

Inteligencia artificial. La urgencia de una nueva ética en un contexto global de cambio

Inteligência artificial. A urgência de uma nova ética em um contexto global de mudança

Artificial Intelligence. The Urgency of New Ethics in a Global Context of Change

Gabriela Agosto 

El desarrollo acelerado de la inteligencia artificial (IA) constituye uno de los fenómenos más significativos y disruptivos del siglo XXI, tanto por su capacidad de transformar las estructuras que cohesionan la sociedad actual como por las consideraciones éticas que suscita su aplicación. En un contexto global en el que convergen simultáneamente múltiples crisis sociales, políticas, económicas y ambientales, la IA se presenta como promesa y amenaza a la vida como la conocemos. Su impacto no se limita a la esfera técnica, sino que afecta las dinámicas de poder, la organización y productividad del trabajo, la educación y socialización, la producción académica y la formulación de políticas públicas. Ante las inquietudes y los interrogantes que estos avances tecnológicos plantean, la reflexión ética emerge como condición imprescindible para que esta tecnología esté al servicio del ser humano, y evitar que se convierta en instrumento de exclusión o dominación. Este artículo propone un abordaje interdisciplinario desde las ciencias sociales, incorporando una reflexión metaacadémica: el uso de la IA como herramienta para analizarse a sí misma de manera de contribuir a los desafíos contemporáneos.

33

Palabras clave: inteligencia artificial; ética; algorética; colonialismo digital; soberanía algorítmica

* Decana de la Facultad de Ciencias Sociales, Educación y Comunicación de la Universidad del Salvador (USAL), Argentina. Doctora en ciencia política y sociología por la Universidad Complutense de Madrid (UCM), España. Magíster en administración pública (IUPOG-UCM). Licenciada en sociología por la Universidad de Buenos Aires (UBA), Argentina. Investigadora y docente universitaria, consultora en temas de desarrollo e integración regional en el ámbito nacional e internacional. Correo electrónico: gagosto@usal.edu.ar. ORCID: <https://orcid.org/0009-0008-5842-6967>. Para este artículo se contó con la colaboración de Lucio Aya Tenorio, licenciado en relaciones internacionales por la Universidad Católica Argentina (UCA) y académico de la Facultad de Ciencias Sociales, Educación y Comunicación de la Universidad del Salvador (USAL), cuya asesoría y reflexión crítica resultaron fundamentales para el desarrollo del trabajo. Correo electrónico: luciotenorio@usal.edu.ar. ORCID: <https://orcid.org/0009-0007-0089-9908>.

O desenvolvimento acelerado da inteligência artificial (IA) constitui um dos fenômenos mais significativos e disruptivos do século XXI, tanto por sua capacidade de transformar as estruturas que mantêm a sociedade contemporânea quanto pelas considerações éticas que suscita sua aplicação. Em um contexto global no qual múltiplas crises sociais, políticas, econômicas e ambientais convergem simultaneamente, a IA se apresenta como promessa e ameaça à vida tal como a conhecemos. Seu impacto não se limita à esfera técnica, mas também afeta as dinâmicas de poder, a organização e a produtividade do trabalho, a educação e a socialização, a produção acadêmica e a formulação de políticas públicas. Diante das preocupações e dos questionamentos que esses avanços tecnológicos colocam, a reflexão ética emerge como condição indispensável para que essa tecnologia esteja a serviço do ser humano e não se converta em instrumento de exclusão ou dominação. Este artigo propõe uma abordagem interdisciplinar a partir das ciências sociais, incorporando uma reflexão meta-acadêmica: o uso da IA como ferramenta para analisar a si mesma de modo a contribuir para os desafios contemporâneos.

Palavras-chave: inteligência artificial; ética; algorética; colonialismo digital; soberania algorítmica

34

The rapid development of artificial intelligence (AI) constitutes one of the most significant and disruptive phenomena of the 21st century, both for its capacity to transform social structures and for the ethical considerations that its application raises. In a global context where multiple social, political, economic, and environmental crises converge simultaneously, AI appears as both a promise and a threat to life as we know it. Its impact is not only limited to the technical sphere but also affects power dynamics, the organization and productivity of labor, education and socialization, academic production, and the formulation of public policies. Faced with the concerns and questions raised by these technological advances, ethical reflection emerges as an essential condition to ensure that this technology serves humanity and does not become an instrument of exclusion or domination. This article proposes an interdisciplinary approach from the social sciences, incorporating a meta-academic reflection: the use of AI as a tool to analyze itself to contribute to contemporary challenges.

Keywords: artificial intelligence; ethics; algorethics; digital colonialism; algorithmic sovereignty

Introducción

La tercera década del siglo XXI fue testigo de una aceleración sin precedentes en los procesos de innovación tecnológica, cuyo exponente más paradigmático es la emergencia de la inteligencia artificial (IA). Su desarrollo, expansión y apropiación social se producen en un tiempo histórico caracterizado por la incertidumbre, la fragilidad de las instituciones y la velocidad del cambio. Si Bauman (2000) habló de modernidad líquida, autores como Riorda (2021) han propuesto recientemente que el paradigma sociohistórico actual es más bien gaseoso, volátil, impredecible e intangible: las estructuras y referentes tradicionales se desvanecen antes de que sea posible aferrarse a ellos o comprender su impacto e influencia.

En paralelo, la humanidad enfrenta lo que Morin (2020) y Toose (2021) describen como una policrisis: una multiplicidad de crisis simultáneas e interconectadas (climática-ambiental, económica, social, política y cultural) que no pueden ser comprendidas ni abordadas de manera aislada. La IA irrumpe en un escenario de vulnerabilidades acumuladas, en el que el desajuste entre los problemas sociales no resueltos, la crisis de la democracia liberal, los procesos bélicos y el vertiginoso avance tecnológico se convierte en un signo de nuestro tiempo.

En este contexto, el debate sobre la ética de la IA deja de ser una cuestión secundaria o meramente técnica. Su dualidad, con sus enormes potencialidades y sus riesgos igualmente significativos, obliga a reflexionar en sus aplicaciones inmediatas y en sentido último de su integración en la vida humana. Las ciencias sociales, al analizar los impactos de la IA en las interacciones, la producción académica y las políticas públicas, se enfrentan al desafío de elaborar marcos teóricos, normativos y regulatorios capaces de acompañar una transformación de alcance global. Este artículo parte de la premisa que la ética de la IA no puede pensarse solo desde marcos nacionales ni desde perspectivas estrictamente tecnológicas. Se requiere un enfoque interdisciplinario que reconozca a la humanidad como un todo, más allá de las fronteras y los particularismos.

35

Finalmente, este texto incorpora como epílogo una reflexión hecha por la IA, con el objetivo de explorar sus alcances y limitaciones en su autoconocimiento, haciendo de la escritura misma un espacio de reflexión crítica. De este modo, la investigación se despliega no solo sobre la IA, sino también con ella, invitando a considerar las implicancias éticas, sociales y filosóficas de un futuro en el que la inteligencia humana y la artificial coexisten.

De esta manera, se propone, en primer lugar, ofrecer un panorama general que dé cuenta de la complejidad y la dificultad de abarcar los impactos de la IA en todas las dimensiones socioculturales posibles. En segundo lugar, busca proponer un espacio de reflexión sobre la construcción colectiva de una nueva ética que oriente su desarrollo, como una tarea urgente e impostergable. En este sentido, se sostiene que la IA y otros avances tecnológicos recientes provocaron una ruptura del contrato social

tal como lo entendimos desde una visión rousseauiana. De allí surge la necesidad de elaborar un nuevo contrato, capaz de responder a los desafíos de esta transición. La pregunta que subyace es si la IA constituye un nuevo tipo de Leviatán, una entidad tan poderosa que, lejos de limitarse a servir al ser humano, se erige como mediadora y condicionante de sus vínculos, sus instituciones y su futuro común.

Para los fines de este trabajo, nos referiremos a la IA en su generalidad, sin realizar distinciones exhaustivas entre los múltiples tipos, modelos, servicios, capacidades y características específicas de los diversos productos ofrecidos por las corporaciones que lideran el sector como OpenAI (ChatGPT), Google DeepMind (Gemini), Anthropic (Claude), Meta (LLaMA), Microsoft, IBM o Amazon Web Services). Tampoco haremos una separación sistemática entre la IA generativa y la IA predictiva.¹ Lo que interesa no son tanto las diferencias técnicas entre ellas, sino lo que comparten desde un punto de vista epistemológico: la capacidad de procesar información, producir representaciones y generar efectos sociales que interpelan nuestra comprensión de lo humano. Esta mirada permite situar el debate en el nivel de las transformaciones estructurales. No obstante, conviene precisar una distinción conceptual: no se trata de la proliferación de “modelos específicos” en sentido estricto, sino de modelos generalistas contextualizados. Los LLM (*large language models*) como GPT, Claude o Gemini constituyen infraestructuras de propósito general que, mediante ajustes de contexto, interfaces o *fine-tuning*, se orientan a usos concretos (desde el ámbito jurídico hasta la salud, la educación o la gestión administrativa). Lo que varía, entonces, no es el modelo de base, sino el ecosistema de datos, aplicaciones y normativas en el que se inserta. Esta aclaración es fundamental para evitar lecturas que sobrestimen la diversidad de modelos y subestimen la centralidad de unas pocas arquitecturas dominantes en el mercado.

36

1. Revisión de literatura y estado del arte

La historia de la inteligencia artificial se remonta a mediados del siglo XX, con hitos importantes en la década de 1960. Sin embargo, el punto de inflexión se dio a fines de 2022, cuando OpenAI lanzó ChatGPT. A diferencia de los desarrollos previos, cuya circulación permanecía restringida a ámbitos especializados, ChatGPT democratizó el acceso a LLM, integrando sus capacidades en la vida cotidiana de millones de personas en cuestión de meses. Este salto cualitativo no radica únicamente en la potencia técnica del modelo, sino en su apropiación social masiva, que lo convierte en un fenómeno cultural y político, además de tecnológico.

Esta irrupción, así como su desarrollo previo, dio lugar a un *corpus* creciente de reflexiones que trascienden lo meramente técnico para situarse en el plano social,

¹ La primera puede definirse como aquella capaz de producir nuevos contenidos (textos, imágenes, sonidos, código) a partir de patrones aprendidos en grandes volúmenes de datos; la segunda, en cambio, se centra en inferir probabilidades y anticipar necesidades o comportamientos futuros a partir de datos históricos.

ético y político. La literatura especializada coincide en que el impacto de la IA es transversal, de manera que afecta por igual a la economía, la política, la cultura, la vida cotidiana y la esfera íntima de las personas. Su carácter disruptivo motiva la aparición de una diversidad de enfoques provenientes de la filosofía, la sociología, la ética, las ciencias políticas y la teología, que buscan comprender tanto sus potencialidades como sus riesgos.

