

UNA GOBERNANZA DE LA IA JUSTA, SOBERANA E INCLUSIVA

Perspectivas y recomendaciones de la sociedad civil y la academia de América Latina y el Caribe en el marco de la Tercera Cumbre Ministerial y de Altas Autoridades sobre la Ética de la IA en ALC con miras al Diálogo Global sobre la Gobernanza de la IA de las Naciones Unidas

Preámbulo

Las organizaciones de la sociedad civil y las instituciones académicas de América Latina y el Caribe (ALC) que suscriben este documento presentamos nuestra posición colectiva ante el primer Diálogo Global sobre la Gobernanza de la Inteligencia Artificial, convocado por las Naciones Unidas en Ginebra los días 6 y 7 de julio de 2026.

Nos pronunciamos desde una región que alberga una diversidad cultural y lingüística excepcional, que concentra algunas de las desigualdades más profundas del mundo, y que históricamente ha ocupado un lugar periférico en los foros donde se definen las reglas de la economía digital global. Asimismo, nos pronunciamos desde una posición de compromiso activo en proveer una contribución sustantiva desde nuestro conocimiento construido por años de trabajo en la intersección entre la tecnología, derechos humanos y democracia.

Esta contribución parte de una convicción compartida: la gobernanza de la inteligencia artificial será legítima y eficaz únicamente si es inclusiva en su diseño, está fundada en los derechos humanos, se orienta al desarrollo sostenible y resulta sensible a las realidades, capacidades y prioridades de las distintas regiones.

Este documento recoge los aportes de 30 organizaciones de la sociedad civil e instituciones académicas de diez países de la región, más organizaciones de alcance regional e internacional con presencia activa en ALC, consultadas a través de una encuesta regional, un webinar preparatorio celebrado el 19 de junio de 2026 y un encuentro presencial el 24 de junio de 2026 en el marco de la tercera Cumbre Ministerial sobre la Ética de la Inteligencia Artificial de ALC. Este proceso de consultas fue organizado por el Centro Regional de Globethics para América Latina y el Caribe.

Finalmente, éste se inscribe en los procesos y referentes regionales e internacionales pertinentes, entre ellos la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial (2021), los Principios de la OCDE (2019), el Convenio Marco del Consejo de Europa sobre Inteligencia Artificial y Derechos Humanos, Democracia y Estado de Derecho (2024), la resolución de la Asamblea General de las Naciones Unidas A/78/L.49 sobre 'Aprovechar las oportunidades de sistemas seguros, protegidos y fiables de inteligencia artificial para el desarrollo sostenible' (2024), el informe Governing AI for Humanity (Grupo Asesor de Alto Nivel sobre Inteligencia Artificial de las Naciones Unidas, 2024), las Declaraciones de Santiago (2023) y Montevideo (2024), todos ellos sobre la base del derecho internacional de los derechos humanos.

Resumen de recomendaciones

Las organizaciones firmantes formulamos las siguientes recomendaciones transversales, dirigidas a los Estados y los organismos internacionales, que sintetizan las perspectivas desarrolladas en cada eje del Diálogo Global:

1. Reorientar la cooperación internacional hacia la soberanía tecnológica, con cuatro líneas medibles: infraestructura de cómputo regional, capacidad de auditoría independiente, formación y retención de talento, y uso de la Metodología RAM de la UNESCO como diagnóstico común.

2. Situar el despliegue de conectividad, la transformación del trabajo y la alfabetización digital en el centro de la agenda de gobernanza con mecanismos medibles y de financiamiento.
3. Reconocer la diversidad lingüística y cultural como derecho digital, con metas progresivas de cobertura para lenguas oficiales e indígenas y crear un Fondo Multilateral para el Desarrollo de Corpus Lingüísticos de Bajo Recurso.
4. Establecer la participación de la sociedad civil y la academia en los procesos de gobernanza de la IA como condición procedimental vinculante, con acceso genuino a la información, protección frente a represalias, y facultad de solicitar auditorías o revisión independiente cuando existan riesgos graves.
5. Garantizar la transparencia algorítmica, la auditoría independiente y la supervisión humana significativa a lo largo de todo el ciclo de vida de los sistemas de IA.
6. Adoptar instrumentos internacionales vinculantes e impulsar la ratificación del Convenio Marco del Consejo de Europa sobre Inteligencia Artificial.
7. Establecer líneas rojas claras, sin excepción por propósito o actor, sobre los sistemas de IA incompatibles con los derechos humanos.
8. Abordar la inteligencia artificial como amplificador de la violencia de género e infantil, con obligaciones de seguridad desde el diseño.
9. Incorporar la sostenibilidad ambiental —huella hídrica, energética y de carbono— como criterio de confiabilidad de la Inteligencia Artificial.
10. Respalda la creación de una Relatoría Especial de las Naciones Unidas sobre Inteligencia Artificial y Derechos Humanos.

CLUSTER 1 · DIÁLOGO GLOBAL

Oportunidades e implicaciones de la IA

Perspectiva regional

Para América Latina y el Caribe, la principal promesa de la inteligencia artificial reside en su capacidad de incidir sobre las brechas estructurales que han condicionado el desarrollo inclusivo y sostenible de la región. Las oportunidades que identificamos con mayor consistencia se concentran en la ampliación del acceso a una educación, una salud de calidad, mediante tutores adaptativos en español, portugués y lenguas originarias, diagnóstico asistido en zonas sin especialistas y vigilancia epidemiológica predictiva, en el fortalecimiento de la gestión pública, la focalización de las políticas sociales, en la adaptación climática y la gestión sostenible de los recursos naturales, y en el aumento de la productividad de las micro, pequeñas y medianas empresas que sostienen el empleo regional. En todos estos ámbitos, el valor de la tecnología se mide por su contribución a la equidad.

