su análisis se queda en el nivel de capacidad moral general (capacidad de responder y ser responsable), de modo que los derechos aparecen como algo externo que se atribuye al sí mismo cuya capacidad moral se presupone ya constituida. Desde esta perspectiva no se logra responder plenamente a la tensión que se suscita al plantear el uso de la inteligencia artificial en contextos migratorios, pues el problema no solo consiste en que dichos algoritmos transgredan derechos por su desconexión con la capacidad moral que los sustenta, sino que suprimen las condiciones de posibilidad desde las cuales los derechos humanos se constituyen en la identidad narrativa del sí mismo como otro.

Asimismo, Elsa Lechner y Leticia Renault enfatizan la importancia de la narración de la experiencia migratoria como vía para reconstruir la identidad narrativa fracturada cuando el migrante no es escuchado.<sup>6</sup> Esa perspectiva destaca el reconocimiento del otro como un acto relacional; sin embargo, la invisibilidad que describen se profundiza en el mismo contexto migratorio con la incorporación del uso de la inteligencia artificial que suprime la particularidad de la persona, desplazando la escucha por una mera clasificación.

Tanto Coeckelbergh, Haker, Lechner y Renault abren discusiones relevantes: el primero sobre las implicaciones de la inteligencia artificial en la responsabilidad narrativa; la segunda sobre los derechos humanos en contextos migratorios desde la ética ricoeuriana; y las últimas dos sobre la experiencia biográfica del migrante. Sin embargo, ninguno aborda de manera directa el problema sobre los procesos algorítmicos de decisión migratoria y su impacto en la constitución de los derechos humanos en la identidad narrativa del sí mismo como otro. El objetivo de este artículo es, entonces, argumentar que dichos procesos, en particular ImmigrationOS, restringen las condiciones de posibilidad desde las cuales los derechos humanos se constituyen como parte de la identidad narrativa del sí mismo como otro.

---

<sup>4</sup>Cfr. HAKER, Hille, "No space. Nowhere. Refugees and the Problem of Human Rights in Arendt and Ricoeur", *Études Ricoeuriennes / Ricoeur Studies*, vol. 8, núm. 2, 2017, pp. 22-45.

<sup>5</sup>HAKER, Hille, "No space. Nowhere. Refugees and the Problem of Human Rights in Arendt and Ricoeur", *Études Ricoeuriennes / Ricoeur Studies*, vol. 8, núm. 2, 2017, p. 33.

<sup>6</sup>LECHNER, Elsa y RENAULT, Leticia, "Migration Experiences and Narrative Identities: Viewing Alterity from Biographical Research", *Critical Hermeneutics*, vol. 2, núm. 1, 2018, p. 13.

## 2. DATOS PRELIMINARES

Para abordar el argumento, se analizará la propuesta ricoeuriana de la identidad narrativa mediante un enfoque hermenéutico aplicado a la constitución de los derechos humanos, en contraste con algunas de las teorías más relevantes dirigidas a su fundamentación. Asimismo, se formulará un análisis crítico de los procesos algorítmicos de inteligencia artificial orientados al control migratorio, tomando como referencia la implementación en 2025 de ImmigrationOS en Estados Unidos, en diálogo con Coeckelbergh, Haker, Lechner y Renault. Finalmente, se expondrán las conclusiones.

## 3. LOS DERECHOS HUMANOS Y SU FUNDAMENTACIÓN

Existen múltiples posiciones teóricas dirigidas a fundamentar los derechos humanos. Estos enfoques parten de distintas perspectivas filosóficas que permiten su identificación en clasificaciones amplias. Entre las propuestas taxonómicas, Enrique Pérez Luño distingue cuatro grupos amplios: el positivismo, el iusnaturalismo, el realismo y el dualismo, todos orientados a alcanzar un fundamento último para los derechos humanos.<sup>7</sup>

El positivismo, ya sea desde su orientación ideológica, el formalismo jurídico, el escepticismo ético o el positivismo conceptual, propone sostener los derechos humanos en su validez jurídica y, para evitar referencias ético-morales fundacionales, establece como único origen su incorporación en ordenamientos jurídicos positivos. Desde esta perspectiva, el sujeto es abstraído de su experiencia narrativa concreta, pues al ser reducido a una categoría normativa, su identidad narrativa resulta irrelevante para el reconocimiento, ejercicio y protección de sus derechos.

Por su parte, el iusnaturalismo, en la forma en que es presentado por Pérez Luño, pretende una fundamentación última de los derechos humanos. Sin embargo, dada la complejidad de este amplio enfoque, resulta necesario recurrir a la distinción propuesta por Rodríguez-Toubes entre teorías sustantivas y formales para identificar sus variantes más relevantes para este texto.<sup>8</sup>

Las teorías morales sustantivas (como el naturalismo, la axiología o el utilitarismo) sitúan el origen último de los derechos en propiedades inmutables, estáticas e inherentes a los humanos, en valores superiores o en la maximización del beneficio de la mayoría, respectivamente.<sup>9</sup> Desde esta perspectiva, los derechos son considerados como objetos con propiedades determinables mediante criterios abstractos, por lo que la dimensión interpretativa y narrativa de la experiencia de quien los vive y padece resulta irrelevante para su determinación.

Las teorías morales formales, como la ética discursiva, el constructivismo o las implicaciones lógicas de la acción, sitúan el origen último de los derechos humanos en consensos basados en condiciones ideales del discurso, en perspectivas prudenciales o

---

<sup>7</sup>Cfr. PÉREZ LUÑO, Antonio-Enrique, *Derechos humanos, Estado de Derecho y Constitución*, 3.ª ed., Madrid, Tecnos, 1990, pp. 134-187.

<sup>8</sup>Cfr. RODRÍGUEZ-TOUBES, Joaquín, *La razón de los derechos*, Madrid, Tecnos, 1995.