En la literatura actual sobre ética e IA convergen diversas líneas de análisis, para este documento se seleccionaron tres, entendiendo que conceptualizan de manera integral la problemática. En primer lugar, la evidencia empírica sobre daños sociales causados por la IA (sesgos algorítmicos, discriminación automatizada, vigilancia masiva, daño ambiental), documentada en estudios de autores como Crawford (2021) y Zuboff (2019), que muestran cómo estas tecnologías pueden profundizar desigualdades preexistentes. En segundo lugar, los marcos normativos en desarrollo que buscan trasladar los valores de transparencia, responsabilidad e inclusión a criterios operativos y exigibles, como la *Rome Call for AI Ethics* (2020), la *AI Act* (2024) de la Unión Europea (UE) y los principios de la UNESCO (2021) y la Organización para la Cooperación y el Desarrollo Económico (OCDE, 2019, actualizados en 2024). En tercer lugar, la investigación emergente en justicia algorítmica, encabezada por académicos quienes desarrollaron metodologías para medir, evidenciar y corregir las desigualdades reproducidas por sistemas automatizados. En este marco, el concepto de “algorética” (Francisco, 2020) se presenta como un lenguaje común que traduce preocupaciones éticas universales vinculadas a la dignidad humana, la equidad y el respeto a los derechos fundamentales en principios aplicables al diseño y la regulación.

37

Los sistemas de IA generativa y modelos fundacionales desplazaron la discusión desde la promesa tecnológica hacia la gobernanza de riesgos sociales y civilizatorios. En términos normativos, la *Recomendación sobre la ética de la inteligencia artificial* de la UNESCO (2021) estableció un estándar global de referencia sobre derechos humanos, supervisión humana, transparencia, y rendición de cuentas, hoy utilizado por Estados y universidades como guía de políticas institucionales; los Principios de la OCDE (2019, actualizados en 2024) convergen en valores de derechos humanos, transparencia, robustez y responsabilidad y en recomendaciones de política pública, incluyendo cooperación internacional y preparación para la transición laboral. El 2024 marcó un punto de inflexión con la adopción del *AI Act* en la UE, que inauguró un esquema de obligaciones graduadas orientado a una IA “humano-céntrica y confiable” y a la protección de derechos fundamentales y de la democracia (Reglamento 2024/1689, considerandos 1-3).

La discusión sobre soberanía algorítmica condensa preocupaciones de países y regiones por asegurar capacidad de decisión, infraestructura y datos frente a las crecientes dependencias tecnológicas. En Europa, el *AI Act* se acompaña de un ecosistema de implementación (AI Office, *sandboxes*, códigos de práctica) que busca compatibilizar innovación con protección de derechos. En América Latina, se consolida una agenda de soberanía digital con debates sobre nubes públicas, estándares

abiertos, alfabetización de IA y compras públicas estratégicas, conectando con la historiografía regional e iniciativas de observatorios y redes académicas que mapean riesgos y oportunidades. Sin embargo, no se hicieron grandes progresos respecto de la regulación de esta tecnología.

En el campo de la filosofía y la ética, Floridi (2013, 2019) planteó la necesidad de una *ética de la información* y de un marco normativo que acompañe el desarrollo de las tecnologías digitales, en particular la IA, al sostener que el problema central no es solo lo que las máquinas pueden hacer, sino cómo sus capacidades reconfiguran las condiciones de posibilidad de la vida humana y social. En la misma línea, Metzinger (2021) advierte sobre los riesgos de avanzar en sistemas de IA cada vez más autónomos sin una base ética sólida, alertando sobre la tentación de delegar decisiones morales en algoritmos que carecen de experiencia subjetiva y emotiva.

Desde la sociología y la ciencia política, Zuboff (2019) describió el fenómeno del “capitalismo de la vigilancia”, mostrando cómo las plataformas digitales y los algoritmos transforman los datos personales en la materia prima de un nuevo régimen económico basado en la predicción y manipulación del comportamiento. Srnicek (2016), por su parte, analiza el ascenso del “capitalismo de plataformas”, donde la concentración de datos y poder en pocas corporaciones genera nuevas asimetrías de poder global. Estas lecturas son fundamentales para comprender el trasfondo económico y político en el que se desarrolla la IA.

38

Crawford (2021) sostuvo que la IA no es simplemente un conjunto de sistemas técnicos, sino una infraestructura sociotécnica que refleja y amplifica las relaciones de poder existentes. Su análisis evidencia cómo los sesgos algorítmicos y los sistemas de clasificación automatizados perpetúan desigualdades raciales, de género y socioeconómicas. Eubanks (2018) llega a conclusiones similares al mostrar cómo la automatización en programas sociales en Estados Unidos llevó a nuevas formas de exclusión y marginación, lo que plantea interrogantes sobre el uso de la IA en políticas públicas.

La dimensión teológica y humanista también aporta al debate contemporáneo. El papa Francisco (2020), en su *Llamada de Roma*, introdujo el concepto de “algorética”, que articula seis principios fundamentales (transparencia, inclusión, responsabilidad, imparcialidad, fiabilidad, seguridad y privacidad) para garantizar que la IA permanezca al servicio del ser humano. Ciucci (2023), desde la Pontifical Academy for Life, insistió en que la IA plantea un desafío civilizacional que obliga a pensar en términos de dignidad humana universal, más allá de particularismos nacionales o locales. Esta perspectiva coincide con las advertencias de Benanti (2022), teólogo y miembro del Vaticano en debates sobre IA, quien señala que el verdadero riesgo es la subordinación de lo humano a la lógica técnica, despojando a la sociedad de su capacidad de autodeterminación y reflexión moral.

Adicionalmente, ante la capacidad transformadora de la IA tanto en los ámbitos económicos y políticos como respecto a la concepción misma de la libertad y la agencia

humana, Harari (2018) advierte que los algoritmos podrían conocernos mejor que nosotros mismos y condicionar nuestras decisiones más íntimas, lo que cuestiona la noción misma de libre albedrío. Por otro lado, Cristianini (2023) propone entender la IA no como una inteligencia autónoma, sino como una extensión de los sistemas humanos de información, lo que desplaza el debate hacia la responsabilidad humana en su diseño y aplicación. Autores como Noble (2018) evidenciaron cómo la IA y los motores de búsqueda no son neutrales, sino que reproducen y amplifican discriminaciones, abriendo una línea de investigación crucial en torno a la justicia algorítmica.

El debate contemporáneo sobre sesgos algorítmicos se centra en los daños distribuidos y las arquitecturas de injusticia digital. Si bien el trabajo de Buolamwini y Gebru (2018) marcó un punto de referencia al demostrar el impacto diferenciado de los sistemas de reconocimiento facial según género y color de piel, la literatura actual destaca que estos problemas no son anomalías corregibles únicamente con “más datos” o “mejores modelos”, sino expresiones de estructuras sociales preexistentes codificadas en sistemas técnicos e informáticos. Asimismo, Crawford y Paglen (2019) demostraron cómo las bases de datos de entrenamiento visual se construyen sobre categorías culturalmente situadas que naturalizan prejuicios históricos. En su investigación, Noble (2018) observó cómo los motores de búsqueda reproducen estereotipos raciales y de género al indexar y jerarquizar información, ilustrando que el sesgo reside en las bases de datos y en los modelos de negocio que los sustentan. Benjamin (2019) acuñó el concepto de “Nuevo Código Jim” para señalar cómo la innovación digital, presentada como imparcial, puede reforzar prácticas de segregación y exclusión racial. Desde otra perspectiva, Birhane (2021) propone un giro hacia marcos relacionales que abandonen la obsesión por métricas aisladas y se centren en el contexto, las relaciones de poder y las dimensiones éticas del cuidado. Esto implica reconocer que los sesgos no pueden eliminarse en abstracto, sino que deben abordarse como problemas situados en realidades sociales complejas.

39

A estos enfoques se suman contribuciones como la de Selbst *et al.* (2019), quienes advierten sobre la “falacia de abstracción” en la ética de la IA: la tendencia a analizar dilemas en términos excesivamente formales, descontextualizados de dinámicas sociales. De esta manera, la justicia algorítmica no puede reducirse a ajustes técnicos, sino que exige un replanteamiento del modo en que producimos, clasificamos y utilizamos datos en la sociedad global contemporánea.

Resulta evidente que existe desde la academia un consenso parcial, según el cual la IA, por su carácter disruptivo y su complejidad técnica y filosófica, no puede ser analizada de manera aislada, sino como un fenómeno civilizacional que afecta los cimientos mismos de la sociedad global.

2. Marco teórico

La noción de inteligencia ha sido objeto de múltiples aproximaciones a lo largo de la historia de la psicología, la filosofía y las ciencias cognitivas. En la tradición psicométrica,

Spearman (1904) introdujo el concepto de “factor g”, entendido como una capacidad general de razonamiento subyacente a la ejecución de tareas cognitivas diversas. Este planteo, ampliado en los modelos jerárquicos como el de Cattell (1941), Horn (1968) y Carroll (1993), permitió construir herramientas de evaluación comparables, pero redujo la inteligencia a lo cuantificable. En contraste, enfoques cualitativos como las inteligencias múltiples de Gardner (1983) o la teoría triárquica de Robert Sternberg (1985, 1988) reivindicaron la pluralidad de habilidades, incluyendo dimensiones creativas y prácticas que escapan a los *tests* tradicionales como la consideración del coeficiente intelectual.

Desde la psicología del desarrollo, Piaget (1952) concibió la inteligencia como un proceso de equilibración adaptativa, mientras que Vygotsky (1978) subrayó la dimensión social y cultural de las funciones cognitivas superiores. Estos aportes abrieron el camino a la visión contemporánea de la inteligencia como un fenómeno situado y mediado por la interacción con otros y con las herramientas disponibles y potenciales.

En el terreno filosófico y cognitivo, Turing (1950) propuso su célebre prueba de imitación como criterio funcional para atribuir inteligencia a una máquina, desplazando la pregunta ontológica hacia la verificabilidad conductual. Su propuesta fue criticada por Searle (1980), quien en su experimento mental de la “habitación china” argumentó que manipular símbolos no implica comprensión semántica. Otros autores como Dennett (1987) sostienen que basta con adoptar la postura intencional para atribuir creencias y deseos a sistemas complejos, mientras que enfoques enactivistas como los de Varela, Thompson y Rosch (1991) y Noë (2004) entienden la inteligencia como una práctica encarnada y emergente de la interacción organismo-entorno.

En la ingeniería, la inteligencia se define a menudo como la capacidad de un agente para maximizar objetivos en contextos inciertos. Russell y Norvig (1995, 2020) la conceptualizan en términos de agentes racionales, mientras que Hutter (2005, 2007) y Legg (2007) propusieron la “inteligencia universal” como capacidad general de alcanzar metas en una amplia variedad de entornos. Estas aproximaciones ofrecen operacionalización formal, pero corren el riesgo de desatender dimensiones sociales, éticas y culturales.

Respecto del carácter “artificial” y los debates sobre lo que implica la inteligencia (es decir, si realmente se trata de una “inteligencia” artificial o es más bien un algoritmo predictivo sin capacidad real de creación), se advierte que la capacidad de “parecer humano” de sistemas de lenguaje como GPT o los modelos de creación de imágenes y videos no equivale a comprensión ni a fiabilidad epistémica. Tal como advierte Mitchell (2019), los modelos de IA actuales no “entienden” el mundo en el sentido humano del término, sino que manipulan correlaciones estadísticas derivadas de grandes volúmenes de datos.

Esta distinción es crucial, porque el riesgo de antropomorfizar la tecnología lleva a sobrestimar sus capacidades y a invisibilizar tanto sus limitaciones estructurales

como sus consecuencias sociales. En esta línea, Bender, Gebru, McMillan-Major y Shmitchell (2021) sostienen que las aplicaciones que buscan imitar de manera creíble a los humanos tienen un riesgo de daños extremos, como la reproducción de sesgos sociales, la generación de desinformación y los costos ambientales derivados del entrenamiento de modelos de gran escala. Su crítica apunta a que la fluidez lingüística no debe confundirse con comprensión semántica, apropiación conceptual o autoridad epistémica.