Partimos, además, de una base sólida. La región aporta una diversidad cultural y lingüística excepcional, una matriz de energías renovables presente en varios países, una hoja de ruta con base en la ética de la inteligencia artificial —expresada en las Cumbres Ministeriales y en las Declaraciones de Santiago (2023), Montevideo (2024) y Santo Domingo (2026)— y un conocimiento situado, construido por su academia y su sociedad civil, imprescindible para una gobernanza legítima. Estas fortalezas constituyen el punto de apoyo desde el cual la región puede convertir la promesa en desarrollo efectivo.

La materialización de estas oportunidades depende de condiciones que hoy presentan avances desiguales. El Índice Latinoamericano de Inteligencia Artificial (ILIA 2025), elaborado por el Centro Nacional de Inteligencia Artificial (CENIA) sobre más de cien subindicadores en diecinueve países de la región, confirma un ecosistema heterogéneo: coexisten países pioneros, adoptantes y exploradores con grados de madurez muy distintos. El ILIA 2025 precisa las brechas que la región se propone cerrar: una inversión que apenas alcanza el 1,12 % del total mundial en inteligencia artificial y una investigación concentrada en pocos países, junto a un talento que se forma de manera más amplia en alfabetización básica que en las capacidades especializadas y de frontera. A ello se añaden la conectividad y la infraestructura de cómputo aún insuficientes, la disponibilidad limitada de datos abiertos y de calidad, y una representación todavía escasa en los organismos donde se definen los estándares técnicos.

Esta heterogeneidad se corrobora en evaluaciones internacionales independientes: el Índice de Inteligencia Artificial y Valores Democráticos 2026 del Center for AI and Digital Policy (CAIDP) clasifica a los países de la región en niveles de desempeño marcadamente dispares —desde Uruguay (10 puntos, Nivel II) y Argentina (9), seguidos de Colombia (8,5) y de Brasil, Chile y Perú (8), hasta República Dominicana (5), Ecuador (4,5) y Venezuela (3)—, lo que ratifica la coexistencia de capacidades de gobernanza muy desiguales en la región.

Estas oportunidades conviven con costos concretos que forman parte de la cadena de valor de la inteligencia artificial y que la región asume de manera desproporcionada. América Latina y el Caribe aportan trabajo de etiquetado de datos y moderación de contenidos en condiciones frecuentemente precarias, sin protección laboral ni acompañamiento psicosocial frente a la exposición a contenido dañino, y albergan, de manera creciente, la infraestructura física de cómputo, con su demanda de energía, suelo y agua para refrigeración en territorios ya tensionados por el estrés hídrico. A esta presión se suman la dependencia de infraestructuras y modelos desarrollados fuera de la región, así como la concentración del poder de mercado. Una valoración honesta de las oportunidades internaliza estos costos y exige condiciones laborales dignas para el trabajo de datos y la evaluación ambiental de la infraestructura de cómputo que se instala en la región.

Sobre esta base, distinguimos entre los riesgos que pueden gestionarse mediante salvaguardas y las prácticas estructuralmente incompatibles con los derechos humanos. Determinadas aplicaciones requieren su prohibición, porque su daño se origina en su propia naturaleza y persiste pese a cualquier ajuste técnico; el conjunto de estas líneas rojas se desarrolla en el Cluster 4.

La dimensión cultural y lingüística define buena parte de la singularidad regional en este debate. América Latina y el Caribe alberga centenares de lenguas indígenas vivas, además de lenguas criollas y variedades propias del español, el portugués, el francés y el inglés caribeño. Los modelos que actualmente dominan los mercados están entrenados predominantemente en inglés y sobre corpus que reflejan contextos del Norte Global, ofrecen un desempeño desigual en las lenguas de la región. Por ello solicitamos otorgar la máxima prioridad a preservar y fortalecer este patrimonio cultural para la humanidad: una intervención deliberada permite que la inteligencia artificial opere como un instrumento de vitalidad lingüística y de afirmación cultural.

Reconocemos, asimismo, que las implicaciones se distribuyen de manera desigual entre la población. Las mujeres y las niñas, los pueblos indígenas, las comunidades afrodescendientes, las comunidades rurales, las personas en situación de pobreza, las personas con discapacidad, las personas migrantes, los adultos mayores y la población LGBTIQ+ se encuentran entre los grupos más expuestos. La niñez y la adolescencia ocupan un lugar singular, como usuarias tempranas e intensivas de estos sistemas; el estándar reforzado de protección que reclamamos para sus datos y derechos se desarrolla en el Cluster 4. En todos los casos, sostenemos que los usos en sectores sensibles —salud, educación, justicia y migración— requieren marcos específicos y garantías reforzadas.

Para diagnosticar capacidades y orientar inversiones, instamos a los Estados de América Latina y el Caribe a aprovechar las herramientas disponibles, en particular la Metodología de Evaluación del Estado de Preparación para la Inteligencia Artificial (RAM) de la UNESCO. El avance regional en esta materia ya es significativo: el Índice de Inteligencia Artificial y Valores Democráticos 2026 del CAIDP documenta que, solo entre abril de 2025 y el cierre del informe, ocho países de la región —entre ellos Perú, Paraguay, Jamaica, Ecuador, Curazao, Cuba, Colombia y Antigua y Barbuda— publicaron sus informes RAM, que se suman a los ya completados por Brasil, Chile, Costa Rica, México, Uruguay y la República Dominicana. Exhortamos a los Estados que aún no la han realizado a llevarla a cabo y, a quienes ya la han completado, a publicar sus resultados e implementar sus recomendaciones, con prioridad en el fortalecimiento de las capacidades institucionales, la gobernanza de datos y la formación de las y los servidores públicos, para un desarrollo y uso de la inteligencia artificial seguro, inclusivo y respetuoso de los derechos humanos.

