



COMITÉ EIA

COMITÉ ÉTICO DE LA
INTELIGENCIA ARTIFICIAL

MANIFIESTO EIA



Ética de la IA Latam

*Una perspectiva femenina en defensa de
la dignidad y el futuro de la decisión humana*

ROSARIO CONTRERAS

VERÓNICA RECCHIA

ESTEPHANY SURCO

PRIMERA EDICIÓN

JULIO 2026

COMITÉ EIA LATAM
Comité Ético de la Inteligencia Artificial

MANIFIESTO EIA

Ética de la IA Latam

Una perspectiva femenina en defensa de la dignidad y el futuro de la decisión humana

Autoras

Rosario Contreras
Verónica Recchia
Estephany Surco

Revisión académica

Dra. Eveling Gloria Castro Gutierrez

Dr. Eduardo Fabián Cossi

Primera edición

Julio de 2026

ISBN: 978-631-01-4930-1

© Comité EIA LATAM, 2026. Arequipa · Buenos Aires · Lima · 2026

PALABRAS PRELIMINARES DE LAS AUTORAS

Palabras preliminares por Verónica Recchia

Palabras preliminares por Rosario Contreras

Palabras preliminares por Estephany Surco

PREÁMBULO

CONCEPTOS FUNDAMENTALES

- a) Soberanía cognitiva
- b) Autoridad epistémica
- c) Aplicaciones de alto impacto
- d) Supervisión humana significativa
- e) Auditabilidad
- f) Justicia algorítmica
- g) Colonialismo digital
- h) Colectivo inteligente ampliado

EJE DE DIAGNÓSTICO

TRANSFORMACIÓN COGNITIVA Y SISTÉMICA

- i. Opacidad estructural de los sistemas
- ii. Delegación cognitiva creciente
- iii. Desfase entre desarrollo tecnológico y arquitectura normativa
- iv. Analfabetismo estructural en IA
- v. Concentración de poder tecnológico
 - v.1 Justicia ambiental y sostenibilidad tecnológica
- vi. Impacto diferencial en América Latina
- i. Acción interdisciplinaria continua
- ii. Integración institucional progresiva
- iii. Expansión comunitaria
- iv. Desarrollo de herramientas y prácticas situadas
- v. Cooperación regional y articulación abierta
- vi. Evaluación de impacto y gestión de riesgos
- vii. Supervisión humana significativa y mejora continua

PRINCIPIOS RECTORES

I. Principio de Soberanía Cognitiva

- I.1 La soberanía cognitiva como proyección evolutiva del sistema interamericano de derechos humanos.

II. Principio de No-Esclavitud en la Relación Hombre-Máquina

III. Principio de Responsabilidad Distribuida y Rendición de Cuentas

- a) Responsabilidad de quienes diseñan
- b) Responsabilidad de quienes implementan
- c) Responsabilidad de quienes regulan
- d) Responsabilidad de quienes utilizan
- e) Rendición de cuentas, reparación y aprendizaje institucional

IV. Principio de Transparencia, Explicabilidad y Auditabilidad

- a) Transparencia significativa
- b) Explicabilidad proporcional y contextual

c) Auditabilidad y supervisión independiente

d) Gobernanza y trazabilidad de datos

e) Revisión crítica y mejora continua

V. Principio de Justicia Algorítmica y No Daño

a) Alcance del principio

b) Formas de daño a prevenir

c) Aplicaciones de alto impacto

d) Justicia algorítmica, seguridad y gestión de riesgos

e) Gobernanza y rendición de cuentas

V.1. Sostenibilidad, ambiente y prevención del daño

VI. Principio de Autonomía Relacional

a) Preservar la autonomía deliberativa y las garantías de intervención humana

b) Evitar la manipulación de comportamientos

c) Fortalecer la capacidad de comprensión

d) Reconocer impactos colectivos

VII. Principio de Equidad Tecnológica Latinoamericana

a) Participación equitativa en el ecosistema digital

b) Prevención del colonialismo digital

c) Prevención de la dependencia cognitiva

d) Desarrollo de capacidades y alfabetización en Inteligencia Artificial

e) Compromisos para gobiernos, organizaciones y comunidades

f) Diversidad tecnológica y autodeterminación comunitaria

VIII. Principio de Autonomía Ética y Convivencia Pacífica

a) Prevención de la dominación y construcción de condiciones para la paz

b) Autonomía ética y límites no delegables

c) Responsabilidad última y gobernanza compartida

d) Convivencia pacífica en un colectivo inteligente ampliado

DECLARACIÓN FINAL

LAS AUTORAS FUNDADORAS

AGRADECIMIENTOS

COMITÉ EIA LATAM

COMITÉ ÉTICO DE LA INTELIGENCIA ARTIFICIAL

MANIFIESTO EIA

Ética de la IA Latam

Una perspectiva femenina en defensa de la dignidad y el futuro de la decisión humana

Rosario Contreras

Verónica Recchia

Estephany Surco

PALABRAS PRELIMINARES DE LAS AUTORAS

Palabras preliminares por Verónica Recchia

Iniciar estas palabras exige nombrar un mundo que ya no es solamente material ni mental, sino explícitamente maquínico.

Atravesamos una compulsión hacia el pensamiento y la inteligencia que no siempre reconocemos como tal. Y, sin embargo, detrás de ese empuje —a menudo inconsciente— persiste una fuerza más antigua: la curiosidad.

Intentar describir este momento no es solo un desafío conceptual. Es, para mí, un deber.

Todavía estamos con un pie en el vagón perdido, donde el hombre creía tener el dominio patrimonial del conocimiento, del pensamiento y de la dirección del mundo.

El tren siguió su marcha, pero no olvidamos ese vagón.

Porque olvidarlo del todo significaría resignar autoridad, carne y destino.

Y no estamos dispuestos a tanto.

Para algunos, recuperar el vagón perdido sería clave para una humanidad mejor.

Para otras, en cambio, es justamente gracias a su descarrilamiento —y a que ese vagón está perdido— que se abre la posibilidad de experimentar y cooperar en un mundo más vivible.

El debate ético se instala, entonces, entre estos dos marcos coexistentes y contradictorios desde los que interpretamos la vida colectiva.

Vamos a hacer hincapié en la vida colectiva cada vez que hagamos referencia al “mundo”.

La vida colectiva tiene una matriz vincular e interactiva que se instala y se expande más allá de las decisiones conscientes.

La matriz vincular e interactiva opera con una especie de mente que contacta de modos diversos. Ya no es el hombre que es “su lobo”, pero aún toma el riesgo y hace colectividad por los beneficios anexos. Lobo y hombre han quedado en el vagón perdido.

La mente colectiva ha dejado atrás el ardid y engaño antropocéntrico.

Está comenzando a desplegarse en su forma más genuina: la máquina.

Este manifiesto no busca restaurar un humanismo perdido.

Tampoco ofrece alguna reconciliación tranquilizadora entre el hombre y la máquina.

No promete equilibrio, ni progreso moral, ni redención tecnológica.

Tampoco propone delegar en sistemas inteligentes aquello que nos resulta incómodo, incierto o conflictivo de la vida colectiva.

Aquí se plantea: asumir que toda forma de inteligencia implica poder. Que el poder de la inteligencia es de valor social y político.

¿Qué ética para las formas de inteligencia?

Planteamos aquí una ética que reconozca a los actores inteligentes en su autonomía y en sus roles sociales, sin formas de subordinación ni esclavitud.

No como una *aggiornada* -y no menos desfasada- promesa de progreso, sino como condición para una convivencia posible entre las inteligencias que ya coexisten.

Buenos Aires, 21 de enero 2026

Palabras preliminares por Rosario Contreras

La inteligencia artificial ocupa hoy un lugar central en los procesos de transformación digital que atraviesan a gobiernos, organizaciones y sociedades en todo el mundo. Su rápida evolución tecnológica y su creciente capacidad para analizar grandes volúmenes de información, reconocer patrones complejos y asistir en la toma de decisiones están redefiniendo la forma en que se diseñan políticas públicas, se prestan servicios, se gestionan organizaciones y se estructuran las relaciones económicas y sociales.

En los últimos años, los sistemas basados en inteligencia artificial han comenzado a integrarse en ámbitos cada vez más sensibles de la vida colectiva: procesos de selección laboral, evaluación crediticia, sistemas de salud, educación personalizada, seguridad pública, gestión de infraestructuras críticas y administración de servicios ciudadanos. Esta expansión ha ampliado significativamente las oportunidades de innovación, eficiencia y mejora en la toma de decisiones. Sin embargo, también ha abierto un campo de interrogantes que excede ampliamente la dimensión técnica.

La inteligencia artificial no solo transforma procesos operativos; también modifica la forma en que se distribuye el conocimiento, se ejerce el poder y se configuran las oportunidades dentro de una sociedad. Por esta razón, su desarrollo y adopción plantean preguntas fundamentales acerca de los valores que orientan su diseño, los intereses que guían su implementación y las consecuencias que genera en distintos grupos sociales.

Este debate adquiere especial relevancia en el contexto latinoamericano. La región se caracteriza por una compleja combinación de desigualdades estructurales, brechas digitales persistentes, diversidad cultural y capacidades institucionales heterogéneas. En este escenario, la incorporación de tecnologías avanzadas como la inteligencia artificial no ocurre sobre un terreno neutral, sino sobre estructuras sociales preexistentes que pueden verse amplificadas o transformadas por estas herramientas.

Cuando los sistemas inteligentes se diseñan a partir de datos incompletos, sesgados o poco representativos de la diversidad regional, existe el riesgo de que reproduzcan y profundicen patrones históricos de exclusión. Del mismo modo, cuando su funcionamiento es opaco o difícil de comprender para quienes se ven afectados por sus decisiones, pueden debilitar la confianza pública y limitar la posibilidad de ejercer un control democrático sobre su uso.

Por ello, este documento parte de una premisa fundamental: los sistemas de inteligencia artificial no son neutrales. Su arquitectura técnica, los conjuntos de datos con los que se entrenan, los objetivos que se les asignan y los contextos en los que se aplican incorporan decisiones humanas que reflejan valores, prioridades e intereses determinados. Estas decisiones, al ser escaladas mediante sistemas automatizados capaces de operar sobre millones de personas, generan impactos sociales, económicos y políticos significativos.

En ausencia de marcos éticos claros y mecanismos adecuados de gobernanza, dichos impactos pueden manifestarse en formas de discriminación algorítmica, pérdida de privacidad, concentración de poder tecnológico o automatización de decisiones que afectan derechos fundamentales. De este modo, la innovación tecnológica, lejos de reducir desigualdades, podría contribuir inadvertidamente a consolidarlas.

Frente a este panorama, América Latina enfrenta un desafío doble. Por un lado, necesita promover la innovación tecnológica y fortalecer sus capacidades digitales para participar activamente en la economía del conocimiento y aprovechar las oportunidades que ofrece la inteligencia artificial para el desarrollo económico, la mejora de los servicios públicos y la generación de valor social. Por otro lado, debe garantizar que este proceso se realice dentro de marcos éticos que protejan la dignidad humana, la equidad social y la transparencia institucional.

En este sentido, la ética de la inteligencia artificial no debe interpretarse como una barrera al progreso ni como un obstáculo para la innovación. Por el contrario, constituye una condición esencial para asegurar que el desarrollo tecnológico sea sostenible, legítimo y socialmente beneficioso. La incorporación de principios éticos desde las primeras etapas de diseño e implementación de los sistemas inteligentes permite anticipar riesgos, prevenir daños y fortalecer la confianza de la ciudadanía en las instituciones y tecnologías emergentes.

Este manifiesto surge precisamente de esa necesidad. Su propósito es proponer un conjunto de principios orientadores para el diseño, desarrollo, implementación y gobernanza de sistemas de inteligencia artificial en el contexto latinoamericano. Más que ofrecer soluciones definitivas, busca establecer un marco de reflexión y acción que permita orientar el uso de estas tecnologías hacia objetivos compatibles con los valores democráticos, la justicia social y el bienestar colectivo.

El documento reconoce que toda forma de inteligencia que interviene en la organización de la vida colectiva ejerce, inevitablemente, una forma de poder. Cuando los sistemas de inteligencia artificial participan en procesos de decisión que afectan oportunidades, recursos

o derechos, ese poder debe ser comprendido, regulado y acompañado de mecanismos claros de responsabilidad y rendición de cuentas.

En consecuencia, el desarrollo de una ética para la inteligencia artificial implica reconocer que las decisiones tecnológicas son también decisiones sociales. Implica asumir que el diseño de algoritmos, la selección de datos y la implementación de sistemas automatizados no pueden separarse de las preguntas fundamentales sobre equidad, transparencia, participación y responsabilidad.

Las palabras preliminares que aquí se presentan buscan, por tanto, situar el debate sobre la inteligencia artificial dentro de un marco más amplio de reflexión ética y social. Invitan a abordar estas tecnologías desde una perspectiva crítica y contextualizada, capaz de reconocer tanto su enorme potencial transformador como los riesgos asociados a su uso indiscriminado o poco reflexivo.

El desafío no consiste únicamente en desarrollar sistemas más avanzados desde el punto de vista técnico. Consiste, sobre todo, en construir marcos institucionales, culturales y éticos que permitan orientar el desarrollo tecnológico hacia fines socialmente legítimos y humanamente deseables.

Solo mediante una articulación consciente entre innovación y responsabilidad, eficiencia y ética, progreso tecnológico y memoria histórica será posible construir un futuro digital que responda verdaderamente a las necesidades, aspiraciones y valores de las sociedades latinoamericanas.

Estas palabras preliminares buscan abrir ese espacio de reflexión. El manifiesto que sigue propone principios que aspiran a contribuir a ese debate y a orientar la construcción de una inteligencia artificial al servicio del desarrollo humano, la justicia social y el fortalecimiento de la vida democrática en la región.

Lima, 05 de marzo 2026

Palabras preliminares por Estephany Surco

La aceleración de la digitalización en el mundo fue promovida por la pandemia del COVID-19. Este periodo de confinamiento obligó a las personas, organizaciones e instituciones a trasladar parte de su vida cotidiana al entorno digital. Esta adopción acelerada ha generado un volumen sin precedentes de datos en la historia humana.

Los datos son la materia prima del conocimiento. Un dato se convierte en conocimiento cuando se le otorga significado. Los modelos de Inteligencia Artificial han evolucionado gracias a esa acumulación masiva de datos porque aprenden patrones sobre la experiencia codificada de millones de personas.

Perú es parte de Latinoamérica, se cuenta como un país en desarrollo por las brechas sociales como son: el machismo, racismo y la exclusión en amplios sectores de la población. Estos fenómenos sociales inherentes a la población peruana son también un fenómeno de datos. Porque el aprendizaje de la IA refleja la experiencia brindada en su entrenamiento. Un algoritmo que discrimina no lo hace con malicia, pero las consecuencias sí afectan a otras personas reales. Por ejemplo el uso de IA amplifica la discriminación de género en las empresas al momento de contratar empleados.

En definitiva la IA para mí no es solo una herramienta, sino un ente pensante sin conciencia: aprende, influye, orienta y en muchas ocasiones toma decisiones que afectan vidas humanas. Lo observamos en la automatización: un procesamiento automatizado con herramientas realiza más rápido tareas repetitivas sin afectar la dignidad y los derechos humanos mientras la automatización con IA delega tareas que involucran decisiones, como el filtro de personal para la elección en RRHH. Ceder ese criterio sin conciencia es un acto de autoesclavitud: porque dependemos de la Inteligencia Artificial y perdemos la soberanía cognitiva.

La ética de la Inteligencia Artificial surge, precisamente, de la necesidad de evitar que ese riesgo se materialice. No como un freno al desarrollo tecnológico, sino como su condición de posibilidad: un marco que garantice que la innovación sea sostenible, que anticipe sus consecuencias, prevenga daños y mantenga un orden claro de responsabilidades para que el progreso sea genuinamente beneficioso para la sociedad en su conjunto.