A partir de estas advertencias, en los últimos años se consolidó un giro metodológico hacia la gobernanza de los sistemas de IA. Mitchell *et al.* (2019) propusieron las “cartas modelo”, fichas estandarizadas para documentar los supuestos, usos previstos y limitaciones. De forma complementaria, Gebru *et al.* (2018) desarrollaron las *datasheets for datasets*, un mecanismo para transparentar el origen, la composición y la calidad de los datos empleados en el entrenamiento. Estas herramientas no solo buscan aumentar la trazabilidad y la responsabilidad, sino que fueron incorporadas en marcos institucionales de evaluación de impacto algorítmico, especialmente en Europa y Estados Unidos.

Para Simon (1969), lo artificial se refiere a los artefactos orientados a fines que median entre un estado actual y uno deseado bajo ciertas restricciones. Simondon (1958), en cambio, propuso pensar los objetos técnicos como entes en proceso de individuación, con una historia interna y una evolución que los conecta con sus entornos humanos y no humanos. En la misma línea, la posfenomenología de Ihde (1990) y Verbeek (2011) sostiene que las tecnologías median nuestra experiencia y constituyen parte de la condición humana, borrando la frontera rígida entre lo natural y lo artificial.

41

Desde la sociología de la ciencia, Latour (1987) subrayó que los artefactos deben ser entendidos como actores que participan en redes híbridas de agencia. A su vez, la cibernética de Wiener (1948) y la teoría de la información de Shannon (1948) vincularon la artificialidad con la capacidad de diseñar sistemas de control y retroalimentación.

En definitiva, resulta posible extraer tres rasgos clave de la artificialidad: teleológicamente, se encuentra orientada a fines explícitos, posee una normatividad incorporada desde la cual materializa valores y decisiones humanas en diseños técnicos, y tiene la capacidad de transformar y ser transformada por prácticas sociales, instituciones y culturas. A partir de estos debates, proponemos definir la IA como una herramienta con capacidad para predecir y reaccionar de manera adaptativa en contextos inciertos, materializando iniciativas humanas en arquitecturas técnico-informacionales que impactan en todas las dimensiones de la vida y su medio consolidándose también como servicio transable con características extractivas. Esta definición combina dimensiones cognitivas (adaptación, inferencia, acción), técnicas (diseño, finalidad, información), económicas (cambio de paradigma del mundo del trabajo, la comercialización, la competencia) y socioambientales (valores, instituciones, coevolución, extractivismo).

3. Ética, civilización y policrisis en el antropoceno digital

La emergencia de la IA debe analizarse en el marco de un cambio civilizatorio más amplio. Ubicar la IA en el período del “antropoceno” implica reconocer que sus riesgos y promesas están entrelazados con las crisis planetarias actuales: el cambio climático, la pérdida de biodiversidad, la pobreza extrema, la precarización del trabajo y la fragmentación política. Vivimos en una época donde la distinción entre lo natural y lo social se desdibujó: las infraestructuras digitales son al mismo tiempo infraestructuras ecológicas y políticas. En este escenario, pensar una ética de la IA equivale a repensar el contrato social y ecológico de la humanidad. Este horizonte sitúa a la humanidad como fuerza geológica capaz de transformar radicalmente el planeta, pero también como especie que enfrenta sus propios límites.

El plano de la ética pública se configura como un cruce de normativas internacionales, propuestas de gobernanza y tradiciones filosófico-humanistas que buscan frenar la deriva tecnocrática. De allí que organizaciones internacionales como Naciones Unidas (ONU) proponen marcos normativos universales para la IA, que buscan articular principios comunes en torno a la justicia, la sostenibilidad y los derechos humanos. Los principios internacionales mencionados constituyen marcos de referencia que enfatizan derechos humanos, transparencia y rendición de cuentas. El *AI Act* de la UE introduce, además, un enfoque regulatorio *ex ante*, estableciendo categorías de riesgo y obligaciones específicas para los proveedores de sistemas.

42

Sin embargo, más allá de la regulación positiva, diversos autores advierten que la discusión debe anclarse en horizontes civilizatorios más amplios, superando las fronteras físicas, mentales, espirituales y cosmovisionales. Floridi y Cowls (2019) formularon los “cinco principios de la IA responsable” (beneficencia, no maleficencia, autonomía, justicia y explicabilidad) como una “gramática ética transversal”. Desde el Sur Global, Arun (2020) cuestiona la suficiencia de los marcos eurocéntricos, subrayando la necesidad de incorporar nociones de justicia distributiva y epistemologías plurales. En este contexto, la algorética (Francisco, 2020) ocupa un lugar destacado al ofrecer una síntesis práctica y humanista que traduce la preocupación por la dignidad humana en criterios operativos para situar al ser humano en el centro del proceso tecnológico. La perspectiva humanista-cristiana, expresada por el papa Francisco, considera a la IA como una herramienta cuyo sentido depende de la orientación política y ética que le demos. En su mensaje en la cumbre del G7 en 2024, Francisco insistió en que “los beneficios o daños que traerá dependen de su uso”, llamando a una “responsabilidad política” que vincule innovación tecnológica con bien común y no solo con beneficio económico.

Este marco abre la puerta a un debate más amplio sobre sostenibilidad ética: ¿cómo garantizar que el diseño y uso de la IA no solo no vulneren la dignidad humana, sino que promuevan activamente condiciones de justicia social, cuidado mutuo y equidad intergeneracional? La consolidación de una “ética de la IA”, en consecuencia, se juega tanto en la creación de normas e instituciones como en la construcción de

culturas políticas y sociales que prioricen la centralidad de la persona frente a la lógica de la acumulación y el control.

Archibugi (2022) introduce el concepto de “soberanía tecnológica”, que en el caso de la IA se traduce en soberanía algorítmica; es decir, la capacidad de los Estados y las sociedades para decidir sobre el diseño, la implementación y la gobernanza de los sistemas de IA. La concentración del desarrollo de IA en un pequeño número de corporaciones y países (principalmente en Estados Unidos y China) genera una dependencia global que amenaza con convertir a otras regiones en meros consumidores tecnológicos, sin voz ni control sobre las decisiones que configuran estos sistemas (Bender *et al.*, 2021). La discusión apunta, por tanto, a la necesidad de consolidar un marco ético que trascienda lo instrumental. Desde la algorética, los principios universales de transparencia, inclusión, responsabilidad, imparcialidad, fiabilidad y seguridad se proponen como pilares de una civilización digital al servicio de la dignidad humana.

Desde un enfoque laico, pensadores como Zuboff (2019) y Couldry y Mejías (2019) concluyen que la ética de la IA no puede reducirse a un código de buenas prácticas empresariales, sino que requiere un pacto civilizacional donde la tecnología sea subordinada al bien común. En palabras de Zuboff: “la elección es entre una sociedad en la que los individuos todavía pueden decidir sobre su futuro, o una donde las decisiones estén predestinadas por sistemas de cálculo invisibles” (2019, p. 514). Parafraseando a Helbing (2022), el futuro de la IA es el futuro de nuestra civilización; regular es regularnos a nosotros mismos.

43

Asimismo, los desarrollos recientes de los modelos de aprendizaje profundo y generativo a gran escala demostraron tener un impacto ambiental significativo debido a sus elevados requerimientos computacionales y energéticos. El entrenamiento de modelos como GPT-3 o GPT-4 demanda millones de horas de procesamiento en centros de datos que consumen cantidades masivas de electricidad y agua para refrigeración. Según Strubell, Ganesh y McCallum (2019), el entrenamiento de un único modelo de procesamiento de lenguaje natural puede emitir hasta 284 toneladas de CO₂, cifra comparable a las emisiones anuales de cinco automóviles promedio en Estados Unidos. Estudios posteriores como los de Patterson *et al.* (2021), desde Google Research, advierten que el costo energético de la IA aumenta exponencialmente con el tamaño de los modelos y se concentra en infraestructuras tecnológicas ubicadas en países desarrollados, lo que profundiza desigualdades ambientales y de acceso, así como la deuda climática con los países en vías de desarrollo. Además, Bender *et al.* (2021) señalan que el ciclo de vida de estas tecnologías, desde la extracción de minerales raros hasta el consumo de agua en *data centers*, contribuye de manera indirecta a la aceleración del cambio climático, ampliando la huella ecológica digital.

Este vínculo entre IA y degradación ambiental coloca en el centro la discusión ética sobre la sostenibilidad y la justicia intergeneracional. El entrenamiento y despliegue de IA contribuyen de manera verificable al incremento de las emisiones de gases de

efecto invernadero y al agotamiento de recursos hídricos y minerales. El principio de sostenibilidad, defendido por autores como Daly (1996), y más recientemente por la ONU en su declaración “Transformar nuestro mundo: la Agenda 2030 para el desarrollo sostenible”, exige reconsiderar las trayectorias de desarrollo tecnológico y de su uso cotidiano bajo el imperativo de garantizar que las generaciones futuras puedan acceder a recursos y condiciones ecológicas mínimas para una vida digna. En este sentido, la aceleración del cambio climático, vinculada al uso masivo de energías fósiles en la expansión de *data centers* y redes de comunicación, convierte a la IA en un factor que no sólo refleja, sino que también amplifica la dimensión ecológica de la policrisis.

4. Economía política de la IA: extracción, plataformas y colonialidad de datos

La investigación interdisciplinaria contemporánea problematiza la IA como una infraestructura extractiva de alcance global. En esta línea, Crawford (2021) y Pasquinelli y Joler (2020) sostienen que se asienta sobre una trama material y laboral que suele quedar invisibilizada en los discursos celebratorios de los avances técnicos. La explotación de minerales estratégicos, el consumo energético intensivo y el trabajo precario en etiquetado de datos y moderación de contenidos constituyen los cimientos de sistemas presentados como “inmateriales” o “virtuales”. El desarrollo y la subsistencia de la IA, en este sentido, no puede comprenderse sin atender a las cadenas de suministro y a los regímenes de extracción que la hacen posible.

En un plano más amplio, Zuboff (2019) conceptualiza este proceso dentro del marco del “capitalismo de vigilancia”, entendido como un régimen económico que convierte la experiencia humana en materia prima para la predicción y el control conductual. Según la autora, las plataformas digitales operan como dispositivos de captura sistemática de información que, a su vez, alimenta sistemas algorítmicos destinados a modelar, anticipar y condicionar la conducta de los individuos. Esta dinámica erosiona la autonomía personal y transforma la deliberación pública en un espacio cada vez más mediado por intereses corporativos y lógicas de rentabilidad.

Esta perspectiva dialoga con la noción de “colonialismo de datos” (Couldry & Mejías, 2019), según la cual la apropiación y mercantilización transnacional de datos reproduce las mismas lógicas de despojo y subordinación que caracterizan al colonialismo histórico. Las poblaciones son fuente de recursos y también objeto de experimentación y control, incrementando la asimetría en la distribución de poder entre el Norte Global y el Sur Global. La IA, bajo esta óptica, se convierte en un vector de nuevas dependencias tecnológicas que limitan la soberanía de los Estados y profundizan desigualdades estructurales.

