RED IBEROAMERICANA DE PROTECCION DE DATOS

Red Iberoamericana
Rede Ibero-Americana

Marco de referencia regional para el diseño, despliegue y supervisión ética de sistemas de Inteligencia Artificial y otras tecnologías emergentes

Principios de Inclusión · Desarrollo Sostenible · Diseño Ético desde la Base
Convergencias y Brechas en el Espacio Iberoamericano · Base Común Regional

*Autor: Luis García Balcarce**
Elaborado en el marco de la Convocatoria "Trabajar en Red" 2025-2026
Septiembre 2026.
Publicado con el apoyo de la Agencia Española de Cooperación Internacional para el Desarrollo AECID

**"Las opiniones, análisis y conclusiones expresados en este documento pertenecen exclusivamente a su autor en calidad de experto independiente y no reflejan necesariamente la postura o criterio de la Red Iberoamericana de Protección de Datos (RIPD). Esta publicación no implica la aprobación de su contenido ni representa la posición institucional de la RIPD, la cual no se hace responsable del uso que pueda darse a la información aquí expuesta."*

Sumario ejecutivo

El presente Marco de Referencia Regional constituye una propuesta desarrollada en el marco del Observatorio Iberoamericano de Protección de Datos de la RIPD, orientada a proveer una base común para el diseño, despliegue y supervisión ética de sistemas de inteligencia artificial y otras tecnologías emergentes en el espacio iberoamericano.

El documento parte de un diagnóstico del estado actual de la regulación de IA en nueve jurisdicciones iberoamericanas seleccionadas, que no agotan el espacio regional, identificando tanto los puntos de convergencia entre los distintos enfoques nacionales como las brechas que persisten. Sobre esa base, propone un conjunto de recomendaciones, categorías de riesgo, obligaciones de gobernanza y mecanismos institucionales que pueden servir como base común para la armonización regional, sin pretender reemplazar los marcos nacionales sino complementarlos y orientarlos.

Las conclusiones principales del análisis son cuatro. Primero, entre las jurisdicciones analizadas existe una convergencia sólida en torno a un núcleo de principios éticos -transparencia, no discriminación, supervisión humana y protección de datos- que puede constituir el fundamento normativo de un marco regional. Segundo, la brecha más significativa se sitúa en los mecanismos de gobernanza institucional: la mayoría de los países analizados carece de autoridades de supervisión de IA con independencia, recursos técnicos y mandato claro. Tercero, los enfoques de diseño ético desde la base -privacy and ethics by design- están presentes en la literatura y en algunos proyectos legislativos nacionales, pero su implementación efectiva es todavía incipiente. Cuarto, el impacto diferencial de la IA sobre poblaciones vulnerables, el mercado laboral y el medioambiente requiere mecanismos específicos de evaluación de impacto que van más allá de los modelos de riesgo centrados únicamente en la seguridad técnica.

La propuesta se organiza en diez recomendaciones, cinco categorías de riesgo adaptadas al contexto iberoamericano, un modelo de gobernanza multinivel, y un conjunto de mecanismos de actualización periódica. Toma como referencia los instrumentos internacionales más relevantes -la Recomendación de la UNESCO sobre Ética de la IA, el AI Act europeo, el Convenio 108+ del Consejo de Europa, y los Principios de la OCDE- y los articula con las particularidades del contexto regional.

1. Introducción: el desafío regulatorio de la IA en Iberoamérica

Los sistemas de inteligencia artificial han dejado de ser una promesa tecnológica distante para convertirse en infraestructura operativa de decisiones con consecuencias directas sobre derechos fundamentales: quién accede a un crédito, quién es contratado, quién recibe atención médica prioritaria, qué información circula en el espacio público. Esta penetración transversal convierte la regulación de la IA en un asunto de constitucionalismo material, no de política tecnológica sectorial.

El espacio iberoamericano presenta una situación paradójica: cuenta con una tradición jurídica sólida en materia de derechos fundamentales, con marcos regionales de protección de datos personales relativamente avanzados -en particular los Estándares Iberoamericanos de la RIPD-, y con una comunidad académica y de sociedad civil activa en el debate sobre ética y tecnología. Al mismo tiempo, la mayoría de los países cuyos marcos se examinan en este documento carece de marcos legales específicos para la IA, y los que existen presentan una gran heterogeneidad en cuanto a enfoque, alcance y mecanismos institucionales. Esta heterogeneidad, si no se gestiona activamente, puede generar fragmentación normativa, arbitraje regulatorio y pérdida de eficacia en la protección de los ciudadanos.

La presente propuesta responde a ese diagnóstico. Su objetivo es identificar una **base común mínima** articulada en torno a principios, obligaciones y mecanismos de gobernanza que los distintos países iberoamericanos puedan adoptar, adaptar y desarrollar según sus propias tradiciones jurídicas y prioridades de política pública.

El documento toma como insumo principal los marcos regulatorios de IA vigentes o en discusión en nueve jurisdicciones seleccionadas: Argentina, Brasil, Chile, Colombia, Ecuador, El Salvador, México, Perú y Uruguay. Considera también los proyectos legislativos en tramitación, las estrategias nacionales de IA y los instrumentos internacionales de referencia más relevantes. Las jurisdicciones seleccionadas constituyen una muestra no exhaustiva del espacio iberoamericano. Otras jurisdicciones iberoamericanas cuentan con marcos normativos recientes y directamente relacionados con las materias analizadas. Es el caso de Andorra, cuya Ley 29/2021, de protección de datos personales, regula, entre otras cuestiones, las decisiones individuales automatizadas, las evaluaciones de impacto, las vulneraciones de seguridad y las funciones y garantías de independencia de la autoridad de control. En consecuencia, las referencias a convergencias, brechas o tendencias mayoritarias se circunscriben a los países y marcos analizados.

2. Referencias internacionales

Antes de analizar el panorama regional, resulta indispensable establecer el mapa de instrumentos internacionales que sirven de referencia para el diseño de cualquier marco regulatorio de IA en el espacio iberoamericano. Estos instrumentos no son directamente vinculantes para la mayoría de los países de la región, pero operan como estándares de legitimidad y como orientadores de las opciones regulatorias disponibles.

2.1. Recomendación de la UNESCO sobre Ética de la IA (2021)

Adoptada el 23 de noviembre de 2021 por los 193 Estados miembros de la UNESCO -incluyendo todos los países iberoamericanos- es el primer instrumento global sobre ética de la IA de alcance universal. Establece cuatro valores fundamentales: respeto, protección y promoción de los derechos humanos y la dignidad; prosperidad del medioambiente y los ecosistemas; vivir en sociedades pacíficas, justas e interrelacionadas; y asegurar la diversidad y la inclusión.

Sus diez principios operativos -proporcionalidad e inocuidad, seguridad y protección, equidad y no discriminación, sostenibilidad, derecho a la intimidad y protección de datos, supervisión y determinación humana, transparencia y explicabilidad, responsabilidad y rendición de cuentas, sensibilización y educación, gobernanza y colaboración adaptativa- constituyen la referencia más comprehensiva disponible y la más apropiada para el contexto iberoamericano, dado que fue diseñada para su aplicación universal incluyendo países en desarrollo.