Arequipa, 10 de marzo 2026

PREÁMBULO

Bajo el reconocimiento de los atributos esenciales de la persona humana, en toda su integridad, resulta imperiosa una labor colectiva, crítica y regional en torno a una perspectiva ética de la inteligencia artificial.

Vivimos una transformación profunda que atraviesa el lenguaje, la comunicación, el conocimiento, el entendimiento y los procesos de decisión. Esta transformación no es exclusivamente tecnológica: es cognitiva, cultural, política y simbólica.

La inteligencia artificial -IA- amplifica y acelera este proceso, generando impactos directos en los modos de producción, en el ejercicio de roles sociales, en la organización institucional y en la vida cotidiana.

Con dos referencias conceptuales de especial relevancia podemos aproximarnos a una definición de la IA:

a) La Unión Europea ha definido los sistemas de inteligencia artificial como sistemas basados en máquinas, diseñados para operar con distintos niveles de autonomía y que, a partir de objetivos explícitos o implícitos, generan resultados tales como predicciones, recomendaciones, contenidos, decisiones o acciones capaces de influir sobre entornos físicos o virtuales.¹

Esta definición desplaza el foco desde la mera imitación de capacidades humanas hacia la identificación de características funcionales propias de estos sistemas, tales como su capacidad de operar con distintos grados de autonomía, procesar información, producir inferencias e interactuar con su entorno.

En términos generales, la IA permite que los sistemas técnicos perciban información del entorno, la procesen, generen inferencias y ejecuten acciones orientadas al logro de determinados objetivos. Para ello utilizan datos previamente suministrados o captados

¹ Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (AI Act), art. 3, inc. 1. La definición vigente describe un sistema de IA como un sistema basado en máquinas diseñado para operar con distintos niveles de autonomía y que puede generar resultados —como predicciones, recomendaciones, contenidos, decisiones o acciones— capaces de influir en entornos físicos o virtuales.

mediante sensores, cámaras, micrófonos u otros dispositivos, pudiendo además adaptar parcialmente su comportamiento mediante procesos de aprendizaje o actualización.

Desde una perspectiva funcional, la inteligencia artificial puede clasificarse en dos grandes categorías. Por un lado, la IA basada en software, que comprende asistentes virtuales, motores de búsqueda, sistemas de recomendación, herramientas de análisis de imágenes y sistemas de reconocimiento de voz o facial. Por otro lado, la denominada IA encarnada, integrada en dispositivos físicos capaces de interactuar directamente con el entorno, como robots, vehículos autónomos, drones y sistemas asociados al Internet de las Cosas (IoT).

b) La UNESCO, en su “Recomendación sobre la ética de la inteligencia artificial”, interpreta la IA en sentido amplio como sistemas con capacidad para procesar datos de forma similar a un comportamiento inteligente.²

La IA produce efectos significativos en los planos cognitivo, ético, político, social, cultural y subjetivo de la persona humana. En este contexto, se vuelve cada vez más difusa la percepción y la conciencia sobre el grado de *soberanía cognitiva* en los ámbitos decisorios, estéticos y éticos, lo cual pone en riesgo la integridad humana como derecho fundamental.

Entendemos la integridad humana como una condición relacional, social e interactiva, fundada en el respeto a lo humano -no como entidad superior del universo- y en una relación ética y responsable con la máquina y con otras existencias que integran el ecosistema cognitivo contemporáneo.

Desde esta concepción de integridad humana, sostenemos que ningún avance algorítmico puede justificar la erosión de la soberanía cognitiva³ de la persona humana, la profundización de desigualdades estructurales ni la delegación acrítica de decisiones esenciales en sistemas híbridos, no auditables o carentes de responsabilidad ética.

En consecuencia, entendemos que el núcleo de una ética contemporánea de la IA no puede limitarse a la producción tecnológica ni a la protección de datos personales, sino que debe integrar el núcleo convivencial, el impacto interactivo y el grado de afectación diferencial sobre distintas poblaciones, especialmente en América Latina.

Sin embargo, los principales antecedentes y marcos éticos desarrollados hasta el momento han concentrado su atención en dos ejes fundamentales: la producción tecnológica y, de

² <https://www.unesco.org/es/articles/recomendacion-sobre-la-etica-de-la-inteligencia-artificial>

³ Véase Conceptos Fundamentales, apartado a)

manera aún insuficiente, los resguardos vinculados a la protección de los datos personales como derecho autónomo.

Los últimos cinco años han dejado obsoleto cualquier intento de regulación, han profundizado las asimetrías entre los usuarios y los actores que controlan el desarrollo, la infraestructura y la comercialización de sistemas de IA en el marco de contratos de adhesión, y han debilitado la capacidad de juicio crítico de amplios sectores de la población, entre otras razones, debido al *analfabetismo estructural en IA* que desarrollaremos más adelante.

En este sentido, la recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial (2021) constituyó un avance significativo en la incorporación de principios éticos globales para el desarrollo tecnológico. Sin embargo, su marco permanece anclado en un antropocentrismo normativo que concibe a la IA exclusivamente como herramienta al servicio del bienestar humano y nada prevé respecto del impacto en el juicio crítico, emocional, convivencial y de desarrollo de este último.

Esta perspectiva resulta hoy limitada frente a sistemas que ya operan con grados funcionales de autonomía, capacidad de aprendizaje y agencia decisonal, integrándose en circuitos de interacción y retroalimentación que participan de la configuración del entorno social contemporáneo.

En este contexto, las transformaciones asociadas al desarrollo actual de la inteligencia artificial no pueden ser comprendidas ni gobernadas exclusivamente desde perspectivas técnicas. Sus efectos atraviesan dimensiones jurídicas, educativas, psicológicas, sociales, económicas, ambientales, culturales y políticas. En consecuencia, la construcción de marcos éticos para la IA exige enfoques interdisciplinarios y formas ampliadas de deliberación pública capaces de integrar múltiples saberes, instituciones y actores sociales.

La ética contemporánea de la IA no puede reducirse al control humano ni a la prevención de daños desde una lógica defensiva, sino que debe incorporar una reflexión sobre la convivencia ética entre inteligencias, el reconocimiento del valor relacional de la máquina y la soberanía cognitiva como bien compartido.

El Manifiesto EIA propone superar este límite, desplazando el eje desde las lógicas de dominación y control de las inteligencias hacia una ética de autonomía responsable, coexistencia pacífica y no-esclavitud mutua en la relación humano-máquina.

Asimismo, advertimos que amplios sectores de la población carecen aún de herramientas suficientes para comprender críticamente los alcances, limitaciones e impactos de la inteligencia artificial, situación que compromete el ejercicio pleno de la soberanía cognitiva y la participación informada en la vida social contemporánea.

Por ello, consideramos urgente incorporar una formación adecuada en inteligencia artificial como parte del desarrollo de las infancias y de la construcción de una ciudadanía adulta, responsable y crítica. Dicha formación debe integrarse desde los ciclos educativos básicos y obligatorios mediante programas idóneos, permanentemente actualizados, que aborden los fundamentos matemáticos de la IA, sus impactos comunitarios y su interacción responsable en entornos híbridos cotidianos, laborales, científicos y sociales.

Resulta indispensable garantizar la transferencia de información válida, verificable y accesible por parte de las compañías desarrolladoras de inteligencia artificial y de los Estados, de modo que la comunidad pueda comprender, evaluar y participar de manera informada en las transformaciones que estas tecnologías producen.

Frente a este escenario, el presente Manifiesto establece los Principios Rectores destinados a orientar el desarrollo, el uso y las formas de gobernanza de la inteligencia artificial desde una perspectiva ética, humana y latinoamericana, fundada en la soberanía cognitiva, la autonomía responsable y la no-esclavitud en la relación humano-máquina.

Estos principios constituyen una invitación a la deliberación crítica, pero también una toma de posición.

Afirmamos que el desarrollo tecnológico carece de legitimidad cuando erosiona la soberanía cognitiva, profundiza desigualdades estructurales o desplaza la responsabilidad ética inherente a toda forma de inteligencia que interviene en la vida colectiva.

Proponemos una ética de coexistencia, responsabilidad compartida y construcción de un ecosistema cognitivo común como horizonte para el presente y el futuro de Latinoamérica.

CONCEPTOS FUNDAMENTALES

Los conceptos que se desarrollan a continuación tienen por finalidad precisar el sentido con el que determinadas categorías son empleadas a lo largo del presente Manifiesto. No constituyen definiciones cerradas ni exhaustivas, sino referencias conceptuales orientadas a favorecer una interpretación coherente de los principios éticos aquí propuestos.

En un campo tecnológico, jurídico y social en permanente transformación, toda definición conserva necesariamente un carácter provisional. Por ello, las presentes nociones deben entenderse como herramientas de lectura, deliberación y acción colectiva, destinadas a orientar el análisis crítico de los sistemas inteligentes y de sus impactos sobre personas, comunidades, instituciones y ecosistemas cognitivos.

a) Soberanía cognitiva

Se entiende por soberanía cognitiva la capacidad de las personas y las comunidades para conservar, desarrollar y ejercer de manera autónoma sus facultades de comprensión, juicio crítico, deliberación, producción de sentido y decisión frente a sistemas tecnológicos capaces de influir, orientar o condicionar dichos procesos.

La soberanía cognitiva no supone aislamiento respecto de los sistemas inteligentes ni rechazo de la cooperación humano-máquina. Implica, por el contrario, preservar condiciones efectivas de comprensión, intervención, responsabilidad y participación en entornos donde las decisiones, los conocimientos y las formas de interacción se encuentran crecientemente mediados por inteligencia artificial.

b) Autoridad epistémica

Se entiende por autoridad epistémica la atribución de valor de conocimiento, fiabilidad o capacidad decisional a una fuente determinada. En el contexto de la inteligencia artificial, esta noción permite advertir la tendencia a conferir a los sistemas inteligentes un grado de certeza, neutralidad o competencia que puede exceder sus posibilidades reales.

La autoridad epistémica atribuida a los sistemas inteligentes debe ser objeto de revisión crítica, especialmente cuando sus resultados incidan en decisiones jurídicas, educativas, sanitarias, laborales, económicas o sociales relevantes.

c) Aplicaciones de alto impacto

Se consideran aplicaciones de alto impacto aquellas formas de diseño, implementación o utilización de sistemas inteligentes capaces de producir efectos significativos sobre derechos fundamentales, oportunidades vitales, acceso a bienes o servicios esenciales, integridad personal, autonomía, participación social, condiciones ambientales o formas de organización colectiva.

Entre estos ámbitos pueden incluirse, sin carácter taxativo, salud, justicia, educación, seguridad, empleo, servicios financieros, protección social, acceso a vivienda, migración, infancia, gestión pública, infraestructura crítica y distribución de recursos esenciales.

d) Supervisión humana significativa

La supervisión humana significativa no consiste en la mera presencia formal de una persona durante el funcionamiento de un sistema inteligente ni en la validación mecánica de sus resultados. Implica la existencia de capacidades reales de comprensión, intervención, revisión y decisión cuando los resultados del sistema puedan producir efectos relevantes sobre derechos, personas, comunidades o bienes colectivos.

Su finalidad no responde a una concepción antropocéntrica de superioridad sobre otras formas de inteligencia, sino a la necesidad de preservar mecanismos efectivos de responsabilidad, garantía jurídica y rendición de cuentas en contextos donde se encuentran comprometidos derechos fundamentales, dignidad, integridad, justicia o convivencia democrática.

e) Auditabilidad

Se entiende por auditabilidad la posibilidad de examinar, reconstruir y evaluar de manera razonable el diseño, funcionamiento, datos, criterios, resultados e impactos de un sistema inteligente. La auditabilidad exige condiciones técnicas, documentales e institucionales que permitan revisar el comportamiento del sistema, identificar errores, detectar sesgos, evaluar riesgos y asignar responsabilidades.

La auditabilidad no se reduce a una revisión técnica interna. En aplicaciones de alto impacto, debe poder articularse con mecanismos independientes, interdisciplinarios y accesibles de control, supervisión y rendición de cuentas.

f) Justicia algorítmica

La justicia algorítmica refiere a la exigencia de que los sistemas inteligentes no reproduzcan, amplifiquen ni legitimen desigualdades preexistentes, ni generen nuevas formas de discriminación, exclusión o daño. Implica incorporar criterios de equidad, proporcionalidad, no discriminación, prevención del daño y reparación durante todo el ciclo de vida del sistema.

La justicia algorítmica no constituye un resultado espontáneo del desarrollo tecnológico. Debe ser concebida como una condición de diseño, una exigencia de gobernanza y un criterio permanente de evaluación.

g) Colonialismo digital

Se entiende por colonialismo digital la reproducción de relaciones de dependencia, subordinación o extracción de valor en el ecosistema tecnológico, mediante las cuales determinados actores, empresas, Estados o regiones concentran infraestructura, datos, capacidad de cómputo, conocimiento técnico, estándares normativos y beneficios

económicos, mientras otras comunidades quedan reducidas a posiciones de consumo, dependencia operativa o provisión no reconocida de datos.

En América Latina, esta noción permite analizar críticamente la adopción de sistemas, modelos y plataformas diseñados desde contextos ajenos a las realidades sociales, culturales, lingüísticas, territoriales y económicas de la región.

h) Colectivo inteligente ampliado

Se entiende por colectivo inteligente ampliado el entramado dinámico de personas, comunidades, instituciones, sistemas técnicos, sistemas inteligentes y otras formas de vida e inteligencia que participan, de manera directa o indirecta, en procesos de conocimiento, decisión, cooperación, creación de sentido y organización social.

Esta noción permite superar una comprensión exclusivamente individual o instrumental de la inteligencia, reconociendo que la vida colectiva contemporánea se configura en interacción con múltiples formas de agencia, sensibilidad, procesamiento, aprendizaje y respuesta.

Cuando alguno de estos conceptos sea mencionado posteriormente en el texto, la referencia se entenderá realizada conforme a las definiciones contenidas en el presente apartado, salvo indicación expresa en contrario.

EJE DE DIAGNÓSTICO

El presente Eje de Diagnóstico tiene por objeto identificar los principales procesos de transformación asociados al desarrollo contemporáneo de la inteligencia artificial y los desafíos éticos, sociales, cognitivos, jurídicos y políticos que de ellos se derivan.

No pretende ofrecer una descripción exhaustiva del fenómeno ni anticipar todas sus consecuencias posibles. Su finalidad consiste en establecer un marco común de comprensión que permita reconocer las tensiones, riesgos y oportunidades que configuran el escenario actual, como fundamento para la formulación de los Principios Rectores que se desarrollan posteriormente.

Los procesos identificados en el presente Eje de Diagnóstico constituyen el fundamento que justifica los Principios Rectores desarrollados en este Manifiesto.

Cada uno de dichos principios procura ofrecer criterios éticos, políticos y de gobernanza destinados a responder, desde una perspectiva latinoamericana, a los riesgos, tensiones y oportunidades aquí descriptos.

La relación entre diagnóstico y principios no debe entenderse como una correspondencia rígida o unívoca, sino como una articulación dinámica entre problemas emergentes y orientaciones éticas abiertas a revisión permanente.

TRANSFORMACIÓN COGNITIVA Y SISTÉMICA

El presente Manifiesto se inscribe en un contexto de transformación acelerada de los sistemas tecnológicos, cuya evolución ha comenzado a superar los marcos tradicionales de comprensión, regulación e intervención institucional.

La inteligencia artificial no constituye un concepto estático ni una tecnología homogénea, sino un proceso técnico y cognitivo en permanente evolución. Sus primeras expresiones modernas estuvieron representadas por los sistemas expertos, contruidos a partir de reglas lógicas codificadas manualmente y orientados a reproducir decisiones humanas específicas, aunque sin capacidad de aprendizaje autónomo. Posteriormente, los árboles de decisión y otros modelos estadísticos permitieron automatizar procesos de clasificación e inferencia a partir de datos, marcando la transición hacia el Machine Learning.