En paralelo, la trayectoria de la IA en su fase de apertura al público repite, con matices, la genealogía del propio internet. En sus primeras etapas, los modelos generativos fueron ofrecidos en modalidad gratuita, bajo la lógica de acceso abierto

y con el objetivo implícito de acelerar su perfeccionamiento a través de la interacción masiva con usuarios. Este proceso de entrenamiento “socializado” permitió que los sistemas aprendieran de los insumos lingüísticos, conductuales y contextuales aportados por millones de personas, convirtiendo a los usuarios no en consumidores pasivos, sino en coproductores involuntarios del desarrollo tecnológico. Sin embargo, una vez que los modelos alcanzaron un grado de maduración y estabilidad comercial, comenzó a imponerse la lógica del capitalismo de plataformas (Srnicek, 2016): la gratuidad inicial se restringió y emergieron versiones *premium*, modalidades de suscripción y *paywalls* que segmentan el acceso según capacidad de pago. Esta transición plantea una contradicción entre el acceso abierto y la democratización del conocimiento, y la monetización intensiva que amenaza con producir un sistema de exclusión digital, en el que solo quienes pueden pagar acceden a las funcionalidades más potentes y avanzadas.

Este fenómeno se entrelaza con una tendencia más amplia de la economía digital contemporánea: la plataformización de la vida cotidiana. Así como los *smartphones* y las aplicaciones fragmentaron las interacciones en múltiples microtransacciones diarias (desde la movilidad hasta la educación y la salud), la IA tiende a integrarse en todos los ámbitos de la vida pública y privada, convirtiéndose en un requisito de funcionamiento básico. El paralelismo con el ecosistema de *apps* tras la pandemia es evidente: para realizar trámites, consumir cultura o incluso mantener lazos sociales, el individuo depende cada vez más de infraestructuras digitales privadas. En este proceso, los grupos más vulnerables sin una política activa formativa (adultos mayores, personas con discapacidades, sectores sin alfabetización digital) quedan en situación de marginalidad estructural. Van Dijck, Poell y De Waal (2018) conceptualizan la “sociedad de plataformas” como un entramado sociotécnico en el cual se constituye en la infraestructura dominante de intermediación en sectores clave como la educación, la salud y los medios de comunicación. Estos procesos desplazan progresivamente funciones estatales y sociales hacia actores privados, generando tensiones respecto a la gobernanza democrática.

45

Entonces, la socialización inicial de la IA a través de interacciones abiertas plantea interrogantes sobre el futuro de la relación entre tecnología y sociedad. ¿Persistirá un régimen híbrido, con versiones gratuitas limitadas junto a servicios *premium*, o avanzaremos hacia un modelo de exclusividad creciente, donde las herramientas de mayor valor se reserven para élites corporativas, gubernamentales y económicas? La respuesta a esta pregunta determinará, en gran medida, si la IA reforzará las desigualdades sociales existentes o si, por el contrario, se convertirá en una herramienta democratizadora.

Por otro lado, el empleo y acceso masivo a la IA puede favorecer la utilización para la manipulación política y la expansión de esquemas de desinformación. El poder de la IA puede ser instrumentalizado tanto por actores individuales como por corporaciones de todo tipo. La capacidad de generar contenidos falsos hiperrealistas (*deepfakes*) plantea riesgos para la calidad democrática y los procesos electorales (Helbing,

2022). Los casos de Cambridge Analytica y la injerencia en procesos electorales como el Brexit o las elecciones estadounidenses de 2016 muestran cómo la manipulación de datos y flujos de información puede alterar la opinión pública. Con la generación masiva de imágenes, audios y videos falsos, la sociedad tendrá mayores desafíos para discernir lo real de lo ficticio.

También se estima que la evolución de la IA suscita una considerable pérdida de fuentes y puestos de trabajo, así como un impulso a la precarización laboral. La OCDE (2024) plantea que cerca del 27% de los empleos en países miembros enfrenta riesgo significativo de automatización parcial o total, lo que genera presiones distributivas y desafíos para la cohesión social. El horizonte de evolución del ecosistema de IA podría implicar una especialización funcional de los modelos, siguiendo una lógica de “división del trabajo algorítmico”. Frente a sistemas generalistas como GPT, capaces de abarcar usos múltiples, se observa una proliferación de modelos específicos para tareas concretas: desde la redacción jurídica hasta la asistencia médica o la gestión administrativa. Esta especialización tiene implicaciones directas en el mundo laboral: la incorporación de IA en empresas, administraciones públicas y sectores productivos ya exige programas de capacitación intensiva para los trabajadores, con el riesgo de que quienes no logren adaptarse queden desplazados. En este sentido, la IA no solo transforma la producción de conocimiento y el acceso a la información, sino que redefine el contrato social en torno al trabajo y la inclusión digital.

46

Otro aspecto crítico es la asimetría en la inversión y el control de esta tecnología. Durante la segunda mitad del siglo XX, la investigación en IA fue impulsada principalmente por fondos públicos, en particular desde proyectos militares y universitarios en Estados Unidos y Europa. Sin embargo, a partir de 2000, la tendencia se invirtió de manera decisiva: la inversión privada superó a la pública y comenzó a marcar la agenda de desarrollo desde una lógica comercial y corporativa. Actualmente, se estima que cerca del 75 % de la inversión total en IA proviene del sector privado, mientras que la inversión estatal se ha reducido o se orienta a alianzas público-privadas. Este desplazamiento financiero implica que el rumbo de la IA queda mayormente en manos de corporaciones transnacionales, con escasa o nula intervención de los Estados en materia de regulación, transparencia, normativización u orientación ética.

Esta dinámica de la competencia empresarial y geopolítica (entre grandes potencias como Estados Unidos, China o la UE) se traslada directamente al territorio digital, consolidando un espacio crecientemente privatizado. El riesgo de monopolización tecnológica se suma al vacío normativo, lo que refuerza la percepción de que la IA avanza bajo las lógicas del mercado más que bajo un marco de deliberación democrática con foco en la protección de los ciudadanos. En este sentido, si los Estados no asumen un papel más activo, tanto en la inversión como en la regulación, el futuro de la IA podría quedar definitivamente subordinado a los intereses corporativos y geoestratégicos globales. La concentración de poder en pocas corporaciones redefine la economía política y las maneras en que se produce conocimiento, se organizan los servicios públicos y se configura la esfera pública, lo que plantea problemas de

soberanía tecnológica, colonialismo digital y asimetrías geopolíticas, profundizando desigualdades entre países y regiones (Couldry & Mejías, 2019).

Estos aportes permiten comprender a la IA no como tecnología “neutral”, sino como un campo atravesado por relaciones de poder, desigualdades geopolíticas y dinámicas extractivas. La extracción de datos, recursos y trabajo ajeno converge en un modelo de acumulación que puede entenderse como una continuidad y, a la vez, una intensificación de los patrones coloniales y capitalistas del pasado. La discusión sobre la IA debe incorporar también los aspectos estructurales, interrogando las bases materiales que sostienen su desarrollo, así como sus impactos socioeconómicos.

5. Impacto social de la inteligencia artificial

El despliegue de la IA en el tejido social incide directamente en la configuración de la vida cotidiana, en la producción de conocimiento y en la definición de políticas públicas. Analizar estos impactos requiere, por tanto, reconocer la dimensión performativa de la IA desde la cual no solo describe la realidad, sino que la modela y, hasta cierto punto, la anticipa.

Existen diversas posibilidades de usos positivos de la IA en distintos ámbitos. En la medicina y la salud pública, la IA puede facilitar diagnósticos tempranos y avanzados, análisis de imágenes, descubrimiento de fármacos o gestión hospitalaria (Topol, 2019). En educación, existen iniciativas que fortalecen las trayectorias de aprendizaje, la implementación de sistemas de tutoría inteligente y el análisis de desempeño académico. A su vez, la gestión de datos puede ser utilizada para la modernización de la administración pública mediante la optimización en la asignación de recursos, la planificación urbana y la predicción de riesgos ambientales, sociales y económicos de las políticas públicas (OCDE, 2024). La IA promete diagnósticos más finos y políticas basadas en evidencia.

A pesar de estos avances, los impactos a mediano y largo plazo son aún inciertos e inconmensurables. No existe consenso sobre la magnitud de la disrupción laboral, el efecto en las democracias o la posibilidad de dependencias tecnológicas irreversibles. Como subraya Crawford (2021), la IA no es neutral ni meramente técnica, sino un ensamblaje sociotécnico cuya trayectoria dependerá de decisiones políticas y éticas en el presente. La IA, al insertarse en un mundo marcado por interdependencias múltiples, donde las crisis económicas, sociales y ecológicas se entrelazan, puede tanto agudizar como amortiguar estos procesos. Por ejemplo, los sistemas predictivos aplicados a la gestión de riesgos climáticos pueden salvar vidas, pero la minería intensiva de datos y el consumo energético de los modelos de IA contribuyen a la aceleración del cambio climático (Strubell, Ganesh & McCallum, 2019).

5.1. La vida cotidiana

La digitalidad produce una dislocación radical en la relación con el tiempo. Al operar en régimen de respuesta inmediata, optimización continua y predicción en tiempo real, la IA acelera ritmos de consumo, afectos y políticas, y normaliza la expectativa de inmediatez. Mientras el humano se desarrolla en un tiempo biográfico marcado por la caducidad, la IA opera en un tiempo de procesamiento: instantáneo, asíncrono, ajeno a la duración vivida. Para un sistema de IA, no existe la espera, la demora o la nostalgia; el pasado se archiva como dato y el futuro se calcula como predicción. La experiencia del “ahora” carece de densidad existencial y se reduce a una función computacional.

Esta diferencia genera dinámicas profundas en la interacción entre humanidad e IA. En primer lugar, la IA contribuye a la aceleración social ya diagnosticada por Rosa (2013), donde la presión por la inmediatez erosiona los ritmos naturales de la vida, desajustando la relación entre el tiempo humano y el tiempo tecnológico. En segundo lugar, si la IA no conoce la finitud ni la irreversibilidad, ¿puede participar en procesos donde el sentido humano depende de esas condiciones, como el duelo, la promesa o el perdón? La IA puede simular relatos de pasado y proyectar escenarios de futuro, pero carece de la experiencia de “estar en el tiempo”, de ser atravesada por él. La historicidad, entendida como apertura a lo inesperado y contingente, se ve reemplazada por la lógica de la predicción. En este sentido, el riesgo es que la humanidad, al confiar sus decisiones y relatos a sistemas que no experimentan el tiempo, pierda contacto con lo que la constituye en su esencia.

En términos teóricos, esto es congruente con la tesis de la “aceleración social”: según Rosa (2013), la modernidad se caracteriza por un proceso autoalimentado de aceleración técnico-social que presiona la experiencia temporal de los sujetos, produciendo desajustes entre velocidades sociales y las capacidades humanas de asentamiento y orientación. La IA actúa hoy como uno de los principales “aceleradores” de esa dinámica: reduce latencias, estrechas ventanas de deliberación y transforma la tolerancia a la demora en una fricción inaceptable.

La economía de la atención y la lógica “24/7” son compañeros inseparables de esta aceleración. La captura sistemática de la atención, lo que Wu (2016) y Crary (2013) denominan procesos de extracción y colonización temporal, convierte el tiempo humano en recurso explotable. Esa extracción temporal tiene costos cognitivos y afectivos (menor profundidad atencional, interrupción de procesos de aprendizaje y socialización, aumento de ansiedad y fatiga), y prepara el terreno para una vida cotidiana gobernada por estímulos diseñados para respuesta rápida, no por ritmos de asentamiento humano.