La Recomendación introduce, además, el concepto de evaluación de impacto en ética (EIE), análogo a las evaluaciones de impacto en protección de datos, como mecanismo preventivo obligatorio para sistemas de IA de alto riesgo. Esta innovación conceptual está siendo incorporada en varios proyectos legislativos iberoamericanos, incluyendo los de Costa Rica, México y Brasil.

2.2. Reglamento de IA de la UE - AI Act (2024)

El Reglamento (UE) 2024/1689, publicado el 12 de julio de 2024 y en vigor desde el 1 de agosto del mismo año, es el primer régimen integral vinculante de alcance general. Su aplicación es escalonada. Desde el 2 de febrero de 2025 rigen las disposiciones generales y las prohibiciones de usos de riesgo inaceptable. Desde el 2 de agosto de 2025 aplican las obligaciones para modelos de IA de propósito general y las disposiciones sobre gobernanza. La fecha de aplicación general es el 2 de agosto de 2026.

Ese calendario fue modificado por el Reglamento (UE) 2026/1744 (Digital Omnibus sobre IA), publicado en el Diario Oficial de la Unión Europea el 24 de julio de 2026 y en vigor desde el 27 de julio de 2026. La reforma difirió la aplicación de las obligaciones para sistemas de alto riesgo. Para los sistemas clasificados como de alto riesgo con arreglo al artículo 6, apartado 2, y al Anexo III, rigen desde el 2 de diciembre de 2027. Para los clasificados con arreglo al artículo 6, apartado 1, y al Anexo I -vinculados a productos sujetos a la legislación de armonización de la Unión-, rigen desde el 2 de agosto de 2028. Las obligaciones para modelos de IA de propósito general no fueron diferidas.

El AI Act establece cuatro categorías de riesgo. Los sistemas de riesgo inaceptable -social scoring, manipulación cognitiva subliminal, identificación biométrica en tiempo real en espacios públicos (salvo excepciones muy específicas)- están prohibidos. Los sistemas de alto riesgo -en salud, educación, empleo, infraestructuras críticas, justicia y aplicación de la ley- están sujetos a requisitos de gestión de riesgos, calidad de datos, transparencia, supervisión humana y registro. Los sistemas de riesgo de transparencia -chatbots, deepfakes, reconocimiento de emociones- tienen obligaciones de divulgación. Los sistemas de riesgo mínimo están prácticamente sin regulación específica.

Para los países iberoamericanos, el AI Act es relevante en dos dimensiones: como referencia técnica para el diseño de marcos nacionales, y como factor de competitividad. Su ámbito territorial alcanza a operadores no establecidos en la Unión en supuestos determinados. Comprende a los proveedores que introduzcan en el mercado o pongan en servicio sistemas de IA en la Unión, o introduzcan en el mercado modelos de IA de propósito general, con independencia de su lugar de establecimiento. Comprende también a los proveedores y responsables del despliegue establecidos en terceros países cuando la información de salida del sistema se utilice en la Unión. Alcanza, además, a importadores, distribuidores, fabricantes de productos que integren sistemas de IA bajo su propio nombre o marca, y representantes autorizados de proveedores no establecidos en la Unión. El ámbito material también está acotado. Quedan excluidos, entre otros, los sistemas destinados exclusivamente a fines militares, de defensa o de seguridad nacional; los desarrollados y puestos en servicio con la finalidad exclusiva de investigación y desarrollo científicos; las actividades de investigación, prueba o desarrollo previas a la introducción en el mercado, salvo las pruebas en condiciones reales; y el uso por personas físicas en actividades puramente personales de carácter no profesional. Las obligaciones concretas dependen del rol del operador y de la categoría de riesgo del sistema. Esa proyección extraterritorial, delimitada por supuestos precisos, genera una propensión hacia la armonización de estándares. Varios proyectos legislativos de la región -los de Brasil (PL 2338/2023), Chile (Boletín 15869-19) y Argentina (Exp. 2130-D-2025)- se inspiran explícitamente en su enfoque basado en riesgos, aunque con adaptaciones significativas.

2.3. Convenio 108+ del Consejo de Europa

El Convenio para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, en su versión modernizada (Convenio 108+), adoptado en el ámbito del Consejo de Europa, constituye el único tratado internacional jurídicamente vinculante en materia de protección de datos personales con vocación global, abierto a la adhesión de Estados no europeos. Países latinoamericanos como Argentina, México y Uruguay son Parte del Convenio 108 y de su Protocolo Adicional, y algunos de ellos han adherido a la versión modernizada han adherido a este instrumento.

El Convenio 108+ refuerza principios como la transparencia, la responsabilidad proactiva (accountability) y la proporcionalidad en el tratamiento de datos, e incorpora previsiones aplicables a tratamientos automatizados, incluyendo garantías frente a decisiones basadas exclusivamente en procesamiento automatizado y la necesidad de evaluar riesgos en operaciones de tratamiento susceptibles de generar impactos significativos.

Su relevancia para América Latina es directa: algunos de los procesos de reforma de los marcos normativos de protección de datos, como el de la Ley 25.326 argentina, toman al Convenio 108+ como referencia de modernización. Asimismo, para los países iberoamericanos que no son miembros del Consejo de Europa, el Convenio opera como estándar técnico a través de su recepción en los Estándares Iberoamericanos elaborados por la Red Iberoamericana de Protección de Datos.

2.4. Principios OCDE sobre IA (2019, actualizados 2024)

Los Principios sobre Inteligencia Artificial de la OCDE, adoptados en 2019 y reafirmados en 2024, constituyen uno de los principales marcos internacionales de referencia en la materia. Estos principios, a los que han adherido numerosos Estados -entre ellos Argentina, Perú, Brasil, Uruguay-, funcionan como guía para países que buscan alinearse con los estándares de la organización. Se trata de un instrumento no vinculante que representa un compromiso político de los países adherentes.

El instrumento establece cinco principios para una IA confiable: crecimiento inclusivo, desarrollo sostenible y bienestar; derechos humanos y valores democráticos, entre ellos la equidad y la privacidad; transparencia y explicabilidad; robustez, seguridad y protección; y rendición de cuentas.

Su relevancia radica especialmente en la articulación entre regulación e innovación: los Principios de la OCDE reconocen que el desarrollo de una IA confiable constituye una condición para la competitividad y el crecimiento económico, y no un obstáculo, lo que resulta clave para equilibrar la protección de derechos con los objetivos de desarrollo tecnológico en el contexto iberoamericano.

2.5. Carta Iberoamericana de Derechos Digitales (CIPDED) y Digital Economy Partnership Agreement (DEPA)

La Carta Iberoamericana de Principios y Derechos en los Entornos Digitales (CIPDED), aprobada por los jefes y jefas de Estado iberoamericanos en la XXVIII Cumbre Iberoamericana en el ámbito de la Secretaría General Iberoamericana, constituye el instrumento de soft law de mayor jerarquía política en la región en materia de derechos digitales. Reconoce y sistematiza principios y derechos aplicables en los entornos digitales, incluyendo garantías vinculadas al uso de tecnologías emergentes, con el objetivo de orientar la adopción y adecuación de legislaciones nacionales y políticas públicas. Su carácter es declarativo y no vinculante.

Por su parte, el Digital Economy Partnership Agreement (DEPA), del que Chile es parte, establece un marco innovador para la gobernanza de la economía digital, promoviendo la interoperabilidad normativa, la libre circulación de datos y la cooperación en tecnologías emergentes, incluida la inteligencia artificial, bajo principios de confianza y facilitación del comercio digital.