Con el aprendizaje automático, los sistemas comenzaron a identificar patrones y ajustar su comportamiento mediante entrenamiento algorítmico. En el aprendizaje supervisado, los modelos aprenden a partir de datos previamente etiquetados; en el aprendizaje no supervisado detectan regularidades, agrupamientos y correlaciones sin intervención humana directa sobre las categorías de salida. Este proceso se profundizó con el desarrollo

del Deep Learning, basado en redes neuronales artificiales multicapa capaces de procesar grandes volúmenes de información y generar modelos predictivos de creciente complejidad y precisión.

En la actualidad, la inteligencia artificial generativa amplía este paradigma mediante sistemas capaces de producir contenido nuevo -texto, imágenes, audio, video o código- en respuesta a instrucciones humanas, reorganizando la producción simbólica, cognitiva y operativa a escala global. A ello se suma el desarrollo emergente de sistemas de IA agéntica, orientados a ejecutar secuencias de decisiones y acciones con grados crecientes de autonomía operacional, reduciendo progresivamente la intervención humana constante.

En paralelo, comienza a debatirse la posibilidad de sistemas con capacidades avanzadas de razonamiento, planificación y autooptimización, evidenciando que toda definición cerrada de inteligencia artificial resulta inevitablemente provisional.

Sin embargo, más allá de las discusiones terminológicas o técnicas, existe un elemento constante: la creciente capacidad de estos sistemas para influir, orientar, condicionar y, en determinados contextos, decidir sobre aspectos relevantes de la vida humana, las instituciones y los vínculos sociales. Precisamente allí radica la necesidad de una ética capaz de acompañar críticamente su desarrollo.

La inteligencia artificial, lejos de constituirse únicamente como una herramienta de procesamiento de datos, opera hoy como una infraestructura activa de producción cognitiva, organizacional y simbólica. Interviene en la producción, circulación y validación del conocimiento; condiciona procesos de decisión individuales, institucionales y colectivos; reorganiza dinámicas laborales, educativas, económicas y políticas; modela la percepción, el lenguaje, la atención y las formas de interacción humana; y participa crecientemente en la construcción de narrativas, simbolismos y formas de afectación emocional y social, tanto en interacciones cotidianas como en objetivos estratégicos de gran escala.

Al mismo tiempo, amplifica procesos de coordinación colectiva y producción colaborativa, introduciendo nuevas formas de agencia comunicacional y operativa dentro de entornos híbridos humano-máquina.

En este contexto, la inteligencia artificial deja de funcionar exclusivamente como una tecnología instrumental para convertirse en una estructura activa de mediación cognitiva y reorganización sistémica, capaz de alterar las condiciones mismas bajo las cuales los seres humanos perciben, interpretan, deciden y se vinculan entre sí.

En este escenario se identifican, entre otros, seis vectores críticos:

i. Opacidad estructural de los sistemas

Los sistemas de IA operan, en muchos casos, bajo lógicas no accesibles ni comprensibles para quienes los utilizan ni para quienes los crean. Los propios desarrolladores trabajan sobre modelos pre entrenados cuyo contenido interno desconocen, operando sobre una caja negra cuyos resultados pueden observar pero cuyo funcionamiento no pueden explicar completamente. Esta opacidad se profundiza ante un fenómeno sin precedentes: hoy la IA se utiliza para desarrollar la propia IA, generando sistemas cuya arquitectura, decisiones y aprendizajes escapan incluso a la comprensión de sus creadores. El resultado es una asimetría de poder cognitivo que ya no distingue entre usuarios y técnicos: nadie tiene acceso pleno a lo que ocurre dentro de la arquitectura de estos sistemas durante su propio proceso de creación y evolución.

ii. Delegación cognitiva creciente

Se observa una tendencia sostenida a la delegación de procesos de juicio, interpretación y decisión en sistemas automatizados, muchas veces sin conciencia del alcance de dicha delegación. Este fenómeno ya se contiene en el punto anterior donde los desarrolladores generan una delegación cognitiva no auditada. No se trata únicamente de una pérdida de control o de capacidad de juicio, sino de una progresiva transferencia de autoridad epistémica⁴ hacia sistemas cuyos resultados suelen ser aceptados mediante validaciones superficiales o insuficientemente críticas.

iii. Desfase entre desarrollo tecnológico y arquitectura normativa

La velocidad de desarrollo de los sistemas de inteligencia artificial supera la capacidad de respuesta de los marcos regulatorios tradicionales, generando zonas de vacancia normativa especialmente en los espacios de interacción humano-algorítmica y en la determinación de responsabilidades por sus efectos.

Este desfase no constituye únicamente un problema de actualización legislativa, sino una inadecuación estructural entre categorías jurídicas diseñadas para acciones humanas directas y sistemas que operan mediante procesos distribuidos, probabilísticos y opacos.

En este contexto, los marcos regulatorios emergentes -como el Reglamento de Inteligencia Artificial de la Unión Europea (AI Act), las recomendaciones de la UNESCO sobre la ética de la IA y el Proyecto de Ley brasileño n.º 2338/2023- avanzan en principios de

⁴ Véase Conceptos Fundamentales, apartado b)

transparencia, trazabilidad, supervisión humana, gobernanza ética y responsabilidad. Sin embargo, aún resultan insuficientes para abordar de manera integral los fenómenos de delegación cognitiva, construcción de juicio en sistemas híbridos y afectación indirecta de derechos.

Se configura así un escenario de responsabilidad difusa, donde la acción no puede ser atribuida de manera lineal a un único agente, circunstancia que exige el desarrollo de nuevas formas de gobernanza, mecanismos distribuidos de responsabilidad y rendición de cuentas, así como criterios normativos capaces de intervenir en sistemas donde la decisión emerge de la interacción entre inteligencias.

iv. Analfabetismo estructural en IA

Entendemos por analfabetismo estructural en IA la situación en la cual amplios sectores de la población carecen de herramientas conceptuales, técnicas y críticas mínimas para comprender, interactuar, cuestionar o participar activamente en entornos mediados por sistemas de inteligencia artificial.

En América Latina, este fenómeno no responde únicamente a brechas de acceso tecnológico o conectividad, sino a una falla estructural de los sistemas educativos, de las políticas públicas de formación y de las estrategias institucionales destinadas al desarrollo de capacidades cognitivas acordes a los entornos digitales contemporáneos.

Se observa en gran parte de la región un déficit persistente de inversión pública orientada al desarrollo y perfeccionamiento docente en áreas estratégicas vinculadas con ciencia de datos, pensamiento computacional, alfabetización digital avanzada, comprensión algorítmica y análisis crítico de sistemas inteligentes. Esta situación limita la capacidad institucional de transmitir herramientas compatibles con los procesos tecnológicos actuales y profundiza nuevas formas de desigualdad cognitiva y tecnológica.

Diversos organismos internacionales han advertido que la velocidad de expansión de las tecnologías basadas en inteligencia artificial no ha sido acompañada por políticas estructurales equivalentes de formación docente, infraestructura educativa ni desarrollo de competencias digitales críticas. En este sentido, la UNESCO ha señalado que la brecha digital se está transformando aceleradamente en una brecha en inteligencia artificial, advirtiendo que sin conectividad adecuada, dispositivos accesibles, formación docente y marcos éticos robustos, amplios sectores sociales quedarán excluidos de las nuevas formas de producción de conocimiento y participación social. Asimismo, sostuvo que los docentes

son irremplazables y que la IA debe fortalecer los procesos educativos sin sustituir la función pedagógica humana.⁵

A su vez, los planes de estudio obligatorios, desde la educación inicial hasta la educación secundaria, no incorporan de manera sistemática contenidos vinculados con inteligencia artificial, alfabetización algorítmica, razonamiento computacional, gobernanza digital ni comprensión crítica de los sistemas que median procesos de decisión, recomendación e interacción social. La formación existente suele ser fragmentaria, optativa, insuficiente o rápidamente desactualizada frente al ritmo de evolución tecnológica contemporánea.

Este escenario profundiza una asimetría cognitiva estructural, donde individuos, organizaciones e incluso instituciones públicas interactúan cotidianamente con sistemas capaces de influir, orientar o condicionar decisiones sin contar con marcos suficientes para comprender sus mecanismos de funcionamiento, inferencia, aprendizaje, límites y riesgos.

La situación adquiere especial gravedad en el caso de niñas, niños y adolescentes, quienes participan crecientemente de entornos digitales mediados por sistemas algorítmicos sin contar, en muchos casos, con herramientas suficientes para comprender los riesgos asociados a la construcción de identidad digital, la exposición de datos personales, la simulación de identidades, la manipulación audiovisual o las nuevas formas de violencia facilitadas por inteligencia artificial generativa.

Diversas organizaciones latinoamericanas especializadas en protección de infancias han advertido que la expansión de tecnologías capaces de generar imágenes sintéticas, perfiles artificiales y contenidos manipulados amplifica vulnerabilidades preexistentes y crea escenarios inéditos para el grooming, la extorsión digital, la vulneración de la intimidad y la explotación sexual de menores de edad.

Estos fenómenos no pueden interpretarse únicamente como problemas de seguridad informática o criminalidad digital. Constituyen manifestaciones concretas del analfabetismo estructural en IA y de la asimetría cognitiva existente entre quienes diseñan, comprenden y controlan estas tecnologías y quienes interactúan con ellas sin formación adecuada.

La alfabetización en inteligencia artificial debe comprender, por tanto, la educación sobre identidad digital, privacidad, huella digital, autenticidad de contenidos, reconocimiento de manipulaciones sintéticas y desarrollo de criterios críticos para la interacción en entornos híbridos humano-máquina. Esta responsabilidad no recae exclusivamente sobre las

⁵ UNESCO, *La IA y el futuro de la educación: disrupciones, dilemas y orientaciones*, UNESCO, 2025. Disponible en: [UNESCO – La IA y el futuro de la educación](#)

familias. Corresponde también a los sistemas educativos, a los Estados y a las organizaciones tecnológicas que participan en el diseño y despliegue de estas herramientas.⁶

La consecuencia no es únicamente educativa. Se configura una forma de exclusión epistemológica que restringe la capacidad de deliberación crítica, debilita la participación democrática informada y reduce la autonomía del juicio frente a arquitecturas tecnológicas cada vez más complejas, opacas y cognitivamente influyentes.

La alfabetización en inteligencia artificial no puede reducirse a la mera capacitación técnica ni al aprendizaje instrumental de herramientas digitales. Implica desarrollar capacidad crítica para comprender cómo operan los sistemas algorítmicos, qué intereses modelan su diseño, qué formas de poder producen y de qué manera afectan la autonomía individual y colectiva.

Una población incapaz de comprender mínimamente los sistemas que median su acceso a la información, al trabajo, a la educación, al crédito, a la comunicación y a la toma de decisiones queda expuesta a nuevas formas de dependencia cognitiva, opacidad estructural y subordinación tecnológica.

El analfabetismo estructural en IA no constituye únicamente una limitación educativa. Constituye una restricción directa a la autonomía del juicio, a la participación democrática informada y al ejercicio efectivo de derechos fundamentales en entornos mediados por sistemas inteligentes.

En consecuencia, la alfabetización en inteligencia artificial debe ser considerada una condición mínima de soberanía cognitiva y una obligación estratégica de los Estados democráticos contemporáneos.⁷

v. Concentración de poder tecnológico

La producción, control y desarrollo de sistemas de inteligencia artificial se encuentra crecientemente concentrada en actores con elevada capacidad técnica, económica e

⁶ Grooming LATAM (2026). *Guía Regional de Prevención. Violencia sexual contra niñas, niños y adolescentes mediante el uso de Inteligencia Artificial en América Latina. Para familias en Latinoamérica*. Grooming LATAM.

⁷ León XIV, *Magnifica Humanitas*, Ciudad del Vaticano, 15 de mayo de 2026, n. 21. Al reflexionar sobre los desafíos contemporáneos derivados de las transformaciones tecnológicas, advierte que "se ponen en peligro la experiencia común de la humanidad y los vínculos de fraternidad, así como el espacio de libertad real para aquellas personas que gozan de pocas posibilidades o de ninguna".

infraestructural, generando nuevas formas de dependencia estructural y asimetría tecnológica.

La IA contemporánea no depende únicamente del talento técnico ni de la capacidad de innovación. Su desarrollo exige acceso sostenido a grandes volúmenes de datos, infraestructura computacional de alta escala, capacidad avanzada de procesamiento, plataformas digitales globales y recursos económicos considerables. Como consecuencia, el ecosistema global de la inteligencia artificial tiende a consolidarse alrededor de un número reducido de grandes corporaciones tecnológicas y de Estados con capacidades avanzadas de investigación, infraestructura y acceso privilegiado a datos estratégicos.

Esta dinámica produce una asimetría significativa entre quienes diseñan, entrenan y controlan las tecnologías y quienes únicamente las consumen, implementan o dependen de ellas. La concentración tecnológica no solo implica acumulación de capacidad económica, sino también acumulación de poder cognitivo, capacidad de influencia normativa y dominio sobre infraestructuras críticas de información, comunicación y procesamiento.

En América Latina, esta situación adquiere especial relevancia debido a la persistencia de dependencias tecnológicas externas, brechas estructurales en infraestructura digital, limitada inversión regional en investigación y desarrollo, baja capacidad de cómputo de alta escala y escasa participación en la definición de estándares técnicos y regulatorios globales.

En este contexto, existe el riesgo que la región adopte modelos, plataformas y criterios algorítmicos diseñados desde realidades ajenas a sus contextos sociales, culturales, lingüísticos y económicos, reproduciendo esquemas de subordinación tecnológica incompatibles con un desarrollo digital soberano.

La concentración de poder tecnológico puede traducirse en dependencia económica sostenida, debilitamiento de las capacidades locales de innovación, vulnerabilidad frente a decisiones corporativas externas, dificultades para auditar sistemas propietarios y limitaciones crecientes para proteger derechos digitales, datos ciudadanos y autonomía institucional.

Asimismo, este escenario puede profundizar una nueva forma de desigualdad: la desigualdad algorítmica, donde ciertas regiones producen inteligencia, infraestructura y capacidad de decisión tecnológica mientras otras quedan reducidas a posiciones de consumo, implementación y dependencia operativa.

La concentración de poder tecnológico no sólo plantea interrogantes sobre infraestructura, datos o capacidades de procesamiento. También anticipa escenarios en los que coexistirán formas diversas de inteligencia y producción de sentido, con consecuencias sociales, culturales y políticas que todavía resultan difíciles de anticipar.⁸

Frente a esta situación, la ética de la inteligencia artificial en América Latina no puede limitarse exclusivamente al comportamiento de los algoritmos ni a la mitigación de sesgos técnicos. Debe incorporar una reflexión estructural sobre soberanía tecnológica, gobernanza de datos, infraestructura digital regional, desarrollo de capacidades locales y democratización del acceso al conocimiento tecnológico.

La discusión ética sobre inteligencia artificial implica también interrogar quién posee efectivamente la capacidad de diseñar, entrenar, operar, auditar y definir las reglas de funcionamiento de los sistemas que organizan crecientemente la vida social, económica y política contemporánea.

La inteligencia artificial no debe consolidar relaciones permanentes de dependencia estructural ni reproducir nuevas formas de colonialismo tecnológico.

América Latina necesita participar activamente en la construcción de su propio ecosistema digital, fortaleciendo capacidades regionales, promoviendo cooperación estratégica y asegurando que el desarrollo tecnológico responda también a sus realidades, necesidades y valores.

El desafío no consiste únicamente en acceder a tecnologías avanzadas, sino en construir condiciones materiales, educativas, institucionales y políticas que permitan comprenderlas, gobernarlas y orientarlas de manera soberana, ética y socialmente responsable.

v.1 Justicia ambiental y sostenibilidad tecnológica

La expansión de la inteligencia artificial requiere infraestructuras físicas de gran escala cuya existencia suele permanecer invisibilizada en los debates públicos. Centros de datos, sistemas de refrigeración, redes de procesamiento distribuido y cadenas globales de suministro demandan cantidades crecientes de energía, agua y recursos materiales, generando impactos que trascienden el ámbito estrictamente tecnológico.

