

Artículos centrales

Colonialismo digital y captura corporativa: Debates actuales sobre los usos de la Inteligencia Artificial en el sector público desde la perspectiva de actores sociales. Estudio Regional.

María Natalia Echegoyemberry^a

Fecha de recepción: 8 de octubre de 2025
Fecha de aceptación: 10 de noviembre de 2025
Correspondencia a: María Natalia Echegoyemberry
Correo electrónico: echegoyemberry2014@mail.com

a. Magister en Salud Pública. Abogada, psicóloga. Directora del Observatorio de Justicia Sanitaria y Climática, Latinoamérica. Vicepresidenta de la Asociación para la promoción y protección de los Derechos Humanos XUMEK.

Resumen:

El uso de inteligencia artificial generativa (IA) en el sector público de América Latina es aún incipiente y se desarrolla en un contexto de desigualdad estructural, brechas digitales y marcos normativos en construcción. Este estudio examina los riesgos y beneficios de su aplicación desde la perspectiva de actores sociales, con el objetivo de formular recomendaciones para mejorar las políticas públicas en IA y derechos humanos. A través de un enfoque cuali-cuantitativo y una triangulación metodológica, se identifican preocupaciones como la opacidad algorítmica, sesgos discriminatorios, inseguridad digital y laboral, así como impactos ambientales y económicos. Se

advierte sobre el colonialismo digital y la captura corporativa del Estado, lo que podría comprometer la soberanía digital y la estabilidad democrática. Se enfatiza la necesidad de un marco normativo sólido, fortalecimiento institucional y un debate político amplio desde una perspectiva crítica del Sur Global.

Palabras clave: Inteligencia Artificial - Colonialismo digital - Derechos Humanos.

Summary

The use of generative artificial intelligence (AI) in the public sector of Latin America is still in its early stages and unfolds within a context of structural inequality, digital gaps, and evolving regulatory frameworks. This study examines the risks and benefits of its application from the perspective of social actors, aiming to formulate recommendations for improving public policies on AI and human rights. Through a qualitative-quantitative approach and methodological triangulation, concerns such as algorithmic opacity, discriminatory biases, digital and labor insecurity, as well as environmental and economic impacts, are identified. The study warns about digital colonialism and corporate capture of the State, which could compromise digital sovereignty and democratic stability. It emphasizes the need for a robust regulatory framework, institutional strengthening, and a broad political debate from a critical Global South perspective.

Key words: Artificial Intelligence; Digital Colonialism; Human Rights.

Introducción

Durante la última década, la inteligencia artificial (IA) ha irrumpido con fuerza en diversos ámbitos de la vida cotidiana y del trabajo, sin que exista aún un consenso claro sobre sus usos, beneficios y riesgos para la ciudadanía y los Estados. El avance acelerado de los sistemas de IA en el sector público plantea desafíos significativos en dimensiones éticas, jurídicas, psicosociales, epistemológicas y ambientales. En América Latina, este fenómeno se desarrolla en contextos atravesados por desigualdades estructurales, capacidades estatales heterogéneas y marcos normativos en construcción. Estas condiciones acentúan la necesidad de examinar críticamente los usos

y el grado de explicabilidad de la IA en la gestión pública. A pesar del creciente interés institucional y académico en regular las tecnologías emergentes, persisten importantes vacíos en la comprensión de cómo se están utilizando, qué percepciones tienen quienes las implementan o gestionan, y cuáles son sus efectos concretos sobre los derechos fundamentales, la democracia y el ambiente.

El presente artículo expone resultados de una investigación¹ orientada a analizar los riesgos y beneficios asociados al uso de la IA en el sector público latinoamericano desde la perspectiva de diversos actores sociales.

1. N de A: investigación Seleccionada en el 28º Congreso Mundial de Ciencias Políticas de la IPSA, Seúl, Corea del Sur, del 12 al 16 de julio de 2025 (wc2025@ipsa.org). Para acceder a la investigación principal consultar: Echegoyemberry, MN, Piccardo, V, Rojo, P, Allassia, L, Galeazzi, C, Brolese, C, Carneiro, A, Duarte, R, Cámara, J (2025). "Políticas públicas para la innovación judicial: Uso, aplicación y explicabilidad de la IA en el sector público. Presentación de Resultados Preliminares. Disponible: <https://justiciasanitaria.org/2025/07/politicas-publicas-para-la-innovacion-uso-aplicacion-y-explicabilidad-de-la-ia-en-el-sector-publico-presentacion-de-resultados-preliminares/>

Las principales preguntas que guiaron el estudio fueron: ¿cuáles son los riesgos y beneficios del uso de la IA en el sector público según los actores sociales de la región?; ¿qué normas regulan su implementación, especialmente en los poderes judiciales?; ¿el marco normativo vigente brinda protección suficiente frente a los riesgos vinculados con la explicabilidad y la toma de decisiones automatizadas?; ¿en qué medida se utiliza la IA en procesos que involucran derechos fundamentales y se garantiza el consentimiento informado de las partes?; y finalmente, ¿cómo puede promoverse la innovación tecnológica en los sistemas judiciales sin comprometer el debido proceso, la confidencialidad y la seguridad jurídica, evitando además que la inversión en IA desvíe recursos de las agendas sociales y ambientales prioritarias?

En los países latinoamericanos, los riesgos derivados del uso de la IA se ven agravados por la persistencia de desigualdades estructurales y por la segregación socioespacial. Tal como ocurre en los sectores de la salud y la educación, los poderes judiciales presentan marcadas brechas en el acceso y uso de tecnologías de la información y la comunicación (TIC) (ACIJ, 2020), lo que puede profundizar las asimetrías existentes (Tejedor-Estuñán, 2025). En este escenario, la expansión de la IA corre el riesgo de reproducir relaciones de colonialidad digital entre el Norte y el Sur global, en un contexto donde los Estados de la región tienen una capacidad limitada de regulación sobre las corporaciones transnacionales que desarrollan tecnología e infraestructuras de datos (Tribunal Constitucional de Colombia, Sentencia T-323, párr. 268).

La incorporación de sistemas de IA en el sector público ha avanzado de facto, sin mediar una deliberación social amplia sobre sus implicancias en términos de derechos humanos, soberanía tecnológica e impactos socioambientales (GIZ, 2025). En el ámbito judicial, diversas experiencias muestran su aplicación para optimizar procesos, mejorar la búsqueda jurisprudencial o redactar documentos, e incluso en decisiones automatizadas como el dictado de sentencias (Tribunal Constitucional de Colombia, Sentencia T-323). Sin embargo, en la mayoría de los casos no se explicita su uso ni se solicita el consentimiento de las partes intervenientes, lo que podría tensionar las garantías del debido proceso.

Mientras algunos autores destacan el potencial democratizador de la IA en el acceso a la información y la mejora de la gestión pública, otros advierten sobre sus dimensiones opacas, asociadas a la falta de transparencia, explicabilidad, gobernanza y sostenibilidad ambiental.

Actualmente, América Latina carece de una normativa específica y vinculante que regule de manera integral el uso y la explicabilidad de la IA en el sector público. No obstante, diecinueve países de la región cuentan con marcos constitucionales que podrían servir de base para su regulación, al incluir principios y derechos tales como la protección de datos, la confidencialidad, el consentimiento informado, el habeas data, la publicidad de los actos de gobierno, la función social de la propiedad y la primacía de los derechos humanos por sobre los intereses comerciales. A pesar de este potencial normativo, su aplicación práctica sigue siendo limitada y, en muchos casos, invisibilizada.