<sup>9</sup>Cfr. RODRÍGUEZ-TOUBES, Joaquín, *La razón de los derechos*, Madrid, Tecnos, 1995, pp. 241-299 y p. 242. Para este autor, una fundamentación sustantiva se da cuando la norma en que se justifican los derechos humanos tiene un contenido material determinado.

en la aceptación de los derechos en condiciones ideales de racionalidad.<sup>10</sup> Si bien estas teorías mantienen contacto parcial con la experiencia de los sujetos que participan en el discurso, dicha experiencia es resultado de una abstracción ideal relativa a condiciones trascendentales del diálogo, de ahí que se llegue a universalismos que no logran conectar del todo con la interpretación y la experiencia concreta que el sujeto tiene sobre los derechos.

El realismo jurídico, para Pérez Luño, fundamenta los derechos humanos en la adopción y ratificación de los instrumentos internacionales que consagran tales derechos. Dado que dicha legislación es resultado de un consenso internacional, se considera irrelevante repensar su justificación, como sugiere Norberto Bobbio.<sup>11</sup> Así, la aceptación fáctica de los derechos reduce la experiencia del sujeto a datos empíricos derivados de la aceptación positiva de los mismos, de modo que los derechos humanos son considerados como datos sociales observables.

Finalmente, el dualismo propone una fundamentación de los derechos humanos a la vez moral y jurídica. Gregorio Peces-Barba representa este esfuerzo al proponer que el origen de los derechos humanos "se encuentra en el plano de la Moral",<sup>12</sup> otorgándoles una naturaleza prejurídica. Sin embargo, sostiene al mismo tiempo que el derecho positivo es su fundamento inmediato, pues solo hasta que pertenecen a un ordenamiento jurídico pueden considerarse propiamente derechos humanos.<sup>13</sup>

La tensión surge al sostener ambas afirmaciones de manera simultánea, si los derechos humanos tienen su origen en la moral, pero "no se completan hasta su positivación",<sup>14</sup> se puede deducir que lo jurídico sería su condición de posibilidad, por lo que su dimensión moral quedaría reducida a un plano secundario. Como señala Rodríguez-Toubes: al proponer un origen simultáneamente moral y jurídico se termina por privilegiar solo uno de ellos.<sup>15</sup>

La contradicción no reside en proponer una doble dimensión moral y jurídica, sino en buscar un fundamento último, único e inmutable, ejercicio propio de la modernidad que, desde la filosofía cartesiana, busca establecer un fundamento racional e incuestionable.

El ejercicio de fundamentar tiene implicaciones antropológicas importantes, pues el cogito ergo sum, al presentarse como fundamento del conocimiento, presupone un sujeto solipsista y ahistórico.

---

<sup>10</sup>Cfr. RODRÍGUEZ-TOUBES, Joaquín, *La razón de los derechos*, Madrid, Tecnos, 1995, pp. 152-240. Las teorías que forman parte de este apartado son sumamente complejas, por lo que un análisis riguroso excedería el propósito de este trabajo; en su lugar, se contrastarán únicamente sus planteamientos con la idea de los derechos como parte constitutiva de la identidad narrativa del sí mismo como otro.

<sup>11</sup>Cfr. BOBBIO, Norberto, *El tiempo de los derechos*, trad. de Rafael de Asís Roig, Madrid, Sistema, 1991, pp. 63-83.

<sup>12</sup>PECES-BARBA, Gregorio, "Sobre el fundamento de los derechos humanos. Un problema de moral y derecho", en MUGUERZA, Javier y PECES-BARBA, Gregorio, coords., *El fundamento de los derechos humanos*, Madrid, Debate, 1989, p. 267.

<sup>13</sup>PECES-BARBA, Gregorio, *Curso de derechos fundamentales. (I) Teoría general*, con la colaboración de R. de Asís y A. Llamas Gascón, Madrid, Eudema, 1983, p. 91.

<sup>14</sup>PECES-BARBA, Gregorio, "Sobre el fundamento de los derechos humanos. Un problema de moral y derecho", en MUGUERZA, Javier y PECES-BARBA, Gregorio, coords., *El fundamento de los derechos humanos*, Madrid, Debate, 1989, p. 267.

<sup>15</sup>Cfr. RODRÍGUEZ-TOUBES, Joaquín, *La razón de los derechos*, Madrid, Tecnos, 1995, p. 136.

Una propuesta distinta a la fundamentación de los derechos humanos, típica de la modernidad, es buscar su comprensión hermenéutica (método que utiliza Paul Ricoeur en su filosofía práctica).<sup>16</sup> La fundamentación y la comprensión tienen propósitos diferentes, mientras que la primera busca encontrar el origen, la razón, la naturaleza, la justificación o el principio de los derechos;<sup>17</sup> la segunda pretende entenderlos de forma integral, con todas sus dimensiones, mediante una interpretación que implica considerar a los derechos como inseparables de la experiencia que se constituye en la identidad narrativa del sí mismo como otro.

De esa manera, Ricoeur quiebra el cogito cartesiano que "lleva implícita una ambición de fundamento último",<sup>18</sup> y con ello abandona la búsqueda de un principio único e inmutable que funja como base de los derechos. En su lugar, articula las dimensiones ética, deontológica y jurídica de los derechos humanos mediante su comprensión hermenéutica. De ahí que el dualismo que articula lo moral y lo jurídico no sea incoherente, sino constitutivo en una comprensión integral, siempre y cuando no se pretenda alcanzar un origen último.

Sin embargo, la comprensión hermenéutica no opera en abstracto, requiere de alguien con identidad narrativa en el que se constituyan los derechos humanos desde sus distintas dimensiones. Si bien, Paul Ricoeur no desarrolla una teoría específica de los derechos humanos, Guido Gorgoni, desde una relectura de Ricoeur, propone su comprensión desde las dimensiones del sí mismo como otro, mostrando que lo jurídico cumple un "papel crucial en la comprensión y construcción de la identidad del sí",<sup>19</sup> no como una abstracción normativa, sino que se integran en el recorrido ético del sí mismo como otro.

De ahí que la identidad narrativa es el lugar donde se constituyen los derechos humanos, por lo que comprenderla es indispensable para entender el impacto de las experiencias de migración forzada y el uso de procesos algorítmicos de inteligencia artificial en dichos contextos.