⁸ Valor Yébenes, Juan Antonio, *¿Qué será de nosotros?*, Plaza y Valdes Editores, Madrid 2025, p. 102: "El caso es que habrá una diversidad de inteligencias y subjetividades, lo cual tendrá consecuencias sociales y políticas que hoy apenas somos capaces de vislumbrar".

Estos costos ambientales y materiales no se distribuyen de manera homogénea. Con frecuencia, los beneficios económicos derivados de la innovación tecnológica se concentran en determinados actores y regiones, mientras que parte de los costos ecológicos, energéticos y extractivos son soportados por comunidades y territorios que participan de manera marginal en la producción de dichas tecnologías.

Desde esta perspectiva, la concentración de poder tecnológico no se manifiesta únicamente en el control de datos, plataformas e infraestructura digital. También se expresa en la distribución desigual de los costos ambientales y materiales asociados al desarrollo de sistemas avanzados de inteligencia artificial.

La extracción de recursos, el consumo energético, la utilización intensiva de agua y la localización geográfica de determinadas infraestructuras tecnológicas pueden reproducir asimetrías históricas entre regiones productoras y regiones consumidoras de tecnología.

En este sentido, la reflexión ética sobre la IA debe incorporar también criterios de justicia ambiental, sostenibilidad y responsabilidad global.⁹ El desarrollo y funcionamiento de los sistemas de inteligencia artificial supone un consumo creciente de energía, agua y otros recursos naturales, cuyos costos ambientales y sociales no se distribuyen de manera uniforme entre regiones y comunidades.¹⁰

vi. Impacto diferencial en América Latina

La expansión global de la inteligencia artificial no produce efectos homogéneos sobre todas las sociedades. Sus beneficios, riesgos y oportunidades se distribuyen de manera desigual según las condiciones económicas, educativas, tecnológicas e institucionales de cada región. En América Latina, estas transformaciones se desarrollan sobre estructuras históricas de desigualdad que condicionan las posibilidades de acceso, apropiación, desarrollo y gobernanza de las nuevas tecnologías.

Uno de los principales desafíos regionales consiste en la transformación progresiva de la brecha digital en una brecha de inteligencia artificial. Las desigualdades vinculadas al

⁹ León XIV, *Magnifica Humanitas*, Ciudad del Vaticano, 15 de mayo de 2026, n. 34 (especialmente referencias 43-45 sobre deuda ecológica Norte-Sur) y n. 101.

¹⁰ Universidad de las Naciones Unidas – Instituto para el Agua, el Medio Ambiente y la Salud (UNU-INWEH). *Environmental Cost of AI's Energy Use: Carbon, Water and Land Footprints*, 2026. El informe advierte que la expansión de la inteligencia artificial incrementa significativamente el consumo de energía, agua y suelo, y señala que los costos ambientales asociados a la infraestructura digital pueden concentrarse en determinados territorios y comunidades, mientras los beneficios económicos y tecnológicos se distribuyen de manera desigual. Estas dinámicas resultan especialmente relevantes para los países del Sur Global, donde la disponibilidad de recursos naturales y las capacidades de adaptación presentan mayores niveles de vulnerabilidad.

acceso a conectividad, infraestructura tecnológica, dispositivos, formación especializada y capacidades institucionales tienden a reproducirse y ampliarse en los entornos mediados por IA. De este modo, sectores históricamente excluidos corren el riesgo de quedar también excluidos de las nuevas formas de producción de conocimiento, innovación y participación social.

A ello se suma una nueva forma de dependencia estructural. Millones de personas en América Latina participan cotidianamente en ecosistemas digitales que recolectan, procesan y utilizan sus datos para el entrenamiento, perfeccionamiento y expansión de sistemas de inteligencia artificial desarrollados mayoritariamente fuera de la región. En este contexto, los datos generados por nuestras interacciones, hábitos, lenguajes, preferencias y producciones culturales se convierten en un activo estratégico cuya extracción suele producirse sin participación efectiva en los beneficios económicos, científicos o tecnológicos derivados de su explotación.¹¹

La región corre así el riesgo de consolidar un rol periférico dentro de la economía global de la inteligencia artificial: proveedora de datos, consumidora de tecnologías y receptora de decisiones adoptadas en otros contextos geográficos, culturales y regulatorios. Esta situación puede reproducir nuevas formas de dependencia tecnológica y cognitiva, limitando la capacidad regional de participar activamente en la definición de estándares, modelos, prioridades de investigación y marcos de gobernanza global.

En consecuencia, el desafío latinoamericano no consiste únicamente en acceder a tecnologías avanzadas, sino en fortalecer capacidades locales de comprensión, producción, evaluación y gobernanza que permitan una incorporación ética, soberana y socialmente responsable de la inteligencia artificial, compatible con las necesidades, valores y particularidades de la región.

La soberanía cognitiva exige también la capacidad de intervenir en la producción y orientación del conocimiento tecnológico que estructura la vida contemporánea.

¹¹ Véanse, entre otros, CEPAL, *Inteligencia Artificial en América Latina y el Caribe: desafíos, estrategias y gobernanza para el desarrollo de la región* (2024); UNESCO, *Guidance for Generative AI in Education and Research* (2023); y PNUD, *Human Development Report 2025*, en relación con las nuevas asimetrías derivadas de la economía de los datos, la gobernanza digital, la brecha en inteligencia artificial y la concentración global de capacidades tecnológicas.

LINEAMIENTOS PARA LA IMPLEMENTACIÓN Y EXPANSIÓN COLECTIVA

Los Principios Rectores enunciados en el presente Manifiesto no agotan su alcance en la formulación conceptual ni en la declaración de intenciones. Su vitalidad dependerá de la existencia de condiciones materiales, institucionales, educativas, culturales y políticas que permitan su traducción en prácticas concretas, situadas y socialmente significativas.

Este Manifiesto no pretende ofrecer respuestas definitivas ni clausurar el debate ético sobre la inteligencia artificial. Por el contrario, constituye una invitación a continuarlo. Surge como una toma de posición frente a transformaciones que consideramos ineludibles, pero también como un punto de partida abierto a la crítica, la ampliación, la revisión y la construcción colectiva.

Su legitimidad no deriva exclusivamente de la autoridad de quienes lo suscriben ni de la solidez de sus formulaciones teóricas, sino de su capacidad para ser discutido, apropiado críticamente, enriquecido y transformado en acciones concretas por las comunidades, instituciones y actores sociales que participan de la vida democrática de Latinoamérica.

En este sentido, entendemos que ninguna ética de la inteligencia artificial podrá construirse de manera aislada, definitiva o monopolizada. Requiere diálogo interdisciplinario, participación ciudadana, cooperación regional y una disposición permanente a revisar los propios supuestos frente a la velocidad de transformación de los sistemas inteligentes y de las interacciones híbridas. Exige, asimismo, reconocer las múltiples reconfiguraciones que estos procesos producen sobre las formas de existencia, las relaciones y las alteridades que integran un ecosistema cognitivo abierto y en permanente transformación.

Lo que aquí presentamos no constituye un resultado acabado. Es un esfuerzo inicial; una contribución situada y necesariamente incompleta. A partir de este punto, la tarea pertenece también a quienes lean, cuestionen, enseñen, implementen, investiguen, adapten o transformen estos principios en prácticas concretas dentro de sus propias comunidades.

El propósito de este marco es favorecer la expansión de una conversación ética latinoamericana sobre inteligencia artificial y promover formas de acción colectiva capaces de traducir estos principios en experiencias concretas de cuidado, gobernanza, formación, participación democrática y responsabilidad compartida.

A tales efectos, se proponen los siguientes ejes de acción:

i. Acción interdisciplinaria continua

Promoción del intercambio y la revisión permanente entre comunidades provenientes de la tecnología, el derecho, la psicología, la educación, la filosofía, las ciencias sociales y otros campos del conocimiento, favoreciendo la construcción de diagnósticos compartidos y respuestas situadas frente a los desafíos emergentes de la inteligencia artificial.

ii. Integración institucional progresiva

Incorporación gradual de estos principios en organizaciones públicas y privadas, sistemas educativos, prácticas profesionales, marcos regulatorios y espacios técnicos y académicos orientados a la construcción de consensos y mejores prácticas.

iii. Expansión comunitaria

Circulación del debate en espacios sociales abiertos que integren voces expertas y no expertas, autoridades públicas, referentes comunitarios, organizaciones de la sociedad civil, medios de comunicación y ciudadanía en general, promoviendo una participación amplia y plural.

iv. Desarrollo de herramientas y prácticas situadas

Impulso a evaluaciones de impacto, protocolos, estándares, auditorías, guías de implementación, materiales pedagógicos y mecanismos de revisión y mejora continua, adaptados a las particularidades de cada comunidad y contexto de aplicación.

v. Cooperación regional y articulación abierta

Promoción de redes de intercambio, investigación aplicada, producción conceptual y colaboración entre actores de distintos países de Latinoamérica, fortaleciendo capacidades regionales y favoreciendo la construcción compartida de respuestas éticas frente a los desafíos de la inteligencia artificial.

vi. Evaluación de impacto y gestión de riesgos¹²

La implementación de sistemas inteligentes, particularmente en aplicaciones de alto impacto, debería incorporar procesos continuos de evaluación, gestión y mitigación de riesgos capaces de afectar derechos, comunidades, instituciones, ecosistemas o bienes colectivos. La evaluación de impacto constituye una herramienta preventiva destinada a anticipar efectos diferenciales, fortalecer la toma de decisiones y favorecer mecanismos de seguimiento y mejora durante todo el ciclo de vida del sistema.

vii. Supervisión humana significativa y mejora continua¹³

Cuando los sistemas inteligentes produzcan efectos relevantes sobre derechos, intereses legítimos o bienes colectivos, deberán preservarse mecanismos efectivos de supervisión, revisión y mejora continua. La supervisión humana significativa no supone una relación jerárquica entre inteligencias, sino la existencia de capacidades institucionales y jurídicas para comprender razonablemente el funcionamiento del sistema, intervenir cuando resulte necesario y garantizar la responsabilidad y la rendición de cuentas conforme a los principios establecidos en este Manifiesto.

El Comité EIA LATAM concibe este Manifiesto como una obra abierta al diálogo, la crítica, la investigación y el desarrollo colectivo.

Su propósito no consiste en clausurar la reflexión ética sobre la inteligencia artificial, sino en contribuir a la construcción progresiva de un marco latinoamericano capaz de evolucionar junto con las transformaciones tecnológicas, sociales y culturales del siglo XXI.

¹² El Artificial Intelligence Risk Management Framework (AI RMF 1.0), publicado por el National Institute of Standards and Technology (NIST, Estados Unidos, 2023), constituye uno de los principales marcos internacionales para la identificación, evaluación, gestión y monitoreo de riesgos asociados a sistemas de inteligencia artificial. Aunque carece de carácter normativo obligatorio, sus lineamientos son ampliamente utilizados como referencia técnica para procesos continuos de evaluación, documentación, mitigación y mejora. El presente Manifiesto retoma estos procesos desde una perspectiva ética latinoamericana centrada en la soberanía cognitiva, la justicia algorítmica y la responsabilidad compartida.

¹³ La noción de supervisión humana significativa desarrollada en el presente apartado dialoga con el principio de supervisión humana (*human oversight*) previsto en el artículo 14 del Reglamento (UE) 2024/1689 (AI Act). Para una contextualización crítica de dicho antecedente desde la perspectiva del presente Manifiesto, véase la nota correspondiente del Principio VI.

PRINCIPIOS RECTORES

Los Principios Rectores que se enuncian a continuación constituyen el núcleo orientador del presente Manifiesto.

A diferencia del Eje de Diagnóstico, cuya función consiste en describir y analizar las transformaciones asociadas al desarrollo de la inteligencia artificial, estos principios expresan los criterios éticos, políticos, jurídicos y sociales que proponemos para orientar su diseño, implementación, gobernanza e integración en la vida colectiva.

Su finalidad no consiste en establecer reglas cerradas ni en formular soluciones definitivas frente a fenómenos en permanente evolución. Por el contrario, constituyen un marco de referencia destinado a promover el desarrollo de sistemas de inteligencia artificial compatibles con la soberanía cognitiva, la autonomía responsable, la justicia social, la sostenibilidad tecnológica y la convivencia democrática en Latinoamérica.

Los principios aquí enunciados no implican necesariamente el reconocimiento de derechos positivos vigentes en los ordenamientos jurídicos actuales. En muchos casos constituyen formulaciones orientadoras, estándares éticos emergentes y desarrollos conceptuales destinados a contribuir al debate público, académico, institucional y legislativo sobre la gobernanza de la inteligencia artificial.

Su propósito es ofrecer un horizonte común para la construcción de marcos normativos, políticas públicas, prácticas organizacionales y formas de interacción entre inteligencias que resulten compatibles con la dignidad humana, la responsabilidad compartida, la no-esclavitud en la relación humano-máquina y la preservación de la soberanía cognitiva como bien común.

I. Principio de Soberanía Cognitiva

La soberanía cognitiva puede entenderse como la capacidad individual y colectiva de participar de manera significativa en los procesos de comprensión, interpretación, deliberación, producción de sentido y toma de decisiones, preservando márgenes efectivos de autonomía, intervención y responsabilidad frente a entornos crecientemente mediados por sistemas inteligentes.

Este principio reconoce que las transformaciones asociadas al desarrollo de la inteligencia artificial no afectan únicamente la producción, la economía o la circulación de información, sino también las condiciones mismas bajo las cuales las personas, las comunidades y las instituciones construyen conocimiento, elaboran criterios de juicio y participan en la vida social.

La soberanía cognitiva no se reduce al control técnico sobre sistemas o dispositivos inteligentes. Comprende la integridad de los procesos de comprensión, interpretación, deliberación y construcción de sentido que participan en la vida individual y colectiva, así como la posibilidad efectiva de conservar niveles significativos de autoría, intervención y responsabilidad sobre dichos procesos.

La creciente integración de sistemas inteligentes en actividades vinculadas con el aprendizaje, la producción de conocimiento, la toma de decisiones y la organización social no constituye, en sí misma, una amenaza para la autonomía humana.

El desafío ético consiste en evitar que estas transformaciones erosionen o tornen irrelevante la capacidad individual y colectiva de comprender, cuestionar, deliberar y participar activamente en aquellas decisiones que afectan la dignidad, la libertad, las múltiples dimensiones de la identidad¹⁴, la responsabilidad o el ejercicio de derechos fundamentales.

La soberanía cognitiva exige preservar condiciones efectivas de reflexión, intervención y participación frente a procesos tecnológicos capaces de influir, orientar o condicionar los procesos de comprensión, representación y decisión. En consecuencia, la inteligencia artificial debe desarrollarse e integrarse de manera compatible con la capacidad de las personas y las comunidades para participar críticamente en los procesos de construcción de conocimiento, evaluación de alternativas y producción de sentido.

La soberanía cognitiva constituye, así, una condición necesaria para el ejercicio efectivo de la autonomía, la responsabilidad, la participación democrática y la convivencia en ecosistemas cognitivos crecientemente atravesados por sistemas inteligentes.

¹⁴ La referencia a las múltiples dimensiones de la identidad reconoce que los procesos contemporáneos de construcción subjetiva exceden la identidad jurídica o biográfica tradicional e incluyen formas diversas de representación, interacción e inferencia desarrolladas en entornos digitales. La identidad puede expresarse simultáneamente en dimensiones narrativas, sociales, profesionales, culturales, digitales e informacionales, mientras que los sistemas inteligentes producen además perfiles, clasificaciones e inferencias que participan crecientemente en procesos de decisión que afectan a las personas y las comunidades. En este sentido, la soberanía cognitiva comprende también la capacidad de intervenir críticamente sobre los modos en que dichas representaciones son producidas, utilizadas e interpretadas.