3. Panorama regional: enfoques nacionales vigentes

A continuación, se sistematiza el estado de la regulación de IA en las nueve jurisdicciones iberoamericanas seleccionadas, distinguiendo entre el marco legal vigente, el enfoque regulatorio adoptado y los aspectos más distintivos de cada sistema.

País	Marco IA	Enfoque regulatorio	Aspectos distintivos
Argentina	Sin ley marco específica. Res. AAIP 2/2023 (IA	Voluntario-principista. Debate entre enfoque	Proceso participativo AAIP con 173 aportes. Programa de Transparencia en

País	Marco IA	Enfoque regulatorio	Aspectos distintivos
	fiable). 50+ proyectos legislativos. Guía AAIP sept. 2024.	basado en riesgos (Exp. 2130-D-2025) y marco voluntario pro-innovación (Exp. 3422-D-2024).	IA. Debate sobre reforma Ley 25.326 como prerrequisito. Mesa Interministerial IA (2023). Proyecto Stargate Argentina (OpenAI, 2025).
Brasil	LGPD (2018, vigente 2020). Estrategia Brasileña de IA (EBIA, 2021). PL 2338/2023 aprobado por Senado dic. 2024.	Basado en riesgos (inspirado en AI Act). Creación de autoridad centralizada designada por el Ejecutivo.	ANPD con roles regulatorios activos. Mayor economía regional. Adecuación UE reconocida. Presidencia de la RIPD 2025-2026. Plano IA para el Bien de Todos (PBIA) con inversión en I+D e infraestructura.
Chile	Ley 21.719 de PDP (vigencia 1/12/2026). Proyecto de ley que modifica la Ley 21.719, ingresado al Senado el 1 de septiembre de 2026 (Boletín 18623-07). Política Nacional de IA (actualizada 2024). Proyecto Boletín 15869-19 en discusión.	Basado en riesgos alineado con RGPD/AI Act. Eje de Gobernanza Ética, Sostenibilidad y articulación internacional.	Ley de PDP. Proyecto 'Algoritmos Éticos' público-privado. Preocupación por impacto ambiental: IA verde como eje estratégico.
Colombia	CONPES 3975 (2019, 14 principios). Hoja de Ruta IA 2024 (5 ejes). Marco Ético IA 2021.	Gobernanza ética + educación/innovación + sectores industriales + datos + privacidad/ ciberseguridad. Soft law predominante.	Circular SIC 002/2024 sobre IA y datos personales. Destaca el seguimiento al impacto en mercado laboral. Énfasis en cooperación académica-industria.
Ecuador	Ley Orgánica de Protección de Datos Personales (LOPDP, 2021). Ley Orgánica para la Transformación Digital y Audiovisual (Tercer Suplemento del Registro Oficial Nº 245, 7/2/2023). Evaluación AIRA con PNUD (2024). EFIA-EC (Acuerdo Ministerial Nº MINTEL-MINTEL-2025-0030, Registro Oficial Suplemento Nº 206, 19/1/2026). Resolución Nº SPDP-SPD-2026-0009-R, reformada por Resolución Nº SPDP-SPD-2026-0037-R	Sin una ley integral específica de inteligencia artificial; con regulación sectorial en materia de protección de datos personales aplicable a sistemas de IA. Sandbox regulatorio con fundamento en el artículo 8 de la Ley Orgánica para la Transformación Digital y Audiovisual, su Reglamento General y los lineamientos del Ministerio de Producción, Comercio Exterior e Inversiones (20/4/2026). Evaluación AIRA alineada con ODS Agenda 2030.	Transformación digital del Estado como motor. Alta aplicación de IA en servicios financieros y atención al cliente. Diagnóstico sectorial en curso. Norma General para la Garantía del Derecho de Protección de Datos Personales en el Uso de Sistemas de Inteligencia Artificial (Resolución Nº SPDP-SPD-2026-0009-R), reformada por Resolución Nº SPDP-SPD-2026-0037-R: exige gestión de riesgos, evaluación de impacto cuando corresponda y auditorías de los sistemas de IA. EFIA-EC como hoja de ruta de la función ejecutiva para 2025-2029.
El Salvador	Ley de Fomento a la Inteligencia Artificial. Ley	Ley específica de IA centrada en el fomento, articulada	Doble institucionalidad: la ANIA para el fomento y la ACE para la supervisión del

País	Marco IA	Enfoque regulatorio	Aspectos distintivos
	para la Protección de Datos Personales (LPDP). Agencia Nacional de Inteligencia Artificial (ANIA) como rectora del fomento de la IA	con la supervisión en materia de datos personales. La Agencia de Ciberseguridad del Estado supervisa el tratamiento de datos personales mediante modelos de IA. Marco Integral de Evaluación de Riesgos a cargo de la ANIA, que concilia innovación tecnológica con seguridad pública y bienestar social, obligatorio para los sistemas que traten datos confidenciales, reservados o personales	tratamiento de datos personales en sistemas de IA. Obligaciones frente a decisiones automatizadas en usos comerciales o de acceso a derechos y servicios: informar si la decisión fue adoptada o impulsada por un sistema de IA, brindar una explicación comprensible y transparente del proceso, y habilitar la impugnación ante una persona natural competente para confirmarla, modificarla o revocarla. La ACE está facultada para elaborar y emitir políticas y guías.
México	Ley General de Protección de Datos Personales en Posesión de Sujetos Obligados y Ley Federal de Protección de Datos Personales en Posesión de Particulares (vigentes 21/03/2025), nuevos ordenamientos que abrogaron y sustituyeron a los anteriores. Proceso de actualización de la Ley Federal en curso. Múltiples proyectos legislativos en discusión.	Sin ley marco de IA. Propuestas: enfoque basado en riesgos con semáforo de 4 niveles (RENSIA), auditoría algorítmica (PNAA), reform. constitucional art. 73.	Extinción del INAI (dic. 2024) en el marco del proceso de Simplificación Orgánica: sus atribuciones fueron transferidas, en el ámbito correspondiente, a la Secretaría Anticorrupción y Buen Gobierno, que supervisa el sector privado a través de la Unidad de Protección de Datos Personales, con autonomía técnica y operativa. No existe una autoridad específica de supervisión de IA: la gobernanza es interinstitucional, con participación de la Agencia de Transformación Digital y Telecomunicaciones y de la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI). Propuesta de Consejo Nacional de IA y Plataforma Nacional de Auditoría Algorítmica.
Perú	Ley 31814 (2023) + Reglamento DS 115-2025-PCM. ENIA 2021-2026. PNTD 2030.	Principios de seguridad basada en riesgos, gobernanza de internet, ética, sostenibilidad, inclusión, privacidad. Revisable cada 2 años.	SGTD como autoridad técniconormativa. Centro Nacional de Innovación Digital e IA. Adhesión en proceso a la OCDE. Alineación con ODS. Riesgo de sobreregulación sectorial fragmentada.
Uruguay	Ley 18.331+Ley 19.670. Hoja de Ruta Ciencia de Datos (2019). Nueva ENIA en elaboración.	Marco de PDP maduro adaptable a IA. Proyecto de ley específico de sistemas de IA en discusión. Gobernanza IA en administración pública como eje.	URCDP activa. Adecuación UE reconocida. Comité Ejecutivo RIPD. Generación de guías técnicas sectoriales. Énfasis en gestión de datos como habilitador de IA.