#### 4. IDENTIDAD NARRATIVA

La identidad ha sido un tema central en la filosofía de Paul Ricoeur, quien la aborda mediante la fenomenología hermenéutica, método en el que confluyen, entre otras influencias, la fenomenología heideggeriana<sup>20</sup> y la hermenéutica de Gadamer.<sup>21</sup> Desde esta perspectiva, la comprensión del sí mismo como otro no es inmediata, pues se da a través del rodeo por los símbolos, los textos y por la narrativa.

Ricoeur propone superar el cógito cartesiano, que parte de un sujeto que se conoce directamente desde la inmediatez, a través de la propuesta del sí mismo como otro (que

<sup>16</sup>Es preciso señalar que Paul Ricoeur no formula una comprensión hermenéutica de los derechos humanos; en este texto se aplica su método fenomenológico hermenéutico para tal efecto.

<sup>17</sup>ABBAGNANO, Nicola, Diccionario de filosofía, voz "fundamento", México, FCE, 1999, pp. 578-580; FERRATER, José, Diccionario de filosofía, voz "fundamento", Barcelona, Ariel, 1994, pp. 1413-1414.

<sup>18</sup>RICOEUR, Paul, *Sí mismo como otro*, México, Siglo Veintiuno Editores, 1996, p. XV.

<sup>19</sup>Cfr. GORGONI, Guido, "Paul Ricoeur filósofo del derecho: el sujeto de derecho y de responsabilidad como hombre capable", *Nuevo Itinerario*, vol. 20, núm. 1, 2024.

<sup>20</sup>Cfr. HEIDEGGER, Martin, *El ser y el tiempo*, México, Fondo de Cultura Económica, 1993.

<sup>21</sup>Cfr. GADAMER, Hans-Georg, *Verdad y Método I*, Salamanca, Ediciones Sígueme, 1999.

solo puede llegar a su comprensión a través de un rodeo). Por un lado, el pronombre reflexivo sí ya implica un rodeo interpretativo, mostrando el dinamismo constitutivo de la identidad.<sup>22</sup> Por otro lado, el sí se abre a la alteridad, pues no es un ente aislado, como el sujeto cartesiano, sino que se constituye en relación con el otro al mantener una "dialéctica del sí y del otro distinto del sí"<sup>23</sup>.

Para el filósofo francés, la identidad se integra de dos momentos ontológicos: la mismidad es aquello que permanece reconocible en el tiempo, y la ipseidad, que no remite a una sustancia fija sino a la capacidad de responder por sí mismo a través del tiempo. Estas dos dimensiones no se excluyen. En el carácter, por ejemplo, el "idem e ipse tienden a coincidir",<sup>24</sup> pues ambos designan disposiciones duraderas en las que se reconoce a una persona, estas se sedimentan entre lo ya adquirido y lo que continuamente se está contrayendo.

La narración constituye la identidad al reconfigurar la experiencia como mediación entre la mismidad y la ipseidad. El relato articula lo que persiste y lo que cambia con el tiempo, pero sin suprimir la tensión que ambos planos suscitan.

La identidad narrativa para Ricoeur es una forma de autocomprensión mediada por el relato que se da en el acto de narrarse, constituyéndose a sí mismo en relación con otros. En este acto, el sí mismo se reconoce como hombre capaz, alguien que puede actuar, narrar, decir e imputarse responsabilidad; sin embargo, es con la imputabilidad, que "la noción de sujeto capaz alcanza su más alta significación",<sup>25</sup> porque en dicha acción se recapitulan los otros modos de capacidad.

Cuando el sí se reconoce como responsable de sus actos, supera toda designación por atributos genéricos, estáticos o sustanciales que responden a qué es el sujeto. La imputabilidad exige formular una pregunta distinta: ¿quién es responsable de sus actos?

Por otra parte, la imputabilidad como acción tiene implícitos predicados ético-morales que emergen en el acto de narrar. En primer lugar, el relato que tiene lugar en la identidad no es solo descripción de lo ocurrido: la narración implica un juicio y una valoración, de ahí que sea el enlace entre "la teoría de la acción y la teoría ética"<sup>26</sup>. En segundo lugar, cuando estos predicados son aplicados reflexivamente a quien es capaz de imputación, se constituye en su identidad narrativa las dimensiones teleológica y deontológica, esta última en sus dimensiones moral y jurídica. Así, cuando el sí se reconoce como sujeto capaz, se reconoce también como responsable ante otros.

Las dimensiones ética y moral de la identidad narrativa se articulan en tres momentos: tender a la vida buena, en la que el sujeto valora sus propias capacidades desde la estima del sí; la solicitud, en la que se reconoce con y para otro desde el cuidado y la reciprocidad; y las instituciones justas, en las que el reconocimiento adquiere forma estructural y universal.<sup>27</sup>

<sup>22</sup>Cfr. RICOEUR, Paul, *Sí mismo como otro*, México, Siglo Veintiuno Editores, 1996, pp. XI-XII.

<sup>23</sup>RICOEUR, Paul, *Sí mismo como otro*, México, Siglo Veintiuno Editores, 1996, pp. XIII-XIV.

<sup>24</sup>RICOEUR, Paul, *Sí mismo como otro*, México, Siglo Veintiuno Editores, 1996, p. 113.

<sup>25</sup>RICOEUR, Paul, *Caminos del reconocimiento*. Tres estudios, trad. Neira, Agustín, México, Fondo de Cultura Económica, 2006, p. 139.

<sup>26</sup>RICOEUR, Paul, *Sí mismo como otro*, México, Siglo Veintiuno Editores, 1996, p. 152.

<sup>27</sup>RICOEUR, Paul, *Sí mismo como otro*, México, Siglo Veintiuno Editores, 1996, p. 176.

Así, la identidad narrativa, en la que el sí se constituye y reconoce como hombre capaz, no se construye en solitario sino a partir de la alteridad. El otro no es un mero testigo externo de la acción ni un simple contraste desde el cual el sí se identifica, sino que es parte constitutiva del sí mismo. En ese sentido, una comprensión hermenéutica de la identidad implica siempre un rodeo por el otro.

El sí mismo como otro se constituye desde la narración que formula de la experiencia vivida, reconociéndose como sujeto capaz. En ese contexto, los derechos humanos no pueden ser reducidos a objetos externos con propiedades específicas ni al resultado de abstracciones, reconocidos, otorgados o garantizados a un sujeto individual, autónomo y abstracto, pues considerarlos así los desvincula de la experiencia concreta. En cambio, cuando se comprenden como parte constitutiva de la identidad del sí mismo como otro que se narra y reconoce como hombre capaz e imputable, adquieren "espesor ético y antropológico"<sup>28</sup>.