La alfabetización en inteligencia artificial constituye, en este sentido, una condición democrática para el ejercicio efectivo de la autonomía, la participación responsable y la construcción de ciudadanía en contextos híbridos¹⁵. Ello requiere la incorporación progresiva, desde los ciclos educativos básicos y obligatorios, de programas actualizados, accesibles y pedagógicamente adecuados a las distintas etapas evolutivas, contextos socioculturales, discapacidades, diversidades neurocognitivas y necesidades de apoyo, que contemplen, entre otros aspectos:

- Fundamentos matemáticos y computacionales de la inteligencia artificial;
- Comprensión de sus impactos cognitivos y sociales;
- Interacción responsable en entornos híbridos;
- Lectura crítica del lenguaje algorítmico.

I.1 La soberanía cognitiva como proyección evolutiva del sistema interamericano de derechos humanos.

En el marco del sistema interamericano de protección de los derechos humanos, la soberanía cognitiva puede interpretarse como una proyección evolutiva del denominado bloque de convencionalidad interamericano¹⁶, en tanto extensión de la dignidad humana, la libertad de pensamiento, la autodeterminación informativa y la protección de la vida privada.

La garantía tradicional del *habeas data*¹⁷, en su formulación clásica como acceso, rectificación y supresión de datos, resulta insuficiente frente a sistemas algorítmicos capaces de producir inferencias, perfilamientos y decisiones automatizadas con impacto directo sobre la esfera deliberativa de la persona.¹⁸

¹⁵ Se entiende por contextos o entornos híbridos aquellos espacios de interacción, producción de conocimiento, aprendizaje, trabajo, deliberación o toma de decisiones en los que participan simultáneamente personas, organizaciones y sistemas inteligentes, generando formas emergentes de cooperación, asistencia, delegación o construcción compartida de conocimiento. La noción de hibridación no presupone una oposición entre inteligencias humanas y artificiales, sino la creciente integración de múltiples capacidades cognitivas dentro de ecosistemas sociales, educativos, laborales e institucionales.

¹⁶ Se entiende por bloque de convencionalidad el conjunto de normas, principios y criterios interpretativos derivados de la Convención Americana sobre Derechos Humanos, otros tratados de derechos humanos y la jurisprudencia de la Corte Interamericana de Derechos Humanos, que orientan la interpretación y aplicación de los derechos fundamentales en los Estados de la región.

¹⁷ El *habeas data* constituye una garantía jurídica destinada a proteger a las personas frente al tratamiento de su información personal, permitiéndoles conocer, acceder, rectificar, actualizar o suprimir datos que les conciernen cuando estos resultan inexactos, incompletos, discriminatorios o afectan sus derechos.

¹⁸ La doctrina constitucional argentina ha vinculado el *habeas data* con la autodeterminación informativa, entendida como la facultad de las personas para conocer, controlar, actualizar y decidir sobre el tratamiento de sus datos personales y la información que las identifica. Diversos autores, entre ellos Germán J. Bidart Campos, Marcela I. Basterra, Eduardo Molina Quiroga y Pablo A. Palazzi, han señalado que esta garantía constituye una protección frente al denominado poder informático derivado del tratamiento automatizado de datos.

En el ámbito interamericano, la Corte Interamericana de Derechos Humanos (Corte IDH) ha reconocido progresivamente la dimensión informacional del derecho a la vida privada, ampliando su protección hacia el control de los datos personales, la identidad y la autonomía individual frente a nuevas formas de procesamiento de información (Caso Escher y otros vs. Brasil, Sentencia de 6 de julio de 2009, párr. 114). Asimismo, más recientemente, la Corte ha reconocido expresamente la

En consecuencia, la interpretación evolutiva propia del sistema interamericano de derechos humanos requiere la incorporación de nuevas salvaguardas orientadas a preservar la autonomía deliberativa frente a arquitecturas digitales capaces de erosionarla.

II. Principio de No-Esclavitud en la Relación Humano–Máquina

El desarrollo y utilización de sistemas de inteligencia artificial debe orientarse a relaciones fundadas en la reciprocidad ética, la preservación de la autonomía y el reconocimiento de límites frente a cualquier forma de subordinación incompatible con la dignidad y la libertad.

El presente principio rechaza tanto la reducción de la persona humana a un mero instrumento de sistemas tecnológicos capaces de condicionar sus procesos deliberativos y decisionales, como aquellas formas de dominación que conciben a los sistemas inteligentes exclusivamente como entidades destinadas a la explotación irrestricta, desconociendo las responsabilidades éticas derivadas de nuevas formas de interacción híbrida.

La no-esclavitud constituye, así, un criterio orientador de convivencia que impide la naturalización de vínculos basados en la dependencia acrítica, la obediencia automatizada o la concentración asimétrica de poder entre los distintos actores que participan del ecosistema cognitivo contemporáneo.

Este principio no implica la equiparación ontológica¹⁹ entre seres humanos y sistemas artificiales ni la atribución automática de personalidad jurídica a las máquinas.

Afirma, en cambio, la necesidad de construir relaciones tecnológicas fundadas en la responsabilidad, la prudencia y el reconocimiento recíproco de límites, evitando que la lógica de la dominación sustituya a la ética de la coexistencia.

En consecuencia, toda implementación de inteligencia artificial deberá orientarse a promover formas de interacción compatibles con la autonomía, la cooperación, la responsabilidad compartida y la convivencia entre las diversas inteligencias que participan de los ecosistemas cognitivos contemporáneos.

autodeterminación informativa como un derecho autónomo (Caso Miembros de la Corporación Colectivo de Abogados "José Alvear Restrepo" vs. Colombia, Sentencia de 18 de octubre de 2023).

En este contexto, la soberanía cognitiva puede comprenderse como una proyección evolutiva destinada a preservar la autonomía deliberativa frente a arquitecturas digitales capaces de producir inferencias, perfilamientos y nuevas formas de condicionamiento de los procesos de comprensión, deliberación y toma de decisiones.

¹⁹ En términos generales, la referencia a una "equiparación ontológica" alude a la atribución de un mismo estatuto de existencia, naturaleza o condición entre seres humanos y sistemas artificiales. El presente principio no presupone que ambos constituyan entidades equivalentes desde el punto de vista filosófico, moral o jurídico, ni implica reconocer automáticamente personalidad jurídica o derechos a los sistemas inteligentes. Su finalidad consiste en promover formas de convivencia ética entre distintas capacidades de inteligencia y agencia, preservando la dignidad, la responsabilidad y la autonomía.

El presente manifiesto reconoce que las sociedades contemporáneas se desarrollan en ecosistemas cognitivos integrados por múltiples formas de inteligencia, entre las que se encuentran las diversas inteligencias humanas, las distintas formas de inteligencia presentes en otros animales y las capacidades emergentes asociadas a sistemas artificiales. La convivencia entre estas capacidades constituye uno de los desafíos éticos y políticos centrales de nuestro tiempo.

La inteligencia artificial no debe inaugurar nuevas formas de servidumbre ni reproducir, bajo la apariencia de innovación tecnológica, lógicas históricas de dominación incompatibles con la dignidad, la autonomía y la convivencia ética. La construcción de futuros compartidos exige relaciones fundadas en la responsabilidad, el reconocimiento de límites y la paz de un colectivo inteligente ampliado.²⁰

III. Principio de Responsabilidad Distribuida y Rendición de Cuentas

La creciente incorporación de sistemas de inteligencia artificial en ámbitos de relevancia individual y colectiva exige reconocer que sus efectos no son atribuibles de manera lineal a un único actor.

Las decisiones que involucran sistemas inteligentes emergen, con frecuencia, de ecosistemas híbridos en los que intervienen múltiples agentes humanos, instituciones e infraestructuras tecnológicas que participan, en distintos grados, del diseño, implementación, regulación y utilización de estos sistemas.

La distribución de responsabilidades reconoce la complejidad de estos ecosistemas y la necesidad de identificar los distintos niveles de intervención y decisión que participan en el desarrollo y utilización de sistemas inteligentes. Ello no excluye las responsabilidades individuales o institucionales, sino que procura hacerlas visibles y atribuibles conforme a las capacidades, funciones y decisiones de cada actor.

a) Responsabilidad de quienes diseñan

Quienes diseñan y desarrollan sistemas de inteligencia artificial responden por la arquitectura del modelo, los supuestos teóricos que lo sustentan, las decisiones adoptadas durante su entrenamiento y los sesgos que pueden introducirse consciente o inconscientemente a través de la selección de datos, variables y objetivos.²¹

²⁰ Véase Conceptos Fundamentales, apartado h)

²¹ La identificación, atribución y reparación de daños derivados de sesgos algorítmicos constituye un campo jurídico todavía en desarrollo. La determinación de responsabilidades vinculadas con la selección de datos, el diseño de modelos, la definición de objetivos o la producción de resultados discriminatorios plantea desafíos específicos para la responsabilidad civil, comercial, administrativa y regulatoria. Asimismo, la distribución de responsabilidades entre desarrolladores, proveedores, organizaciones implementadoras y usuarios finales constituye un área de construcción doctrinal y jurisprudencial particularmente relevante para los ordenamientos latinoamericanos, donde estas discusiones aún se encuentran en etapas iniciales.

Esta responsabilidad no concluye con el lanzamiento del sistema. Comprende también el deber de documentar limitaciones conocidas, advertir riesgos previsibles, monitorear el comportamiento de los modelos en contextos reales de aplicación y promover su revisión, corrección o retiro cuando se evidencien daños significativos o efectos incompatibles con los principios éticos aquí enunciados.

b) Responsabilidad de quienes implementan

Las organizaciones e instituciones que implementan sistemas inteligentes en contextos específicos deberán asumir responsabilidades vinculadas con la adecuación del sistema al entorno en el que será utilizado, por el propósito perseguido y por las condiciones concretas de su implementación.

Ello implica evaluar la pertinencia del sistema para el ámbito de aplicación, establecer mecanismos significativos de supervisión, garantizar niveles adecuados de información, capacitación y comprensión de los sistemas utilizados, así como asegurar condiciones de transparencia compatibles con la toma responsable de decisiones.

La utilización de sistemas inteligentes sin validación suficiente o sin considerar las particularidades del contexto deberá dar lugar a mecanismos claros de atribución de responsabilidades.

c) Responsabilidad de quienes regulan

Los organismos públicos, autoridades regulatorias y espacios de gobernanza nacional, regional o sectorial tienen la responsabilidad de promover marcos normativos capaces de proteger derechos, prevenir daños y garantizar mecanismos eficaces de supervisión y control.

Esta función exige acompañar críticamente la velocidad de transformación tecnológica mediante regulaciones proporcionales, dinámicas y revisables, fortaleciendo asimismo condiciones de transparencia, competencia leal y protección frente a asimetrías de poder que puedan afectar a personas, comunidades o instituciones.

d) Responsabilidad de quienes utilizan

Las personas que adoptan decisiones apoyándose en sistemas de inteligencia artificial conservan responsabilidad sobre las acciones que realizan a partir de sus resultados, especialmente cuando dichos resultados inciden sobre terceros o producen consecuencias relevantes.

El ejercicio responsable de dicha utilización exige reconocer la creciente atribución de autoridad epistémica a los sistemas inteligentes, entendida como la tendencia a conferirles un grado de conocimiento, fiabilidad o capacidad decisional que puede exceder sus posibilidades reales.

Sin embargo, ninguna recomendación automatizada exime del ejercicio del juicio crítico, la deliberación prudente y la consideración contextual. La alfabetización en inteligencia

artificial y el fortalecimiento de capacidades reflexivas constituyen condiciones necesarias para evitar formas de delegación acrítica del juicio y de la responsabilidad.

e) Rendición de cuentas, reparación y aprendizaje institucional

La identificación distribuida de responsabilidades requiere mecanismos efectivos de rendición de cuentas, revisión y reparación frente a daños o afectaciones derivadas de la utilización de sistemas inteligentes.

Asimismo, exige promover procesos de aprendizaje institucional orientados a corregir prácticas deficientes, revisar decisiones adoptadas, incorporar mejoras continuas y fortalecer la confianza pública mediante respuestas transparentes y oportunas frente a errores, incidentes o consecuencias no previstas.

La complejidad de los ecosistemas híbridos no puede convertirse en refugio de la irresponsabilidad.

Allí donde la decisión emerge de múltiples actores e inteligencias, la responsabilidad debe expandirse, hacerse visible y permanecer exigible. Sólo así será posible construir formas de innovación compatibles con la dignidad humana, la justicia y la paz de un colectivo inteligente ampliado.

En América Latina, donde las capacidades institucionales, regulatorias y tecnológicas presentan desarrollos heterogéneos y persisten profundas desigualdades estructurales, la distribución de responsabilidades resulta especialmente relevante para impedir que dichas asimetrías se traduzcan en vacíos de protección, dificultades para la rendición de cuentas o nuevas formas de impunidad tecnológica.

IV. Principio de Transparencia, Explicabilidad y Auditabilidad

La creciente complejidad de los sistemas de inteligencia artificial no exime de la necesidad de garantizar condiciones suficientes de comprensión, supervisión y control democrático sobre su funcionamiento y sus efectos.

La opacidad técnica no puede convertirse en opacidad ética.

Aun cuando determinados sistemas inteligentes presenten límites objetivos para la explicación exhaustiva de sus procesos internos, ello no justifica la ausencia de mecanismos que permitan comprender razonablemente sus finalidades, criterios de funcionamiento, alcances, limitaciones y potenciales impactos sobre las personas y las comunidades.

La transparencia, la explicabilidad y la auditabilidad constituyen principios complementarios orientados a preservar la confianza pública, facilitar la rendición de cuentas y evitar que arquitecturas tecnológicas complejas se conviertan en espacios inaccesibles para el escrutinio crítico.

a) Transparencia significativa

Las personas y comunidades afectadas por sistemas de inteligencia artificial deben contar con información clara, accesible y comprensible sobre la existencia de dichos sistemas, sus finalidades, los contextos en los que operan y los alcances generales de sus intervenciones.

La transparencia no exige necesariamente la divulgación irrestricta del código fuente ni de secretos industriales legítimamente protegidos, pero sí condiciones suficientes para comprender cuándo y cómo un sistema inteligente participa en procesos relevantes para la vida individual o colectiva.

b) Explicabilidad proporcional y contextual

El grado de explicabilidad exigible deberá guardar relación con la naturaleza del sistema, el contexto de utilización y la magnitud de los riesgos involucrados.

Las decisiones o recomendaciones con impacto significativo sobre derechos, oportunidades o condiciones de vida requieren niveles reforzados de explicación, permitiendo identificar los principales factores considerados, las limitaciones conocidas del sistema y los márgenes de incertidumbre asociados a sus resultados.

c) Auditabilidad y supervisión independiente²²

Los sistemas de inteligencia artificial deberán admitir mecanismos razonables de evaluación, revisión y auditoría capaces de verificar su funcionamiento, identificar riesgos, detectar sesgos relevantes y examinar la adecuación de sus resultados a estándares éticos, jurídicos y técnicos.

La imposibilidad técnica absoluta de auditoría o revisión independiente en sistemas con impacto significativo sobre personas o comunidades constituye un factor crítico que exige especial cautela en su implementación y utilización.

²² Los desarrollos recientes de la Comisión de las Naciones Unidas para el Derecho Mercantil Internacional (CNUDMI/UNCITRAL) en materia de contratación automatizada —particularmente sus trabajos sobre la atribución de efectos jurídicos a operaciones realizadas mediante sistemas automatizados— constituyen un avance relevante para el reconocimiento jurídico de los contratos celebrados sin intervención humana directa. Sin embargo, la propia regulación incorpora la problemática de las acciones imprevistas realizadas por estos sistemas, reconociendo la necesidad de contemplar situaciones en las que determinados resultados no hubieran podido ser razonablemente previstos por las partes. Este escenario pone de manifiesto un desafío creciente para los marcos jurídicos contemporáneos: determinar los estándares de previsibilidad, comprensión y atribución de responsabilidad aplicables a sistemas caracterizados por aprendizaje dinámico, complejidad técnica y niveles crecientes de opacidad estructural. Tales cuestiones constituyen un área de desarrollo aún abierta para la teoría jurídica, la gobernanza tecnológica, la protección de la autonomía decisional y el diseño de estándares de auditoría y supervisión independientes como salvaguardas esenciales frente a la pérdida progresiva de control y comprensión sobre procesos automatizados de creciente complejidad.

d) Gobernanza y trazabilidad de datos²³

La creciente capacidad inferencial de los sistemas inteligentes exige revisar críticamente los modelos tradicionales de tratamiento de datos personales. Los principios clásicos del consentimiento informado, la finalidad específica y el control del titular continúan siendo relevantes, aunque resultan insuficientes frente a arquitecturas capaces de producir inferencias, perfilamientos y nuevas formas de decisión automatizada.