Desarrollo

La inteligencia artificial generativa ha sido objeto de estudio desde múltiples perspectivas disciplinares, incluyendo la informática, la filosofía, el derecho, la comunicación, la ética aplicada, la economía, la epistemología, los estudios culturales y los estudios de ciencia, tecnología y sociedad (CTS). Desde la informática, se han desarrollado arquitecturas como los modelos funcionales (foundation models) y técnicas avanzadas de procesamiento de lenguaje natural (Brown et al., 2020; Bommasani et al., 2021). La filosofía y la ética aplicada, por su parte, han explorado los dilemas morales que plantea la IA generativa, los desafíos, la responsabilidad ética, así como los riesgos asociados a esta tecnología y las posibles vías para su mitigación (Bender et al., 2021; Floridi & Cowls, 2019).

Bender et al. (2021) proponen considerar los costos ambientales y financieros de estos sistemas, y alicantan líneas de investigación que trascienden el desarrollo de modelos lingüísticos cada vez más amplios. Por su parte, Floridi y Cowls (2019) investigaron los principios éticos que rigen en materia de IA. En el ámbito jurídico, se ha examinado el impacto de la IA generativa en los derechos humanos (UNESCO, 2024), el acceso a la justicia, el debido proceso, la privacidad, confidencialidad de datos en salud y aspectos problemáticos de la propiedad intelectual (GIZ, 2025, Sentencia T 373/2024 Tribunal Constitucional de Colombia). Asimismo, se han analizado los modelos híbridos de regulación de la Inteligencia Artificial adoptados por la Unión Europea basado principalmente en la seguridad y la estandarización de los productos y la protección de los derechos fundamentales (Gstrein, Haleem & Zwitter, 2024). Fuster (2022) plantea la necesidad de avanzar desde una gobernanza

reactiva hacia una proactiva en materia de IA. Las ciencias sociales y los estudios sobre CTS, han problematizado el papel de la IA generativa en la producción de conocimiento, la creación cultural, la desinformación y la amplificación estructural de desigualdades (Crawford, 2021). También se ha abordado el fenómeno del capitalismo de vigilancia, caracterizado por una concentración extrema del poder y del conocimiento fuera de cualquier control democrático (Zuboff, 2019); también se han analizado los procesos de discriminación algorítmica (CIPDH, Waisbrot, 2024). Por su parte, grupos de estudios latinoamericanos se encuentran investigando los efectos geopolíticos de la IA (CLACSO, 2025).

Existe un campo de estudio que aborda el colonialismo e imperialismo digital y las reconfiguraciones geopolíticas a partir del uso de tecnologías emergentes. El concepto de colonialismo digital ha surgido como una herramienta analítica para comprender cómo las dinámicas de poder históricas se reproducen en el ámbito digital. Diversos autores han explorado esta noción desde múltiples perspectivas, abordando dimensiones como la infraestructura tecnológica, la extracción de datos, la epistemología, la normativa y las condiciones laborales. Couldry y Mejías (2019) introducen el concepto de colonialismo de datos, describiendo cómo la extracción masiva de datos personales sin consentimiento reproduce dinámicas coloniales al convertir la vida humana en una fuente de valor económico.

Por su parte, Kwet (2019) encontró que las corporaciones tecnológicas estadounidenses ejercen una forma de imperialismo digital en el Sur Global mediante el control de infraestructuras críticas como software, hardware y conectividad de red. También algunos autores utilizan el concepto de tecnofeudalismo para dar cuenta con ello de otra modalidad que asume el capitalismo de las tecnologías digitales (Varoufakis, 2024). En sentido contrario, Morozov (2022) formula una crítica a la utilización del concepto de tecno feudalismo, y en su lugar plantea dinámicas de refeudalización donde hay una clase parasitaria que no interviene en el proceso productivo y se apropiá del producto. Ruha Benjamin (2019) encuentra que los sistemas tecnológicos replican desigualdades raciales y analiza las dimensiones epistémica-cultural del colonialismo digital.

A los fines de esta investigación se entiende por colonialismo digital un patrón contemporáneo de dominación tecnológica, informacional y simbólica, mediante el cual actores concentrados, principalmente corpora-

ciones tecnológicas y gobiernos del Norte Global, ejercen control sobre las infraestructuras digitales, los flujos de datos, los algoritmos y las condiciones de acceso al conocimiento. De esta manera, se reproducen lógicas históricas de subordinación, dependencia y extractivismo en el ámbito digital, especialmente sobre países, comunidades y personas en situación de vulnerabilidad, con pérdida de autodeterminación tecnológica y de soberanía. Para ello utilizan diferentes formas de prácticas empresariales contrarias al interés público, como captura algorítmica y articulaciones públicas privadas que colocan a ciertos actores en posición de preferencia generando violación o elusión normativa (Echegoyemberry, 2025).

Crawford (2021) advierte que los sistemas de IA en su diseño tienden a reproducir intereses dominantes preexistentes, en este sentido señala que “no son autónomos, racionales ni capaces de discernir algo sin un entrenamiento extenso e intensivo”. En esta línea, Strubell, Ganesh y McCallum (2019) exploran nuevas perspectivas de investigación vinculadas a la sostenibilidad ambiental, enfocándose en los costos ecológicos del aprendizaje profundo aplicado al procesamiento de lenguaje natural (PLN), y destacando las exigencias energéticas y materiales reales de estos sistemas.

A pesar del creciente uso de estas tecnologías en el sector público, persisten importantes brechas en la comprensión de su funcionamiento, la interacción entre componentes y los mecanismos mediante los cuales se generan los resultados, así como la organización funcional y la estructura emergente que resulta del entrenamiento automatizado de estos sistemas. Tal escenario ha impulsado la emergencia de un campo de investigación centrado en lograr mayor transparencia, orientado a construir sistemas más explicables, confiables, seguros y menos opacos (Ribeiro et al, 2016).

En este contexto, adquiere relevancia el campo interdisciplinario de la explicabilidad de la IA (XAI, por sus siglas en inglés), enfocado en los sistemas basados en aprendizaje automático (ML), desarrollando estrategias, líneas de investigación, taxonomías (Barredo Arrieta et al, 2020) y métodos específicos para abordar dicha explicabilidad (Käster y Crook, 2024). El campo de la explicabilidad de la IA articula saberes provenientes de la informática, la filosofía y la psicología (Käster y Crook, 2024). Según Käster y Crook (2024) pueden identificarse dos estrategias principales para abordar la explicabilidad. La primera, llamada “divide y vencerás”, busca explicar el funcionamiento de los modelos en contextos

específicos. La segunda, denominada “interpretabilidad mecanicista” (IM), propone una comprensión holística del funcionamiento sistémico. Desde esta perspectiva, se plantea “abrir la caja negra”²² para posibilitar un uso seguro de la IA en escenarios de alto riesgo (Kästner y Crook, 2024:52). Estas discusiones teóricas tienen lugar en un contexto de aceleración tecnológica, pero también de profundas desigualdades sociales, brechas digitales y limitada accesibilidad a tecnologías de la información y la comunicación (TIC). Por tanto, aún no es posible dimensionar con claridad su impacto en los derechos humanos y las democracias contemporáneas. A su vez, gran parte de los debates académicos se dan a espaldas de la ciudadanía y de grupos en situación de vulnerabilidad.

Resultados y discusión

El presente estudio se diseñó con un enfoque descriptivo, exploratorio, transversal y de tipo cuali-cuantitativo. Se empleó una triangulación metodológica para potenciar la validez y confiabilidad de los hallazgos, combinando fuentes primarias y secundarias. En la encuesta participaron personas de 15 países de América Latina. Adicionalmente, participaron integrantes de 16 laboratorios de innovación, centros de capacitación judicial y de investigación de Argentina, Brasil, Colombia, Bolivia y Nicaragua (Ver apartado metodológico en la investigación principal).