FODA regional

<p>Fortalezas (interno +)</p> <ul style="list-style-type: none"> • Diversidad cultural y lingüística como activo • Matriz de energías renovables en varios países • Hoja de ruta regional en ética de IA (Cumbres; Declaraciones de Santiago, Montevideo y Santo Domingo) 	<p>Oportunidades (externo +)</p> <ul style="list-style-type: none"> • IA en educación, salud, gestión pública, adaptación climática y productividad de MIPYMES • Institucionalidad regional (eLAC/CEPAL) • IA como infraestructura pública y bienes públicos digitales
<p>Debilidades (interno –)</p> <ul style="list-style-type: none"> • Conectividad e infraestructura insuficientes • Inmadurez regulatoria y acceso limitado a cómputo • Embudo de formación avanzada de talento • Datos poco disponibles; financiamiento restringido 	<p>Amenazas (externo –)</p> <ul style="list-style-type: none"> • Homogeneización cultural y riesgo de extinción digital de lenguas • Trabajo precarizado de etiquetado y moderación • Huella ambiental en territorios con estrés hídrico • Dependencia de infraestructuras y modelos externos; concentración de mercado

1.Recomendaciones al Diálogo Global sobre el cluster 1

Dirigimos a los Estados y los organismos internacionales las siguientes recomendaciones.

- 1.1. Reconocer la inteligencia artificial como infraestructura pública y estratégica, y gobernarla como bien público digital mediante inversión pública, acceso equitativo, estándares abiertos, gobernanza democrática y orientación al interés público.
- 1.2. Recomendamos reconocer la diversidad lingüística y cultural como derecho digital, operacionalizado mediante: (1) un observatorio regional que mida el desempeño de la

IA por lengua; (2) reporte obligatorio de cobertura lingüística por parte de los desarrolladores de modelos de frontera; (3) metas progresivas de cobertura para lenguas oficiales e indígenas de mayor población hablante; y (4) operación mínima de los servicios públicos digitales con IA en las lenguas oficiales e indígenas predominantes de cada territorio, en el marco del Decenio Internacional de las Lenguas Indígenas 2022–2032.

- 1.3. Establecer un régimen de soberanía de datos indígenas y comunitarios: uso de patrimonios lingüísticos y culturales en IA sujeto a consentimiento libre, previo e informado (Convenio 169 OIT, DNU DPI), Principios CARE de gobernanza de datos, y reparto de beneficios cuando generen valor económico. Prohibir su recolección sin consentimiento y adoptar un modelo regional de licencia comunitaria gestionado por las propias comunidades.
- 1.4. Incorporar perspectivas interseccionales —con énfasis en género, condición socioeconómica, origen étnico y discapacidad— en todas las evaluaciones de impacto y en el diseño de los marcos de gobernanza.
- 1.5. Establecer normas globales para la reutilización responsable de datos que comprometan a los desarrolladores de modelos comerciales a ofrecer una atribución clara, verificar la procedencia de la información y sostener las infraestructuras de las que se extrae el conocimiento.
- 1.6. Situar la transformación del trabajo en el centro de la agenda de gobernanza, reconociendo los derechos laborales en toda la cadena de valor de la IA, asegurando condiciones dignas y acompañamiento psicosocial para el trabajo de etiquetado y moderación de datos, garantizando la transparencia y la rendición de cuentas de los sistemas de gestión algorítmica, y preservando la supervisión humana de los sistemas que inciden en el empleo.
- 1.7. Promover estrategias regionales de educación y alfabetización digital que fortalezcan el uso responsable, ético y seguro de las tecnologías en los distintos sectores sociales, con especial atención a las comunidades históricamente excluidas del acceso a la formación.
- 1.8. Generar reportes periódicos y compartir lecciones aprendidas sobre los avances de la cooperación gubernamental y empresarial en materia de IA en los países de la región, de modo que la experiencia acumulada se traduzca en aprendizaje colectivo.

CLUSTER 2 · DIÁLOGO GLOBAL

Reducir las brechas de IA

Perspectiva regional

Para que América Latina y el Caribe aprovechen las oportunidades de la inteligencia artificial, la región necesita cerrar un conjunto de brechas estructurales que identificamos con notable coincidencia: la conectividad y la infraestructura digital aún insuficientes, los marcos regulatorios en proceso de maduración, la escasez de talento especializado, el acceso

limitado a capacidades de cómputo, la disponibilidad y calidad de los datos, y las restricciones de financiamiento. El ILIA muestra que la investigación se concentra en unos pocos países y que los datos, aunque abundantes, permanecen poco disponibles debido a la limitada apertura y estandarización. Esta heterogeneidad se repite dentro de los propios países del índice, líderes y en aceleración por igual: capital científico y académico consolidado (el CONICET en Argentina, la ANID en Chile, la CONACYT de Paraguay o el MESCYT en República Dominicana) convive con brechas persistentes en infraestructura de cómputo, financiamiento sostenido y marcos regulatorios integrales. Esa combinación de fortalezas y rezagos, presente incluso en los casos más avanzados de la región, es la que la cooperación regional debe abordar de manera prioritaria.