El panorama, que incluye normas vinculantes, proyectos legislativos y políticas públicas, evidencia cuatro patrones distintos entre las jurisdicciones analizadas: (i)países con marcos legislativos específicos de IA ya vigentes o en vigencia diferida (Perú, El Salvador, Chile proyectado); (ii) países con proyectos legislativos avanzados en tramitación (Brasil, Argentina, México); (iii) países que regulan la IA predominantemente a través de sus marcos de protección de datos y estrategias nacionales de soft law (Uruguay, Ecuador, Colombia); y (iv) países con marcos incipientes que están en fase de diagnóstico y consulta.

4. Convergencias y brechas entre enfoques nacionales

El análisis comparado permite identificar con precisión las convergencias y las brechas que separan los enfoques nacionales de las jurisdicciones analizadas.

Dimensión	Convergencia regional	Brecha principal
Fundamento en derechos humanos	Todos los marcos nacionales analizados invocan los derechos humanos como fundamento de la regulación de IA y rechazan explícitamente el uso de IA para discriminar o erosionar derechos fundamentales.	Distinta jerarquía normativa: algunos países consagran el fundamento en la Constitución, otros solo en leyes ordinarias o resoluciones administrativas. Ausencia de mecanismos de enforcement constitucional específicos para IA.
Transparencia y explicabilidad	Principio presente en todos los instrumentos regionales e internacionales de referencia y en la mayoría de los proyectos legislativos nacionales de las jurisdicciones analizadas.	Dispar especificidad técnica: algunos marcos se limitan a enunciar el principio; pocos establecen obligaciones operativas.
Supervisión humana significativa	Reconocida como condición indispensable para sistemas de alto riesgo en todos los marcos analizados.	Falta de definición de los modelos de supervisión (human-in-the-loop, human-on-the-loop, human-out-of-the-loop) y de sus umbrales de aplicación.
Enfoque basado en riesgos	Predominante en los marcos más recientes de las jurisdicciones analizadas.	Heterogeneidad en la clasificación: algunos adoptan 4 niveles, otros 3, otros se limitan a alto/bajo. Ausencia de armonización de criterios para la asignación de categorías.
Privacidad por diseño	Incorporado como principio en algunas normas de las jurisdicciones analizadas.	Su implementación efectiva es incipiente en la mayoría de los países: carece de especificación técnica, mecanismos de verificación y obligaciones de documentación. No existe estándar regional de referencia para su aplicación práctica.

Dimensión	Convergencia regional	Brecha principal
Inclusión y no discriminación	Presente como principio en todos los marcos nacionales de las jurisdicciones analizadas y en todos los instrumentos internacionales de referencia.	Escasa regulación operativa: muy pocos marcos de las jurisdicciones analizadas establecen obligaciones de auditoría de sesgos, métricas de equidad, evaluaciones de impacto diferencial o representación de grupos históricamente discriminados en la gobernanza.
Sostenibilidad ambiental	Incorporada explícitamente por algunas legislaciones.	La mayoría de los marcos analizados omite el impacto ambiental de la IA. Ausencia de requisitos de eficiencia energética, ciclo de vida sostenible o compensación de huella de carbono en sistemas de IA.
Gobernanza institucional	Todos los marcos analizados reconocen la necesidad de una autoridad de supervisión.	Heterogeneidad en independencia, recursos y alcance. La mayoría de las autoridades de PDP de las jurisdicciones analizadas asumen funciones de supervisión de IA sin recursos ni mandato expreso. Propuestas de nuevos organismos sin evaluación de capacidad institucional preexistente.
Regulación de decisiones automatizadas	Algunas normativas contemplan disposiciones específicas.	La mayoría de las jurisdicciones analizadas carecen de regulación específica. Ausencia de armonización regional sobre los estándares mínimos del derecho a la explicación y la intervención humana.
Mecanismos de actualización periódica	Hay normativas con revisión cada 2 años y revisiones reglamentarias.	La mayoría de los marcos iberoamericanos carece de mecanismos formales de actualización periódica. No existe un proceso regional de revisión de estándares coordinado que habilite respuestas rápidas ante tecnologías emergentes.

5. Marco de recomendaciones regionales

Las diez recomendaciones que siguen a continuación surgen a partir del análisis de coincidencias y brechas observadas. Se presentan como sugerencias de optimización a partir de los criterios interpretativos identificados regionalmente.

1. Centralidad humana

Los sistemas de IA deben estar diseñados, desplegados y supervisados de manera que la persona humana permanezca en el centro de las decisiones que le afectan. La IA es una herramienta al servicio del desarrollo humano, no un sustituto de la agenda y la responsabilidad humanas. Toda decisión que afecte derechos, libertades o intereses significativos de las personas debe tener una intervención humana significativa y revisable.

2. Derechos humanos como límite no negociable

Los derechos humanos reconocidos en los sistemas constitucionales nacionales y en los instrumentos internacionales -incluyendo la dignidad, la igualdad, la privacidad, la no discriminación y el debido proceso- constituyen límites absolutos que no pueden ser negociados en ningún análisis de riesgo-beneficio. El enfoque basado en riesgos no puede autorizar la lesión de derechos fundamentales bajo ninguna circunstancia.

3. Transparencia y explicabilidad

Los responsables del desarrollo y despliegue de sistemas de IA deben ser capaces de explicar, de manera comprensible para las personas afectadas, la lógica, los criterios y los datos que determinan las decisiones de sus sistemas. Esta obligación es gradual: más intensa cuanto mayor sea el impacto de la decisión sobre el titular. La transparencia no es solo información técnica dirigida a reguladores, sino comunicación efectiva dirigida a ciudadanos.

4. No discriminación y equidad algorítmica

Los sistemas de IA no deben reproducir, amplificar ni generar nuevas formas de discriminación por razón de origen étnico o racial, género, edad, discapacidad, condición socioeconómica, orientación sexual, creencias religiosas o cualquier otra categoría protegida. Los responsables deben adoptar medidas activas de identificación y mitigación de sesgos, incluyendo auditorías periódicas de equidad algorítmica.

5. Privacidad y ética desde el diseño

La protección de la privacidad y los derechos de las personas no debe ser un añadido posterior, sino un criterio de diseño desde la fase inicial de desarrollo del sistema. El principio de ética por diseño exige que las consideraciones éticas -incluyendo equidad, transparencia y sostenibilidad- sean integradas en el ciclo de vida completo del sistema, desde la selección de datos de entrenamiento hasta el monitoreo postdespliegue.

6. Responsabilidad y rendición de cuentas

Debe existir en todo momento una persona o entidad jurídica identificable que sea legalmente responsable de las decisiones y los daños causados por un sistema de IA. La cadena de responsabilidad debe ser clara a lo largo del ciclo de vida del sistema: desarrolladores, proveedores, operadores y desplegadores tienen obligaciones proporcionales a su capacidad de influir en el diseño y uso del sistema.

7. Seguridad, robustez y confiabilidad

Los sistemas de IA deben ser seguros, técnicamente robustos y resistentes a manipulaciones, ataques y fallos imprevistos. La seguridad incluye tanto la ciberseguridad como la seguridad funcional: el sistema debe comportarse de manera predecible y dentro de los márgenes especificados, incluyendo en condiciones de uso inadecuado razonablemente previsible.