A) Sistemas de inteligencia artificial utilizados en el sector público

Los datos recabados indican una incorporación incipiente pero creciente de sistemas de Inteligencia Artificial (IA) en el sector público. Su aplicación se concentra en funciones auxiliares y administrativas, aunque se extiende progresivamente a la toma de decisiones automatizadas de alto riesgo, como sentencias judiciales o diagnósticos médicos. Esta evolución plantea exigencias específicas en términos de transparencia, supervisión humana y evaluación ética y legal. El uso de la IA en el sector público se caracteriza por su naturaleza predominantemente informal y no institucionalizada.

En el sector salud, la IA está transformando las prácticas clínicas y el diagnóstico, permitiendo la detección precisa de patologías a partir de grandes volúmenes de datos visuales. También se observa su incorporación en la gestión hospitalaria inteligente, para la previsión de demanda de recursos y organización de turnos. Sin embargo, las personas entrevistadas en el sector salud advierten unánimemente sobre los riesgos derivados de la falta de verificación y control humano en el proceso de diagnóstico y tratamiento, así como problemas relacionados con la confidencialidad de la información, destacando la necesidad de una regulación más precisa. Los sistemas de IA más usados provienen principalmente de grandes desarrolladores globales, como ChatGPT (OpenAI), Copilot (Microsoft) y Gemini (Google DeepMind), junto con herramientas de código abierto (DeepSeek) y soluciones de empresas como Anthropic (Claude) e IBM (Watson Health/Merative). Existen también IA especializadas en medicina (Google DeepMind, Aidoc, Nuance Dragon Medical One) y en el ámbito jurídico (Prometea, ROSS Intelligence, CaseText). En Argentina, la IA se integra especialmente en la justicia, con herramientas como Prometea y JustiBot (CABA), Arapy (Corrientes), LEXA (Mendoza) y Jurisemia (Córdoba), siguiendo patrones similares a otras iniciativas regionales (Pretoria en Colombia, Sócrates en Brasil).

Se identificaron múltiples beneficios asociados al uso de sistemas de inteligencia artificial (IA) en el sector público. Asimismo, se destaca su contribución para afrontar cuellos de botella históricos en áreas críticas como la salud, la administración pública y el sistema judicial, especialmente en contextos caracterizados por la sobrecarga de trabajo y la escasez de recursos. Los resultados sugieren que, bien implementadas, estas tecnologías pueden fortalecer el ejercicio efectivo de derechos y mejorar la calidad de la gestión pública. Estos beneficios, no obstante, requieren de condiciones institucionales, normativas y éticas para materializarse. Interpelan a pensar el modelo de gestión, el modelo de atención y el modelo de financiamiento de las políticas de IA.

2. N de A: El término “caja negra” y “zonas grises” es problematizado por el equipo de investigación del Observatorio de Justicia Sanitaria y climática partiendo de una perspectiva crítica del lenguaje y de derechos humanos, porque consideramos que refuerza una asociación simbólica donde lo oscuro, negro o gris se vincula con lo desconocido, incomprendible o problemático, mientras que lo blanco suele asociarse con claridad, transparencia y positividad. Esta dicotomía, presente en múltiples expresiones del lenguaje, refleja construcciones culturales que históricamente han atribuido valores negativos a lo negro y positivos a lo blanco, reproduciendo sesgos que pueden tener implicaciones raciales. En el contexto de la inteligencia artificial, cuestionar el uso de este término no implica negar su utilidad técnica o que sea una solución de conveniencia, sino reconocer que el lenguaje no es neutral y que ciertas metáforas pueden perpetuar estructuras de pensamiento discriminatorias en términos raciales y por lo tanto es necesario visibilizarlas.

B) Percepción de conocimiento, capacitación y explicabilidad de la IA

Los resultados muestran un bajo nivel de capacitación formal en IA entre los encuestados: el 79,7 % no ha recibido formación específica, mientras que solo el 19,3 % indicó haber recibido algún tipo de capacitación laboral, y apenas un 1 % participó en experiencias alternativas como conferencias o formación autodidacta. De quienes recibieron capacitación, el 62,5 % consideró que esta les permite actuar de manera segura y confiable, y el 37,5 % la percibió como insuficiente. Estos hallazgos coinciden con estudios previos de la UNESCO (2024), que destacan un uso frecuente de herramientas de IA sin la formación adecuada.

En términos de explicabilidad, transparencia y conocimiento técnico, se observa un nivel general bajo. Casi la mitad de los encuestados desconoce qué datos se emplean para entrenar los sistemas de IA, y el 51 % percibe un nivel bajo o nulo de explicabilidad. Solo el 15 % conoce el tipo de datos utilizados, mientras que el 47 % indica no tener información al respecto, y un 36 % considera que no se emplean datos en su sector. La percepción predominante es que la IA es opaca, aunque mayoritariamente se percibe como una tecnología neutra; apenas un 4 % considera que los sistemas son completamente explicables. El bajo conocimiento técnico se refleja en que el 43 % no entiende el proceso de entrenamiento de IA, el 34 % tiene nociones básicas, y solo un 2 % posee conocimientos avanzados o experiencia directa, generalmente vinculada a laboratorios de innovación. La falta de comprensión sobre entrenamiento, parámetros, patrones y retroalimentación limita la capacidad de evaluar críticamente los sistemas y los riesgos asociados, particularmente en entornos públicos donde se manejan datos sensibles y decisiones de alto impacto.

Se identifica además un fenómeno de invisibilización del uso de IA, derivado de la externalización de su desarrollo y del desconocimiento del personal sobre el funcionamiento de los modelos y el uso de datos de retroalimentación para entrenamiento y mejora continua. Esta situación plantea desafíos significativos para la implementación responsable, la trazabilidad de decisiones automatizadas y la rendición de cuentas. Los resultados evidencian que la formación insuficiente, el bajo nivel de conocimiento técnico y la limitada explicabilidad de los sistemas representan barreras críticas para la adopción segura y ética de la

IA en el sector público, subrayando la necesidad de estrategias de capacitación, transparencia y gobernanza robusta.

C) Riesgos percibidos de la IA en el sector público

Los riesgos identificados por los actores sociales en torno al uso de la inteligencia artificial en el sector público no constituyen simplemente un listado de problemas técnicos o administrativos a resolver. Lo que emerge de las narrativas es un entramado complejo de tensiones que atraviesan dimensiones tecnológicas, institucionales, sociales y subjetivas, y que interullan de manera profunda el rol del Estado, la configuración del poder institucional y la experiencia ciudadana en los procesos públicos. El análisis de entrevistas, encuestas y foros con actores sociales de distintos sectores de América Latina permitió identificar un conjunto amplio de riesgos y problemáticas vinculados al uso de la inteligencia artificial (IA) en el sector público, que se distribuyen en siete ejes interrelacionados:

En primer lugar, se identifican problemas inherentes al diseño, funcionamiento y gobernanza de los sistemas de IA, tales como la opacidad algorítmica, los sesgos estructurales, la generación de resultados erróneos, y la débil fiscalización humana, en un contexto marcado por la ausencia de estándares comunes y marcos de gobernanza anticipatoria.

En segundo lugar, se reconocen problemas estructurales del ecosistema digital, donde la concentración corporativa, la falta de marcos regulatorios específicos, la pérdida de soberanía digital y el impacto ambiental y económico de la IA configuran un entorno global asimétrico que condiciona su adopción soberana y sostenible.

En un tercer eje, se destacan limitaciones meso-institucionales vinculadas a la escasa formación técnica y ética en el sector público, la inexistencia de protocolos y controles adecuados, y una infraestructura tecnológica obsoleta, lo que afecta la capacidad estatal para integrar la IA con criterios de legalidad, eficacia y equidad.