A ello se añade el reto de formar y retener talento especializado, ya que la alfabetización básica en IA avanza con mayor amplitud que las capacidades técnicas y de frontera. Estas brechas reflejan asimetrías de poder en la economía digital global y condicionan la posición de la región en cualquier discusión sobre gobernanza.

Estas brechas se asientan sobre desigualdades anteriores a la inteligencia artificial que conviene atender de manera simultánea: la educación digital aún incipiente, el acceso desigual a la conectividad en las zonas rurales y periféricas, y la necesidad de políticas de formación con perspectiva de género y de derechos humanos. La región enfrenta, asimismo, la violencia de género facilitada por la tecnología, cuyo abordaje sustantivo desarrollamos en el Cluster 4.

Frente a este diagnóstico, sostenemos que la respuesta más eficaz combina los esfuerzos nacionales con una cooperación regional articulada. Proponemos avanzar hacia una infraestructura compartida de cómputo de alto rendimiento, accesible para los países con menor capacidad instalada; un fondo regional de innovación con gobernanza multiactor y criterios de equidad territorial; programas de formación, intercambio y retención de talento; el desarrollo colaborativo de modelos abiertos entrenados con datos y realidades de la región; y redes de centros de excelencia con mandato de transferencia de conocimiento. De manera transversal, valoramos el rol de la diplomacia digital en la medida que la coordinación de posiciones comunes aumenta el peso negociador de la región en los foros internacionales.

Esta cooperación se apoya en la institucionalidad regional ya existente, como la Agenda Digital para América Latina y el Caribe (eLAC), coordinada por la CEPAL, y en los espacios de integración subregional. Recomendamos articular las iniciativas de cómputo, datos y talento con estos marcos para ganar escala y continuidad, evitar la duplicación de esfuerzos y dotar a la región de una gobernanza multiactor estable, sostenida más allá de los ciclos de financiamiento externo.

Los modelos abiertos constituyen una contribución de primer orden para que la región participe en el desarrollo de la inteligencia artificial. Su pleno aprovechamiento depende de una inversión simultánea en infraestructura, capacidades humanas, gobernanza comunitaria y sostenibilidad de los bienes públicos digitales, de modo que la apertura se acompañe de las condiciones materiales que la hacen efectiva. El éxito y la sostenibilidad de estos modelos descansan en mecanismos sólidos de gobernanza, mantenimiento y financiamiento, y en el fortalecimiento de las comunidades que los desarrollan mediante esquemas de gobernanza participativa, democrática y orientada al bien común.

Asimismo, cerrar las brechas exige modelos de financiamiento previsibles y de largo plazo. Proponemos combinar fondos públicos regionales, banca de desarrollo y compromisos de cofinanciamiento por parte de los desarrolladores de modelos de frontera que se benefician de los datos y los mercados de la región. La infraestructura de cómputo que se construya conviene que sea energéticamente sostenible y aproveche la matriz de energías renovables de varios países de ALC; los criterios ambientales que rigen esa infraestructura se desarrollan en el Cluster 3.

La agenda de reducción de brechas se proyecta de manera directa sobre la soberanía tecnológica de la región y su participación en la gobernanza internacional. Para ello, identificamos las barreras que conviene superar para incidir en la definición de los estándares técnicos: la representación institucional aún escasa de América Latina y el Caribe en los principales organismos de estandarización —que el ILIA 2025 constata como un factor que reduce la influencia de la región en las reglas globales—, la coordinación regional todavía incipiente, las asimetrías de recursos frente a los países que lideran el desarrollo, la complejidad de procesos técnicos que requieren continuidad, y la documentación disponible casi exclusivamente en inglés. Frente a este escenario, sostenemos que la soberanía tecnológica se construye a través de la cooperación Sur-Sur y de la creación conjunta de conocimiento entre los países del Sur Global, que permita compartir capacidades, datos y aprendizajes en condiciones de horizontalidad, sostener una posición negociadora común y desarrollar marcos e infraestructuras propios. La cooperación internacional resulta más valiosa cuando se orienta a la autonomía de la región y cuando el conocimiento generado en la región llega a moldear los estándares internacionales.

2.Recomendaciones al Diálogo Global sobre el cluster 2

Dirigimos a los Estados y los organismos internacionales las siguientes recomendaciones:

- 2.1. Crear un Fondo Multilateral para el Desarrollo de Corpus Lingüísticos de Bajo Recurso que comprometa a los desarrolladores de modelos de frontera a cofinanciar la digitalización y la preservación de las lenguas indígenas y criollas, con gestión a cargo de instituciones locales y liderazgos comunitarios.
- 2.2. Garantizar mecanismos de participación efectiva de América Latina y el Caribe en los organismos y paneles de gobernanza técnica de estándares de inteligencia artificial, como ISO, UIT e IEEE, mediante financiamiento público para representación sostenida, apoyo lingüístico, transparencia en los procesos de decisión y presencia de sociedad civil y academia de la región en los grupos técnicos correspondientes.
- 2.3. Reorientar la cooperación internacional hacia la soberanía tecnológica con cuatro líneas medibles: (1) financiamiento de infraestructura de cómputo regional; (2) capacidad de auditoría y evaluación independiente de sistemas de IA; (3) programas de formación, intercambio y retención de talento en ética y gobernanza de la IA, con becas y esquemas de retorno que mitiguen la fuga de talento; (4) Usar la Metodología de Evaluación del Estado de Preparación

(RAM) de la UNESCO como diagnóstico común para identificar y priorizar las brechas de cada país.