8. Inclusión digital y acceso equitativo

El desarrollo y despliegue de IA debe contribuir a reducir -y no ampliar- las brechas digitales existentes. Los sistemas de IA en servicios públicos esenciales (salud, educación, justicia, seguridad social) deben garantizar accesibilidad para personas con discapacidad, adultos mayores, minorías y personas en zonas de baja conectividad. La gobernanza de la IA debe asegurar la representación de grupos históricamente marginados.

9. Sostenibilidad ambiental

El ciclo de vida de los sistemas de IA -desde el entrenamiento de modelos hasta la inferencia y el mantenimiento de infraestructura- tiene impactos ambientales significativos, incluyendo consumo energético, huella de carbono y uso de recursos hídricos. Los responsables de sistemas de IA de escala deben evaluar y mitigar estos impactos, priorizando el uso de energías renovables y el diseño eficiente.

10. Gobernanza multiactor y cooperación regional

La regulación efectiva de la IA requiere la participación activa de múltiples actores -gobierno, industria, academia, sociedad civil y comunidades afectadas- en el diseño, monitoreo y actualización de los marcos normativos. A nivel regional, la cooperación entre países iberoamericanos es indispensable para evitar la fragmentación normativa y construir capacidades técnicas compartidas.

6. Categorías de riesgo adaptadas al contexto iberoamericano

El enfoque basado en riesgos es el más adoptado en los marcos regulatorios de referencia y en los proyectos legislativos más avanzados de las jurisdicciones analizadas. Ante esta tendencia se propone una clasificación en cinco categorías -en lugar de las cuatro del AI Act de la UE- para abarcar los riesgos sistémicos propios del contexto regional, particularmente la concentración de poder tecnológico y el impacto sobre poblaciones vulnerables.

Se aclara que los criterios de R0, R1 y R3-R4 adaptan las categorías del artículo 5 y Anexo III del Reglamento (UE) 2024/1689, con la salvedad de la cláusula abierta prevista para R1.

Categoría R0 - Riesgo inaceptable (uso prohibido)

Son usos que vulneran derechos fundamentales de manera irreversible o que son incompatibles con los valores democráticos y el Estado de derecho. Su prohibición es absoluta y no admite excepciones bajo ningún marco de autorización.

- Sistemas de puntuación social (social scoring) por autoridades públicas o privadas con efectos sobre derechos, acceso a servicios o libertad de movimiento.
- Manipulación cognitiva subliminal que explota vulnerabilidades de personas para modificar su comportamiento sin su conciencia.
- Identificación biométrica masiva y en tiempo real en espacios públicos sin orden judicial previa, salvo amenaza grave e inminente a la seguridad.
- Predicción de delincuencia o evaluación de riesgos penales basada exclusivamente en rasgos inmutables (etnia, origen, género).
- Categorización de personas en función de rasgos de personalidad inferidos de datos de comportamiento para negar acceso a derechos.
- Sistemas de IA destinados a la vigilancia masiva de actividades políticas, sindicales o religiosas lícitas.
- Deepfakes de contenido sexual no consentido y deepfakes de figuras públicas en contexto electoral sin etiquetado.

Categoría R1 - Riesgo alto (obligaciones intensas)

Son sistemas con potencial impacto significativo sobre derechos fundamentales, acceso a servicios esenciales o decisiones de alto impacto individual. Requieren evaluación de impacto previa, supervisión humana cualificada, documentación técnica completa, y registro ante la autoridad de supervisión.

- Sistemas de IA en administración de justicia penal, civil o familiar: predicción de reincidencia, evaluación de pruebas, asistencia en resoluciones judiciales.
- Sistemas de IA en procesos de selección, evaluación y despido laboral.

- Sistemas de IA en decisiones de acceso a crédito, seguros, subsidios públicos y prestaciones de seguridad social.
- Sistemas de IA en diagnóstico médico, priorización de atención sanitaria y determinación de tratamientos.
- Sistemas biométricos de identificación o verificación en contextos de aplicación de la ley, migración o control fronterizo.
- Sistemas de IA en educación: evaluación automatizada de estudiantes, admisiones, certificaciones.
- Infraestructuras críticas con control automatizado: redes eléctricas, agua, transporte.
- Sistemas de detección de emociones en contextos laborales, educativos o de aplicación de la ley.

También se considerarán de riesgo alto aquellos sistemas de inteligencia artificial que, aun cuando no se encuentren expresamente comprendidos en los supuestos anteriores, por su finalidad, contexto de utilización, escala, grado de autonomía o naturaleza de los datos tratados puedan producir efectos jurídicos o impactos significativos sobre los derechos y libertades fundamentales de las personas. Esta cláusula alcanza en particular a los sistemas que involucren personas menores de edad, grupos en condición de vulnerabilidad, datos personales sensibles, vigilancia sistemática, perfilamiento o decisiones que determinen o condicionen el acceso a derechos, servicios esenciales u oportunidades.

Categoría R2 - Riesgo de concentración sistémica (nuevas obligaciones regionales)

Esta categoría, específica para la región, aborda los riesgos derivados de la concentración de poder en grandes plataformas tecnológicas extrarregionales que no están cubiertas por la jurisdicción de los países iberoamericanos, pero afectan masivamente a sus ciudadanos. Requieren obligaciones de transparencia reforzada, interoperabilidad y cooperación con las autoridades nacionales.

- Modelos de IA de propósito general (LLMs y sistemas fundacionales) con alcance masivo en la región: obligaciones de documentación técnica, evaluación de riesgos sistémicos y acceso de autoridades regulatorias.
- Plataformas de redes sociales con sistemas de recomendación que influyen sobre el discurso público, las elecciones y la formación de opinión.
- Sistemas de publicidad conductual basados en perfilado de usuarios a escala regional.

- Infraestructuras de cómputo concentradas en pocas jurisdicciones con impacto sobre la soberanía de datos de los países iberoamericanos.

Categoría R3 - Riesgo limitado (obligaciones de transparencia)

Son sistemas con impactos moderados y circunscritos que requieren principalmente obligaciones de información y transparencia para garantizar que las personas sepan que están interactuando con IA y puedan ejercer sus derechos.

- Chatbots, asistentes virtuales y sistemas de atención al cliente automatizados: deben identificarse como IA.

- Sistemas de recomendación de contenido, publicidad personalizada y filtrado de información: deben divulgar la lógica básica de personalización.

- Sistemas de monitoreo y evaluación del rendimiento de trabajadores en contextos no relacionados con decisiones de alto impacto.

- Contenido generado o modificado con IA de manera significativa: debe etiquetarse como tal.

Categoría R4 - Riesgo mínimo (autorregulación orientada)

Son sistemas con impactos mínimos sobre derechos fundamentales, donde la regulación prescriptiva generaría más costos que beneficios. Se promueve la autorregulación sectorial, la adopción voluntaria de marcos de buenas prácticas, y la incorporación de cláusulas de responsabilidad en contratos de uso.

Esta categoría incluye la gran mayoría de aplicaciones de IA en productividad personal, entretenimiento, juegos, filtros de spam y herramientas de asistencia creativa sin impacto sobre terceros.

La aplicación de estas categorías responde a una regla de proporcionalidad: a mayor capacidad de un sistema de IA para afectar derechos o producir consecuencias significativas sobre una persona, mayor debe ser el nivel exigible de supervisión humana, transparencia, evaluación de riesgos, protección de datos, seguridad y rendición de cuentas.