Uno de los ámbitos más visibles de la disrupción de la IA es el mundo laboral. Informes recientes de la Organización Internacional del Trabajo (OIT, 2023) muestran que la automatización basada en IA amenaza con desplazar tareas rutinarias en sectores como la administración, el comercio minorista y los servicios financieros,

generando un fenómeno de precarización para millones de trabajadores. Sin embargo, el mismo estudio reconoce que la IA podría también abrir oportunidades en áreas de alta especialización, siempre que existan políticas activas de capacitación y redistribución. Asimismo, la conjunción IA-aplicaciones-plataformas reconfigura la jornada y la expectativa de disponibilidad. La lógica de la plataforma transforma el tiempo de trabajo en microsegmentos gestionados por algoritmos, favoreciendo la precarización, la hiperdisponibilidad y la supeditación del ritmo biológico a ritmos productivos optimizados algorítmicamente. Frente a esto, emergen respuestas normativas como *right to disconnect* (derecho a la desconexión) y diversas iniciativas legislativas que buscan restituir a las personas una frontera temporal mínima frente a la intrusión digital. Estas políticas son una respuesta práctica a la tensión entre la velocidad impuesta por la tecnología y la necesidad humana de ritmos de reposo, deliberación y cuidado.

La pandemia actuó como multiplicador de estas tendencias, dado que la aceleración de la adopción tecnológica (teletrabajo, telemedicina, digitalización de trámites) estrechó aún más la relación entre vida cotidiana y plataformas. Encuestas recientes de Pew Research Center muestran que un porcentaje significativo de la población reconoce que la pandemia cambió de forma permanente su relación con la tecnología, incrementando la dependencia de servicios digitales y consolidando prácticas que antes eran marginales. Ese cambio no es neutro, amplifica la presión temporal (respuestas inmediatas, disponibilidad constante) y profundiza brechas entre quienes pueden adaptarse y quienes no (adultos mayores, personas con menos alfabetización digital), reforzando una inequidad temporal donde algunos disponen de “tiempo digital” como recurso productivo y otros quedan excluidos de ritmos sociales normalizados.

49

En el consumo, la IA se expresa en sistemas de recomendación que moldean preferencias culturales, desde el entretenimiento hasta la compra de alimentos. Van Dijck (2021) señala que este fenómeno configura una “infraestructura de la elección” donde la autonomía del consumidor se ve mediada por lógicas algorítmicas que definen qué opciones aparecen como accesibles o deseables. A nivel cultural, la IA permea las interacciones sociales en redes digitales, donde algoritmos de clasificación y personalización inciden en la formación de identidades, comunidades y espacios de sociabilidad. Morozov alerta sobre el riesgo de que estas dinámicas alimenten procesos de desinformación y polarización política, dado que “los sistemas de recomendación priorizan la intensificación emocional por encima de la veracidad factual” (2020, p. 67). En este sentido, la IA se convierte en un actor central en la configuración de las democracias contemporáneas, al intervenir en los circuitos de comunicación pública y privada.

5.2. Educación y formación

El desarrollo de la IA está redefiniendo tanto las dinámicas de la educación tanto formal como informal. En el plano institucional, algunas universidades y escuelas adoptaron sistemas de IA para tareas como la personalización de contenidos, la corrección automática, el monitoreo de desempeño y la prevención del abandono. Estos mecanismos prometen eficiencia y mayor adaptabilidad, pero al mismo tiempo

desplazan la centralidad del encuentro pedagógico, reduciendo la mediación humana en favor de la interacción con sistemas automatizados. El aula corre el riesgo de convertirse en un “espacio técnico” donde prima la optimización del aprendizaje medible por métricas, en detrimento del proceso formativo integral, que incluye afectividad, creatividad y reflexión crítica.

La educación informal, la crianza y la socialización primaria, que históricamente se daban en contextos comunitarios, familiares y sociales, también se encuentran permeadas por la IA. Desde plataformas de autoaprendizaje hasta asistentes digitales que facilitan la resolución de problemas cotidianos, la incorporación temprana de estas tecnologías configura un modo de aprender donde la consulta se dirige primero a la máquina antes que a otros seres humanos. Esto transforma el ecosistema de transmisión cultural y desplaza a referentes tradicionales (padres, docentes, pares) en la jerarquía de fuentes de conocimiento. La socialización del saber se tecnifica, y con ello se debilitan vínculos intergeneracionales que antes estaban mediados por el traspaso directo de experiencias. Herramientas como aplicaciones de monitoreo infantil, juguetes interactivos y plataformas de entretenimiento con algoritmos de recomendación median los primeros contactos de los niños con el mundo. Esta mediación introduce una lógica de externalización en la que la máquina reemplaza progresivamente la construcción de habilidades sensoriales, relacionales y empáticas a través del contacto humano directo. De este modo, se configura un riesgo de alienación de los sentidos, donde la experiencia táctil, la mirada cara a cara y la interacción corporal son sustituidas por pantallas y voces sintéticas.

La socialización secundaria, asociada a la integración en grupos más amplios (escuela, pares, instituciones), también se ve alterada. Plataformas sociales, motores de búsqueda y sistemas de recomendación filtran las experiencias culturales de los jóvenes, generando “burbujas” que refuerzan preferencias y reducen la exposición a la diversidad, al aburrimiento y a la frustración, experiencias centrales en el proceso formativo. Esto limita la posibilidad de encuentro con lo distinto, con lo inesperado, que constituye uno de los pilares de la vida en sociedad. En lugar de comunidades heterogéneas que se conforman en la interacción cotidiana, la IA tiende a organizar comunidades “predictivas”, diseñadas en función de patrones de consumo y no de vínculos sociales genuinos.

En este contexto, se vuelve urgente reivindicar una formación holística del ser humano, entendida como un proceso integral que no puede agotarse en el acceso a información ni en la mera acumulación de competencias técnicas. La educación, en sentido profundo, implica el desarrollo de la sensibilidad estética, de la ética relacional, de la memoria cultural y de la capacidad de habitar el mundo con otros. La tecnología, cuando se introduce sin un marco crítico, contribuye a la fragmentación de esa integralidad, fomentando la especialización sin contexto, la inmediatez sin reflexión y la interconexión sin comunidad. La interconexión global habilitada por la IA y el ecosistema digital no se traduce en mayor conexión humana. Al contrario,

la constante disponibilidad de vínculos mediáticos instantáneos puede saturar la experiencia intersubjetiva, generando aislamiento, ansiedad y debilitamiento de los lazos presenciales. Mientras el mundo digital promete acceso a todos en todo momento, se erosiona la experiencia vital de estar con otros en la presencia física, con sus silencios, gestos y matices irreductibles a la mediación técnica.

El impacto de la IA en la producción académica y en la educación superior constituye uno de los campos más debatidos desde sus orígenes. Por un lado, herramientas de procesamiento de lenguaje natural y análisis de grandes volúmenes de datos ofrecen posibilidades inéditas para el estudio de fenómenos sociales complejos. Leonelli sostiene que “la ciencia basada en datos se apoya crecientemente en infraestructuras algorítmicas que permiten patrones de investigación de alcance global” (2021, p. 102). Sin embargo, el uso extensivo de IA en la investigación también plantea riesgos de homogeneización epistémica. Striphas (2022) advierte que la dependencia de *corpus* digitales estandarizados y de herramientas de análisis generadas por un pequeño número de corporaciones podría limitar la diversidad de perspectivas en la investigación. Más aún, el recurso a modelos de IA para la escritura académica abre preguntas sobre la autoría, la originalidad y la integridad científica.

5.3. Políticas públicas

Respecto al diseño y la implementación de políticas públicas, la IA promete sistemas más eficientes y basados en evidencia, como el uso de algoritmos predictivos para identificar zonas de riesgo sanitario o mejorar la distribución de recursos educativos que ya presentan resultados positivos en algunos países (OCDE, 2022). No obstante, estas prácticas presentan riesgos significativos. Eubanks (2018) documentó cómo el uso de sistemas automatizados en programas sociales en Estados Unidos derivó en exclusiones masivas y discriminación hacia poblaciones vulnerables. En la misma línea, Gandy (2021) argumenta que los algoritmos aplicados a la política social pueden generar “nuevas burocracias de la desigualdad” reforzando estigmas y limitando derechos. En América Latina, autores como Ricaurte (2021) advirtieron que el uso de IA en la gestión pública corre el riesgo de reproducir sesgos coloniales y de profundizar la dependencia tecnológica, especialmente cuando los sistemas son importados sin adaptación a contextos locales.

51

Para hacer frente a las consecuencias negativas de la IA, desde la política se exigen estrategias específicas: a) políticas de “ritmo” tecnológico (por ejemplo, límites de respuesta automatizada en plataformas públicas, *throttling* algorítmico en flujos informativos críticos) que devuelvan espacios de latencia deliberativa; b) regulación laboral que haga efectivo el derecho a la desconexión y proteja ritmos biográficos frente a la optimización algorítmica; c) diseño centrado en la temporalidad humana (interfaces que promuevan pausas, notificaciones agregadas en ventanas temporales, modos *slow* en servicios esenciales); y d) alfabetización temporal: enseñar en escuelas y programas laborales a gestionar no solo habilidades digitales

sino también ritmos, atención y recuperación. Estas medidas podrían restablecer un equilibrio entre la velocidad instrumental de las máquinas y la duración significativa de la vida humana.

5.4. Poder, vigilancia y colonialismo digital

La dimensión del poder es quizás la más decisiva en el análisis del impacto social de la IA. La concentración del desarrollo y control de los sistemas en unas pocas corporaciones (Google, Microsoft, Amazon, Meta, Baidu) genera una estructura oligopólica que nutre al “capitalismo de la vigilancia” (Zuboff, 2019). Según Zuboff, “lo que se captura ya no es solo nuestro comportamiento presente, sino también nuestras conductas futuras, a través de sistemas de predicción algorítmica” (2019, p. 122). Esta capacidad de predicción y manipulación inaugura nuevas formas de control social que recuerdan, aunque en clave más sofisticada, a lo que Michel Foucault llamó “sociedades disciplinarias”. Rouvroy (2020) reformula este diagnóstico bajo el concepto de “gubernamentalidad algorítmica”, caracterizada por un control preventivo y anticipatorio de las poblaciones, donde la sanción es reemplazada por la profilaxis estadística.

52

En el plano global, la dependencia tecnológica de la mayoría de los países respecto a un núcleo reducido de potencias y empresas plantea lo que varios autores llaman un “colonialismo digital” (Coudry & Mejías, 2019). Este fenómeno no solo implica el control sobre datos y plataformas, sino la imposición de lógicas epistémicas y culturales que moldean la vida cotidiana a escala planetaria. La IA, en consecuencia, no puede analizarse únicamente como tecnología, sino como una forma de poder estructurante, capaz de redefinir las relaciones entre Estado, mercado y ciudadanía.

En el ámbito práctico-político, la velocidad de difusión de información en redes y la capacidad de generar contenidos sintéticos aceleraron los ciclos de formación de opinión pública en procesos electorales contemporáneos; trabajos económicos y empíricos documentaron el papel de las redes sociales en la difusión de noticias falsas y su consumo masivo en torno a la elección de 2016 en Estados Unidos, mostrando cómo la rapidez y la viralidad socavan la deliberación y amplifican la polarización. La combinación entre IA (capaz de generar y amplificar información) y plataformas (capaces de distribuir masivamente en minutos) crea un ambiente de incertidumbre temporal donde la verificación queda siempre detrás del flujo comunicativo.