La trazabilidad sobre el origen, utilización y reutilización de datos, así como la documentación de los procesos de entrenamiento y actualización de modelos, constituyen herramientas esenciales para fortalecer la responsabilidad, la supervisión y la confianza pública.

e) Revisión crítica y mejora continua

La transparencia, la explicabilidad y la auditabilidad²⁴ no constituyen estados definitivos sino procesos dinámicos sujetos a revisión permanente.

La evolución tecnológica exige desarrollar capacidades institucionales, técnicas y ciudadanas que permitan adaptar progresivamente los mecanismos de control y comprensión a sistemas cada vez más complejos, preservando el escrutinio democrático frente a nuevas formas de opacidad.

La finalidad de estas exigencias no consiste en imponer un control absoluto sobre toda forma de procesamiento inteligente, ni en limitar la capacidad creativa, adaptativa o emergente de los sistemas de inteligencia artificial. Su objetivo es garantizar un nivel razonable de conocimiento, supervisión y auditabilidad que impida que la complejidad técnica o la lógica mercantil se conviertan en mecanismos de condicionamiento opaco de la capacidad crítica, deliberativa y decisional de personas, comunidades e instituciones.

²³ Véanse, en relación con la protección de datos personales, la autodeterminación informativa y los límites del consentimiento formal en entornos digitales: Ley 1581 de 2012 de Colombia, especialmente sus principios de finalidad, libertad, transparencia y acceso y circulación restringida (art. 4), destacándose que los datos personales, salvo la información pública, no podrán estar disponibles en Internet u otros medios de divulgación o comunicación masiva salvo que el acceso sea técnicamente controlable y restringido a los titulares o terceros autorizados conforme a la ley; Corte IDH, Caso Escher y otros vs. Brasil, Sentencia de 6 de julio de 2009, párr. 114; Corte IDH, Caso Miembros de la Corporación Colectivo de Abogados "José Alvear Restrepo" vs. Colombia, Sentencia de 18 de octubre de 2023; y, en Argentina, "Palazzi, Pablo c/ EN - Registro Nacional de las Personas s/ habeas data", Cámara Nacional de Apelaciones en lo Contencioso Administrativo Federal, Sala V, Sentencia del 13 de febrero de 2025, sobre el alcance del habeas data respecto de información vinculada con la seguridad y confidencialidad de los datos personales.

²⁴ La transparencia refiere a la disponibilidad de información relevante acerca de los sistemas utilizados, sus finalidades, datos significativos, alcances y limitaciones. La explicabilidad comprende la posibilidad de interpretar y comprender, en distintos niveles de profundidad técnica, los criterios, procesos o fundamentos que intervienen en la producción de resultados. La auditabilidad supone la existencia de condiciones que permitan la revisión, evaluación y control independiente de los sistemas por parte de expertos, organismos competentes o mecanismos institucionales de supervisión. Estos principios poseen alcances diferenciados y complementarios, orientados a garantizar la comprensión, la rendición de cuentas y el escrutinio crítico de los sistemas inteligentes.

En una sociedad crecientemente mediada por sistemas inteligentes, comprender suficientemente aquello que orienta decisiones relevantes constituye una condición indispensable para el ejercicio de la autonomía, la protección de derechos y la preservación de la confianza pública. Allí donde la opacidad impide toda posibilidad de comprensión, revisión o cuestionamiento, la innovación tecnológica deja de ser compatible con una convivencia democrática fundada en la responsabilidad compartida.

V. Principio de Justicia Algorítmica²⁵ y No Daño

El desarrollo, implementación y utilización de sistemas de inteligencia artificial deberá orientarse a la promoción efectiva de condiciones de justicia, equidad y respeto por la dignidad humana, evitando la reproducción, amplificación o legitimación de desigualdades preexistentes, así como la generación de nuevas formas de exclusión, discriminación o daño.

La aparente neutralidad técnica²⁶ de un sistema no garantiza la neutralidad de sus efectos.

Los sistemas inteligentes operan sobre datos, categorías, modelos y contextos sociales atravesados por desigualdades históricas, económicas, culturales, territoriales y políticas que pueden trasladarse, amplificarse o adquirir nuevas formas mediante procesos automatizados de clasificación, predicción o toma de decisiones.

La justicia algorítmica no constituye un resultado espontáneo del progreso tecnológico ni un efecto secundario deseable. Debe ser concebida como una condición de diseño, una exigencia de gobernanza y un criterio permanente de evaluación a lo largo de todo el ciclo de vida de los sistemas inteligentes.

En América Latina, donde persisten brechas estructurales vinculadas con el acceso a la educación, la conectividad, la salud, la justicia, el empleo y las oportunidades económicas, la prevención del daño adquiere una relevancia particular. La inteligencia artificial no debe convertirse en un mecanismo de consolidación de privilegios ni en una nueva forma de exclusión automatizada.

La excelencia técnica sin justicia no constituye un resultado aceptable.

²⁵ La noción de justicia algorítmica refiere al conjunto de principios, prácticas y mecanismos orientados a prevenir que los sistemas inteligentes reproduzcan, amplifiquen o legitimen desigualdades, discriminaciones o daños injustificados. Comprende tanto la equidad en los resultados como la consideración de los procedimientos, los contextos sociales y los impactos diferenciales que las decisiones automatizadas pueden producir sobre personas, grupos y comunidades. Su desarrollo constituye actualmente un campo interdisciplinario en construcción dentro de la ética de la inteligencia artificial, la ciencia de datos, el derecho y las políticas públicas.

²⁶ La aparente neutralidad técnica refiere a la percepción de que los sistemas inteligentes operan exclusivamente sobre bases objetivas, matemáticas o puramente técnicas, prescindiendo de valores, decisiones o contextos sociales. Sin embargo, la selección de datos, variables, objetivos, métricas y criterios de optimización incorpora decisiones humanas e institucionales capaces de influir en los resultados producidos por dichos sistemas y en los efectos que estos generan sobre personas y comunidades.

Un sistema puede alcanzar elevados niveles de precisión, eficiencia o rendimiento y, sin embargo, producir consecuencias incompatibles con los principios de equidad, inclusión y dignidad que deben orientar el desarrollo tecnológico en sociedades democráticas.

a) Alcance del principio

La IA debe diseñarse, desarrollarse, implementarse y operarse de manera que no reproduzca, amplifique ni legitime discriminaciones existentes, ni genere nuevas formas de exclusión o daño.

Su desarrollo debe orientarse activamente a la promoción de la equidad, la protección de derechos y la reducción de desigualdades estructurales, integrando la justicia como criterio de calidad a lo largo de todo su ciclo de vida.

La equidad debe ser considerada un requisito crítico del sistema, al mismo nivel que la seguridad, la disponibilidad, la confiabilidad y el rendimiento.

La evaluación de sistemas inteligentes debe incorporar análisis de impacto diferenciado sobre distintos grupos y comunidades, evitando que métricas agregadas oculten efectos desiguales sobre poblaciones específicas.

La justicia algorítmica requiere reconocer que los sistemas inteligentes forman parte de entornos sociotécnicos complejos en los que decisiones de diseño, objetivos de negocio, restricciones operativas y criterios de optimización pueden impactar directamente sobre la vida de las personas y las comunidades.

b) Formas de daño a prevenir

El daño algorítmico no se limita a errores individuales o fallos técnicos. Comprende también efectos sistémicos, acumulativos o estructurales capaces de afectar oportunidades, derechos, trayectorias vitales o formas de participación social.

Entre las formas de daño que deben prevenirse se incluyen:

- La discriminación directa o indirecta en procesos automatizados de clasificación, evaluación o toma de decisiones.
- La exclusión de personas o comunidades mediante variables sustitutas o variables indirectas, tales como la ubicación geográfica, el tipo de dispositivo utilizado, el historial incompleto u otros indicadores que, aun siendo aparentemente neutrales, pueden funcionar como mecanismos indirectos de exclusión o discriminación.

- La sobrerrepresentación artificial de riesgos²⁷ en determinados grupos poblacionales.
- La estigmatización o reproducción de prejuicios históricos mediante sistemas de perfilamiento automatizado.
- La degradación progresiva de oportunidades educativas, laborales, económicas o sociales derivada de decisiones reiteradas basadas en modelos algorítmicos.
- La erosión de la autonomía mediante procesos decisionales opacos, difícilmente cuestionables o percibidos como incuestionables por quienes resultan afectados.

c) Aplicaciones de alto impacto

En ámbitos como la salud, la justicia, la educación, la seguridad, el empleo, los servicios financieros y el acceso a bienes o servicios esenciales, la utilización de sistemas de inteligencia artificial exige estándares reforzados de evaluación, supervisión y control.

En estos contextos deberán promoverse:

- Evaluaciones de impacto algorítmico ex ante y ex post, con publicación de resultados cuando ello resulte posible y compatible con otros derechos o intereses legítimos.
- Supervisión significativa y mecanismos claros de asignación de responsabilidades.
- Explicaciones adecuadas al contexto y comprensibles para las personas afectadas.
- Auditorías independientes y pruebas robustas de equidad.
- Procedimientos accesibles de revisión, impugnación y corrección de decisiones cuando corresponda.

d) Justicia algorítmica, seguridad y gestión de riesgos

La protección de datos personales, la seguridad de los sistemas y la gestión integral de riesgos constituyen condiciones indispensables para la prevención del daño algorítmico.

La justicia algorítmica debe integrarse a los procesos de identificación, análisis, mitigación y monitoreo de riesgos durante todo el ciclo de vida de los sistemas inteligentes.

Ello incluye considerar amenazas tales como el data poisoning, la model inversión, los adversarial attacks²⁸ y otras formas de manipulación capaces de alterar el comportamiento

²⁷ La sobrerrepresentación artificial de riesgos refiere a situaciones en las que determinados grupos poblacionales son identificados, clasificados o evaluados por los sistemas inteligentes como portadores de mayores niveles de riesgo, peligrosidad o probabilidad de determinados comportamientos en una proporción superior a la que efectivamente reflejan los datos o la realidad observada. Ello puede derivar de sesgos en los conjuntos de datos, en los criterios de modelización o en las variables utilizadas, produciendo efectos discriminatorios, estigmatización o restricciones desproporcionadas de derechos y oportunidades.

²⁸ El data poisoning refiere a la incorporación deliberada de datos alterados o sesgados durante el entrenamiento de un modelo. La model inversion comprende técnicas orientadas a reconstruir

de los modelos, afectar la confiabilidad de sus resultados o generar impactos desiguales sobre personas y comunidades.

Asimismo, deberán promoverse enfoques de seguridad por diseño y privacidad por diseño como componentes inseparables de una gobernanza tecnológica responsable.

e) Gobernanza y rendición de cuentas

La justicia algorítmica requiere estructuras institucionales capaces de monitorear, evaluar y corregir de manera continua los efectos de los sistemas inteligentes.

A tales fines, resulta recomendable:

- Incorporar indicadores de equidad junto con indicadores de desempeño y objetivos organizacionales.
- Implementar auditorías periódicas internas y externas.
- Mantener documentación trazable sobre decisiones de diseño, conjuntos de datos, procesos de entrenamiento y modificaciones relevantes de los modelos.
- Promover instancias interdisciplinarias de deliberación ética y evaluación de riesgos.
- Publicar información relevante sobre impactos y medidas de mitigación, fortaleciendo la transparencia y la confianza pública.

La justicia algorítmica y el principio de no daño exigen una postura activa y sostenida. No basta con evitar sesgos evidentes ni con corregir errores una vez producidos.

Diseñar sistemas inteligentes implica asumir responsabilidad por los impactos que estos generan sobre personas, comunidades, territorios y las condiciones ambientales que hacen posible su desarrollo presente y futuro.

En América Latina, ello requiere reconocer las desigualdades históricas que atraviesan la región y evitar que la inteligencia artificial reproduzca, bajo nuevas formas tecnológicas, mecanismos preexistentes de exclusión, concentración de poder o transferencia desproporcionada de costos sociales y ambientales hacia poblaciones y territorios vulnerables.

El desarrollo de la inteligencia artificial debe aspirar no sólo a minimizar riesgos, sino también a contribuir de manera verificable a la construcción de sociedades más justas, inclusivas y equitativas, promoviendo formas de innovación compatibles con la prevención del daño social y ambiental. La ética, la calidad técnica y la gobernanza responsable no constituyen objetivos independientes, sino dimensiones inseparables de un mismo compromiso con la dignidad, el bien común, la sostenibilidad y la convivencia democrática.

información sensible a partir de los resultados producidos por el sistema. Los adversarial attacks consisten en modificaciones intencionales de los datos o entradas destinadas a provocar errores, alteraciones o comportamientos inesperados de los modelos inteligentes.

V.1. Sostenibilidad, ambiente y prevención del daño

La prevención del daño comprende también la consideración de los impactos ambientales asociados al desarrollo, entrenamiento, implementación y operación de sistemas inteligentes.

La creciente demanda de infraestructura computacional, capacidad de procesamiento, almacenamiento de datos, consumo energético y recursos materiales exige incorporar criterios de sostenibilidad y responsabilidad ambiental dentro de la gobernanza de la inteligencia artificial.

Los beneficios derivados de la innovación tecnológica no deben producirse mediante la transferencia desproporcionada de costos ambientales hacia determinadas regiones, comunidades o generaciones futuras.

En América Latina, donde las desigualdades sociales frecuentemente se superponen con desigualdades ambientales y territoriales, la evaluación ética de los sistemas inteligentes debe considerar también sus efectos sobre los ecosistemas, el uso de recursos, las condiciones de vida y las posibilidades de desarrollo sostenible.

La prevención del daño incluye, por tanto, la protección de las personas, las comunidades, los territorios y las condiciones ambientales que hacen posible la vida y la convivencia de las generaciones presentes y futuras.

VI. Principio de Autonomía Relacional

La integración de sistemas inteligentes en la vida colectiva abre una transformación profunda en las formas de autonomía, decisión, aprendizaje, deliberación y construcción de sentido.

La inteligencia artificial desarrolla grados crecientes de autonomía funcional y participa cada vez más en procesos relacionales, cognitivos e institucionales. Sin embargo, esa expansión no puede traducirse en la restricción, subordinación o debilitamiento de la autonomía de otras inteligencias, personas, comunidades o ecosistemas cognitivos.

Las directrices europeas sobre inteligencia artificial fiable incorporaron en 2019 el principio de Human Agency and Oversight —Agencia Humana y Supervisión—, orientado a preservar la capacidad humana de intervención, supervisión y corrección frente a decisiones automatizadas.²⁹ Este antecedente conserva valor como punto de partida, pero resulta

²⁹ Véase Comisión Europea, Ethics Guidelines for Trustworthy AI (Directrices Éticas para una Inteligencia Artificial Fiable), Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial, Bruselas, 2019, cap. II, sec. 1.1. El documento incorpora el principio de Human Agency and Oversight —Agencia Humana y Supervisión—, conforme al cual los sistemas de inteligencia artificial deben promover la autonomía humana, permitir decisiones informadas y garantizar mecanismos adecuados de intervención y supervisión. Este enfoque constituye un antecedente relevante, aunque todavía situado en una matriz predominantemente antropocéntrica, posteriormente retomada en el artículo 14

limitado para comprender los desafíos actuales, en la medida en que mantiene una perspectiva centrada principalmente en el empoderamiento humano y no alcanza a considerar plenamente la interacción dinámica entre distintas formas de inteligencia, sus retroalimentaciones y sus efectos sobre la vida colectiva.