- 2.4. Fortalecer los mecanismos de cooperación Sur-Sur que permitan que el conocimiento sobre gobernanza de la IA generado en la región fluya de manera horizontal y contribuya a moldear los estándares internacionales.
- 2.5. Incentivar la investigación multidisciplinaria y abrir espacios para la construcción de tecnología por parte de las comunidades indígenas y de los distintos sectores sociales, desde su propia perspectiva.
- 2.6. Acompañar la inversión en inteligencia artificial con políticas de educación y alfabetización digital con perspectiva de género y de derechos humanos, que aborden de manera integral las desigualdades estructurales de la región.
- 2.7. Articular el apoyo internacional con la institucionalidad regional existente (eLAC/CEPAL e integración subregional) y orientar la nueva infraestructura de cómputo hacia criterios de sostenibilidad energética e hídrica, que se desarrollan en el Cluster 3.
- 2.8. Desarrollar programas regionales y nacionales de alfabetización en IA para la ciudadanía, docentes, funcionarios públicos, sociedad civil, periodistas y comunidades vulnerables, con enfoque de derechos humanos, género, accesibilidad, pertinencia cultural y lenguas locales.
- 2.9. Crear marcos regulatorios que incentiven y constituyan mecanismos sólidos de gobernanza, mantenimiento y financiamiento de proyectos para el desarrollo de sistemas de código abierto encaminados a la consecución de objetivos de soberanía y cierre de brechas de todos y cada uno de los países de la región.

CLUSTER 3 · DIÁLOGO GLOBAL

IA segura, confiable e interoperable

Perspectiva regional

Para América Latina y el Caribe, la confiabilidad de un sistema de inteligencia artificial debe ser una propiedad que se verifica en su contexto de despliegue y a lo largo de todo su ciclo de vida. Consideramos que esta confiabilidad descansa en un conjunto de principios — transparencia, privacidad, no discriminación, rendición de cuentas, supervisión humana y explicabilidad— y que su evaluación adquiere sentido pleno cuando incorpora las realidades lingüísticas, sociales y ambientales de la región y la participación de las comunidades afectadas.

La región fortalece su confiabilidad cuando desarrolla capacidad propia de prueba, evaluación y certificación. Proponemos avanzar hacia bancos de prueba con datos representativos de la región y desagregados por idioma y por género, y hacia laboratorios de auditoría con reconocimiento mutuo entre países. Esta capacidad permite que la confiabilidad de un sistema se determine con evidencia generada en la región y reduce la dependencia de la

autoevaluación de quienes desarrollan los modelos, de modo que la verificación recaiga en instancias independientes y con conocimiento del contexto.

La seguridad y la soberanía de los datos constituyen una dimensión central de esta agenda, y la evidencia regional confirma su urgencia. Los sistemas de inteligencia artificial amplían la superficie de riesgo frente a la filtración de datos, la desinformación, los *deepfakes*, la clonación de voz, la suplantación de identidad y el fraude. En América Latina y el Caribe, el fraude de identidad creció un 137 % durante 2024, impulsado por la proliferación de *deepfakes*, y los incidentes con *deepfakes* en la región aumentaron un 255 % entre 2023 y 2024, según el Índice Global de Fraude de Sumsb. A escala global, los *deepfakes* ya representan uno de cada cinco intentos de fraude biométrico. Reivindicamos, por ello, controles de ciberseguridad robustos, exigencias de calidad de los datos, mecanismos de cumplimiento normativo y una protección reforzada de la información sensible. La adopción de estándares técnicos reconocidos —entre ellos la familia ISO 42000 y los marcos del NIST— ofrece una base común para estas garantías, siempre que su definición y aplicación cuenten con la participación de la sociedad civil y la academia de la región.

La sostenibilidad ambiental y climática es una dimensión transversal de la confiabilidad. La presión que la infraestructura de cómputo ejerce sobre el agua, la energía y las emisiones de carbono forma parte de la valoración de un sistema y exige requisitos de divulgación verificables. Por ello sostenemos que el Diálogo Global articule su trabajo con los procesos de gobernanza ambiental —entre ellos la Convención Marco de las Naciones Unidas sobre el Cambio Climático (UNFCCC), el Programa de las Naciones Unidas para el Medio Ambiente (PNUMA) y la Comisión de Ciencia y Tecnología para el Desarrollo (CSTD)—, de manera que la gobernanza de la inteligencia artificial y la gobernanza planetaria avancen de forma coordinada. La región parte, además, de una ventaja: una matriz de energías renovables que permite orientar la infraestructura de cómputo hacia criterios de sostenibilidad.

La interoperabilidad que la región promueve eleva el umbral de protección. Entendemos la armonización internacional como un proceso que fortalece las salvaguardas y toma el derecho internacional de los derechos humanos como referencia común, de modo que ningún proceso de convergencia rebaje las garantías vigentes, en particular para los países del Sur Global. La cooperación internacional adquiere su mayor valor cuando reconoce que distintas tradiciones jurídicas y contextos de desarrollo pueden coexistir y enriquecerse mutuamente; el Diálogo Global representa una oportunidad relevante para tender puentes entre ellos preservando esa diversidad. En la misma línea, recomendamos su articulación con el Pacto Digital Global y con el proceso WSIS+20, para que las distintas agendas digitales avancen de forma convergente y eviten la fragmentación.

