Propuesta de un Marco Ético y Normativo para una Inteligencia Artificial Responsable en América Latina

Catherine Muñoz
Daniel Rodríguez Maffioli
Danielle Zaror
Ricardo Baeza-Yates

Abril, 2023

OptIA, **Chile**

Resumen Ejecutivo

El presente documento propone una estructura teórica para la construcción de un marco ético y normativo que guíe el desarrollo e implementación responsable de tecnologías basadas en inteligencia artificial, en adelante “IA”, dentro del contexto latinoamericano. El propósito es guiar el desarrollo de políticas, estrategias y regulaciones nacionales en torno a la IA o servir de base para la elaboración de políticas internas de IA en el sector privado.

Con dicho fin, este trabajo se ha estructurado de la siguiente manera después de la introducción: en el Capítulo 2 se describen las bases fundamentales para un marco ético en materia de IA en Latinoamérica, a través de un análisis exploratorio comparativo de estándares y principios éticos establecidos por reconocidos organismos internacionales. En el Capítulo 3, se identifican y describen algunas propuestas normativas en discusión y regulaciones vigentes de diferentes jurisdicciones, cuya implementación busca, directa o indirectamente, establecer una regulación vinculante para el uso responsable de la IA. El análisis de los apartados 2 y 3 se realiza tanto dentro como fuera de Latinoamérica. En el Capítulo 4, y con base en los hallazgos derivados de los capítulos anteriores, se describe una propuesta general de marco ético y normativo para una inteligencia artificial responsable en Latinoamérica.

Sobre los Autores

- Catherine Muñoz (Santiago, Chile), abogada, magíster en Derecho Internacional, Inversiones y Comercio por la Universidad de Chile y Master of Laws en Leyes Internacionales por la Universidad de Heidelberg, Alemania, cuya tesis trata de la regulación de IA en la Unión Europea. Cuenta con un Diplomado en Política Comercial por la Universidad de Chile y cursos avanzados sobre propiedad intelectual. Consultora desde hace más de 11 años en el área de propiedad intelectual, regulación de tecnologías y políticas públicas, centrando su estudio en los últimos 4 años en materias relacionadas con impactos sociales y regulación de sistemas de inteligencia artificial y desarrollos digitales avanzados. Cofundadora de OptIA.

- Daniel Rodríguez Maffioli (San José, Costa Rica), máster en Regulación por la Universidad Carlos III de Madrid y certificado en Protección de Datos y regulación algorítmica de la London School of Economics. Es director ejecutivo de la Fundación Privacidad y Datos (PRIDAT) de Costa Rica, abogado del área de TMT de la firma Écija y Consultor en regulación digital. Miembro del comité científico de la consultora española Éticas Consulting, e investigador ad honorem en diferentes centros internacionales sobre política tecnológica y derechos digitales.

- Danielle Zaror Miralles (Santiago, Chile) es abogada, Magíster en Derecho Económico y Doctora en Derecho. Cuenta además con un Magíster en Derecho Económico y es diplomada en Regulación General de la Comisión de Mejora Regulatoria y también diplomada en Regulación de Mercado Eléctrico y de Telecomunicaciones, tiene además una Especialidad en Legislación Racional de la U. de Girona. Fue asesora en el Ministerio de Economía de Chile entre los años 2007 a 2012, en el Instituto Nacional de Estadísticas de Chile entre 2014 y 2017, y consultora del BID entre los años 2018 a 2020. Actualmente se desempeña como investigadora del Centro de Estudios en Derecho Informático de la Facultad de Derecho de la U. de Chile. Cofundadora de OptIA.