En el plano micro, *el cuarto eje* aborda los impactos laborales y en la prestación de servicios públicos, donde emergen procesos de despersonalización, precarización, afectación de la salud mental y pérdida del juicio profesional humano.

El quinto eje refiere a problemas de percepción social y subjetividad, con fuerte presencia de desconfianza, temor, baja alfabetización digital y una percepción generalizada de deshumanización del vínculo con el Estado.

En el eje seis, se subrayan desafíos éticos, legales y de derechos humanos, entre ellos la afectación de garantías fundamentales, el uso no consentido de datos sensibles, y la automatización de decisiones públicas sin control democrático.

Por último, *el séptimo eje* incorpora riesgos epistemológicos y cognitivos, vinculados a la opacidad del conocimiento algorítmico, los sesgos epistémicos, la falta de pluralismo y la descontextualización del saber producido, lo cual resulta particularmente crítico en sectores sensibles como justicia, salud y protección social. En conjunto, estos hallazgos permiten visibilizar los múltiples planos en los que la IA impacta. Las preocupaciones y riesgos percibidos por los actores sociales de Latinoamérica coinciden con estudios previos, en los que se identifican la opacidad de las redes neuronales artificiales (Burrell, 2016), la falta de transparencia algorítmica, los sesgos discriminatorios, la equidad, la privacidad de la información (Hueso, 2020), la seguridad, y la sostenibilidad ambiental y económica.

D) Vacíos normativos y débil conocimiento normativo por parte de actores sociales

El desarrollo normativo en el ámbito de la Inteligencia Artificial (IA) evidencia avances a nivel internacional, regional y nacional, planteando interrogantes sobre los riesgos asociados y los límites para la garantía de derechos fundamentales. La Asamblea General de Naciones Unidas (ONU, Res/A /78) advierte a los Estados el deber de abstenerse de usar sistemas de IA incompatibles con el derecho internacional o que supongan riesgos indebidos para los derechos humanos.

Así, en el ámbito jurídico, se delinean dos grandes posturas: por un lado, una visión que promueve un Estado activo, con capacidad de intervenir mediante normativas vinculantes (*hard law*); por otro, una perspectiva de *laissez-faire* que advierte sobre los riesgos de una soberregulación que inhibía la innovación, y que propicia marcos flexibles (*soft law*) o basados en directrices, guías o a partir de la autorregulación del sector privado y de las empresas transnacionales, por ejemplo, a través de sus “políticas de uso” o “términos de uso.”³. Incluso dentro de la visión reguladora, persisten tensiones en torno a la pertinencia de diseñar marcos legales anticipatorios —capaces de prever distintos escenarios futuros— frente a propuestas que sugieren establecer principios generales lo suficientemente amplios como para adaptarse a la evolución vertiginosa de estas tecnologías emergentes. Algunos autores afirman que en la región es aún incipiente el desarrollo de los marcos normativos específicos, estos oscilan entre tendencias regulatorias que convergen con estándares europeos, por un lado, y la generación de caminos propios, por el otro (Contreras, 2022:482; Muñoz, 2024). Internacionalmente, Estados Unidos y China mantienen un enfoque de *laissez-faire*, mientras que la Unión Europea y Corea del Sur adoptan estrategias duales, que combinan principios éticos con obligaciones legales, aunque con distintos énfasis (Lee Kwang Hyung, 2025). La Ley Básica de IA de Corea del Sur es una legislación flexible, centrada en los impactos de la IA más que en sus riesgos, a diferencia del enfoque eminentemente precautorio o defensivo que caracteriza a la regulación europea. La Unión Europea clasifica los usos de acuerdo al nivel de riesgo, considerando usos como de alto riesgo aquellos que tienen impacto en los derechos fundamentales y la vida de las personas (Reglamento (UE) 2024/1689, art. 6). Dicho reglamento establece requisitos y condiciones específicas para estos usos, incluyendo sistemas de gestión de riesgo, documentación técnica, conservación de registros de eventos, transparencia, vigilancia humana, ciberseguridad y delimitación de responsabilidades de proveedores. Aspectos que aún no se han asumido institucionalmente en los países de América Latina.⁴

3. N de A: Algunos autores proponen la autorregulación corporativa para no frenar la innovación; sin embargo, creemos que dejar la regulación al mercado no protege adecuadamente los derechos humanos. OpenAI, por ejemplo, limita el uso de ChatGPT como única fuente de asesoramiento jurídico o financiero, en contextos médicos -diagnóstico y tratamientos- y en decisiones gubernamentales de alto riesgo -incluidas la aplicación de la ley, la justicia penal, la migración o el asilo-. (Anexo C, políticas de uso OpenAI).

4. N de A: La Comisión Europea enunció las prácticas de alto riesgo en la Ley de Inteligencia Artificial 2024. Para los sistemas de alto riesgo se dispone de medidas estrictas para implementarlos. Disponible: <https://digital-strategy.ec.europa.eu/es/policies/regulatory-framework-ai>

Mientras que, en América Latina, la gobernanza de la IA es heterogénea y presenta vacíos regulatorios (Contreras, 2024). Este autor explica esta heterogeneidad a partir de considerar tres factores: 1) débiles lazos de integración, 2) baja densidad normativa en la regulación internacional y 3) estrategias unilaterales basadas en intereses de integración y comercio (Contrera, 2024: 471). El análisis de los marcos normativos sobre Inteligencia Artificial (IA) en 19 países latinoamericanos revela un escenario plural, tanto por el nivel de elaboración de normas sobre IA, como por la perspectiva o enfoque que adoptan, por lo que no puede afirmarse que exista un enfoque regional uniforme. En este panorama, se identifican países con un desarrollo incipiente en la materia, otros países, por su parte, se limitan a disponer de **estrategias nacionales sin una regulación específica vinculante**. A su vez también existe diversidad en el enfoque (de riesgos o de impacto) y la forma de regular. Este escenario incluye marcos regulatorios duales o híbridos, que combinan normativas vinculantes con principios éticos no vinculantes, lo que puede generar ambigüedad en su aplicación e interpretación.

Así, si bien se encontraron en la mayoría de ellos normas sobre protección de datos personales (Colombia, Argentina, Chile, Costa Rica, México y Uruguay) son muy pocos los que tienen legislación específica en materia de IA, algunos regularon en determinados campos (como el electoral o defensa del consumidor) o aprueban guías, recomendaciones éticas en la materia.

En el caso de Perú y El Salvador cuentan con normativa específica en materia de IA. Por su parte países como Brasil y Chile han adoptado un enfoque de riesgos similar al Reglamento Europeo de IA. En el caso de Chile se posiciona como pionero en la región al incorporar, a nivel constitucional, la regulación de los neurodatos y la protección de tráfico transfronterizo de datos (cross-border privacy rules) reflejando un esfuerzo por anticipar desafíos asociados a la gobernanza y seguridad de la IA. Algunos países, como Argentina, Chile, Panamá y República Dominicana, han optado por desarrollar estrategias nacionales de inteligencia artificial, aunque estas no tienen carácter vinculante y suelen ser elaboradas por el Poder Ejecutivo. Otros, como Honduras, han incorporado la IA dentro de planes nacionales de gobierno digital. Por su parte, países como Costa Rica, Colombia y México aún carecen de estrategias nacionales específicas en esta materia (CENIA, 2024). Mientras que la mayoría de los países relevados cuentan con proyectos de ley en tratamiento legislativo (Argenti-

na, Ecuador, Costa Rica, Panamá y Venezuela).