Esta agenda se apoya en un acervo normativo del que la región es partícipe activa: la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial (2021), los Principios de la OCDE (2019), la Carta Iberoamericana de Principios y Derechos en los Entornos Digitales, el Convenio Marco del Consejo de Europa (2024), la resolución de la Asamblea General de las Naciones Unidas sobre una IA segura y confiable y el informe *Governing AI for Humanity*. Estos instrumentos ofrecen la base sobre la cual la interoperabilidad puede construirse sin diluir las garantías.

Finalmente, la región orienta su atención ante todo a los riesgos presentes y ya materializados, sin desatender los riesgos emergentes de los modelos de frontera y de los sistemas agénticos, que operan con escasa intervención humana. Esta priorización ordena los esfuerzos hacia los daños que ya afectan a las personas, a la vez que mantiene la previsión necesaria frente a las capacidades en desarrollo y los riesgos de uso dual.

3.Recomendaciones al Diálogo Global sobre el cluster 3

Dirigimos a los Estados y los organismos internacionales las recomendaciones siguientes.

- 3.1. Impulsar un instrumento internacional vinculante de gobernanza de la IA con un órgano de seguimiento y un ciclo periódico de reporte estatal, que tome el derecho internacional de los derechos humanos y la Recomendación de la UNESCO como línea de base. Todo proceso de convergencia o reconocimiento mutuo debe regirse por una cláusula de no regresión: tomar como referencia el estándar de protección más alto aplicable y nunca rebajar salvaguardas, con especial atención a los países del Sur Global.
- 3.2. Implementar normas vinculantes en materia de transparencia algorítmica que funcionen como mecanismo de auditoría de seguridad y rendición de cuentas, que permita a investigadores independientes y a la sociedad civil examinar los sistemas en busca de sesgos, vulnerabilidades o fallos, garantizando un registro verificable y reproducible.
- 3.3. Desarrollar una capacidad regional de prueba, evaluación y certificación, con bancos de prueba que utilicen datos representativos y desagregados, y laboratorios de auditoría con reconocimiento mutuo entre los países, que coopere con la Red Internacional de Institutos de Seguridad de la IA.
- 3.4. Adoptar estándares técnicos y de ciberseguridad con participación de la sociedad civil y la academia, de manera que la confiabilidad descansa en la verificación independiente y no en la autoevaluación de los desarrolladores. Estos controles deben cubrir expresamente toda la cadena de suministro de la IA, incluyendo la procedencia y licenciamiento de los datos, los modelos preentrenados, las bibliotecas y dependencias, las APIs, los servicios de nube, las actualizaciones, los proveedores externos y los mecanismos de retiro seguro, ya que la seguridad de un sistema de IA depende de todo su ciclo de vida.
- 3.5. Incorporar la sostenibilidad ambiental como criterio de confiabilidad, con requisitos vinculantes de divulgación de la huella hídrica, energética y de carbono de los sistemas, en coordinación con la UNFCCC, el PNUMA y la CSTD.
- 3.6. Establecer la supervisión humana significativa como principio de confiabilidad a lo largo del ciclo de vida, con marcos de rendición de cuentas adecuados para los sistemas de frontera y los sistemas agénticos, así como para los riesgos de uso dual.
- 3.7. Articular el Diálogo Global con el Pacto Digital Global, el proceso WSIS+20 y los procesos de gobernanza ambiental, de modo que las agendas digitales y ambientales avancen de forma coordinada y se evite la fragmentación normativa.

CLUSTER 4 · DIÁLOGO GLOBAL

Derechos humanos, transparencia y rendición de cuentas

Perspectiva regional

Los derechos humanos son el marco vinculante de toda la gobernanza de la inteligencia artificial. Para América Latina y el Caribe, el desarrollo y el uso de estos sistemas adquieren legitimidad cuando se subordinan al derecho internacional de los derechos humanos y a las obligaciones que los Estados ya han asumido. Desde esta convicción, sostenemos que la gobernanza de la IA traduce esas obligaciones en garantías concretas de transparencia, rendición de cuentas y supervisión humana.

Situamos en el centro de la gobernanza a las personas y comunidades sobre las que estos sistemas inciden con mayor fuerza: las mujeres y las niñas, los pueblos indígenas, las comunidades afrodescendientes, las comunidades rurales, las personas en situación de pobreza, las personas con discapacidad, las personas migrantes y la población LGBTIQ+. Una gobernanza centrada en quienes asumen los mayores riesgos protege, por extensión, a toda la sociedad.

Reconocemos el valor de los marcos internacionales existentes citados previamente y proponemos perfeccionarlos hacia instrumentos vinculantes, con mecanismos efectivos de cumplimiento y contruidos con la participación de la región. La región cuenta, además, con un acervo propio y vinculante: el Sistema Interamericano de Derechos Humanos —la Convención Americana, la jurisprudencia de la Corte Interamericana y los estándares de la Comisión y de su Relatoría Especial para la Libertad de Expresión— constituye una fuente directa de obligaciones para los Estados, y no una referencia secundaria.

Sobre esta base, distinguimos con claridad las prácticas estructuralmente incompatibles con los derechos humanos, cuyo daño se origina en su propia naturaleza y persiste pese a cualquier salvaguarda técnica. En ese contexto, solicitamos su prohibición mediante el establecimiento de líneas rojas las cuales comprenden:

- i. las armas autónomas letales;
- ii. la vigilancia biométrica masiva y la identificación biométrica remota en los espacios públicos;
- iii. los sistemas de puntuación social;
- iv. la actuación policial predictiva basada en la elaboración de perfiles;
- v. la categorización biométrica por raza, género, orientación sexual, creencias, estado de salud u otras categorías protegidas;
- vi. el reconocimiento de emociones y la inferencia psicométrica en el trabajo, la educación, la migración y la aplicación de la ley;
- vii. los sistemas que adoptan decisiones determinantes sobre el crédito, el empleo, las prestaciones sociales, la migración o la justicia sin supervisión humana significativa, transparencia ni posibilidad de recurso efectivo.