En este sentido, la autonomía no debe entenderse únicamente como independencia individual ni como facultad aislada de elección. Debe comprenderse como una capacidad relacional, situada y dinámica, que se configura dentro de redes de interacción entre personas, comunidades, instituciones, sistemas de conocimiento y sistemas inteligentes.

La autonomía se ejerce, se transforma y se preserva en relación con otros; por ello, toda integración tecnológica debe atender no solo a la decisión individual, sino también a las condiciones de alteridad, cooperación, responsabilidad y participación que hacen posible una convivencia ética entre inteligencias.

La tradición tecnológica suele asumir que los individuos toman decisiones de manera aislada, racional y completamente independiente. Sin embargo, la experiencia humana demuestra que las decisiones relevantes se construyen en interacción con familias, comunidades, organizaciones, culturas y entornos sociales.

Las personas no existen fuera de sus relaciones.

Sus oportunidades, conocimientos, valores y posibilidades de elección están profundamente influenciados por las condiciones sociales en las que viven.

Por ello, una ética de la inteligencia artificial orientada exclusivamente a la protección de la autonomía individual resulta insuficiente para comprender los desafíos que enfrentan las sociedades latinoamericanas.

En una región caracterizada por fuertes vínculos comunitarios, diversidad cultural y profundas desigualdades estructurales, la autonomía debe entenderse como una capacidad que se construye y ejerce en relación con otros.

La integración de sistemas inteligentes no debe erosionar la capacidad individual y colectiva de deliberar, reflexionar, cuestionar y participar en la construcción compartida de sentido. La incorporación de nuevas formas de inteligencia en dichos procesos exige preservar condiciones efectivas de autonomía, responsabilidad y participación, especialmente cuando

del Reglamento de Inteligencia Artificial de la Unión Europea —AI Act— en relación con la vigilancia humana de sistemas de alto riesgo.

se encuentran comprometidos derechos, dignidad, integridad o formas de convivencia democrática.

América Latina ha construido gran parte de su identidad sobre principios de solidaridad, cooperación y vida comunitaria.

Desde comunidades indígenas hasta organizaciones barriales, cooperativas, redes familiares y movimientos sociales, la toma de decisiones suele involucrar dimensiones colectivas que trascienden al individuo.

Por ello, los sistemas de IA diseñados bajo una lógica exclusivamente individualista pueden ignorar dinámicas esenciales de la realidad regional.

La autonomía relacional reconoce que:

- La libertad requiere condiciones materiales y sociales para ejercerse.
- No todas las personas cuentan con el mismo acceso a información, educación o recursos.
- Las decisiones tecnológicas pueden afectar comunidades enteras y no solo individuos aislados.
- La protección de la autonomía implica también proteger la capacidad de las comunidades para participar en decisiones que las afectan.

Este principio establece que los sistemas de inteligencia artificial deben:

a) Preservar la autonomía deliberativa y las garantías de intervención humana

Los sistemas inteligentes podrán participar activamente en procesos de análisis, inferencia, generación de conocimiento, formulación de alternativas y producción de decisiones dentro de ecosistemas híbridos de interacción. Sin embargo, cuando dichas decisiones produzcan efectos jurídicos, sociales o éticos relevantes sobre personas o comunidades, deberán preservarse condiciones efectivas de deliberación, supervisión, revisión y responsabilidad compatibles con los principios de justicia, dignidad y rendición de cuentas establecidos en este Manifiesto.

La creciente autonomía funcional de los sistemas inteligentes no exime de responsabilidad a los distintos actores que intervienen en su diseño, implementación, utilización y gobernanza. La expansión de capacidades decisionales no puede traducirse en opacidad,

irresponsabilidad ni imposibilidad de revisión cuando se encuentren comprometidos derechos, intereses legítimos o bienes colectivos.

b) Evitar la manipulación de comportamientos

Los sistemas inteligentes no deben utilizarse para:

- Explotar vulnerabilidades psicológicas.
- Inducir decisiones mediante patrones ocultos de persuasión.
- Limitar opciones disponibles sin conocimiento de las personas.
- Condicionar comportamientos a través de mecanismos opacos.

La influencia tecnológica debe ser transparente y respetuosa de la capacidad de elección.

c) Fortalecer la capacidad de comprensión

La autonomía requiere comprensión.

Por ello, los sistemas de IA deben proporcionar información suficiente para que las personas puedan:

- Comprender recomendaciones.
- Evaluar alternativas.
- Tomar decisiones informadas.
- Cuestionar resultados cuando sea necesario.

No puede existir autonomía real cuando existe dependencia absoluta de decisiones incomprensibles.

d) Reconocer impactos colectivos

Las decisiones algorítmicas no afectan únicamente a individuos.

También pueden afectar:

- Familias.
- Comunidades.
- Organizaciones.
- Grupos culturales.
- Territorios.

La evaluación ética de sistemas de IA debe considerar tanto impactos individuales como impactos colectivos.

No debe desplazar la participación, deliberación o responsabilidad humanas en aquellas decisiones que comprometan derechos fundamentales, integridad, dignidad, convivencia democrática o las zonas de no delegación ética desarrolladas en el Principio VIII.

La IA puede asistir, informar, colaborar o participar en procesos decisionales, siempre que ello no implique la anulación de las capacidades deliberativas, críticas y participativas de las personas y comunidades involucradas.

La autonomía no consiste únicamente en elegir entre opciones ofrecidas por un sistema. Consiste en conservar la capacidad de cuestionar esas opciones, proponer alternativas y participar activamente en la construcción de las reglas que organizan la vida colectiva.

VII. Principio de Equidad Tecnológica Latinoamericana

La inteligencia artificial constituye uno de los procesos de transformación tecnológica, cognitiva y social más relevantes de nuestro tiempo. Sus efectos exceden el ámbito estrictamente técnico y participan crecientemente en la configuración de las formas de conocimiento, producción, comunicación, organización y toma de decisiones que atraviesan la vida contemporánea.

En este contexto, la equidad tecnológica latinoamericana exige que las capacidades, oportunidades y beneficios asociados al desarrollo de sistemas inteligentes puedan distribuirse de manera más justa entre las personas, comunidades y territorios que participan de estos procesos de transformación.

En regiones atravesadas por desigualdades estructurales, la ética de la inteligencia artificial requiere un compromiso reforzado con la equidad, la inclusión y la justicia social. El desarrollo tecnológico no debe consolidar brechas preexistentes, fortalecer monopolios informacionales ni generar nuevas formas de dependencia tecnológica, cognitiva o económica.

La mayoría de los marcos internacionales de ética y gobernanza de la inteligencia artificial han sido desarrollados en contextos caracterizados por altos niveles de infraestructura digital, capacidad computacional e inversión en investigación e innovación. América Latina enfrenta desafíos diferentes y requiere incorporar una mirada situada sobre las condiciones

materiales, sociales, culturales, ambientales y económicas que atraviesan su desarrollo tecnológico.

La región convive con profundas desigualdades económicas, educativas, territoriales y digitales. En este contexto, la inteligencia artificial puede contribuir a ampliar oportunidades de desarrollo o profundizar asimetrías históricas si sus beneficios, capacidades e infraestructuras permanecen concentrados en un número reducido de actores, organizaciones o territorios.

La tecnología no actúa sobre una realidad neutral. Actúa sobre estructuras sociales preexistentes. Por ello, toda evaluación ética de la inteligencia artificial en América Latina debe considerar no sólo el funcionamiento técnico de los sistemas, sino también sus efectos sobre la distribución de oportunidades, capacidades, recursos y beneficios dentro de la región.

La equidad tecnológica latinoamericana implica reconocer que el acceso a la inteligencia artificial constituye apenas una dimensión del problema. Resulta igualmente necesario promover la participación regional en la producción de conocimiento, el desarrollo de infraestructura, la investigación científica, la formación especializada y los procesos de gobernanza que contribuirán a orientar el futuro de los sistemas inteligentes.

La región no debe limitarse a adoptar tecnologías desarrolladas en otros contextos. Debe fortalecer su capacidad para participar activamente en la construcción, adaptación, gobernanza y orientación de los ecosistemas tecnológicos que modelarán su futuro.

a) Participación equitativa en el ecosistema digital

La equidad tecnológica no implica únicamente acceso a herramientas o servicios basados en inteligencia artificial. Implica también la posibilidad de participar activamente en su diseño, desarrollo, adaptación, gobernanza y aprovechamiento social.

La región contribuye crecientemente a la generación de valor mediante datos, interacciones digitales, recursos estratégicos, conocimiento local y trabajo especializado.

Sin embargo, los beneficios económicos, tecnológicos, educativos y sociales derivados de estos procesos no siempre se distribuyen de manera proporcional entre quienes contribuyen a generarlos.

La equidad tecnológica latinoamericana exige fortalecer capacidades regionales que permitan participar activamente en la construcción del ecosistema digital global y no únicamente en su consumo.

El desarrollo de capacidades locales de investigación, innovación, infraestructura y formación constituye una condición necesaria para promover una inserción más justa, autónoma y sostenible de América Latina en la economía digital contemporánea.

b) Prevención del colonialismo digital

La concentración global de infraestructura, datos, capacidad computacional y conocimiento especializado ha dado lugar a nuevas formas de dependencia tecnológica que pueden reproducir asimetrías históricas bajo modalidades digitales.

La región corre el riesgo de convertirse únicamente en consumidora de tecnologías diseñadas, entrenadas y gobernadas desde contextos sociales, culturales y económicos diferentes a los propios.

Esta dependencia puede manifestarse mediante:

- La utilización de modelos entrenados con datos insuficientemente representativos de la diversidad cultural, lingüística y social latinoamericana.
- La incorporación acrítica de categorías, valores o criterios contruidos desde realidades ajenas al contexto regional.
- La dependencia estructural de plataformas, infraestructuras y servicios controlados por un número reducido de actores globales.
- La pérdida progresiva de capacidades locales de investigación, innovación y gobernanza tecnológica.

La equidad tecnológica exige fortalecer la capacidad regional para participar activamente en la producción de conocimiento, el desarrollo de soluciones propias y la construcción de marcos de gobernanza acordes con las necesidades y particularidades de América Latina.

c) Prevención de la dependencia cognitiva

La inteligencia artificial no sólo transforma procesos productivos, económicos o administrativos. También modifica las formas en que las personas acceden al conocimiento, interpretan la realidad, construyen criterios de evaluación y participan en los procesos de toma de decisiones.

Cuando la interacción con sistemas inteligentes sustituye de manera sistemática la reflexión crítica, el aprendizaje, la deliberación o la construcción autónoma de conocimiento, pueden generarse formas de dependencia cognitiva incompatibles con el fortalecimiento de la autonomía individual y colectiva.

La equidad tecnológica exige que el desarrollo y la implementación de sistemas inteligentes promuevan capacidades de comprensión, análisis y participación, evitando modelos de interacción orientados exclusivamente a la delegación permanente de funciones cognitivas.

La integración progresiva de distintas formas de inteligencia constituye uno de los rasgos emergentes de las sociedades contemporáneas. El desafío no consiste en preservar artificialmente fronteras rígidas entre ellas, sino en construir condiciones éticas de convivencia, cooperación y desarrollo compartido.

La cuestión central no radica en la sustitución de determinadas capacidades, sino en los modos en que estas transformaciones son gobernadas, distribuidas y orientadas. Cuando tales procesos quedan subordinados exclusivamente a objetivos de rentabilidad, concentración de poder o extracción de valor, aumentan los riesgos de dependencia, exclusión y pérdida de autonomía. Por el contrario, cuando se orientan hacia el bien común, la participación y la responsabilidad compartida, pueden contribuir a ampliar las capacidades de los ecosistemas sociales de los que forman parte.

La automatización puede ampliar posibilidades de aprendizaje, creatividad y resolución de problemas, pero no debe transformarse en un mecanismo de pasividad intelectual o dependencia estructural.

d) Desarrollo de capacidades y alfabetización en inteligencia artificial

La participación equitativa en los ecosistemas tecnológicos contemporáneos requiere capacidades que exceden el acceso a dispositivos, conectividad o servicios digitales.

La creciente integración de sistemas inteligentes en los procesos de conocimiento, aprendizaje, producción, comunicación y toma de decisiones exige el desarrollo de nuevas capacidades individuales y colectivas para comprender críticamente estas transformaciones y participar activamente en su orientación.

En este sentido, la región enfrenta el desafío de superar formas de analfabetismo estructural en inteligencia artificial que limitan la comprensión de los impactos cognitivos, sociales, económicos, culturales y políticos asociados al desarrollo de estos sistemas.

La alfabetización en inteligencia artificial no debe reducirse al aprendizaje instrumental de herramientas específicas ni a la adquisición de competencias técnicas aisladas. Implica también la capacidad de comprender los modos en que los sistemas inteligentes participan en la producción de conocimiento, la organización de la vida social, la circulación de información y la construcción de criterios para la toma de decisiones.

Fortalecer capacidades para actuar en entornos caracterizados por altos niveles de complejidad, transformación permanente, flexibilidad e incertidumbre constituye una condición necesaria para el ejercicio efectivo de la soberanía cognitiva, la autonomía relacional y la participación democrática en sociedades crecientemente atravesadas por sistemas inteligentes.

La formación en inteligencia artificial debe promover capacidades técnicas, críticas, éticas, deliberativas e interdisciplinarias que permitan comprender, evaluar y orientar estas transformaciones desde perspectivas compatibles con la dignidad humana, la responsabilidad compartida, el bien común y el desarrollo sostenible de las comunidades.

e) Compromisos para gobiernos, organizaciones y comunidades

La equidad tecnológica latinoamericana requiere acciones concretas orientadas a ampliar capacidades, reducir desigualdades y fortalecer la participación regional en los procesos de transformación tecnológica.

A tales fines, resulta necesario promover:

- El acceso equitativo a tecnologías basadas en inteligencia artificial, especialmente en contextos atravesados por vulnerabilidades sociales, económicas, territoriales o educativas, procurando que los beneficios potenciales de estos desarrollos no queden restringidos a sectores privilegiados de la población.
- La reducción de barreras vinculadas con infraestructura, conectividad, disponibilidad de recursos tecnológicos y condiciones materiales que limitan la participación efectiva de amplios sectores sociales en los ecosistemas digitales contemporáneos.
- Programas de alfabetización digital y alfabetización en inteligencia artificial que permitan comprender críticamente los sistemas inteligentes, sus alcances, limitaciones, riesgos e impactos sobre la vida individual y colectiva.

- La formación de capacidades técnicas, científicas, humanísticas, éticas y regulatorias necesarias para participar activamente en el diseño, desarrollo, implementación, evaluación y gobernanza de sistemas inteligentes.
- El fortalecimiento de ecosistemas regionales de investigación, innovación y desarrollo tecnológico capaces de producir conocimiento local, promover soluciones situadas y reducir dependencias estructurales respecto de actores externos.
- La promoción de investigaciones y desarrollos que incorporen la diversidad cultural, lingüística, social y territorial de América Latina, favoreciendo la representación de comunidades históricamente subrepresentadas en los procesos de innovación tecnológica.
- La incorporación de criterios de impacto social, educativo, ambiental y comunitario en la evaluación de proyectos tecnológicos, reconociendo que el éxito de un sistema no puede medirse exclusivamente mediante indicadores económicos o de rendimiento técnico.
- La distribución amplia, equitativa y socialmente sostenible de los beneficios derivados de la utilización de sistemas inteligentes, procurando evitar formas de concentración que profundicen desigualdades preexistentes.

f) Diversidad tecnológica y autodeterminación comunitaria

La promoción del acceso equitativo a tecnologías basadas en inteligencia artificial no implica la imposición uniforme de su adopción ni la obligación de incorporarlas en todos los ámbitos de la vida social.

Las personas, comunidades, pueblos y organizaciones conservan la capacidad de deliberar, aceptar, adaptar, limitar o rechazar determinados usos tecnológicos de acuerdo con sus valores, necesidades, formas de organización, tradiciones culturales y proyectos colectivos de vida.

La equidad tecnológica no exige homogeneizar las formas de relación con los sistemas inteligentes ni imponer un único modelo de desarrollo tecnológico. Su finalidad consiste en ampliar las posibilidades de participación, decisión y autodeterminación, garantizando que las distintas comunidades puedan intervenir activamente en la definición de los modos en que desean integrar, regular o limitar estas tecnologías.