Comprenderlos desde la identidad narrativa implica un contenido ético y moral que se despliega desde el recorrido por los tres momentos de la pequeña ética ricoeuriana. Así, los derechos humanos se manifiestan en la experiencia concreta del sujeto y a través de la narración se constituyen como parte de la identidad del sí mismo como otro, dotándoles de sentido ético, deontológico y jurídico, mediante un relato que nunca es neutral.

Para Ricoeur, el sujeto capaz de actuar, narrar y decir se realiza plenamente en su capacidad de imputación reflexiva como sujeto de derecho. El filósofo francés utiliza el lenguaje jurídico formal, pero supera la idea del sujeto de derecho como una abstracción jurídica, resaltando sus dimensiones ética y moral constituidas en su identidad narrativa.

El reconocimiento como sujeto de derechos tiene dos objetivos: la norma, con validez universal, y el otro, con su singularidad. Dicha conjunción enriquece las capacidades, pues la dimensión jurídica que se desprende de ser sujeto de derechos se constituye en el sí mismo como otro enriqueciendo su identidad narrativa.<sup>29</sup>

Sin embargo, ¿cómo se materializan las capacidades en derechos reales? No basta con ser sujeto de derechos en el plano jurídico formal, sino que debe darse un proceso continuo de actualización de las capacidades desde la relación con otros y las estructuras jurídicas, políticas y sociales, para que los derechos humanos sean poderes reales. Esta actualización es resultado de la indignación, actualizada por el sentido de lo justo,<sup>30</sup> debido a la falta de reconocimiento y de las luchas concretas por alcanzarlo.<sup>31</sup>

La vulnerabilidad que expresa la indignación por la falta de reconocimiento es complementaria con la autonomía del sujeto de derechos que aspira a constituirse como hombre capaz, tarea que debe rehacerse siempre como parte de la praxis humana. El sujeto vulnerable solicita reconocimiento para volverse autónomo, pero al volverse

---

<sup>28</sup>GORGONI, Guido, "Paul Ricoeur filósofo del derecho: el sujeto de derecho y de responsabilidad como *homme capable*", Nuevo Itinerario, vol. 20, núm. 1, 2024, p. 4.

<sup>29</sup>RICOEUR, Paul, Caminos del reconocimiento. Tres estudios, trad. Neira, Agustín, México, Fondo de Cultura Económica, 2006, p. 250.

<sup>30</sup>Cfr. VILLA SÁNCHEZ, J. A., "Sobre el sentido de lo justo y la ley", En-Claves del Pensamiento, núm. 37, 2025, pp. 216-236.

<sup>31</sup>GORGONI, Guido, "Paul Ricoeur filósofo del derecho: el sujeto de derecho y de responsabilidad como *homme capable*", Nuevo Itinerario, vol. 20, núm. 1, 2024, p. 5.

autónomo no abandona su fragilidad.<sup>32</sup> Esta dinámica expone la insuficiencia de fundamentar los derechos humanos en la autonomía abstracta del sujeto individual moderno, pues el sí mismo como otro los constituye en su identidad narrativa al vivíros, experimentarlos, padecerlos y reclamarlos desde su fragilidad y vulnerabilidad.

## 5. EXPERIENCIA MIGRATORIA

La vulnerabilidad y fragilidad que experimenta el sí mismo como otro en la migración forzada (aquella que no deriva de la libre elección sino que es obligada debido a condiciones económicas adversas, de violencia, de persecución, entre otras razones) no comienza con el proceso de huida, sino que es precisamente en el lugar de origen donde se interrumpe inicialmente la constitución de derechos en su identidad narrativa, fractura que lo fuerza a buscar en otro lugar las condiciones de posibilidad para reconstruirse como sujeto capaz.<sup>33</sup>

Hille Haker retoma la crítica que Hannah Arendt formula contra el régimen internacional de derechos humanos, que no repara en aquellos que "no están protegidos en sus propios países".<sup>34</sup> a pesar de que dichos derechos se revistan de universalidad al proclamarse como "derechos iguales e inalienables de todos los miembros de la familia humana"<sup>35</sup>. Según la Declaración Universal, los derechos humanos deberían ser reconocidos incluso en países en los que no se cuenta con ciudadanía; sin embargo, en la práctica, sin ciudadanía o permiso de acceso al interior de un país no existe reconocimiento, y lo más común es que quienes se ven forzados a migrar carezcan del estatus jurídico que les permita ingresar a otro país.<sup>36</sup>

Mediante la ciudadanía los derechos humanos se constituyen de distintas formas. Esto implica que un ciudadano tiene derechos de tránsito y residencia, de identidad y personalidad jurídica, derechos culturales y lingüísticos, derecho al libre desarrollo de la personalidad, derechos a la familia y a la vida privada, derechos civiles y políticos, laborales, y a la educación y a la salud; en otras palabras: "el derecho a tener derechos o el derecho de cada individuo a pertenecer a la Humanidad"<sup>37</sup>.

La migración forzada implica la pérdida del hogar, es decir, todo el entramado de relaciones sociales, pertenencias, lengua materna e instituciones, que dan reconocimiento al sí mismo como otro y sostienen la identidad narrativa. Sin embargo, un problema más profundo que el desplazamiento en sí, es la imposibilidad de ser aceptado en ningún

<sup>32</sup>GORGONI, Guido, "Paul Ricoeur filósofo del derecho: el sujeto de derecho y de responsabilidad como *homme capable*", *Nuevo Itinerario*, vol. 20, núm. 1, 2024, p. 9.

<sup>33</sup>Incluso dentro de los países de origen se puede vivir un desplazamiento forzado; sin embargo, en este texto se enfatiza la migración a otros países.

<sup>34</sup>HAKER, Hille, "No space. Nowhere. Refugees and the Problem of Human Rights in Arendt and Ricoeur", *Études Ricoeuriennes / Ricoeur Studies*, vol. 8, núm. 2, 2017, p. 24.

