

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 real¹. 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 responsabilidad².

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 humanos³.

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

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

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

³ORGANIZACIÓN 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 derecho⁴.

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 fundamentales⁵.

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.

⁴CONSEJO 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.

⁵MARTÍN-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 ajenos⁶.

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 técnica⁷. 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,

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

⁷DANAHER, 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 estadístico⁸. 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 comparables⁹. 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.¹⁰

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.¹¹ 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

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

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

¹⁰ORGANIZACIÓN 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.

¹¹EUBANKS, 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 contenido¹².

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 tecnológica¹³.

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 Políticos¹⁴. 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 adecuado¹⁵.

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

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

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

¹⁴ORGANIZACIÓN 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.

¹⁵ZUBOFF, 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 desaparecerán¹⁶. 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 humana¹⁷.

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 informado¹⁸. 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 implicaciones¹⁹.

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"²⁰. 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

¹⁶ZUBOFF, Shoshana. *The Age of Surveillance Capitalism...* op. cit.

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

¹⁸UNIÓN 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.

¹⁹ZUBOFF, Shoshana. *The Age of Surveillance Capitalism...*, op. cit.

²⁰HAN, 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?²¹ 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 enfrentar²².

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 humanos²³.

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

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

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

²³GREENWALD, 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 ser²⁴.

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 femenino²⁵.

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 individual²⁶.

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 impugnarlos²⁷. 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 vulnerabilidad²⁸. 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.),

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

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

²⁶O'NEIL, Cathy. *Weapons of Math Destruction...*, op. cit.

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

²⁸EUBANKS, 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 formal²⁹.

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 disponibles³⁰.

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 objetivo³¹. 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

²⁹O'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.

³⁰O'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.

³¹O'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 decisiones³².

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 derecho³³.

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 algorítmico³⁴.

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 jurídico³⁵.

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

³³ZUBOFF, 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.

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

³⁵PÉREZ 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.³⁶

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 reducirlas³⁷.

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

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

³⁷COMISIÓN 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 técnica³⁸.

³⁸ORGANIZACIÓN 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 posterior³⁹.

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 discriminatorios⁴⁰.

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

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

⁴⁰UNIÓN 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].
- 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].
- ZUBOFF, Shoshana. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Nueva York: PublicAffairs, 2019.