7. Diseño ético y privacidad desde la base

El principio de privacy and ethics by design -diseño ético y privacidad desde la base- es una de las implementaciones operativas más importantes en el enfoque preventivo en la regulación de IA. Su incorporación en la legislación es necesaria pero insuficiente: requiere especificación técnica, mecanismos de verificación y una cultura organizacional de responsabilidad proactiva.

Para ello se debe entender el diseño ético como un proceso estructurado que abarca cinco dimensiones del ciclo de vida del sistema de IA:

7.1. Gobernanza de datos de entrenamiento

Los datos con los que se entrenan los sistemas de IA determinan sus capacidades, sus limitaciones y sus sesgos. El diseño ético exige que la gobernanza de datos de entrenamiento contemple, como mínimo:

- Auditoría de origen y representatividad de los datos: documentación del origen, verificación de la legalidad de su recolección, y evaluación de si el dataset representa adecuadamente a todas las poblaciones que serán afectadas por el sistema.

- Identificación y mitigación de sesgos históricos: análisis de si los datos reflejan patrones de discriminación histórica que el modelo podría perpetuar o amplificar.

- Minimización y proporcionalidad: uso solo de los datos estrictamente necesarios para el propósito declarado del sistema.

- Trazabilidad: documentación de todas las operaciones realizadas sobre el dataset, incluyendo transformaciones, aumentos y limpiezas.

- Bases legítimas para el uso de datos públicos en entrenamiento: criterios claros para determinar cuándo el interés legítimo o el interés público pueden sustentar el uso de datos disponibles públicamente, incluyendo salvaguardas técnicas como anonimización y privacidad diferencial.

7.2. Evaluación de impacto ético previa al despliegue

Todo sistema de IA que pertenezca a la categoría R1 o R2 debe someterse a una Evaluación de Impacto Ético (EIE) antes de su despliegue. Esta evaluación debe ser realizada por un equipo multidisciplinario -que incluya no solo técnicos sino también juristas, especialistas en ética, trabajadores sociales y representantes de los grupos potencialmente afectados- y debe cubrir, como mínimo:

- Descripción del sistema y su propósito declarado.
- Identificación de todas las poblaciones afectadas, con especial atención a grupos vulnerables.
- Evaluación de riesgos de discriminación, violación de privacidad, erosión de derechos y concentración de poder.
- Medidas de mitigación adoptadas y residuales remanentes.
- Plan de monitoreo postdespliegue con indicadores específicos.
- Mecanismo de impugnación disponible para las personas afectadas.

7.3. Diseño de la interfaz de supervisión humana

La supervisión humana efectiva requiere un diseño deliberado de la interfaz entre el sistema de IA y el operador humano que garantice que la intervención humana sea genuinamente significativa y no meramente ceremonial. Se distinguen tres modelos de supervisión cuya aplicación debe ser determinada en la EIE:

- Humano en el circuito (human-in-the-loop): el sistema asiste, pero la decisión es adoptada íntegramente por el humano. Requerido para decisiones de alto impacto irreversible (justicia penal, privación de libertad, decisiones de vida o muerte).
- Humano sobre el circuito (human-on-the-loop): el sistema decide autónomamente pero el humano supervisa y puede intervenir. Apropiado para decisiones de impacto significativo con volumen alto (crédito, selección laboral, servicios sociales).
- Humano fuera del circuito (human-out-of-the-loop): el sistema decide autónomamente sin supervisión activa. Solo admisible para decisiones de impacto mínimo y con mecanismos de auditoría retroactiva robustos.

7.4. Explicabilidad diferenciada por audiencia

En el derecho a la explicación se pueden distinguir tres niveles de explicabilidad que el sistema debe ser capaz de proveer:

- Explicación ciudadana: en lenguaje claro, comprensible para la persona afectada, qué criterios determinaron la decisión y cómo puede impugnarla.
- Explicación operacional: para los operadores y auditores internos, descripción del funcionamiento del modelo, los datos más influyentes en la decisión y los márgenes de incertidumbre.
- Explicación técnica: para la autoridad de supervisión y auditores externos, documentación completa del modelo, su arquitectura, datos de entrenamiento, métricas de rendimiento y evaluaciones de equidad.

7.5. Monitoreo continuo postdespliegue

Los sistemas de IA requieren monitoreo continuo porque su comportamiento puede deteriorarse por desviación del modelo (model drift), cambios en la distribución de los datos de entrada, o interacciones imprevistas con el entorno social. Se exige para sistemas de categoría R1 y R2:

- Indicadores de rendimiento ético predefinidos: métricas de equidad, tasas de error diferenciadas por grupos demográficos, tasa de impugnaciones.
- Alertas automáticas cuando los indicadores superen umbrales predefinidos.
- Revisión periódica por equipo multidisciplinario (mínimo anual para R1; semestral para R2).
- Canal de retroalimentación ciudadana: mecanismo accesible para que las personas afectadas reporten experiencias adversas.
- Protocolo de suspensión de emergencia: procedimiento para desactivar el sistema si se detectan riesgos graves no previstos.

8. Inclusión, desarrollo sostenible y derechos de grupos vulnerables

La inclusión y el desarrollo sostenible son condiciones de legitimidad en la regulación de IA en el contexto iberoamericano. Una región con niveles significativos de desigualdad, con brechas digitales profundas y con poblaciones históricamente marginadas requiere una IA que no amplíe esas brechas. Al mismo tiempo, la IA ofrece oportunidades reales para cerrarlas -si se despliega con intencionalidad inclusiva.

8.1. Grupos de protección especial en contextos de IA

Se pueden identificar seis grupos que requieren protecciones específicas en el diseño y despliegue de sistemas de IA, más allá de las protecciones generales:

- Niñas, niños y adolescentes: prohibición de uso de IA para perfilado conductual con fines comerciales en servicios dirigidos a menores; evaluación de impacto obligatoria en sistemas educativos; prohibición de sistemas de reconocimiento de emociones en entornos educativos.

- Pueblos indígenas y comunidades afrodescendientes: evaluación obligatoria de representación en datasets; prohibición de uso de IA para decisiones que afecten derechos colectivos sin consulta previa, libre e informada; reconocimiento de la dimensión colectiva del derecho a la protección de datos.

- Personas con discapacidad: accesibilidad de las interfaces de explicación y los mecanismos de impugnación; evaluación de accesibilidad universal en sistemas de servicios públicos con IA.

- Adultos mayores: accesibilidad de la información sobre el uso de IA en servicios esenciales; protección reforzada contra sistemas de manipulación conductual.

- Trabajadores en sectores con alta exposición a automatización: derecho a información sobre el uso de IA en procesos de contratación, evaluación y despido; obligación de negociación colectiva sobre introducción de sistemas de IA con impacto en el empleo.

- Migrantes y solicitantes de asilo: prohibición de uso de IA en decisiones de deportación sin intervención humana cualificada; evaluación de riesgos específicos de los modelos entrenados en datos que no representan a poblaciones migrantes.