Complementariamente, se registran países que han adherido a las recomendaciones sobre Ética de la Inteligencia Artificial de la UNESCO, aunque no dictaron normativa interna a tal efecto.

A pesar de estos avances parciales, un hallazgo relevante de este estudio es el escaso conocimiento de los marcos normativos, guías y protocolos específicos sobre IA por parte de los actores públicos, incluso en aquellos países que ya cuentan con legislación formal. Por otro lado, de manera casi unánime, los actores consultados omitieron referenciar marcos constitucionales protectores que, aunque no estén explícitamente diseñados para la IA, podrían ser aplicables en áreas fundamentales como la protección de datos, la confidencialidad, la necesidad de consentimientos informados y el habeas data. Esta omisión sugiere una subutilización del andamiaje jurídico existente que podría ofrecer bases para una gobernanza más robusta de la IA en la región. Los marcos constitucionales preexistentes y las leyes de protección de datos personales (presentes en muchos países de la región) constituyen una primera línea de defensa y una base sólida sobre la que asentar la arquitectura de gobernanza de la IA

E) Captura algorítmica del Estado, espacio digital y los impactos geopolíticos de la IA

Uno de las preocupaciones normativas identificados en esta investigación es en materia de regulación de monopolios, licitaciones y compras públicas nacionales dada las particularidades de la economía de la Inteligencia Artificial (datos como activo, efectos de red, altos costos de desarrollo e infraestructura), especialmente cuando se trata de tecnologías estratégicas críticas desarrolladas por actores privados con posiciones monopólicas dominantes. Dos aspectos resultan necesarios abordar: 1) captura algorítmica del Estado y monopolios; 2) elusión normativa en las articulaciones público-privadas (APP).

Un reciente estudio constitucional comparado encontró que en 19 países de Latinoamérica se encuentran reguladas de manera casi uniforme las prácticas antimonopólicas (GIZ, 2025). Sin embargo, la legislación actual se muestra insuficiente para prevenir prácticas anticompetitivas que puedan anticipar y regular la acumulación de poder inherente a los datos, los algoritmos propietarios y la infraestructura de *cloud computing* que sustentan estos sistemas. En este sentido, se puede mencionar que las leyes antimonopolio tradicionales no fueron diseñadas para activos

intangibles como los datos o la lógica algorítmica. Esto implica ir más allá de los criterios económicos convencionales para la selección de proveedores del Estado. Esto supone incorporar evaluaciones de impacto algorítmico, auditorías de sesgos, ciberseguridad, la exigencia de transparencia en las soluciones de IA ofrecidas por estos desarrolladores, así como evaluaciones de impacto en derechos humanos y estándares técnicos obligatorios en materia de interoperabilidad y explicabilidad. También implica incorporar el requisito de legalidad en el uso de los datos en toda la cadena de valor de la IA (desde el entrenamiento hasta la utilización), así como la obligación de explicitar la procedencia de datos y la utilización de datos sintéticos para el entrenamiento de modelos. Además, se deben considerar incentivos fiscales que promuevan el desarrollo local, fondos de inversión pública para startups de IA, cláusulas de transferencia de tecnología en contratos con ETN, o la creación de *sandboxes regulatorios* para que innovadores locales prueben soluciones. De esta manera, se fomenta la competencia, la innovación y la diversificación de la oferta en tecnologías estratégicas, sin que la necesidad de soluciones avanzadas genere una consolidación de facto de monopolios tecnológicos de empresas transnacionales.

También está presente el riesgo de captura regulatoria del Estado o elusión normativa a partir de las articulaciones públicas-privadas (APP) en el ámbito tecnológico. Esta captura adquiere nuevas y más sutiles dimensiones. Al presentarse como “proyectos de colaboración” o “pruebas piloto”, pueden evitar los rigurosos procesos de licitación pública diseñados para garantizar la transparencia, la competencia y la rendición de cuentas. Las APP pueden ser diseñadas para operar en un “espacio gris” normativo. Esto permite a las empresas establecerse como proveedores preferenciales sin pasar por un escrutinio estatal adecuado. La elusión no es solo de normas de contratación, sino también de aquellas relacionadas con protección de datos, sesgos algorítmicos o responsabilidad civil bajo el argumento de la experimentación. Este fenómeno facilita una forma de arbitraje regulatorio donde las empresas privadas, aprovechando su asimetría en conocimiento y capacidad técnica en el ámbito de la Inteligencia Artificial (IA), establecen posiciones dominantes. En última instancia, estas estrategias empresariales no solo comprometen la competencia y la rendición de cuentas, sino que también impiden o retrasan siste-

máticamente el desarrollo de capacidades soberanas que el propio Estado podría desplegar de manera autónoma. Esto genera una dependencia tecnológica -*vendor lock-in*- que, al consolidar el poder en manos de unos pocos proveedores, configura una captura algorítmica del Estado, debilitando su autonomía funcional y su capacidad de control sobre infraestructuras digitales críticas. Las políticas anti-monopolio y de compras públicas deben priorizar la interoperabilidad, la desagregación de servicios, la ciberseguridad y el uso de estándares abiertos para evitar la “captura algorítmica” del Estado y reducir la dependencia de proveedores únicos, a la vez que permite al Estado mantener el control sobre los datos y la lógica de sus procesos críticos. El fomento de la investigación aplicada y el desarrollo de capacidades internas es fundamental para construir soberanía tecnológica que prevenga la dependencia y la asimetría que condicionan el desarrollo soberano de los países del sur global.

El colonialismo digital se manifiesta en múltiples dimensiones, reflejando y perpetuando estructuras de poder históricas en el entorno digital. Los datos constituyen el principal insumo de la economía digital, y el uso de IA en el sector público implica decisiones políticas sobre qué problemas se priorizan, qué datos se procesan, con qué fines y bajo qué marcos regulatorios. Estas decisiones pueden fortalecer o debilitar la capacidad del Estado para responder equitativamente a las necesidades sociales y desviar recursos de otras agendas prioritarias. Por ello, consideramos que es necesario poner énfasis en la dimensión geopolítica que atraviesa el desarrollo y despliegue de tecnologías de IA. La concentración del poder tecnológico y económico en un reducido grupo de países y corporaciones transnacionales plantea riesgos de dependencia digital, colonialismo de datos y pérdida de soberanía tecnológica para América Latina.

Para ello, se considera que es fundamental promover modelos de desarrollo tecnológico soberanos que permitan a los países de la región ejercer un control democrático sobre las infraestructuras digitales, tecnologías y algoritmos que condicionan sus políticas públicas, sus sistemas de justicia y los derechos fundamentales. En este contexto, se torna necesario reflexionar sobre en qué medida la dependencia de infraestructura tecnológica y consecuente inversión puede afectar el sistema de frenos y contrapesos de

los poderes del Estado. De ahí la importancia de revisar críticamente las actuales salvaguardas, a fin de preservar la independencia de los poderes del Estado. Desde una mirada crítica, los hallazgos revelan que los riesgos asociados a la IA no se limitan a problemas técnicos, sino que reflejan decisiones de diseño, gobernanza y prioridades políticas que pueden profundizar desigualdades, consolidar sesgos y amenazar derechos fundamentales.

Este artículo destaca la necesidad de adoptar una gobernanza prospectiva, capaz de construir escenarios futuros que contemplen tanto las amenazas como las oportunidades de la IA con un enfoque intergeneracional. Frente a la opacidad de los sistemas de IA no alcanza con regulaciones que propongan la supervisión humana, el monitoreo y la evaluación de impacto. Se requiere un rol activo del Estado y de la academia. Latinoamérica debe implementar taxonomías y métodos de explicabilidad de los sistemas de IA, garantizando que el Estado no renuncie a la comprensión y control de las tecnologías que utiliza. La regulación de neurodatos y su comercialización transfronteriza plantea nuevos desafíos, que exigen estrategias de investigación para comprender la organización funcional de sistemas mayoritariamente desarrollados por corporaciones transnacionales.