<sup>35</sup>ORGANIZACIÓN DE LAS NACIONES UNIDAS, *Declaración Universal de Derechos Humanos*, ONU, 1948, Preámbulo, p. 5.

<sup>36</sup>Existe una tradición de discusión abierta por Kant, en la que propone un programa jurídico-político para un derecho cosmopolita en el que operan nociones como la ciudadanía mundial y la hospitalidad universal; asimismo, autores como Habermas, Benhabib, Pogge, entre otros, abordan esa discusión desde distintas perspectivas; sin embargo, dicho tema no será abordado en este texto porque excede sus propósitos.

<sup>37</sup>ARENDT, Hannah, *Los orígenes del totalitarismo*, trad. Solana, Guillermo, Madrid, Taurus, 1998, pp. 248 y 249.

otro lugar para reconstituirse, situación que no obedece a un problema de "espacio sino de organización política"<sup>38</sup>.

Una persona no pierde su estatus de sujeto de derechos humanos por su condición migratoria; sin embargo, dicha condición sí tiene repercusiones en el reconocimiento y la protección de esos derechos. Esto lleva a preguntarse: ¿de qué forma se interrumpe la constitución de los derechos humanos en la identidad narrativa del sí mismo como otro en la migración forzada? La interrupción de las condiciones de posibilidad ocurre cuando la persona es desplazada fuera de las prácticas sociales, de sus relaciones simétricas, de su cultura y de su lengua materna, lo que afecta su propio relato y el reconocimiento tanto de sí misma como por los otros. La lengua, los vínculos y la cultura de los que se ve forzada a renunciar no son elementos externos sino constitutivos de sí; los derechos humanos no crean esas prácticas, pero son su condición de posibilidad, pues al ser reconocidas y protegidas el sujeto puede narrarse y constituirse en sus dimensiones ética, moral y jurídica. Es decir, el sí mismo como otro constituye en su identidad narrativa los derechos humanos y aquello que protegen de manera simultánea. Si esas condiciones de posibilidad le son negadas tanto en el país de origen como en el de destino, la ruptura se profundiza. La identidad narrativa no desaparece, pero se genera una ruptura en su base constitutiva, ya que la cultura y las relaciones simétricas que la sostenían mediante el reconocimiento ya no forman parte del presente de la persona, porque han sido negadas las condiciones de posibilidad para reconstituirlas.

Los derechos, de esa forma, no son entendidos como datos empíricos —capacidades, necesidades, objetos externos—, ni como meros preceptos o normas positivas, ni como el resultado de consensos abstractos o lógicas prudenciales. Son, en cambio, condiciones de posibilidad para que las capacidades y las necesidades se constituyan en la identidad narrativa del sí mismo como otro; y al mismo tiempo, son constitutivos del sí desde sus dimensiones ética, moral y jurídica: no desde el solipsismo, sino desde la intencionalidad ética, la solicitud y las instituciones justas que se construyen en la alteridad mediante la comunicación.

La migración forzada afecta el relato del sí mismo como otro desde distintas dimensiones. En primer lugar, presenta una situación paradójica: por un lado, es producto de la victimización de quienes se ven obligados a migrar; por otro, deriva en la reiterada criminalización "del movimiento de los migrantes vulnerables y desfavorecidos"<sup>39</sup>. Desde ambas vertientes, el contexto migratorio forzado afecta la narrativa desde la intencionalidad ética en la estima del sí, pues el migrante se asume simultáneamente como víctima y como infractor.

Desde la disimetría en la relación con el otro, la narrativa con la que se comprende a los migrantes suele tratarlos como una "categoría homogénea, descritos a menudo por características fijas, como la vulnerabilidad y las dificultades de adaptación"<sup>40</sup>. Es común que no se les permita relatar sus historias concretas, lo que provoca una pérdida

<sup>38</sup>HAKER, Hille, "No space. Nowhere. Refugees and the Problem of Human Rights in Arendt and Ricoeur", *Études Ricoeuriennes / Ricoeur Studies*, vol. 8, núm. 2, 2017, p. 24.

<sup>39</sup>LECHNER, Elsa y RENAULT, Leticia, "Migration Experiences and Narrative Identities: Viewing Alterity from Biographical Research", *Critical Hermeneutics*, vol. 2, núm. 1, 2018, p. 3, Traducción propia.

<sup>40</sup>LECHNER, Elsa y RENAULT, Leticia, "Migration Experiences and Narrative Identities: Viewing Alterity from Biographical Research", *Critical Hermeneutics*, vol. 2, núm. 1, 2018, p. 13, Traducción propia.

irreparable en su propia narración y genera invisibilidad ante los otros. Sin embargo, dicha invisibilidad frecuentemente es producto de las "políticas migratorias de securitización y vigilancia",<sup>41</sup> que suelen suprimir por completo la particularidad del migrante, considerándolo un dato o una amenaza. En ningún caso, la negación de los derechos, relacionados con la migración y la ciudadanía, puede reducirse a una mera determinación burocrática o administrativa,<sup>42</sup> como suele ser justificada por las instituciones gubernamentales, pues implica la interrupción de las condiciones de posibilidad desde las cuales el sí mismo constituye su identidad narrativa.

Haker muestra, desde una relectura de la agencia moral ricoeuriana, que a pesar de todo lo que el migrante se ve forzado a dejar, no pierde su capacidad de responder. Sin embargo, esa capacidad se erosiona, queda sin base y se limita su ejercicio cuando se le niegan las condiciones narrativas, relacionales e institucionales, por lo que es necesario que las autoridades políticas abandonen la retórica de los derechos humanos y "cargar con el peso de su realización"<sup>43</sup>.

## **6. INTELIGENCIA ARTIFICIAL E IDENTIDAD NARRATIVA**

En la actualidad, el uso de procesos algorítmicos de inteligencia artificial relacionados con la migración es una muestra evidente de la abstracción que se hace del migrante particular y concreto, lo que deriva en su invisibilidad. Así, la inteligencia artificial reduce a las personas a datos y a clasificaciones estadísticas con la finalidad de formular predicciones y optimizar los procedimientos migratorios.<sup>44</sup> El más claro ejemplo de ello es la implementación en 2025 de ImmigrationOS, desarrollado por Palantir.