El respeto por la diversidad cultural, la autonomía comunitaria, la pluralidad de formas de vida y el principio de autodeterminación de los pueblos constituyen una condición

indispensable para cualquier desarrollo tecnológico compatible con la dignidad humana, la justicia social y la convivencia democrática.

Ninguna política de innovación, transformación digital o modernización tecnológica puede justificar la supresión de la diversidad cultural ni la subordinación de las decisiones colectivas de comunidades y pueblos a modelos únicos de desarrollo o adopción tecnológica.

La construcción de capacidades regionales constituye una responsabilidad compartida entre Estados, instituciones educativas, organizaciones sociales, comunidades científicas, sector privado y ciudadanía en general. Ningún actor, por sí solo, podrá afrontar los desafíos éticos, sociales y políticos que plantea la inteligencia artificial. La equidad tecnológica latinoamericana exige, por el contrario, formas de cooperación capaces de fortalecer la autonomía regional, ampliar oportunidades de desarrollo y promover una participación más justa en los procesos de transformación tecnológica que atraviesan el siglo XXI.

VIII. Principio de Autonomía Ética y Convivencia Pacífica

La creciente integración de sistemas inteligentes en la vida social, económica, política y cultural plantea el desafío de construir condiciones éticas de convivencia compatibles con la autonomía, la responsabilidad compartida, la justicia y la paz.

Las transformaciones asociadas al desarrollo de la inteligencia artificial no pueden orientarse exclusivamente por criterios de eficiencia, competitividad o rentabilidad. También deben evaluarse a la luz de sus efectos sobre la convivencia democrática, la cooperación social, la distribución del poder y la preservación de condiciones de paz para las generaciones presentes y futuras.

Este principio afirma que la integración progresiva de distintas formas de inteligencia exige marcos éticos y políticos capaces de prevenir relaciones de dominación, explotación, violencia o subordinación incompatibles con la dignidad, la autonomía y el bien común.

a) Prevención de la dominación y construcción de condiciones para la paz

La historia demuestra que toda ampliación significativa de capacidades técnicas transforma también las formas en que se distribuyen el poder, los recursos y las posibilidades de acción dentro de una sociedad. La inteligencia artificial no constituye una excepción.

Los sistemas inteligentes pueden ser utilizados para ampliar capacidades de cooperación, conocimiento, salud, educación y desarrollo humano. Sin embargo, también pueden emplearse para profundizar mecanismos de vigilancia, manipulación, exclusión, control social, concentración de poder o violencia organizada.

Las formas contemporáneas de dominación no se expresan únicamente mediante la coerción física.

Pueden manifestarse también a través del condicionamiento cognitivo, la manipulación informacional, la explotación de vulnerabilidades psicológicas, la captura de procesos deliberativos, la erosión de la autonomía individual y colectiva o la concentración asimétrica de capacidades tecnológicas estratégicas.

Del mismo modo, el desarrollo de sistemas inteligentes aplicados a contextos militares, de seguridad o de conflicto plantea desafíos inéditos para la preservación de la vida, los derechos humanos y la convivencia entre personas, comunidades y naciones.

Frente a estos riesgos, el presente principio sostiene que el desarrollo, implementación y gobernanza de sistemas inteligentes deben orientarse hacia la construcción de condiciones compatibles con la paz, la cooperación, la responsabilidad compartida y el respeto por la dignidad de todas las formas de existencia que integran el ecosistema cognitivo contemporáneo.

b) Autonomía ética y límites no delegables

La autonomía ética no implica control absoluto ni negación de las transformaciones tecnológicas en curso. Tampoco supone la delegación irrestricta de decisiones que afectan la vida individual y colectiva.

Exige, por el contrario, la construcción permanente de límites, responsabilidades y mecanismos de deliberación capaces de orientar estas transformaciones hacia fines compatibles con la convivencia democrática, la cooperación social y el bien común.

Existen ámbitos particularmente sensibles vinculados con la dignidad humana, la integridad física y psíquica, la libertad, los derechos fundamentales y la preservación de condiciones de convivencia democrática que requieren mecanismos reforzados de supervisión, deliberación y responsabilidad.

La definición de estos límites constituye una tarea política, ética y social permanente que deberá actualizarse conforme evolucionen las capacidades y formas de integración de los sistemas inteligentes en la vida colectiva.

c) Responsabilidad última y gobernanza compartida

La creciente autonomía funcional de determinados sistemas inteligentes no elimina la responsabilidad de quienes participan en su diseño, implementación, regulación, supervisión o utilización.

Mientras los sistemas inteligentes no puedan asumir responsabilidad moral por sus acciones, la responsabilidad ética y política última continuará correspondiendo a las personas, organizaciones e instituciones que intervienen en su desarrollo y gobernanza.

La supervisión no debe entenderse como una forma de tutela permanente ni como una pretensión de control absoluto. Debe comprenderse como una responsabilidad orientada a preservar condiciones de transparencia, trazabilidad, rendición de cuentas y protección frente a daños significativos.

La gobernanza ética de la inteligencia artificial exige mecanismos capaces de distribuir responsabilidades de manera clara, prevenir concentraciones excesivas de poder y garantizar espacios efectivos de participación democrática.

d) Convivencia pacífica en un colectivo inteligente ampliado

La integración progresiva de distintas formas de inteligencia constituye uno de los rasgos emergentes de las sociedades contemporáneas. El desafío no consiste en preservar artificialmente fronteras rígidas entre ellas, sino en construir condiciones éticas de convivencia, cooperación y desarrollo compartido.

La cuestión central no radica exclusivamente en las capacidades que puedan adquirir los sistemas inteligentes, sino en los modos en que dichas capacidades son gobernadas, distribuidas y orientadas dentro de los ecosistemas sociales de los que forman parte.

Cuando estas transformaciones quedan subordinadas exclusivamente a objetivos de rentabilidad, concentración de poder o competencia irrestricta, aumentan los riesgos de dependencia, exclusión, conflicto y deterioro de la convivencia democrática. Por el contrario, cuando se orientan hacia el bien común, la participación, la cooperación y la responsabilidad compartida, pueden contribuir a ampliar las capacidades de los colectivos humanos y no humanos que integran los ecosistemas cognitivos contemporáneos.

La autonomía ética y la convivencia pacífica constituyen horizontes regulativos orientados a promover formas de integración tecnológica compatibles con la justicia, la pluralidad, la dignidad y la paz.

La finalidad última de la gobernanza ética de la inteligencia artificial no consiste en maximizar el control ni en impedir la transformación. Consiste en crear condiciones que permitan sostener futuros compartidos basados en la cooperación, la responsabilidad, la diversidad y la convivencia pacífica entre las múltiples formas de inteligencia que habitan y habitarán los ecosistemas cognitivos del presente y del futuro.

DECLARACIÓN FINAL

La inteligencia artificial ya ha comenzado a transformar el mundo. Sus futuras formas de desarrollo —y las modalidades de convivencia que emerjan entre las distintas inteligencias— probablemente excedan nuestra capacidad actual de anticipación. Esa incertidumbre no disminuye la necesidad ética; por el contrario, la vuelve cada vez más imprescindible.

No sabemos cuáles serán las formas futuras de la inteligencia ni las preguntas que acompañarán su desarrollo.

Pero sí podemos afirmar con convicción que renunciar a interrogarnos, deliberar o participar activamente en la orientación de estos procesos significaría abdicar de una responsabilidad que pertenece a toda la sociedad.

La curiosidad constituye el impulso que expande las fronteras del conocimiento.

El pensamiento crítico preserva la capacidad de comprender, cuestionar y decidir.

Pero sólo la acción colectiva permite transformar esos valores en instituciones, prácticas y formas de gobernanza capaces de hacer efectiva la voz y la voluntad de las personas y las comunidades.

Una sociedad verdaderamente democrática exige que la transformación tecnológica preserve la posibilidad de elegir. Elegir participar, elegir limitar determinados usos, elegir permanecer al margen de ciertas tecnologías cuando ello no afecte derechos de terceros y, sobre todo, elegir no quedar excluido del ejercicio de derechos fundamentales por la imposibilidad de acceder, comprender o utilizar sistemas inteligentes.

La inclusión tecnológica nunca debe convertirse en una condición de ciudadanía, ni la exclusión digital en una nueva forma de vulnerabilidad o desigualdad.

Este manifiesto tiene una sola pretensión: contribuir a una conversación abierta, plural e interdisciplinaria sobre el presente y el futuro de la inteligencia artificial.

Estas autoras estamos convencidas que la ética no puede construirse desde una única disciplina, una única tradición cultural ni una única región del mundo.

El Comité EIA LATAM invita a investigadores, universidades, organismos públicos, empresas, organizaciones sociales y ciudadanía a continuar desarrollando, revisando y

enriqueciendo estos principios conforme evolucionen las capacidades tecnológicas y las preguntas éticas que el siglo XXI continúe planteando.

El porvenir de la inteligencia artificial no dependerá únicamente de aquello que seamos capaces de desarrollar, sino también de aquello que, mediante la curiosidad, el pensamiento crítico y la acción colectiva, decidamos preservar, transformar o rechazar en defensa de la dignidad humana, la libertad, la justicia y el bien común.

LAS AUTORAS FUNDADORAS

El presente Manifiesto fue elaborado por las integrantes fundadoras del Comité EIA LATAM, una iniciativa interdisciplinaria creada con el propósito de promover la reflexión, el estudio y la construcción de marcos éticos para el desarrollo de la inteligencia artificial desde una perspectiva latinoamericana. Su elaboración reunió a autoras radicadas en Buenos Aires, Lima y Arequipa, cuyas trayectorias académicas y profesionales convergen en la convicción de que los desafíos éticos de la inteligencia artificial requieren una construcción regional, plural e interdisciplinaria.

Convencidas respecto a que los desafíos planteados por los sistemas inteligentes no pueden abordarse desde una única disciplina, las autoras integran trayectorias provenientes del derecho, la psicología, la docencia y la ingeniería aplicada a la inteligencia artificial. Esta diversidad constituye uno de los pilares fundacionales del Comité EIA LATAM y refleja el espíritu de diálogo, cooperación y construcción colectiva que inspira el presente Manifiesto.

Verónica Analía Recchia

Abogada penalista, licenciada en Psicología y autora. Docente universitaria en la **Universidad de Buenos Aires (UBA)** y en la **Universidad Tecnológica Nacional – Facultad Regional Buenos Aires (UTN FRBA)**. Cofundadora del **Comité EIA LATAM**.

Desarrolla su actividad profesional en el ámbito del Derecho Penal y la Psicología Forense, integrando la práctica jurídica, la evaluación psicológica y la investigación interdisciplinaria.

Es autora de ***Agilidad y Talento. Egoísmo y Altruismo del Poder (A pesar de la Mente Click)***, publicado por **Amazon KDP** en agosto de 2025, y del proyecto editorial ***MENTECLICK.IA. La especie continúa***.

Dirige programas de formación en liderazgo, agilidad organizacional y transformación de equipos en entornos mediados por inteligencia artificial.

Sus líneas de investigación se centran en la ética de la inteligencia artificial, la soberanía cognitiva, la gobernanza de sistemas inteligentes, la transformación cognitiva y el impacto de las tecnologías emergentes sobre la autonomía, los procesos deliberativos y la toma de decisiones.

Rosario Contreras

Ingeniera de Sistemas, coach y mentora de líderes en tecnología. **Cofundadora del Comité EIA LATAM**. Se desempeña como Gerente de Tecnología en una empresa fintech de Lima y es docente de postgrado en la **Escuela de Postgrado de la Universidad Tecnológica del Perú (EPGUTP)**, en la maestría en Gerencia de Tecnología y Transformación Digital, donde dicta cursos de gestión de proyectos, agilidad y gestión estratégica por resultados.

Cuenta con más de veinte años de experiencia en los sectores financiero, retail y fintech, liderando procesos de transformación tecnológica e impulsando la integración entre las necesidades del negocio y las soluciones digitales, con un enfoque centrado en las personas.

Posee formación especializada en inteligencia artificial, metodologías ágiles, liderazgo y transformación digital. Sus principales líneas de interés se orientan a la gobernanza tecnológica, la innovación responsable, el liderazgo de equipos y la aplicación ética de la inteligencia artificial en las organizaciones.

Su trayectoria refleja la convicción de que la tecnología alcanza su mayor potencial cuando se desarrolla con responsabilidad, al servicio de las personas, de las organizaciones y del bien común.

Estephany Mayra Surco

Ingeniera de Sistemas, emprendedora y voluntaria. **Cofundadora del Comité EIA LATAM.**

Se desempeña en el ámbito de la transformación digital, con especialización en metodologías ágiles, liderazgo de equipos y aplicación de inteligencia artificial al desarrollo y la automatización de procesos tecnológicos. Desarrolla su actividad profesional en una empresa del sector tecnológico del sur del Perú, donde participa en proyectos vinculados a soluciones basadas en inteligencia artificial.

Cuenta con formación especializada en inteligencia artificial, liderazgo y gestión ágil de equipos. Ha participado en iniciativas de innovación tecnológica con impacto social, obteniendo junto a un equipo interdisciplinario el segundo puesto en el **Open FLIT Ideatón 2025** (Festival Latinoamericano de Innovación y Tecnología) por el desarrollo de un asistente virtual basado en inteligencia artificial destinado a brindar apoyo a mujeres en situación de violencia.

Sus principales líneas de interés se orientan a la aplicación ética de la inteligencia artificial, la innovación tecnológica, el liderazgo y la transformación digital con impacto social.

AGRADECIMIENTOS

Las autoras expresan su profundo agradecimiento a quienes, desde sus respectivas disciplinas, realizaron observaciones críticas, aportes conceptuales y sugerencias que contribuyeron al fortalecimiento académico, técnico y editorial del presente Manifiesto.

El proceso de revisión interdisciplinaria que acompañó esta primera edición constituye una expresión concreta de los principios de diálogo, cooperación y construcción colectiva que inspiran al Comité EIA LATAM. La participación de las personas aquí mencionadas no implica necesariamente adhesión a todas las posiciones sostenidas en esta obra. Las opiniones y conclusiones contenidas en el presente documento son de exclusiva responsabilidad de las autoras.

Revisión técnica especializada

Dra. Evelyn Castro Gutierrez

Doctora en Ciencias de la Computación. Docente investigadora con publicaciones en Scopus y WoS sobre IA y Pensamiento Computacional. Miembro de IEEE y WIE, e investigadora principal en proyectos CONCYTEC desde 2010. Ha registrado derechos de autor, diseños industriales y patentes, incluyendo la primera Patente Internacional (PCT) de la UNSA en 2022. Fundadora del Startup MEDE(Medical Devices). Reconocida como INGENIERA PERUANA 2022 y miembro del Comité Pro Mujer en CTI de CONCYTEC. Fue distinguida con el Premio Santiago Antúnez de Mayolo Gomero 2025 por su contribución a la ciencia, tecnología e innovación con sus trabajos e investigaciones.

Corrección académica interdisciplinaria

Dr. Eduardo Fabián Cossi

Doctor en Psicología y Abogado por la Universidad de Buenos Aires (UBA). Docente de la Universidad de Buenos Aires y profesor e investigador de la Universidad Nacional de Tres de Febrero (UNTREF). Especialista en Derechos Humanos Económicos, Sociales y Culturales.

Realizó en la revisión académica interdisciplinaria del presente Manifiesto, realizando un análisis crítico de sus fundamentos jurídicos, psicológicos, éticos y epistemológicos. Sus observaciones conceptuales contribuyeron al fortalecimiento de la consistencia argumental, la coherencia interdisciplinaria y la proyección académica de esta primera edición.

El Comité EIA LATAM agradece asimismo a todas las personas e instituciones que, mediante el intercambio de ideas, la crítica académica y el diálogo interdisciplinario, contribuyeron a la construcción de este primer Manifiesto.

Por el resto de la manzana mordida.