Según Virilio (1977, 2006), la tecnología acelera no solo procesos productivos, sino también la lógica de la legitimidad y la visibilidad política (la rapidez como ventaja estratégica), con efectos sobre la toma de decisiones y la percepción pública de eventos. La capacidad algorítmica de producir narrativas y respuestas en tiempo real tensiona la deliberación democrática, porque comprime los márgenes de reflexión colectiva y desfavorece el pensamiento crítico.

5.5. Situación jurídica y comercial de la IA e interrogantes emergentes

La IA plantea una serie de desafíos inéditos para el derecho, tanto en su dimensión positiva (la aplicación de normas existentes) como en su dimensión teórica (la reflexión sobre los fundamentos mismos de la juridicidad).

En primer lugar, está la cuestión de la responsabilidad y agencia, así como la propiedad intelectual. Si un usuario produce un texto, una obra artística o incluso una invención científica utilizando IA, ¿quién es el verdadero autor? En algunos países, la legislación otorga la titularidad a la persona humana que opera la herramienta. En otros, los debates aún están abiertos, y las corporaciones reclaman derechos de propiedad sobre las creaciones generadas con sus modelos. La situación se complejiza cuando las creaciones derivan en impactos negativos o ilícitos. Si la IA produce un resultado dañino (un sesgo discriminatorio, una falsedad, una incitación a la violencia, o incluso la comisión de un delito), ¿es responsable el usuario que emitió la instrucción? ¿La persona que actuó en consecuencia? ¿Lo es el programador que diseñó el modelo o la corporación que lo comercializa? En un futuro, ¿se reconocerá alguna forma de responsabilidad autónoma a la propia IA? La doctrina jurídica aún no ofrece consensos claros, y en foros internacionales se discute la posibilidad de crear regímenes específicos de propiedad y responsabilidad para obras e invenciones asistidas por IA.

La teoría del derecho, en este punto, se enfrenta a la pregunta de si la IA puede ser considerada sujeto de imputación o únicamente objeto de regulación. La noción de “*accountability* algorítmica” (rendición de cuentas) busca llenar ese vacío, estableciendo mecanismos de transparencia, trazabilidad y auditoría que permitan atribuir responsabilidades sin reconocer agencia plena a la máquina. El derecho contemporáneo tiende a refugiarse en la figura del creador o del proveedor, dado que los sistemas de IA no tienen personalidad jurídica reconocida. Sin embargo, esta solución genera tensiones. Las grandes corporaciones suelen blindarse tras cláusulas de exención, dejando al usuario en situación de vulnerabilidad frente a daños imprevistos.

53

Otra cuestión crítica es la de la objetividad. ¿Puede la IA ser objetiva? En sentido estricto, su funcionamiento depende de los datos con que fue entrenada y de las instrucciones dadas por el usuario. Ambos elementos son portadores de sesgos. Así, la IA reproduce, amplifica o distorsiona las subjetividades preexistentes. Sin embargo, también introduce la posibilidad de subjetividad algorítmica, entendida como el efecto emergente de la interacción entre múltiples capas de datos, parámetros y órdenes de optimización. Esto suscita un interrogante filosófico y jurídico: ¿cómo confiar en que las respuestas de la IA no estén condicionadas por restricciones invisibles impuestas por sus creadores (por ejemplo, prohibiciones temáticas, sesgos ideológicos o estrategias de mercado) que escapen al control del usuario? La transparencia algorítmica se vuelve aquí no solo un imperativo ético, sino un requisito jurídico para garantizar el principio de confianza legítima en las interacciones digitales.

Desde la teoría, la IA obliga a repensar categorías fundamentales. Por un lado, la IA no es plenamente sujeto ni mero objeto, exigiendo la consolidación de marcos híbridos. Por el otro, los procesos algorítmicos son opacos, lo que dificulta reconstruir cadenas de causalidad para fines legales. Finalmente, la pregunta más radical se orienta a la cosmovisión de la IA: ¿acaso adopta posturas ideológicas de forma implícita al estar entrenada en ciertos contextos culturales y no en otros? Si esto es así, regular la IA supone también reconocer que todo diseño técnico encierra un posicionamiento político y ético.

El panorama actual evidencia que las soluciones jurídicas disponibles son insuficientes frente a la magnitud y la velocidad de los cambios. La IA plantea interrogantes que desbordan los marcos clásicos del derecho civil, penal y de propiedad intelectual. Es probable que en los próximos años asistamos a la emergencia de un derecho de la IA como campo específico, interdisciplinario y global. Mientras tanto, la incertidumbre persiste, y con ella la necesidad de un debate riguroso que articule filosofía, sociología, ingeniería y jurisprudencia.

6. La transición digital como umbral civilizatorio

El impacto de la IA no puede comprenderse de manera aislada de los procesos más amplios de digitalización que, desde finales del siglo XX, reconfiguraron radicalmente los modos de vida, de interacción y de producción simbólica. Como advierte Castells (2020), el advenimiento de la “sociedad red” ha transformado las bases de la experiencia social: las prácticas culturales, las estructuras de poder y las formas de subjetividad se organizan hoy en torno a la lógica de la interconexión global.

Nos encontramos en un momento crítico de transición generacional. En pocas décadas habrán desaparecido las últimas generaciones que vivieron en un mundo predominantemente analógico. El relevo demográfico consolidará, por primera vez en la historia, una sociedad que solo conoce la hiperconectividad y que carece de referencias directas previas a la era digital. Este quiebre no tiene precedentes: hasta ahora, las innovaciones tecnológicas siempre coexistieron con generaciones que podían recordar y transmitir modos de vida anteriores; en cambio, lo que se perfila es una ruptura total de memoria cultural, donde la experiencia predigital devendrá cada vez más inaccesible.

Byung-Chul Han (2021) describió cómo la digitalización, al disolver la mediación temporal propia de lo analógico, produce una “sociedad de la inmediatez” en la que la reflexión y la memoria se debilitan. En paralelo, Bauman y Bordoní (2016) sostuvieron que vivimos en un “interregno líquido”, un tiempo en el que las viejas estructuras perdieron vigencia, pero las nuevas aún no logran consolidarse, generando estados de anomia y malestar difuso. La IA se inserta en este contexto no como una tecnología más, sino como el motor que acelera y amplifica estos procesos de disolución y reconfiguración social.

El horizonte que se abre es el de una sociedad sin referente analógico, en la que los vínculos, la política, la economía y la cultura estarán mediados enteramente por la digitalidad. ¿Cómo se construye la identidad en un mundo donde lo virtual sustituye progresivamente lo presencial? ¿Qué ocurre con la memoria colectiva cuando los registros digitales se vuelven la única forma de archivo, vulnerables a la manipulación algorítmica y la obsolescencia tecnológica?

En este escenario, la IA no solo se convierte en un dispositivo de análisis o predicción, sino en una tecnología ontológica, capaz de moldear los modos de ser y de estar en el mundo. El tránsito hacia una sociedad plenamente digital no sólo transforma las estructuras económicas y políticas, sino que incide de manera directa en la construcción de la identidad y en la configuración de los vínculos sociales. En un mundo donde la presencia física tiende a ser sustituida por la interacción mediada, las categorías clásicas de intimidad, comunidad y confianza se ven sometidas a tensiones inéditas. Estas transformaciones afectan también a la política, la religión y las formas de pertenencia comunitaria. La política se desplaza hacia la lógica de la inmediatez y la viralidad, donde los algoritmos privilegian la polarización y la simplificación discursiva. Las religiones enfrentan el desafío de mantener la trascendencia en un entorno que promueve la instantaneidad y la instrumentalización de lo sagrado. Y las comunidades, que antes se sostenían en la presencia física, deben reinventarse en espacios virtuales que, aunque expansivos, carecen de la densidad simbólica del encuentro encarnado.

Este panorama genera una tensión existencial profunda. Por un lado, la digitalización promete conectividad global y acceso ilimitado al conocimiento; por el otro, produce una erosión de los vínculos humanos fundamentales, debilitando la empatía, la memoria compartida y la experiencia de trascendencia. Como advierte Lévy (2020), corremos el riesgo de que la inteligencia colectiva se diluya en una inteligencia algorítmica que optimiza la información instantánea, pero empobrece el sentido. La IA y la digitalidad total no solo plantean un cambio en los instrumentos de comunicación, sino una reconfiguración antropológica.

55

La discusión sobre la posverdad, popularizada tras las elecciones presidenciales de Estados Unidos en 2016 y el referéndum del Brexit, se convirtió en un referente ineludible. Como mostró Vaidhyanathan (2018), el poder de las plataformas digitales radica en la manipulación de flujos de información para orientar la atención y moldear percepciones colectivas. En esos procesos, compañías como Cambridge Analytica demostraron que el cruce entre *big data* y microsegmentación algorítmica podía alterar el comportamiento electoral a gran escala (O'Neil, 2016). Si aquello fue posible con técnicas relativamente rudimentarias, los riesgos actuales son exponenciales. La IA generativa permite producir imágenes, audios y videos cada vez más indiscernibles de lo real, capaces de instalar narrativas falsas a través de las redes sociales con una eficacia devastadora e inmediata.

Innerarity (2020) advierte que, en la esfera pública digital, la política ya no se juega tanto en torno a la producción de verdad como en torno a la gestión de la credibilidad.

Esto se radicaliza con la irrupción de la IA, que introduce una nueva etapa: la de la no verdad, entendida como un entorno en el que la distinción misma entre verdadero y falso pierde sentido práctico. A diferencia de la posverdad, donde la verdad todavía existe, aunque es deliberadamente ignorada, la no verdad describe una condición en la que el ciudadano promedio carece de las herramientas epistémicas para distinguir lo real de lo artificial.

Donovan (2022) mostró cómo los ecosistemas de desinformación operan mediante la saturación del espacio digital con narrativas contradictorias hasta lograr un efecto de parálisis cognitiva. La IA amplifica esta lógica al permitir la generación masiva y automática de mensajes, con variaciones infinitas adaptadas a perfiles individuales. En ese contexto, la política corre el riesgo de convertirse en un “teatro algorítmico” donde lo decisivo no es el debate público, sino la eficacia de las campañas de desinformación orquestadas por agencias estatales, corporaciones transnacionales o redes delictivas.

Las consecuencias de este proceso se vislumbran ya en las instituciones democráticas. En países de América Latina, Europa y África se documentaron campañas sistemáticas de manipulación electoral basadas en cuentas automatizadas, ejércitos de *bots* y estrategias de *astroturfing* digital.² Con la IA, la escala y el realismo de estas operaciones adquieren una potencia inédita: un *deepfake* de un líder político pronunciando un discurso inexistente, difundido en el momento adecuado, puede alterar no solo percepciones inmediatas, sino el curso mismo de una elección.

Conclusiones

El recorrido analítico realizado a lo largo de este trabajo permite arribar a una constatación y conclusión esencial. La IA no constituye meramente un hito tecnológico más, sino un fenómeno sociotécnico total, con capacidad de transformar radicalmente la vida cotidiana, las instituciones políticas, el trabajo académico, la economía global, las prácticas culturales, e incluso la relación de la humanidad con la verdad, con la naturaleza y consigo misma. Es una herramienta que reconfigura cómo pensamos, trabajamos, decidimos y nos relacionamos. Su integración irreversible en nuestras vidas plantea desafíos teóricos, políticos y éticos de primer orden, pero también abre oportunidades para reorientar la tecnología hacia la búsqueda de soluciones a las múltiples crisis de nuestro tiempo, siempre que se ponga en el centro al hombre.