F) Discriminaciones, sesgos de género y falta de interseccionalidad

Las personas encuestadas identificaron riesgos en los sistemas de IA vinculados a sesgos y discriminaciones algorítmicas en el uso de la IA. Los resultados de este estudio concuerdan con estudios previos que visibilizan discriminaciones algorítmicas. En efecto, según ONU Mujeres (2024) una investigación del Centro Berkeley Haas de Equidad, Género y Liderazgo encontró que, de un total de 133 sistemas de IA analizados, casi la mitad (44%) presentaba sesgos sexistas y un 25% sesgos sexistas y raciales de manera conjunta.

En este mismo sentido, diversos estudios han evidenciado que los sistemas de inteligencia artificial pueden reproducir y amplificar desigualdades estructurales de género y discriminación hacia grupos históricamente marginados. Perdomo Reyes (2024) y Rivas Vallejo (2022) advierten sobre la injusticia epistémica y los sesgos discriminatorios que surgen cuando se desarrollan algoritmos sin una perspectiva de derechos. A su vez, Perdomo Reyes (2024) men-

ciona que la IA se presenta como una herramienta neutral, a pesar de su opacidad. Según Guzmán Gómez (2025), la IA es un espejo de la sociedad actual, de modo que reproduce los sesgos y prejuicios existentes en relación al género, la etnicidad, la raza, la clase social, entre otros. Estos riesgos evidencian la necesidad de contar con enfoques de interseccionalidad de género en las estrategias de implementación de IA en el sector público, a fin de evitar formas de discriminación y de violencia epistémica hacia grupos en situación de vulnerabilidad.

G) Impacto en la división de poderes del Estado

De las entrevistas realizadas a personas expertas surge que la implementación de sistemas de inteligencia artificial (IA) en el ámbito estatal plantea desafíos significativos en términos de gobernanza democrática, especialmente cuando uno de los poderes del Estado —en este caso, el Ejecutivo— concentra el diseño, desarrollo y control de las estrategias de IA sin la participación o liderazgo del Poder Judicial. Esta situación puede generar un desbalance institucional, afectando el principio de división de poderes sobre el que se sustenta el Estado de Derecho. Si el Poder Judicial no es dueño ni co-responsable de la estrategia de IA que regula o condiciona sus propios procesos, su autonomía funcional y decisional puede verse comprometida. Sin embargo, esta discusión no está en el centro de la escena.

La dependencia tecnológica del Poder Judicial respecto de soluciones diseñadas, parametrizadas o controladas por el Poder Ejecutivo puede traducirse en una afectación de la independencia de los poderes, riesgo que también está presente en relación a las corporaciones transnacionales.

En términos más amplios, la concentración de la estrategia de IA en manos del Poder Ejecutivo puede erosionar el sistema de frenos y contrapesos que garantiza el equilibrio entre los poderes del Estado. La tecnología, en lugar de fortalecer la institucionalidad, podría transformarse en un instrumento de recentralización y control político si no existen salvaguardias normativas, organizativas y participativas que aseguren su uso conforme a estándares democráticos y de derechos humanos. Por ello, resulta imperativo que el Poder Judicial asuma un rol activo en la definición de los marcos éticos, técnicos y jurídicos de la IA, garantizando su independencia y el respeto a la autonomía de cada poder, ejes del Estado de

derecho. Entre las recomendaciones identificadas se encuentra la necesidad de establecer presupuestos específicos para el desarrollo e implementación de herramientas de IA bajo control del propio Poder Judicial, evitando la dependencia financiera y tecnológica del Ejecutivo y de sectores privados. Asimismo, es necesario promover la creación de unidades técnicas internas especializadas en IA y ética digital, capaces de evaluar y auditar las tecnologías en uso. La participación de los tres poderes en mesas interinstitucionales de gobernanza de IA, con representación de la sociedad civil y expertos independientes, puede ser otro paso fundamental para garantizar un enfoque colaborativo y respetuoso de los valores democráticos.

H) Uso de la IA en el Poder Judicial

La implementación de la Inteligencia Artificial (IA) en los sistemas judiciales presenta tanto oportunidades significativas para mejorar el acceso a la justicia, las garantías de defensa en juicio y el debido proceso, como riesgos inherentes a la amplificación de la desconfianza en los Poderes Judiciales, si su uso carece de transparencia, responsabilidad, y *compliance judicial*.

La desconfianza en el Poder Judicial es un fenómeno significativo en América Latina, manifestado en percepciones negativas sobre su independencia y eficacia (ACIJ, 2013; ATAJO, 2016). La incorporación de la IA en este ámbito, si bien promete eficiencia, consistencia y velocidad, simultáneamente plantea desafíos complejos que podrían exacerbar esta desconfianza. Según el Índice de Confianza en la Justicia (ICJ, 2023), un 91% de las personas entrevistadas percibe a la Justicia como “nada” o “poco confiable” en términos de capacidad y eficiencia (Di Tella, FORES, 2023). La comparación entre la actuación judicial humana y los sistemas de Inteligencia Artificial (IA) en la administración de justicia revela diferencias sustanciales en la naturaleza y gestión de sus limitaciones. Si bien ambos son susceptibles de error, los fallos humanos están sujetos a revisión institucional, sanciones, incluso la remoción y el escrutinio público. En contraste, los errores de un sistema de IA suelen ser opacos a nivel técnico, lo que complejiza la identificación de responsabilidades y la corrección de sesgos estructurales o “alucinaciones”. Además, el impacto y la propagación de un error en un sistema de IA pueden escalar rápidamente, con un margen de tiempo reducido para su contención. La eficiencia, consistencia y velocidad que promete

la IA en el ámbito judicial deben ponderarse frente a la necesidad de salvaguardas procesales inherentes al debido proceso. La garantía de audiencia, la motivación explícita de las decisiones y la posibilidad de una revisión efectiva son irrenunciables. Sin estas protecciones, la integración de la IA en la justicia podría comprometer derechos fundamentales y debilitar el sistema judicial.

La Inteligencia Artificial (IA) ofrece una posibilidad para mejorar la aplicación de criterios judiciales de manera más uniforme entre distintas jurisdicciones. Podría, por ejemplo, facilitar la consistencia en decisiones sobre casos complejos como los relacionados con medicamentos de alto costo y tecnologías sanitarias. Adicionalmente, la IA tiene el potencial de optimizar la aplicación de tratados internacionales, la jurisprudencia de la Corte Interamericana de Derechos Humanos, asegurando una mayor uniformidad en su incorporación. Un entrenamiento adecuado permitiría a estos sistemas incluso proponer argumentaciones que ayuden a mitigar sesgos y prejuicios, incluyendo aquellos de género. El elemento central para estas posibilidades reside en el acceso a los datos de entrenamiento (transparencia algorítmica), la identidad de quienes los entrena y los criterios utilizados para el desarrollo de estos sistemas de IA.