Estas prohibiciones deben regir sin excepción por el propósito que se invoque ni por el actor que las opere. La urgencia de proscribir la vigilancia y la identificación biométrica encuentra respaldo en las evaluaciones por país del Índice de Inteligencia Artificial y Valores Democráticos 2026 del CAIDP, que constatan una preocupación recurrente en la región por el despliegue de sistemas de reconocimiento facial —advertida, entre otros, en los casos de Argentina, Uruguay y la República Dominicana— por su potencial para vulnerar los derechos individuales en ausencia de salvaguardas adecuadas.

El derecho a la privacidad, la autodeterminación informativa y la protección de los datos personales constituyen un umbral crítico de esta agenda. La inteligencia artificial intensifica la capacidad de inferir información sensible y de moldear el comportamiento, con efectos que alcanzan la dimensión psicosocial de las personas. Por ello reivindicamos límites firmes al extractivismo de datos psicométricos y emocionales, y garantías que preserven la autonomía de las personas sobre su información.

La protección de las niñas, los niños y los adolescentes merece el estándar más alto. Su condición de usuarios tempranos e intensivos de estos sistemas exige el nivel reforzado de protección de sus datos, restricciones estrictas a la elaboración de perfiles y a la publicidad dirigida, y mecanismos de supervisión y recurso adaptados a su edad, conforme al principio del interés superior del niño.

La gobernanza aborda, asimismo, la inteligencia artificial como amplificador de la violencia de género y de la violencia contra la niñez, un fenómeno documentado y en plena expansión. Las herramientas actuales permiten generar este material a partir de muy pocas imágenes y en cuestión de minutos. En paralelo, el Reporte Internacional de Seguridad (2025 p.64) indica que cerca del 96 % de los videos deepfake en circulación que corresponden a contenido sexual no consentido están dirigidos de manera abrumadora hacia mujeres, lo que confirma su carácter de violencia de género. Frente a esta realidad, reivindicamos prohibiciones claras, obligaciones de seguridad desde el diseño para quienes desarrollan modelos y operan plataformas, y mecanismos accesibles de denuncia, protección y reparación.

Sostenemos, además, que la responsabilidad no recae únicamente en los Estados. Las empresas tienen la responsabilidad de respetar los derechos humanos conforme a los Principios Rectores de las Naciones Unidas sobre las Empresas y los Derechos Humanos, lo que comprende procesos de debida diligencia que identifiquen, prevengan y mitiguen los impactos de sus sistemas a lo largo de todo el ciclo de vida. La innovación tecnológica convive con las obligaciones estatales de proteger frente a los abusos, investigar, sancionar y garantizar el acceso a una reparación efectiva, que conservan plena vigencia.

La rendición de cuentas hace operativos estos principios. Reivindicamos evaluaciones de impacto algorítmico y de derechos humanos previas al despliegue, auditorías independientes, registros públicos de los sistemas que utiliza el Estado, vías accesibles de denuncia y recurso, regímenes de responsabilidad claros y la facultad de suspender o retirar los sistemas que vulneren derechos.

Finalmente, la participación de la sociedad civil y la academia es una condición de legitimidad de toda esta arquitectura. Esa participación adquiere sentido pleno cuando es formal, permanente y dotada de capacidad real —con funciones concretas de auditoría, evaluación

de impacto y documentación—, cuando cuenta con las condiciones materiales que la hacen efectiva —capacidades técnicas, financiamiento estable y acceso genuino a la información— y cuando ofrece canales protegidos y confidenciales para las personas y organizaciones que operan en contextos de espacio cívico restringido.

4.Recomendaciones al Diálogo Global sobre el cluster 4

Dirigimos a los Estados y los organismos internacionales las siguientes recomendaciones.

- 4.1. Adoptar instrumentos internacionales vinculantes anclados en el derecho internacional de los derechos humanos, e impulsar la ratificación del Convenio Marco del Consejo de Europa sobre Inteligencia Artificial, abierto a todos los Estados, como referente de obligaciones exigibles.
- 4.2. Establecer evaluaciones de impacto algorítmico y de derechos humanos obligatorias y previas al despliegue, considerando la debida diligencia de los Principios Rectores de la ONU sobre Empresas y Derechos Humanos, con especial rigor en los sectores sensibles de la salud, la educación, la justicia y la migración.
- 4.3. Establecer líneas rojas claras sobre los sistemas de IA incompatibles con los derechos humanos. Esto se materializa en la prohibición legal expresa y vinculante, sumada a un régimen sancionatorio efectivo.
- 4.4. Crear mecanismos regulatorios vinculantes que combinen la exigencia de requisitos de transparencia con mandatos institucionales claros para la supervisión. Los reguladores independientes deberían contar con autoridad legal explícita, recursos suficientes, facultades legales para auditar y sancionar, y mecanismos para auditorías de terceros y lideradas por la comunidad. Estos instrumentos deberán contar con plantillas de reporte y evaluación estandarizadas y plazos de remediación exigibles, así como mecanismos de impugnación significativos, de modo que la transparencia, la supervisión y la exigibilidad trabajen conjuntamente para proteger los derechos y la confianza pública.
- 4.5. Abordar la inteligencia artificial como amplificador de la violencia de género e infantil, con prohibiciones claras, obligaciones de seguridad desde el diseño para desarrolladores y plataformas, y mecanismos accesibles de denuncia, protección y reparación para las víctimas.
- 4.6. Establecer la participación de la sociedad civil y la academia como condición procedimental vinculante, con financiamiento independiente, acceso genuino a la información y canales protegidos y confidenciales para los contextos de espacio cívico restringido, protección frente a represalias, y facultad de solicitar auditorías o revisión independiente cuando existan riesgos graves.
- 4.7. Crear registros públicos de los sistemas de inteligencia artificial que utiliza el Estado y mecanismos híbridos y accesibles de denuncia, recurso y reparación integral del daño algorítmico.
- 4.8. Respaldar la creación de una Relatoría Especial de las Naciones Unidas sobre Inteligencia Artificial y Derechos Humanos, con mandato horizontal y