En este punto, es importante destacar que la migración forzada interrumpe las condiciones relacionales, culturales e institucionales en las que se sostiene la identidad narrativa del sí mismo como otro, es decir, el migrante pierde las condiciones desde las cuales se narra y era reconocido, pero no pierde la posibilidad de narrarse, de actuar, de decir y de imputarse actos, aunque son afectadas no desaparecen. El uso de la inteligencia artificial, en cambio, tiene implicaciones más profundas, pues suprime las condiciones de posibilidad para constituir los derechos humanos en la identidad narrativa del sí mismo como otro al reducir al sujeto a un dato que sirve para la clasificación y la estadística, ya que anula su capacidad de imputación y niega las condiciones relacionales e institucionales, negando la posibilidad de ser reconocido como sujeto narrativo en el proceso algorítmico.

De acuerdo con su justificación oficial, ImmigrationOS tiene tres objetivos: "Targeting and Enforcement Prioritization, Self-Deportation Tracking e Immigration Lifecycle

---

<sup>41</sup>LECHNER, Elsa y RENAULT, Leticia, "Migration Experiences and Narrative Identities: Viewing Alterity from Biographical Research", *Critical Hermeneutics*, vol. 2, núm. 1, 2018, p. 3, Traducción propia.

<sup>42</sup>Un ejemplo ilustrativo es el caso del Departamento de Estado de los Estados Unidos, que requiere a todo ciudadano extranjero que desee ingresar a dicho país contar con una visa, documento que a su vez exige pasaporte vigente y una entrevista consular en la que se deben comprobar distintos requisitos. Cfr. U.S. DEPARTMENT OF STATE, U.S. Visas, Bureau of Consular Affairs, 2024.

<sup>43</sup>HAKER, Hille, "No space. Nowhere. Refugees and the Problem of Human Rights in Arendt and Ricoeur", *Études Ricoeuriennes / Ricoeur Studies*, vol. 8, núm. 2, 2017, p. 41, Traducción propia.

<sup>44</sup>Cfr. COECKELBERGH, Mark, "Narrative responsibility and artificial intelligence. How AI challenges human responsibility and sense-making", *AI & Society*, vol. 38, 2023, pp. 2437-2450.

Process",<sup>45</sup> es decir, la priorización de identificación y aplicación de la ley, el seguimiento de autodeportaciones en tiempo real y el proceso del ciclo de vida migratorio.

En su aplicación práctica, el sistema está diseñado para agilizar la identificación y aprehensión de personas sujetas a deportación, el rastreo con visibilidad en tiempo real a quienes abandonan voluntariamente Estados Unidos y gestionar el ciclo completo desde la identificación hasta la deportación. Es importante destacar que el documento oficial, al momento de indicar sus objetivos, no utiliza expresiones como: persona, individuo, humano, ni siquiera migrante o cualquier otra con carga antropológica. La sustitución del sujeto por procesos, datos y operaciones, refleja la supresión de la identidad narrativa del sí mismo como otro.

En julio de 2025, Amnistía Internacional se pronunció, entre otras cosas, sobre ImmigrationOS de Palantir, señalando que dicho sistema tiene capacidades automatizadas de análisis de información pública que permiten "el monitoreo, la vigilancia y las evaluaciones constantes"<sup>46</sup>. El monitoreo en tiempo real sobre las publicaciones en redes sociales de distintas plataformas, datos de fuentes públicas y privadas, y bases de datos gubernamentales permite el reconocimiento de patrones y la detección de amenazas, "lo que puede llevar a deportaciones arbitrarias y violaciones al derecho a la privacidad"<sup>47</sup>.

Amnistía Internacional denuncia que la aplicación de herramientas de vigilancia con inteligencia artificial como ImmigrationOS, en conjunción con el abuso de poder y discriminación, en la revocación de visas, las redadas sin una orden judicial, las detenciones y las deportaciones, tiene como objetivo la intimidación contra migrantes por parte de las autoridades estadounidenses, lo que trae por consecuencia la violación de sus derechos humanos: a la privacidad, libertad de expresión, acceso a la información, libre tránsito, igualdad, no discriminación y a la protesta.

La continua aplicación de la inteligencia artificial en contextos migratorios, pese a las denuncias de Amnistía Internacional, expone el intento por reducir los derechos humanos a una "referencia retórica a los 'derechos naturales'; no son un concepto político-jurídico precisamente porque son derechos que no pueden hacerse valer en un espacio político"<sup>48</sup>.

El problema que advierte Haker adquiere, en el contexto de ImmigrationOS, una nueva dimensión, pues no solo es en el espacio territorial de Estados Unidos donde el migrante no puede hacer valer sus derechos humanos al negársele la membresía ciudadana, sino también en el espacio algorítmico que clasifica al sujeto como dato, negándole sus capacidades de actuar, decir, narrar e imputarse actos.

---

<sup>45</sup>U.S. IMMIGRATION AND CUSTOMS ENFORCEMENT, Limited-Sources Justification J&A-25-0092, Office of Acquisition Management, ICE Acquisition Manual 3006.301-90, Washington, D.C., 17 de abril de 2025, p. 1.

<sup>46</sup>AMNESTY INTERNATIONAL, Palantir: Response from Palantir to AMNESTY INTERNATIONAL, AMR 51/0211/2025, Londres, AMNESTY INTERNATIONAL, 21 de agosto de 2025, p. 3, Traducción propia.

<sup>47</sup>AMNESTY INTERNATIONAL, Palantir: Response from Palantir to AMNESTY INTERNATIONAL, AMR 51/0211/2025, Londres, AMNESTY INTERNATIONAL, 21 de agosto de 2025, p. 3, Traducción propia.

<sup>48</sup>Haker, desde una relectura de Arendt, sostiene que esta reducción efectivamente se consuma; sin embargo, en este texto se habla de intento porque algunas dimensiones de los derechos humanos como parte constitutiva de la identidad narrativa son irreductibles y los procesos algorítmicos o los regímenes políticos no pueden capturarlos por completo. Cfr. HAKER, Hille, "No space. Nowhere. Refugees and the Problem of Human Rights in Arendt and Ricoeur", *Études Ricoeuriennes / Ricoeur Studies*, vol. 8, núm. 2, 2017, pp. 22 y 31, Traducción propia.