Para este estudio se elaboraron criterios diferenciadores entre el juez humano y un sistema de IA en la administración de justicia incluyen la inmediatez judicial, la competencia, la independencia, la imparcialidad, los plazos, la potencialidad de lawfare, la presencia de sesgos y prejuicios, la responsabilidad funcional y el riesgo de captura corporativa. Aunque la intervención humana no está exenta de errores, sesgos y prejuicios, o de la posibilidad de lawfare como herramienta de persecución selectiva, existen contrapesos. La autonomía judicial es susceptible a la influencia de grupos de interés o élites transnacionales, cuyas prácticas de captura corporativa pueden afectar la administración de justicia. Los sesgos y prejuicios del juez, vinculados a su sistema de creencias, clase, género, raza, ideología y trayectorias personales, pueden influir en las decisiones. La dependencia de algoritmos entrenados con datos sesgados o datos sintéticos plantea riesgos de parcialidad y amplificación de problemas sistémicos. La responsabilidad funcional, claramente definida para el/juez/a humano, se diluye en el caso de la IA

entre desarrolladores y operadores, lo que dificulta la atribución de responsabilidades.

I) La IA en el sector sanitario y la protección de datos

Uno de los riesgos recurrentes señalados por las personas encuestadas del sector salud se refieren a la utilización para entrenamiento de sistemas de IA de datos privados y sensibles que fueron obtenidos sin consentimiento explícito o se le da un uso indebido a información consentida (reutilización de datos prohibidos). Los riesgos percibidos por las personas encuestadas en materia de salud coinciden con los principales debates académicos en la temática y con el abordaje que se está haciendo de estos por parte de los organismos internacionales e intergubernamentales que ponen el acento en la ciberseguridad, la protección de datos sensibles, la responsabilidad en toda la cadena de valor de la IA. Estos aspectos se tornan aún más críticos teniendo en cuenta que algunos autores mencionan que una vez que ingresan los datos a los sistemas de IA, es casi imposible borrarlos (Kondrup, 2025; Seth Neel, 2024). Corrales Compagnucci (2024) señala que un reciente Dictamen 28/2024 del Comité Europeo de Protección de Datos (CEPD) implica un hito para la protección de datos que utilizan los modelos de IA. Este instrumento eleva los estándares para el desarrollo de IA en Europa, especialmente en salud, aunque no es de carácter vinculante para los países, puede guiar la interpretación e inspirar nueva normativa. En este escenario, se menciona que no alcanza con la anonimización de historias clínicas en IA, pues existen técnicas que permiten reidentificar datos, sino que tienen que cumplir una serie de pasos tanto en el entrenamiento de modelos como en el uso. El instrumento exige: anonimización real de los modelos de IA; justificación de la necesidad legítima del uso de datos; garantía de que los derechos de las personas no se vean comprometidos y comprobación de que, tanto durante el entrenamiento como en su aplicación, los beneficios superen los riesgos, los cuales deben ser “razonablemente improbables”.

Así, si un modelo de IA fue entrenado con datos obtenidos o procesados ilegalmente, su implementación posterior también puede considerarse ilícita. Este dictamen enfatiza que la privacidad debe estar integrada desde el diseño, especialmente en áreas como genética o enfermedades raras, donde los riesgos de reidentificación son mayores. Los modelos

deben ser sometidos a pruebas técnicas, evaluaciones de impacto sobre la protección de datos y supervisión constante. El uso de datos ilícitos compromete toda la cadena de valor de la IA, desde el modelo hasta los sistemas donde se aplique (Corrales Compagnucci, 2024).

Hasta aquí podemos señalar que los riesgos identificados en torno al uso de la inteligencia artificial en el sector público no constituyen simplemente un listado de problemas técnicos o administrativos a resolver. Lo que emerge de las narrativas es un entramado complejo de tensiones que atraviesan dimensiones tecnológicas, institucionales, sociales y subjetivas, y que interpelan de manera profunda el rol del Estado, la configuración del poder institucional y la experiencia ciudadana en los procesos públicos.

En términos institucionales, la incorporación de la IA pone en evidencia limitaciones estructurales del aparato estatal para gobernar tecnologías de alta complejidad, no sólo en relación al propio sector público sino en relación a las empresas transnacionales y el sector privado. La falta de marcos regulatorios adecuados, la debilidad en los sistemas de control y fiscalización, y la escasa capacidad de implementación crítica revelan que la innovación tecnológica puede avanzar más rápido que las condiciones institucionales necesarias para su apropiación y desarrollo democrático. En lugar de fortalecer al Estado, la adopción acrítica de la IA puede profundizar su dependencia de actores corporativos transnacionales, erosionar su autonomía y poner en riesgo su legitimidad como garante de derechos.

Desde una perspectiva estatal más amplia, los riesgos también hablan de disputas por la soberanía digital, el control del conocimiento y la justicia distributiva en el acceso y uso de tecnologías.

La centralización del desarrollo de IA en el norte global, la falta de estándares y el carácter opaco de muchos sistemas amenazan con reproducir lógicas coloniales y extractivistas en el plano informacional. Frente a esto, se vuelve urgente repensar modelos de gobernanza tecnológica que prioricen la transparencia, la rendición de cuentas y la participación ciudadana, así como una infraestructura digital pública, soberana y sustentable. No es posible la sustentabilidad digital sin sustentabilidad social, económica y ambiental.

A nivel subjetivo, la percepción de deshumanización, desconfianza, miedo y pérdida de autonomía revela que el problema no radica únicamente en lo que hace la IA, sino cómo transforma la manera en que las personas se vinculan con el Estado, con sus trabajos y consigo mismas.

Es necesario generar futuras líneas de investigación del campo de la psicología ambiental, organizacional, comunitaria o clínica que aborden aspectos subjetivos en entornos digitales emergentes y el impacto en la salud mental. La automatización de decisiones públicas puede deteriorar el vínculo entre ciudadanía e instituciones, erosionar la dimensión relacional de los servicios públicos y sustituir el juicio humano por procesos algorítmicos que, aunque eficientes, resultan ajenos, impersonales y poco explicables. Pensar los riesgos de la IA en el sector público requiere entonces un abordaje integral, que no se limite a soluciones técnicas sino que promueva una reflexión profunda sobre los modelos de Estado, de conocimiento y de subjetividad que están en juego. La pregunta clave no es si usar o no IA, sino cómo, para qué y bajo qué condiciones institucionales, normativas y éticas se la pone al servicio del interés público, la justicia social y el fortalecimiento democrático.

Resulta necesario contar con un enfoque de gobernanza anticipatoria y estudios prospectivos. En este sentido, los observatorios y laboratorios de innovación en políticas públicas se constituyen como espacios privilegiados para impulsar este tipo de iniciativas.

Conclusión

Este artículo analiza críticamente los beneficios y riesgos del uso de la inteligencia artificial (IA) en el sector público latinoamericano, a partir de procesos participativos con actores estatales, académicos y sociales. Se constata que la IA tiene un alto potencial para mejorar eficiencia, optimizar recursos y ampliar el acceso a servicios públicos, pero su adopción ocurre en contextos de brechas digitales, desigualdades estructurales y marcos normativos insuficientes, predominando el uso informal de herramientas desarrolladas por corporaciones transnacionales, sin capacitación ni protocolos adecuados. Se identifican riesgos de captura algorítmica del Estado, dependencia tecnológica, concentración de poder en datos y algoritmos, y debilitamiento de la soberanía digital. Es necesario fortalecer políticas de compras públicas, interoperabilidad, competencia local y estándares abiertos, así como promover marcos regulatorios específicos para monopolios, licitaciones y protección de datos, garantizando transparencia, explicabilidad y auditorías de sesgos.

En el sector público es necesario asegurar trazabilidad, explicabilidad, equidad, consentimiento informado y protección de datos sensibles, especialmente para poblaciones vulnerables. También se destaca el rol de la academia en desarrollar metodologías que reduzcan la opacidad de los sistemas y promuevan un uso ético y democrático de la IA. Los hallazgos muestran un escenario de oportunidades, pero con tensiones éticas, legales y sociales, donde la adopción acrítica de tecnologías externas puede erosionar la soberanía digital y desplazar el control de decisiones públicas hacia actores privados transnacionales.