función de reporte independiente, que contribuya a definir las líneas rojas internacionales.

- 4.9. Fortalecer los marcos de protección de datos personales, privacidad y autodeterminación informativa, incorporando principios de minimización, finalidad, seguridad, consentimiento, límites a inferencias sensibles, protección reforzada de datos biométricos y prohibición del extractivismo de datos psicométricos y emocionales.

Seguimiento de las recomendaciones

Proponemos la creación de un mecanismo regional de seguimiento liderado por la sociedad civil y la academia en colaboración con los Estados Miembros. Este mecanismo debe contar con indicadores públicos y verificables, así como con reportes periódicos que documenten los avances hacia el Diálogo Global, la próxima Cumbre Ministerial y los procesos del Pacto Digital Global y WSIS+20. Solicitamos asimismo que el Diálogo Global contemple espacios formales para la recepción de estos informes y para responder a ellos de manera sustantiva.

Reconocimiento

Este Documento de Posición, presentado ante la Secretaría del Diálogo Global sobre la Gobernanza de la Inteligencia Artificial de las Naciones Unidas, fue coordinado y editado por Paola Galvez Callirgos, Gerente de IA Ética de Globethics. De manera particular, reconocemos y agradecemos los aportes, comentarios y recomendaciones formulados durante las etapas de redacción y revisión por Adriana Baravalle (Universidad Austral), Abdul Abner Lugo Jiménez, Amalia Toledo (Wikimedia Foundation) Carlos Cortés (Linterna Verde), Jéssica Elizabeth Morales Bonilla (Kichwa Institute of Science, Technology and Humanities, KISTH), Julio Herrera (Red Ciudadana), Natalia Alarcón Rueda (Center for AI and Digital Policy - CAIDP), Osvaldo Larancuent (Internet Society - República Dominicana), Syra Taveras Pineda (Centro de Investigación para la Acción Femenina (CIPAF), Vladimir Cortés (Digital Action), César Peña (Librodesa International República Dominicana), cuyo compromiso y expertise enriqueció este trabajo.

Respaldo

La presente versión final es respaldada por las siguientes organizaciones y profesionales:

CENIA (Centro Nacional de Inteligencia Artificial de Chile) / Chile

Center for AI and Digital Policy (CAIDP) / con presencia en América Latina

Centro de Investigación para la Acción Femenina (CIPAF) / República Dominicana

Digital Action / México

Diversa / Ecuador

Fórum Integra / Brasil

Fundación KISTH - Kichwa Institute of Science Technology and Humanities / Ecuador

Fundación Huaira / Ecuador

Globethics - Centro Regional para América Latina y el Caribe / Argentina
Grupo de Cooperación Internacional de Universidades Brasileñas (GCUB) / Brasil
Instituto Panamericano de Derecho y Tecnología (IPANDETEC) / Panama
ISOC capítulo República Dominicana (isoc.do) / República Dominicana
Librodesa Internacional / República Dominicana
ONG Amaranta / Chile
Red Ciudadana / Guatemala
Semillero de Inteligencia Artificial y Derecho de la Pontificia Universidad Javeriana / Colombia
Universidad Austral / Argentina
Usuarios Digitales / Ecuador

Asimismo, la presente versión final es respaldada por los siguientes profesionales:

Merve Hickok, Presidenta y Directora de Políticas de CAIDP
Natalia Alarcón Rueda, Directora y fundadora del Semillero de Inteligencia Artificial y Derecho de la Pontificia Universidad Javeriana y becaria de CAIDP/ Colombia
Marc Rotenberg, Fundador y Director Ejecutivo de CAIDP
Christabel Randolph, Directora Asociada de CAIDP
Ren Bin Lee Dixon, Asociado/a del Grupo de Políticas de CAIDP
Dra. Grace Thomson, Directora de la Clínica de Políticas de IA de CAIDP
Maria Eugenia Barroso, Gerente Regional del Centro para América Latina y el Caribe de Globethics / Argentina
Rossana Valéria de Souza e Silva, Directora Ejecutiva del GCUB / Brasil
Paola Galvez Callirgos, Gerente de la Ética de la IA de Globethics, Miembro del MSCG de UNESCO Women for Ethical AI, Miembro del Comité Directivo de UN Women AI School / Perú
Jessica Morales, CEO y Presidente de Fundación KISTH / Ecuador
Dayanara Antonia Reyes, Especialista en Diseño de Formación Profesional y Gobernanza de la Inteligencia Artificial de Novaltech / República Dominicana