Sin embargo, la transgresión directa de esos derechos no es el problema más profundo, pues la creación de sistemas de vigilancia dirigidos al "monitoreo masivo y toma de decisiones automatizada",<sup>49</sup> como ImmigrationOS, suprime las condiciones de posibilidad desde las cuales los derechos humanos se constituyen en la identidad narrativa del sí mismo como otro: al reducir al migrante a un dato clasificable, anula su reconocimiento como sujeto capaz de imputación y niega la posibilidad de narrar su propia experiencia. Considerar los derechos humanos no como objetos externos que se atribuyen a un sujeto cuya capacidad moral se presupone ya constituida, sino como parte constitutiva de la identidad narrativa, implica reconocer que su realización es siempre particular y concreta, y que no se puede sustituir ese proceso desde ningún sistema algorítmico.

La identidad narrativa no se reduce a la capacidad moral general. Desde esa perspectiva reduccionista, los derechos humanos aparecerían como objetos externos que se atribuyen a un sujeto ya constituido en su capacidad de responder y ser responsable. Sin embargo, el sí mismo como otro no es una capacidad aislada, pues tiene una historia, vive relaciones, habita instituciones y se narra a sí mismo en el tiempo.

Esas capacidades no preexisten a las condiciones que las hacen posibles, sino que se constituyen desde los tres momentos de la pequeña ética ricoeuriana, que corresponden, a su vez, a las dimensiones ética, deontológica y jurídica.

Por ello, los derechos humanos no son algo externo que se le otorga al sujeto ya constituido. Se constituyen en la identidad narrativa porque se experimentan desde las capacidades de narrar, actuar, decir, responder e imputarse los propios actos, mismas que solo se despliegan en la continuidad entre esas tres dimensiones. Los derechos humanos participan así de una constitución narrativa compleja e irreductible que, como toda acción humana, nunca termina de hacerse, pero que parte de un horizonte enriquecido por la experiencia vivida y su respectiva narración.

Coeckelbergh identifica el riesgo que los sistemas algorítmicos de inteligencia artificial representan para la responsabilidad narrativa, como el ser "organizados en historias creadas por quienes obtienen beneficios de la tecnología o tienen objetivos que no concuerdan con los propios"<sup>50</sup>. Sin embargo, el riesgo más profundo es que esa co-escritura tecnológica puede implicar una ruptura narrativa, que comienza con la manera en que el sí mismo se autopercibe, infractor y víctima simultáneamente, en la forma en la que se sostienen las relaciones con el otro, desde la invisibilidad y la amenaza, y en la manera en que las instituciones se habitan (la retórica vacía del régimen de los derechos humanos, la lógica de eficiencia de los sistemas migratorios y la anulación de la identidad narrativa del sí mismo como otro).

Si la identidad narrativa se suprime mediante procesos algorítmicos de inteligencia artificial dirigidos a la migración, y con ella los derechos humanos que son parte de su constitución, el problema no solo es hermenéutico, sino también ético, moral y jurídico. ImmigrationOS no solo impide la responsabilidad narrativa del migrante, sino que niega

---

<sup>49</sup>AMNESTY INTERNATIONAL, Palantir: Response from Palantir to AMNESTY INTERNATIONAL, AMR 51/0211/2025, Londres, AMNESTY INTERNATIONAL, 21 de agosto de 2025, p. 5, Traducción propia.

<sup>50</sup>COECKELBERGH, Mark, "Narrative responsibility and artificial intelligence. How AI challenges human responsibility and sense-making", *AI & Society*, vol. 38, 2023, p. 2444 Traducción propia.

las condiciones de posibilidad para desarrollar las capacidades desde las cuales el sí mismo se constituye como sujeto de derechos.

Elsa Lechner y Leticia Renault proponen, como forma de superar el impacto en la identidad narrativa con la migración, la aplicación de talleres biográficos de personas migrantes para “producir relatos en primera persona sobre su experiencia migratoria”<sup>51</sup>. Según esta propuesta, la narrativa de quienes padecen la migración forzada requiere no solo del relator sino de alguien que lo escuche.

Esta escucha implica no solo una acción pasiva, sino una responsabilidad social relacionada con la hospitalidad como condición de posibilidad del reconocimiento; sin embargo, ejemplos como ImmigrationOS, en los que se exponen los esfuerzos de las autoridades por clasificar, detectar y optimizar procesos, llevan implícita la no escucha y el no reconocimiento de los migrantes como sujetos capaces. De ahí la pertinencia de la propuesta de Coeckelbergh de no delegar a la inteligencia artificial la “responsabilidad moral y relacional con el papel hermenéutico de los seres humanos”<sup>52</sup>.

## 7. CONCLUSIONES

El análisis expuesto permite sostener que la tendencia a fundamentar los derechos humanos, propia de la modernidad, los reduce a objetos externos con un origen último; en cambio, a través de una comprensión hermenéutica es posible articular las dimensiones ética, deontológica y jurídica de los derechos, los cuales se constituyen en la identidad narrativa del sí mismo como otro. Desde la perspectiva ricoeuriana, la identidad es mediada en sus dos momentos, *idem* e *ipse*, por la narración, formulada, nunca de manera neutral, por el hombre capaz de actuar, narrar, decir e imputarse sus propios actos, a partir de los tres momentos de la pequeña ética y desde sus distintas dimensiones normativas.

Los derechos humanos como experiencia y narración no pueden ser considerados como objetos externos atribuidos a un sujeto abstracto, sino que se constituyen en el acto de narrarse. La migración forzada y el rechazo de los países a dar hospitalidad no solo tienden a la transgresión de los derechos humanos de los migrantes, sino que interrumpen las condiciones de posibilidad desde las cuales, quienes se ven forzados a migrar, se constituyen en la identidad narrativa. Los procesos algorítmicos de inteligencia artificial orientados al control migratorio, como ImmigrationOS, tienen un alcance más profundo, pues al reducir al migrante a un dato clasificable, suprimen su particularidad narrativa, anulan su capacidad de imputación y niegan las condiciones relacionales e institucionales que hacen posible su reconocimiento.