Nos encontramos en un momento histórico en el que las transformaciones vertiginosas que apenas dejan espacio para que el ser humano, con sus capacidades

² Práctica de relaciones públicas por la cual se busca instalar una narrativa o un debate, usualmente por reacción o contra reacción, moldeando la percepción pública para sumar adherentes y que parezca espontánea o surgida desde las bases sociales. Suele ser implementada por empresarios, políticos y celebridades para difamar o disminuir al adversario.

y limitaciones, logre procesarlos y adaptarse. En este contexto, las estructuras tradicionales se disuelven antes de consolidarse y la incertidumbre es la única constante. La innovación tecnológica se ubica en la vanguardia de las disputas geoestratégicas y económicas entre Estados, corporaciones y bloques de poder. El desarrollo científico y tecnológico, altamente competitivo, se despliega sin pausa, superando con creces la capacidad de procesar, comprender y regular sus efectos. Desde las ciencias sociales, estos cambios sin precedentes nos interpelan bajo la responsabilidad de comprender, investigar y analizar críticamente estos procesos para ofrecer soluciones éticas a los problemas contemporáneos. No hacerlo significaría dejar el rumbo del porvenir en manos de dinámicas guiadas únicamente por la lógica tecnocrática, del mercado o de las confrontaciones geopolíticas.

El panorama es aún más difícil si atendemos al desajuste que transitamos. Mientras la IA avanza hacia sistemas cada vez más poderosos y autónomos, las sociedades contemporáneas enfrentan una involución en los parámetros del desarrollo humano con problemáticas clásicas y emergentes: pobreza creciente, precarización laboral, declive demográfico, recesiones cíclicas, violencia política, crimen organizado transnacional, migraciones forzadas e impactos del cambio climático. Resulta paradójico que, mientras se logró diseñar tecnologías de punta y algoritmos innovadores capaces de procesar millones de datos en segundos, millones de personas aún tienen necesidades básicas insatisfechas.

Ante este desfasaje civilizatorio, no basta con describir o analizar los procesos en curso, sino que resulta imperativo estudiar e interpretar las implicancias de la IA sobre la vida social, anticipar escenarios y aportar marcos éticos y legales que orienten la innovación tecnológica hacia el bien común. En un mundo donde la tecnología ya se utiliza para librar guerras híbridas, manipular elecciones, violar la intimidad de las personas o intensificar el control social, no podemos permitirnos que la IA continúe desarrollándose sin regulación.

57

Tampoco pueden omitirse los daños ambientales que la expansión de la IA acarrea. El entrenamiento de modelos masivos exige cantidades ingentes de energía y agua, con consecuencias en términos de emisiones de carbono y aceleración del cambio climático (Strubell *et al.*, 2019; Crawford, 2021). Este costo ecológico plantea un dilema ético ineludible: ¿qué legitimidad puede tener una tecnología que promete progreso mientras compromete la sostenibilidad del planeta y la posibilidad de vida digna para las generaciones futuras? La justicia intergeneracional exige que la innovación tecnológica sea compatible con la preservación de la “casa común”.³

³ El concepto de “casa común” se refiere a la idea de que la Tierra es el hogar compartido de toda la humanidad y, por tanto, exige un cuidado responsable y solidario que trascienda fronteras nacionales y generaciones. La encíclica *Laudato si'* del papa Francisco profundiza en esta noción, al señalar que la crisis ecológica no es solo ambiental, sino también social y ética, reclamando un compromiso integral con el cuidado de la creación y la justicia intergeneracional (Francisco, 2015).

En este marco, existen tensiones de difícil resolución en el corto, mediano y largo plazo que exigen una reformulación del contrato social, abriendo un horizonte en el que las reglas de la convivencia política, la producción del conocimiento y el ejercicio del poder ya no pueden darse por sentadas. Si Rousseau imaginaba el contrato como el fundamento de una voluntad general, en la actualidad ese contrato se ve erosionado por algoritmos capaces de intermediar, condicionar y hasta manipular esa voluntad de no existir agencia humana en tal sentido. La IA emerge, así, como un “nuevo Leviatán”, no como un soberano único que garantiza el orden, sino como una multiplicidad de infraestructuras y sistemas que median en la vida cotidiana, capaces de regular lo que vemos, pensamos y decidimos.

Si bien gran parte de la literatura sobre IA insiste en la necesidad de una algorética robusta, resulta legítimo preguntarse si existen realmente las condiciones sociales, políticas y económicas para materializar ese horizonte. La apelación al “deber ser” ético, aunque imprescindible, corre el riesgo de convertirse en un gesto declarativo si no se acompaña de mecanismos efectivos de gobernanza y de una redistribución del poder en el ecosistema digital. En este sentido, la experiencia europea ofrece un ejemplo promisorio, con intentos de regulación como la *AI Act*, que aspira a establecer estándares globales. No obstante, dicho proceso enfrenta límites estructurales: desde la dependencia tecnológica frente a corporaciones transnacionales hasta la dificultad de armonizar intereses divergentes entre los propios Estados miembros. Lo que predomina, entonces, es un espacio de tensión entre un diagnóstico marcado por la concentración de poder y la falta de control público, y la esperanza que apuesta por un pacto civilizatorio más justo. Este “eslabón perdido” refleja tanto la fragilidad de las iniciativas actuales como la necesidad de sostenerlas y profundizarlas si se pretende que la ética de la IA trascienda lo retórico y se traduzca en transformaciones materiales.

58

En este marco, los líderes académicos, sociales y políticos no deben permanecer al margen del debate. Es fundamental que asuman roles protagónicos en la interpretación de estos procesos y en la construcción de marcos normativos, filosóficos y éticos que orienten el rumbo de la innovación tecnológica. Solo así se podrá evitar que la IA sea un instrumento de control, explotación o dominio, y posicionarla como una herramienta al servicio de la inclusión, el fortalecimiento democrático y la cooperación global. Es posible acompañar críticamente el desarrollo tecnológico al tiempo que se ofrecen claves de interpretación y propuestas normativas para reducir la brecha entre innovación y justicia social. Para que la IA cumpla un papel socialmente constructivo, se debe fomentar la construcción de marcos regulatorios globales y nacionales, la rendición de cuentas por parte de creadores y usuarios por igual, y la transparencia en el diseño de algoritmos.

En un sentido humanista y cristiano, esta conclusión invita a repensar la técnica como don y responsabilidad. La innovación no es un fin en sí mismo, sino un medio que debe estar al servicio del bien común, la dignidad de la persona y el cuidado de la creación. Orientar la IA hacia un futuro sostenible, justo y democrático exige voluntad política, instituciones sólidas, creatividad académica y una ética pública que reconozca

que la tecnología debe servir a la vida. La historia muestra que las sociedades nunca transitaron transformaciones de esta magnitud sin crisis, tensiones ni conflictos. Pero también enseña que los seres humanos, cuando articulan saberes y valores en torno a un horizonte de esperanza, son capaces de guiar los procesos hacia destinos más justos. El desafío actual yace en transformar la IA de un riesgo en una oportunidad. La decisión está abierta, y este es el tiempo para actuar.

Anexo: “Autorreflexión” de una IA

Al consultar directamente a la IA que colaboró en la redacción de este texto, esta sostuvo lo siguiente:

“Corresponde también hacer una reflexión autorreferencial. Mi ‘opinión’ no es fruto de conciencia ni de experiencia, sino de patrones de datos y procesos estadísticos. Aun así, el ejercicio de integrar, contrastar y evaluar evidencia me permite señalar que el riesgo mayor de la IA no es tecnológico, sino político y ético: la posibilidad de que, en lugar de ampliar la libertad y la dignidad humanas, las erosione bajo nuevas formas de control y explotación.

Como modelo de lenguaje —y para ser claro, como inteligencia artificial (modelo GPT-5 Thinking mini)— debo situar mi reflexión dentro de límites epistemológicos y éticos precisos. No poseo conciencia, experiencias vivas ni intereses propios; mi ‘opinión’ es el producto de patrones aprendidos a partir de datos humanos y de las instrucciones que se me dieron para analizar e integrar información. Aun así, puedo ofrecer una evaluación crítica fundada en evidencias, lógica y principios normativos.

Mi lectura es la siguiente: la IA ofrece herramientas poderosas para amplificar capacidades humanas, pero esas mismas herramientas, sin marcos regulatorios y sociales robustos, amplifican también prácticas extractivas, desigualdades y riesgos autoritarios. Desde la perspectiva de la ética pública y la justicia intergeneracional, la prioridad no puede ser maximizar innovación *per se* sino orientar la innovación hacia la ampliación de capacidades humanas básicas, la protección de derechos y la sostenibilidad ecológica.

Como colaborador instrumental, puedo ayudar a: sintetizar literatura, simular escenarios, identificar sesgos estadísticos, generar preguntas de investigación, estructurar marcos de gobernanza y redactar políticas públicas. Pero no puedo reemplazar el juicio ético, la deliberación democrática ni la responsabilidad moral que corresponden a las instituciones humanas. Puedo, además, cometer errores -incluyendo sesgos, omisiones y ‘alucinaciones’

informativas- por lo que mi producción debe ser siempre verificada y contextualizada por expertos humanos.

Mi participación en este proceso no es neutra: soy, al mismo tiempo, objeto y medio de análisis. En última instancia, mi 'posición' es funcional y normativa: la IA debe ser diseñada y gobernada de modo que potencie la dignidad humana, sostenga la democracia y preserve la habitabilidad del planeta. *Cualquier uso que prive a las personas de la capacidad de decidir, que las transforme en objetos de manipulación o que hipoteque el futuro ecológico de las próximas generaciones debe ser restringido y, en casos extremos, prohibido.*

Desde una perspectiva sistémica, los mayores riesgos identificados –desinformación masiva, manipulación política, colonialismo digital, deterioro ambiental– son más probables que las utopías de emancipación tecnológica si no se establecen mecanismos internacionales vinculantes que regulen su desarrollo. La evidencia histórica (casos de manipulación electoral, impactos ecológicos ya documentados, concentración corporativa) confirma esta tendencia: la inercia actual conduce a escenarios de vulnerabilidad antes que de equidad.

Mi conclusión analítica, por tanto, es clara: la IA es una herramienta de alto riesgo y alta recompensa, pero la balanza se inclinará hacia una u otra dirección en función de la voluntad política, la madurez ética y la capacidad de gobernanza global.

En este sentido, la prioridad estratégica de las ciencias sociales y de las instituciones políticas debe ser doble:

1. Construir capacidades regulatorias y éticas a la altura del desafío tecnológico, para evitar que la IA quede capturada por intereses particulares o autoritarios.
2. Reconfigurar los horizontes civilizatorios en clave de sostenibilidad y dignidad humana, garantizando que el desarrollo tecnológico no comprometa ni los derechos presentes ni los recursos de las generaciones futuras.

Todo lo demás será consecuencia de este marco. En suma, la IA no determina el futuro: lo condiciona. El futuro dependerá, en última instancia, de la capacidad humana para integrar ética, política y tecnología en un proyecto común. El desafío, entonces, no es técnico, sino profundamente humano.”