8.2. Impacto en el mercado laboral

El impacto de la IA sobre el empleo constituye una de las preocupaciones más transversales en las jurisdicciones analizadas. Este eje se encuentra explícitamente incorporado en instrumentos de política pública como los CONPES en Colombia, la Estrategia Brasileña de Inteligencia Artificial (EBIA) y la Política Nacional de Inteligencia Artificial (PNIA) de Chile, así como en diversos proyectos legislativos y discusiones regulatorias en Argentina, donde el abordaje aún se encuentra en desarrollo. En este ámbito hay tres aspectos específicos a considerar:

- Evaluación de impacto laboral previa: para proyectos de automatización que afecten a más de 50 trabajadores, evaluación del impacto sobre el empleo, la naturaleza del trabajo y las condiciones laborales, con participación de los representantes sindicales.

- Derecho a la información y formación: los trabajadores tienen derecho a ser informados sobre los sistemas de IA que les evalúan o supervisan, y a recibir formación para adaptarse a los cambios tecnológicos.

- Seguimiento agregado de impacto: las autoridades laborales deben establecer sistemas de seguimiento del impacto de la IA en el mercado de trabajo, con publicación de informes periódicos.

8.3. Brecha digital e inclusión en el acceso

El diseño ético desde la base exige que los sistemas de IA desplegados en servicios públicos esenciales no requieran conectividad de alta velocidad, dispositivos de última generación o habilidades digitales avanzadas para ser utilizados. Los Estados iberoamericanos deben garantizar que la transición hacia servicios públicos asistidos por IA incluya canales alternativos no digitales para quienes no puedan acceder a las interfaces digitales.

8.4. Sostenibilidad ambiental de los sistemas de IA

El entrenamiento de grandes modelos de lenguaje y otros sistemas de IA a escala requiere volúmenes de energía y agua significativos. La instalación de infraestructura de cómputo en la región -incluyendo el proyecto Stargate Argentina (OpenAI/Sur Energy) anunciado en 2025- plantea preguntas sobre el impacto ambiental que los marcos regulatorios deben abordar.

Al efecto se podrían considerar los siguientes requisitos para centros de datos e infraestructura de cómputo dedicada a IA de escala:

- Declaración de impacto ambiental obligatoria, incluyendo consumo energético, fuente de energía y consumo de agua.

- Preferencia por el uso de energías renovables, con metas progresivas de descarbonización.
- Reporte anual público de la huella de carbono de las operaciones.
- Para sistemas de IA de propósito general con entrenamiento masivo: divulgación de la estimación del impacto ambiental del ciclo de entrenamiento.

9. Gobernanza institucional

La efectividad de cualquier marco regulatorio de IA depende de la calidad de su arquitectura institucional. El análisis de las jurisdicciones seleccionadas evidencia que la dispersión de competencias, la subordinación de los organismos de supervisión al poder ejecutivo, y la falta de recursos técnicos especializados son las principales preocupaciones. Lo ideal es un modelo de gobernanza multinivel articulado en tres escalas.

9.1. Nivel nacional

Una autoridad de supervisión de IA que tenga, como mínimo:

- Independencia funcional, orgánica y presupuestaria respecto del Poder Ejecutivo.
- Mandato explícito para la supervisión de sistemas de IA, diferenciado -aunque articulado- con el mandato de supervisión de protección de datos personales de las Autoridades de Protección de Datos.
- Capacidad técnica multidisciplinaria: la autoridad debe contar con equipos que integren expertise en ciencia de datos, ingeniería de sistemas, derecho, ética, ciencias sociales y economía.
- Facultades efectivas: emisión de normas vinculantes, investigación de oficio, sanción proporcional, acceso a sistemas y datos para auditoría, y habilitación de mecanismos de impugnación ciudadana.
- Cooperación internacional: mandato explícito de participar en espacios de debate internacional.
- Rendición de cuentas parlamentaria: informes periódicos ante el Congreso, sin subordinación jerárquica.

La articulación entre la autoridad de supervisión de IA y la autoridad de protección de datos personales admite distintos modelos institucionales: autoridad única con doble mandato, autoridades separadas con coordinación formal, o estructura de red con autoridades sectoriales coordinadas por un organismo central. Más allá del modelo elegido, la coordinación debe ser efectiva, sin vacíos ni superposición de competencias.

9.2. Nivel sectorial

Dado que la IA penetra en múltiples sectores con particularidades técnicas y jurídicas propias, la supervisión de IA en sectores de alto riesgo no puede recaer exclusivamente en una autoridad general. Ante esto, es una buena práctica la creación de unidades de IA dentro de los reguladores sectoriales existentes -salud, finanzas, competencia, educación, justicia- con competencias específicas para supervisar el uso de IA en su ámbito, articuladas con la autoridad nacional de supervisión a través de protocolos de coordinación.

9.3. Nivel regional

En este nivel resulta necesario el monitoreo de la intersección entre IA y derechos fundamentales:

- Mapa actualizado de marcos regulatorios de IA en la región, con análisis comparado de convergencias y brechas.
- Sistema de alertas tempranas sobre tecnologías emergentes con impacto en derechos fundamentales.
- Banco de buenas prácticas de evaluación de impacto ético, auditoría de sesgos y supervisión humana.
- Apoyo técnico a los países con menor capacidad institucional para el desarrollo de sus marcos regulatorios.
- Desarrollo de estándares técnicos regionales para la implementación de privacy y ethics by design.
- Coordinación con organismos internacionales (UNESCO, OCDE, Consejo de Europa, UE) para la armonización de estándares.

9.4. Mecanismos de participación ciudadana

La gobernanza de la IA debe institucionalizar espacios de participación ciudadana a tres niveles:

- Diseño de marcos regulatorios: consultas públicas obligatorias con plazos adecuados, accesibilidad para organizaciones de la sociedad civil y representantes de grupos vulnerables.
- Supervisión de sistemas específicos: mecanismos de retroalimentación ciudadana sobre el funcionamiento de sistemas de IA en servicios públicos, con canales accesibles y respuestas obligatorias.
- Gobernanza de la autoridad de supervisión: consejos consultivos multisectoriales con representación de academia, sociedad civil, industria y comunidades afectadas.

10. Mecanismos de actualización periódica

El ritmo de evolución de la IA hace que cualquier marco regulatorio estático quede obsoleto con rapidez. Ante esta situación existen, entre otros, cinco mecanismos complementarios de actualización que pueden integrarse en los marcos nacionales.

1. Revisión obligatoria quinquenal de los marcos nacionales, con publicación de un informe de evaluación que incluya el estado de implementación, los riesgos emergentes no cubiertos y las propuestas de modificación.
2. Habilitaciones reglamentarias a las autoridades de supervisión para actualizar la clasificación de riesgos, incorporar nuevas categorías de sistemas y establecer requisitos técnicos específicos sin necesidad de reforma legislativa.
3. Cláusulas de caducidad tecnológica para disposiciones específicamente vinculadas a tecnologías concretas: revisión obligatoria en un plazo de tres años.
4. Sandbox regulatorio: habilitación legal para la experimentación controlada con sistemas de IA innovadores bajo supervisión de la autoridad competente, con evaluación de los aprendizajes antes de la regulación definitiva.
5. Informe anual regional con un análisis del estado de la regulación de IA en el espacio iberoamericano, que sirva de insumo para la revisión de los marcos nacionales y de los Estándares Iberoamericanos.

11. Recomendaciones para la armonización regional

Para los legisladores nacionales

1. Adoptar las diez recomendaciones de este documento como cláusulas generales de aplicación supletoria en sus leyes marco de IA, garantizando coherencia interpretativa a nivel regional.