Finalmente, es posible sostener que los procesos algorítmicos de inteligencia artificial suprimen las condiciones de posibilidad desde las cuales los derechos humanos se constituyen como parte de la identidad narrativa del sí mismo como otro; para hacer frente a estos desafíos se vuelven condiciones éticas ineludibles (tal como se advierte

---

<sup>51</sup> LECHNER, Elsa y RENAULT, Leticia, “Migration Experiences and Narrative Identities: Viewing Alterity from Biographical Research”, *Critical Hermeneutics*, vol. 2, núm. 1, 2018, p. 3, Traducción propia.

<sup>52</sup> COECKELBERGH, Mark, “Narrative responsibility and artificial intelligence. How AI challenges human responsibility and sense-making”, *AI & Society*, vol. 38, 2023, p. 2448, Traducción propia.

en la propuesta de Lechner, Renault y Coeckelbergh) la escucha, la hospitalidad a las personas migrantes y la responsabilidad narrativa no delegada a la inteligencia artificial.

## **8. REFERENCIAS BIBLIOGRÁFICAS**

- ABBAGNANO, Nicola, Diccionario de filosofía, voz "fundamento", México, FCE, 1999, pp. 578-580.
- AMNESTY INTERNATIONAL, Palantir: Response from Palantir to AMNESTY INTERNATIONAL, AMR 51/0211/2025, Londres, AMNESTY INTERNATIONAL, 21 de agosto de 2025. [Consultada: 18 de abril de 2026]. [Disponible en: <https://www.amnesty.org/en/documents/amr51/0211/2025/en/>].
- ARENDT, Hannah, Los orígenes del totalitarismo, trad. Solana, Guillermo, Madrid, Taurus, 1998.
- BOBBIO, Norberto, El tiempo de los derechos, trad. de Rafael de Asís Roig, Madrid, Sistema, 1991.
- COECKELBERGH, Mark, "Narrative responsibility and artificial intelligence. How AI challenges human responsibility and sense-making", *AI & Society*, vol. 38, 2023, pp. 2437-2450. [Consultada: 8 de abril de 2026]. [Disponible en: <https://doi.org/10.1007/s00146-021-01375-x>].
- FERRATER, José, Diccionario de filosofía, voz "fundamento", Barcelona, Ariel, 1994, pp. 1413-1414.
- GADAMER, Hans-Georg, Verdad y Método I, Salamanca, Ediciones Sígueme, 1999.
- GORGONI, Guido, "Paul Ricoeur filósofo del derecho: el sujeto de derecho y de responsabilidad como *homme capable*", *Nuevo Itinerario*, vol. 20, núm. 1, 2024, pp. 1-14. [Consultada: 15 de abril de 2026]. [Disponible en: <https://doi.org/10.30972/nvt.2017586>].
- HAKER, Hille, "No space. Nowhere. Refugees and the Problem of Human Rights in Arendt and Ricoeur", *Études Ricoeuriennes / Ricoeur Studies*, vol. 8, núm. 2, 2017, pp. 22-45. [Consultada: 6 de abril de 2026]. [Disponible en: <https://ricoeur.pitt.edu/ojs/index.php/ricoeur/article/view/412>].
- HEIDEGGER, Martin, *El ser y el tiempo*, México, Fondo de Cultura Económica, 1993.
- LECHNER, Elsa y RENAULT, Leticia, "Migration Experiences and Narrative Identities: Viewing Alterity from Biographical Research", *Critical Hermeneutics*, vol. 2, núm. 1, 2018, pp. 1-25. [Consultada: 10 de abril de 2026]. [Disponible en: <https://doi.org/10.13125/CH/3436>].
- ORGANIZACIÓN DE LAS NACIONES UNIDAS, Declaración Universal de Derechos Humanos, ONU, 1948.
- PECES-BARBA, Gregorio, "Sobre el fundamento de los derechos humanos. Un problema de moral y derecho", en MUGUERZA, Javier y PECES-BARBA, Gregorio, coords., *El fundamento de los derechos humanos*, Madrid, Debate, 1989, pp. 265-278.
- PECES-BARBA, Gregorio, *Curso de derechos fundamentales. (I) Teoría general*, con la colaboración de R. de Asís y A. Llamas Gascón, Madrid, Eudema, 1983.
- PÉREZ LUÑO, Antonio Enrique, *Derechos humanos, Estado de Derecho y Constitución*, 3.ª ed., Madrid, Tecnos, 1990.
- RICOEUR, Paul, *Si mismo como otro*, México, Siglo Veintiuno Editores, 1996.
- RICOEUR, Paul, *Caminos del reconocimiento. Tres estudios*, trad. Neira, Agustín, México, Fondo de Cultura Económica, 2006.
- RODRÍGUEZ-TOUBES, Joaquín, *La razón de los derechos*, Madrid, Tecnos, 1995.
- U.S. DEPARTMENT OF STATE, U.S. Visas, Bureau of Consular Affairs, 2024. [Consultada 20 de abril de 2026]. [Disponible en: <https://travel.state.gov/content/travel/en/us-visas.html>].
- U.S. IMMIGRATION AND CUSTOMS ENFORCEMENT, Contract modification P00009, No. 70CTD022FR0000170, U.S. Department of Homeland Security, 2025. [Consultada 19 de abril de 2026]. [Disponible en: [https://www.usaspending.gov/award/CONT\\_AWD\\_70CTD022FR0000170\\_7012\\_GS35F0086U\\_4730](https://www.usaspending.gov/award/CONT_AWD_70CTD022FR0000170_7012_GS35F0086U_4730)].
- U.S. IMMIGRATION AND CUSTOMS ENFORCEMENT, Limited-Sources Justification J&A-25-0092, Office of Acquisition Management, ICE Acquisition Manual 3006.301-90, Washington, D.C., 17 de abril de 2025.

VILLA SÁNCHEZ, J. A., "Sobre el sentido de lo justo y la ley", *En-Claves del Pensamiento*, núm. 37, 2025, pp. 216-236. [Consultada: 16 de abril de 2026]. [Disponible en: <https://doi.org/10.46530/ecdp.v0i37.701>].