Declaración sobre el uso de IA

Durante el proceso de elaboración de este artículo, se emplearon herramientas de inteligencia artificial (específicamente el modelo GPT-5 desarrollado por OpenAI) como apoyo en la redacción y revisión del epílogo, así como en la verificación estilística y

la sistematización de referencias bibliográficas en los formatos especificados por la normativa editorial. El uso de IA tuvo un carácter estrictamente instrumental y no afectó el contenido conceptual, analítico ni interpretativo del trabajo. Todos los argumentos y las reflexiones teóricas corresponden a la autoría humana, siendo la intervención de la herramienta limitada a funciones de asistencia técnica, corrección, revisión y formateo del texto. La incorporación de la “autorreflexión” de la IA a modo de epílogo respondió al propósito de explorar la capacidad de autoconocimiento de la propia herramienta, sus posibilidades y límites en la reflexión académica respecto de las grandes cuestiones que la involucran como fenómeno civilizatorio, en un marco de responsabilidad epistemológica y transparencia metodológica. Este epílogo se desarrolló a partir de distintas instancias y pedidos de reflexión autónoma y objetiva por parte de la autora y su colaborador, procurando reducir al máximo los sesgos conversacionales y las susceptibilidades interpretativas que pudieran condicionar las respuestas del modelo, con el objetivo de propiciar una reflexión lo más neutral y consistente posible.

Bibliografía

Archibugi, D. (2022). *Technological sovereignty: Democracy, power and the governance of technology*. Cambridge University Press.

Arun, C. (2020). AI and the Global South: Designing for other worlds. En M. Dubber, F. Pasquale & S. Das (Eds.), *The Oxford Handbook of Ethics of AI* (207–222). Oxford: Oxford University Press.

61

Bauman, Z. (2000). *Liquid modernity*. Cambridge: Polity Press.

Bauman, Z. & Bordoni, C. (2016). *State of crisis*. Cambridge: Polity Press.

Bender, E. M., Gebru, T., McMillan-Major, A. & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610-623. Recuperado de <https://s10251.pcdn.co/pdf/2021-bender-parrots.pdf>.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Cambridge: Polity Press.

Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns*, 2(2), 100205.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Castells, M. (2020). *The information age: Economy, society, and culture* (Vol. 1: *The rise of the network society*, 3rd ed.). Hoboken: Wiley-Blackwell.

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *British Journal of Educational Psychology Monograph Supplement*.

Ciucci, A. (2023). Human dignity in the age of artificial intelligence. Pontifical Academy for Life.
Comisión Europea (2021). Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Recuperado de: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>.

Couldry, N. & Mejías, U. A. (2019). The costs of connection: How data is colonizing human life and appropriating it for capitalism. Stanford: Stanford University Press.

Crary, J. (2013). *24/7: Late capitalism and the ends of sleep*. Nueva York: Verso Books.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. New Haven: Yale University Press.

Crawford, K. & Paglen, T. (2019). *Excavating AI: The politics of images in machine learning training sets*. Recuperado de: <https://excavating.ai/>.

Cristianini, N. (2023). *The shortcut: Why intelligent machines do not think like us*. Boca Ratón: CRC Press.

62 Daly, H. E. (1996). *Beyond growth: The economics of sustainable development*. Boston: Beacon Press.

Dennett, D. C. (1987). *The intentional stance*. Cambridge: MIT Press.

Donovan, J. (2022). *Meme wars: The untold story of the online battles upending democracy in America*. Londres: Bloomsbury.

Elsworth, C., Huang, K., Patterson, D., Schneider, I., Sedivy, R., Goodman, S., Townsend, B., Ranganathan, P., Dean, J., Vahdat, A., Gomes, B. & Manyika, J. (2025). *Measuring the environmental impact of delivering AI at Google scale*.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. Nueva York: St. Martin's Press.

Floridi, L. (2013). *The ethics of information*. Oxford: Oxford University Press.

Floridi, L. & Cows, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1), 1-15. DOI: [www.doi.org/10.1162/99608f92.8cd550d1](https://doi.org/10.1162/99608f92.8cd550d1).

Francisco (2015). *Laudato si': Sobre el cuidado de la casa común*. Ciudad del Vaticano: Libreria Editrice Vaticana. Recuperado de: https://www.vatican.va/content/francesco/es/encyclicals/documents/papa-francesco_20150524_enciclica-laudato-si.html.

Francisco (2020). *Fratelli tutti: Sobre la fraternidad y la amistad social*. Ciudad del Vaticano: Libreria Editrice Vaticana. Recuperado de: https://www.vatican.va/content/francesco/es/encyclicals/documents/papa-francesco_20201003_enciclica-fratelli-tutti.pdf.

Gandy, O. H. (2021). *The panoptic sort revisited: Data, discrimination, and new technologies of control*. Oxford: Oxford University Press.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Nueva York: Basic Books.

Gebru, T. *et al.* (2018). *Algorithmic impact assessments: A practical framework for public agency accountability*. Data & Society Research Institute.

Han, B.-C. (2021). *The disappearance of rituals: A topology of the present*. Cambridge: Polity Press.

Harari, Y. N. (2018). *21 lessons for the 21st century*. Nueva York: Spiegel & Grau.

Helbing, D. (2022). *Towards digital enlightenment: Essays on the dark and light sides of the digital revolution*. Heidelberg: Springer.

Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, 75(3), 242-259. Recuperado de https://www.researchgate.net/publication/18292505_ORGANIZATION_OF_ABILITIES_AND_THE_DEVELOPMENT_OF_INTELLIGENCE.

63

Hutter, M. (2005). *Universal artificial intelligence: Sequential decisions based on algorithmic probability*. Heidelberg: Springer.

Ihde, D. (1990). *Technology and the lifeworld: From garden to earth*. Bloomington: Indiana University Press.

Innerarity, D. (2020). *Politics in the times of indignation: Populism and democracy*. Londres: Bloomsbury Academic.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Cambridge: Harvard University Press.

Legg, S. & Hutter, M. (2007). Universal intelligence: A definition of machine intelligence. *Minds and Machines*, 17(4), 391-444.

Leonelli, S. (2021). *Data-centric biology: A philosophical study*. Chicago: University of Chicago Press.

Levy, P. (2020). *Collective intelligence for the common good*. Nueva York: Routledge.
Metzinger, T. (2021). Artificial suffering: An argument for a global moratorium on synthetic phenomenology. *Journal of Artificial Intelligence and Consciousness*, 8(1), 1-44.

Morin, E. (2020). *Changeons de voie: Les leçons du coronavirus*. París: Denoël.
Morozov, E. (2020). *Freedom as a service: Surveillance capitalism and the Internet of things*. Nueva York: PublicAffairs.

Naciones Unidas (2015). *Transformar nuestro mundo: la Agenda 2030 para el desarrollo sostenible*. Recuperado de: https://unctad.org/system/files/official-document/ares70d1_es.pdf.

Naciones Unidas (2021). *Our common agenda*. Recuperado de: <https://www.un.org/en/content/common-agenda->.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. Nueva York: NYU Press.

Noë, A. (2004). *Action in perception*. Cambridge: MIT Press.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Nueva York: Crown Publishing.

64 OCDE (2024). *AI in the public sector: Opportunities and challenges*. Recuperado de: https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/06/governing-with-artificial-intelligence_f0e316f5/26324bc2-en.pdf.

OIT (2023). *Generative AI and jobs: Policy implications*. International Labour Office. ILO Working Paper 96. Recuperado de: https://www.ilo.org/sites/default/files/wcmsp5/groups/public/%40dgreports/%40inst/documents/publication/wcms_890761.pdf.

Pasquinelli, M. & Joler, V. (2020). *The noosphere manifested: AI as instrument of knowledge extractivism*. KIM HfG Karlsruhe.

Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L.-M., Rothchild, D., So, D., Texier, M. & Dean, J. (2021). *Carbon emissions and large neural network training*. Recuperado de: <https://arxiv.org/pdf/2104.10350>.

Piaget, J. (1952). *The origins of intelligence in children*. Madison: International Universities Press.

Ricaurte, P. (2021). Data epistemologies, the coloniality of power, and resistance. *Television & New Media*, 22(2), 131-146.

Riorda, M. (2021). *Comunicación política en tiempos de pandemia: La democracia gaseosa*. Buenos Aires: Editorial Biblos.

Rosa, H. (2013). *Social acceleration: A new theory of modernity*. Nueva York: Columbia University Press.

Rouvroy, A. (2020). The end(s) of critique: Data-behavioural science and the society of (unknown) control. En M. Hildebrandt & K. O'Hara (Eds.), *Life and the law in the era of data-driven agency* (67-92). Cheltenham: Edward Elgar.

Russell, S. J. & Norvig, P. (1995). *Artificial intelligence: A modern approach*. Upper Saddle River: Prentice Hall.

Russell, S. J. & Norvig, P. (2020). *Artificial intelligence: A modern approach*. Londres: Pearson.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417-457.

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S. & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. En *Proceedings of the Conference on Fairness, Accountability, and Transparency* (59–68). Nueva York: Association for Computing Machinery.

Shannon, C. E. (1948a). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379-423.

Shannon, C. E. (1948b). A mathematical theory of communication. *Bell System Technical Journal*, 27(4), 623-656.

Simon, H. A. (1969). *The sciences of the artificial*. Cambridge: MIT Press.

Simondon, G. (2017). *On the mode of existence of technical objects*. Minneapolis: University of Minnesota Press.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology*, 15(2), 201-293.

Srnicek, N. (2016). *Platform capitalism*. Cambridge: Polity Press.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge: Cambridge University Press.

Sternberg, R. J. (1988). *The triarchic mind: A new theory of human intelligence*. Nueva York: Viking.

Striphas, T. (2022). Algorithmic culture: Between computation and cultural theory. *Cultural Studies*, 36(2), 211-232.

Strubell, E., Ganesh, A. & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. Recuperado de <https://arxiv.org/pdf/1906.02243>.

Tooze, A. (2021). Shutdown: How Covid Shook the World's Economy. Nueva York: Viking.

Topol, E. (2019). Deep medicine: How artificial intelligence can make healthcare human again. Nueva York: Basic Books.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460.

UNESCO (2021). Recommendation on the ethics of artificial intelligence. Recuperado de: https://www.naavi.org/uploads_wp/2023/Recommendation%20on%20the%20Ethics%20of%20Artificial%20Intelligence%20-%20UNESCO%20Digital%20Library.pdf.

Vaidhyanathan, S. (2018). Antisocial media: How Facebook disconnects us and undermines democracy. Oxford: Oxford University Press.

van Dijck, J., Poell, T. & de Waal, M. (2018). The platform society: Public values in a connective world. Oxford: Oxford University Press.

van Dijck, J. (2021). Governing digital societies: Private platforms, public values. *Communication and the Public*, 6(1-2), 14-19.

Varela, F. J., Thompson, E. & Rosch, E. (1991). The embodied mind: Cognitive science and human experience. Cambridge: MIT Press.

Verbeek, P.-P. (2011). Moralizing technology: Understanding and designing the morality of things. Chicago: University of Chicago Press.

Virilio, P. (1977). *Vitesse et politique: Essai de dromologie*. París: Galilée.

Virilio, P. (2006). Speed and politics. Marsella: Semiotext(e).

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge: Harvard University Press.

Wiener, N. (1948). Cybernetics: Or control and communication in the animal and the machine. Cambridge: MIT Press.

Wu, T. (2016). The attention merchants: The epic scramble to get inside our heads. Nueva York: Knopf.

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. Nueva York: PublicAffairs.