2. Establecer obligaciones de evaluación de impacto ético (EIE) para sistemas de categorías R1 y R2, con metodologías compatibles con los estándares del Observatorio Iberoamericano.

3. Garantizar la independencia de la autoridad de supervisión de IA. Evitar modelos de subordinación jerárquica al Poder Ejecutivo.

4. Incorporar las dimensiones de inclusión digital, impacto laboral y sostenibilidad ambiental como obligaciones de evaluación en los sistemas de categoría R1 y R2.

5. Establecer mecanismos formales de revisión periódica de los marcos regulatorios, con participación ciudadana institucionalizada.

6. Habilitar la creación de sandboxes regulatorios para la experimentación controlada, con protocolos de coordinación con la autoridad de supervisión.

Para las autoridades de control de datos y supervisión de IA

1. Adoptar el presente documento como referencia para el desarrollo de guías sectoriales, instrucciones y criterios de supervisión en el ámbito de la IA.

2. Desarrollar metodologías de evaluación de impacto ético compatibles con los estándares internacionales (UNESCO, AI Act, OCDE) y adaptadas al contexto iberoamericano.

3. Establecer mecanismos de alerta temprana sobre tecnologías emergentes con impacto en los derechos, en coordinación con el Observatorio Iberoamericano.

4. Fortalecer las capacidades técnicas de los equipos de supervisión, incluyendo formación en auditoría algorítmica, evaluación de sesgos y ciberseguridad de sistemas de IA.

Para la RIPD y el Observatorio Iberoamericano

1. Desarrollar un modelo de evaluación de impacto ético estándar para el espacio iberoamericano, con metodología abierta y herramientas de apoyo para organizaciones de distinto tamaño.

2. Establecer un protocolo regional de intercambio de información sobre sistemas de IA de alto riesgo en operación y sobre incidentes graves.

3. Crear un programa de asistencia técnica regional para países con menor capacidad institucional, especialmente en el diseño de marcos regulatorios y en la formación de equipos de supervisión.

4. Desarrollar una guía específica para la incorporación del principio de ethics by design en los procesos de contratación pública de sistemas de IA en los países de la región.

Para el sector privado

1. Considerar las diez recomendaciones de este documento.

2. Incorporar la evaluación de impacto ético como parte de los procesos de desarrollo de productos y servicios basados en IA, articulándola con los marcos existentes de evaluación de impacto en protección de datos.

3. Desarrollar capacidades internas de auditoría de sesgos y monitoreo continuo de métricas de equidad, especialmente para sistemas de categorías R1 y R2.

4. Participar activamente en los procesos de consulta pública sobre marcos regulatorios de IA, aportando evidencia técnica y comprometiéndose con la transparencia sobre el funcionamiento de los sistemas.

12. Referencias

Instrumentos internacionales

- Recomendación sobre la Ética de la Inteligencia Artificial. UNESCO, París, 23 de noviembre de 2021.
- Reglamento (UE) 2024/1689 – AI Act. Diario Oficial de la Unión Europea, 12 de julio de 2024.
- Reglamento (UE) 2026/1744 del Parlamento Europeo y del Consejo, de 8 de julio de 2026 (Digital Omnibus sobre IA). Diario Oficial de la Unión Europea, 24 de julio de 2026.
- Convenio 108+ del Consejo de Europa para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal (versión modernizada, 2018).
- Principios de la OCDE sobre Inteligencia Artificial, 2019 (actualizados 2024).
- Estándares de Protección de Datos Personales para los Estados Iberoamericanos, RIPD, 2017 y 2026.
- Plan Estratégico 2026–2030 de la RIPD, aprobado en el XXII Encuentro, Cartagena de Indias, junio 2025.
- Carta Iberoamericana de Principios y Derechos en los Entornos Digitales (CIPDED), SEGIB, 2023.
- Acuerdo de Asociación de Economía Digital (DEPA): Chile, Singapur y Nueva Zelanda.

Legislación y documentos nacionales

- Argentina: Ley 25.326; Resolución AAIP 2/2023; Guía para el Uso Responsable de IA (AAIP, sept. 2024); Exp. 2130-D-2025 (Gollán); Exp. 3422-D-2024 (Yeza); Proyecto Ley Marco Gobernanza IA; Res. 282/2025 Lineamientos CTI 2025-2027.
- Brasil: LGPD (Lei 13.709/2018); Estratégia Brasileira de Inteligência Artificial (EBIA, 2021); PL 2338/2023.

- Chile: Ley 21.719 (2024); Boletín 18623-07 (proyecto que modifica la Ley 21.719, ingresado el 1 de septiembre de 2026); Política Nacional de IA (actualizada 2024); Boletín 15869-19.

- Colombia: CONPES 3975 (2019); Marco Ético para la IA (2021); Hoja de Ruta IA (2024); Circulares SIC 002 y 003 (2024).

- Ecuador: Ley Orgánica de Protección de Datos Personales (LOPDP), Quinto Suplemento del Registro Oficial Nº 459, de 26 de mayo de 2021; Ley Orgánica para la Transformación Digital y Audiovisual, Tercer Suplemento del Registro Oficial Nº 245, de 7 de febrero de 2023; Evaluación AIRA-PNUD (2024); Estrategia para el Fomento del Desarrollo y Uso Ético y Responsable de la Inteligencia Artificial en el Ecuador (EFIA-EC), Acuerdo Ministerial Nº MINTEL-MINTEL-2025-0030, Registro Oficial Suplemento Nº 206, de 19 de enero de 2026.

- Ecuador: Resolución Nº SPDP-SPD-2026-0009-R de la Superintendencia de Protección de Datos Personales, por la que se expidió la Norma General para la Garantía del Derecho de Protección de Datos Personales en el Uso de Sistemas de Inteligencia Artificial, reformada por Resolución Nº SPDP-SPD-2026-0037-R.

- El Salvador: Ley de Fomento a la Inteligencia Artificial; Ley para la Protección de Datos Personales (LPDP).

- México: Ley General de Protección de Datos Personales en Posesión de Sujetos Obligados (2025); Ley Federal de Protección de Datos Personales en Posesión de Particulares (2025); Diseños institucionales para abordar la IA (análisis RIPD, 2026).

- Perú: Ley 31814 (2023); DS 115-2025-PCM (Reglamento); ENIA 2021-2026.

- Uruguay: Ley 18.331; Ley 19.670 (2018); Decreto 64/020; Hoja de Ruta Ciencia de Datos (2019).

- República de Corea: Ley Básica sobre Inteligencia Artificial (Ley N° 20676, vigente 22 enero 2026).

Bibliografía académica

- Arellano Toledo, W. (Dir.) (2023). Derecho, ética e inteligencia artificial. Valencia: Tirant lo Blanch.
- Rocha Espíndola, M.; Sansó-Rubert Pascual, D.; Rodríguez Dos Santos, N. (coords.) (2023). Inteligencia artificial y derecho: reflexiones jurídicas para el debate sobre su desarrollo y aplicación. Madrid: Dykinson.
- Granero, H. R. (2025). Inteligencia artificial en el derecho: entre la innovación y la protección de derechos. Buenos Aires: Eldial.com.
- Guilera, S., Pomares, J., Guillén, M. B., & Luvini, P. (2024). Marcos regulatorios de la IA: tendencias emergentes en la agenda legislativa argentina. CEPE Di Tella / GIDE Latam. Informe de política pública.