Bibliografía

- ACIJ (2020). *Reporte de resultados situación de acceso a la justicia en Latinoamérica durante la pandemia de covid-19 desde la perspectiva de actores sociales*. Informe elaborado por MN Echegoyemberry. Disponible: <https://acij.org.ar/reporte-acceso-a-la-justicia-en-latinoamerica-durante-la-pandemia-de-covid-19/>
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://link.springer.com/article/10.1007/s13194-024-00614-4#citeas>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? En *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (FAccT '21), (pp.610–623). <https://doi.org/10.1145/3442188.3445922>
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Liang, P. (2021). *On the opportunities and risks of foundation models*. arXiv <https://doi.org/10.48550/arXiv.2108.07258>.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, pp.1877–1901. <https://doi.org/10.48550/arXiv.2005.14165>
- Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1). doi. org/ 10.1177/ 20539 51715 622512
- Centro Nacional de Inteligencia Artificial (Cenia) (2024). Índice Latinoamericano de Inteligencia Artificial (ILIA) Índice latinoamericano de IA. <https://indicelatam.cl/>
- Crawford, K (2022) *Atlas de inteligencia artificial: poder, política y costos planetarios*. Fondo de Cultura Económica
- Couldry, N, Mejías U A. (2019). *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*.Stanford University Press.
- Contreras, P (2024). *Convergencia internacional y caminos propios: regulación de la inteligencia artificial en América Latina*. <https://revista-ajij.com/convergencia-internacional-y-caminos-propios-regulacion-de-la-inteligencia-artificial-en-america-latina/>
- Corrales Compagnucci (2024). *Europe Tightens Data Protection Rules for AI Models—And It's a Big Deal for Healthcare and Life Sciences* <https://petrieflom.law.harvard.edu/2025/02/24/europe-tightens-data-protection-rules-for-ai-models-and-its-a-big-deal-for-healthcare-and-life-sciences/>
- Echegoyemberry, MN, Piccardo, V, Rojo, P, Allassia, L Galeazzi, C, Brolese, C, Carneiro, A, Duarte, R, Cámarra, J (2025). *Políticas públicas para la innovación judicial: Uso, aplicación y explicabilidad de la IA en el sector público. Resultados Preliminares*.<https://justiciasanitaria.org/2025/07/politicas-publicas-para-la-innovacion-uso-aplicacion-y-explicabilidad-de-la-ia-en-el-sector-publico-presentacion-de-resultados-preliminares>
- Fuster, G. (2022).Regulación de la IA de propósito general y la Ley de IA de la Unión Europea. *Internet Policy Review*, 13(3). <https://policyreview.info/articles/analysis/general-purpose-ai-regulation-and-ai-act>
- Floridi, L., Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). doi.org/10.1162/99608f92.8cd550d1.
- GIZ (2025). *El Derecho a la Salud en Latinoamérica Desde la Perspectiva de Actores Sociales*. Documento elaborado por Natalia Echegoyemberry. Disponible: <https://dirapjus.org/es/publicaciones>
- GIZ (2025). *Lineamientos para una comprensión regional del Derecho a la Salud: Propuesta metodológica para un análisis constitucional comparado desde una perspectiva de diseño normativo*. <https://dirapjus.org/es/repositorio-desca/salud>
- Governança da Inteligência Artificial no BRICS: Cooperação e Desenvolvimento para a Inclusão Social. (s/f) <https://brics.br/pt-br/noticias/artigos/governanca-da-inteligencia-artificial-no-brics-cooperacao-e-desenvolvimento-para-a-inclusao-social>
- Guzmán Gómez, M. E. (2025). Inteligencia artificial y equidad de género: una perspectiva histórica de los sesgos culturales y su impacto en la relación humana con las tecnologías de la información y comunicación. *Sintaxis*, (14), 14–30. <https://doi.org/10.36105/stx.2025n14.02>

- Gstrein, O.J., Haleem, N., Zwitter, A. (2024). General-purpose AI regulation and the European Union AI Act. *Internet Policy Review*, 13(3) doi.org/10.14763/2024.3.1790
- Hueso, L. C. (2020). Inteligencia artificial, big data y aplicaciones contra la COVID-19: privacidad y protección de datos, en *IDP. Revista de Internet, Derecho y Política* (Nº 31), 1-17. <https://idp.uoc.edu/articles/abstract/10.7238/idp.v0i31.3244/>
- Kästner, L., Crook, B. (2024). Explicación de la IA mediante la interpretabilidad mecanicista. *Revista Europea de Filosofía de la Ciencia*, 14(4), 52. <https://doi.org/10.1007/s13194-024-00614-4>
- Kwet, M (2019). Digital Colonialism: US Empire and the New Imperialism in the Global South https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3232297
- Kondrup, E (2025). *Informed Consent, Redefined: How AI and Big Data Are Changing the Rules*. <https://petrieflom.law.harvard.edu/2025/04/11/informed-consent-redefined-how-ai-and-big-data-are-changing-the-rules/>
- Kwang-Hyung, Lee (2025). *La importancia de la Ley Básica de IA*. <https://koreajoongangdaily.joins.com/news/2025-01-21/opinion/columns/The-importance-of-the-AI-Basic-Act/2226170>
- Morozov, Evgeny (2022). Crítica al tecnofeudalismo. *New Left Review*, 135, 99-140, <https://newleftreview.es/issues/133/articles/critique-of-techno-feudal>
- Muñoz, V. (2024). *Inteligencia artificial: potenciando el futuro de América Latina y el Caribe*. Análisis Carolina, (18). Fundación Carolina. https://doi.org/10.33960/AC_18.2024
- Organización de las Naciones Unidas (ONU). (2023). *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development* (Res. A/78/L.49). <https://digitallibrary.un.org/record/4040897>
- Perdomo Reyes, I. (2024). Injusticia epistémica y reproducción de sesgos de género en la inteligencia artificial. *Revista Iberoamericana de Ciencia, Tecnología y Sociedad* 19(56), 89–100. <https://doi.org/10.52712/issn.1850-0013-555>.
- Ribeiro, M. T., Singh, S., Guestrin, C. (2016). “Why Should I Trust You?”: Explaining the predictions of any classifier. *Proceedings of the Conference of the North American Chapter of the ACL* 97–101. <https://doi.org/10.18653/v1/N16-3020>
- Rivas Vallejo, P. (2022). Sesgos de género en el uso de inteligencia artificial para la gestión de las relaciones laborales: análisis desde el derecho antidiscriminatorio. *e-Revista Internacional de la Protección Social*, 7(1), 61–84. <https://doi.org/10.12795/e-RIPS.2022.i01.04>
- Seth Neel. By Rachel Layne (2024). *How to Make AI Forget' All the Private Data It Shouldn't Have*. <https://www.library.hbs.edu/working-knowledge/qa-seth-neel-on-machine-unlearning-and-the-right-to-be-forgotten>
- Strubell, E., Ganesh, A., McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645–3650. <https://doi.org/10.18653/v1/P19-1355>.
- Tejedor-Estupiñán, J. M. (2025). Impacto de la inteligencia artificial en la producción, la productividad y el aumento de la desigualdad en América Latina. *Revista Finanzas y Política Económica*, 17, 1–3. <https://doi.org/10.14718/revfinanzpolitecon.v17.2025.0>
- Varoufakis, Yanis (2024). *Tecnofeudalismo: El sigiloso sucesor del capitalismo*. https://www.lecturalia.com/libro/111771/tecnofeudalismo#google_vignette
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs. <https://www.hbs.edu/faculty/Pages/item.aspx?num=56791>