- **Ricardo Baeza-Yates** (Palo Alto, EE. UU.), ingeniero eléctrico y magíster en ciencia de la computación de la Universidad de Chile. Dr. en Ciencia de la Computación de la Univ. de Waterloo, Canadá. Director de Investigación del Instituto de Inteligencia Artificial Experiencial de Northeastern University, EE. UU. Es también catedrático a tiempo parcial en los Departamentos de Tecnologías de la Información y de las Comunicaciones de la Universitat Pompeu Fabra (UPF) en Barcelona y Ciencia de la Computación de la Universidad de Chile en Santiago. Actualmente es miembro del comité de políticas tecnológicas de la ACM, miembro del comité de ética tecnológica de la IEEE y miembro experto del grupo de trabajo de IA responsable de GPAI. Experto en algoritmos, tecnologías de búsqueda, minería de datos e IA responsable. Cofundador de OptIA.

Sobre OptIA

El Observatorio Público para la Transparencia e Inclusión Algorítmica se construye por un grupo profesionales interdisciplinarios que comparten la preocupación por el impacto de la implementación y el uso de algoritmos, particularmente sistemas de inteligencia artificial y tecnologías digitales sobre la sociedad chilena, su impacto sobre grupos vulnerables, históricamente marginados y excluidos, compartiendo asimismo, la necesidad de ser un agente colectivo de cambio para la generación de políticas públicas justas e inclusivas en relación con estas tecnologías. Conozca más sobre nosotros en https://optia.cl/

Contenido

1. Introducción

El acelerado avance de la inteligencia artificial ("IA") está generando cambios sustanciales en la forma en como interactuamos, nos desarrollamos y crecemos como sociedad y como individuos. La tecnología en cuestión conlleva la promesa implícita de ser una aliada para el bienestar de las personas y un apoyo en la búsqueda de soluciones a los grandes desafíos que nos enfrentamos actualmente.

Si bien es posible aprovechar las ventajas y los beneficios de la IA, es importante tomar conciencia que, como parte de su desarrollo e implementación, deben existir procesos de toma de decisiones responsables y reflexivos. Estos procesos deben ponderar los límites de esta tecnología y las potenciales consecuencias negativas que su implementación puede generar.

A nivel mundial se han documentado los daños reales y los riesgos imprevistos generados por un uso indiscriminado e irreflexivo de esta tecnología. Es por esta razón que se han elaborado diversas iniciativas a nivel internacional, tanto públicas como privadas, traducidas en instrumentos vinculantes y no vinculantes, que tienen por finalidad considerar, mitigar y contrarrestar los daños y peligros mencionados.
.
El presente informe examina los diferentes instrumentos deónticos y jurídicos que se han propuesto a nivel comparado para abordar los riesgos apuntados y los contrasta con la realidad y las necesidades de la región latinoamericana. A partir de dicho análisis, se propone un marco ético y regulatorio propio que guíe a los países de Latinoamérica en la construcción de sus propias políticas públicas, con el fin de que el aprovechamiento de las bondades de estos sistemas sociotécnicos sea responsable, transparente, inclusivo, y contextualizado a nuestra región.

Para este propósito, el Capítulo 2 se inicia con un estudio de los principales marcos y principios éticos de adhesión voluntaria sobre IA que han propuesto distintas organizaciones internacionales. Se ha hecho especial énfasis en los Principios de la OCDE sobre Inteligencia Artificial y la Recomendación sobre la Ética de la IA elaborada por UNESCO en el año 2021.

Adicionalmente, en forma de anexo, se realizó una identificación e individualización de los principales tratados internacionales de Derechos Humanos aplicables en la región, que son pertinentes al desarrollo e impacto de la IA y se exploraron las potenciales vulneraciones a Derechos Humanos derivadas del despliegue de sistemas de IA, a la luz del modelo de marcos éticos del Consejo de Europa.

En el Capítulo 3 se identifican leyes y propuestas legislativas sobre IA que existen o se discuten a nivel internacional. El diagnóstico y, en general, el análisis comparado de marcos éticos y regulatorios contenidos en los acápites 2 y 3, permitieron elaborar en el Capítulo 4 una propuesta de marco ético y legal para el despegue de un ecosistema responsable de la IA en Latinoamérica.

Antes de profundizar en el contenido de este informe, es importante brindar un marco conceptual básico en el que se defina, en primer lugar, qué es o qué no es la "inteligencia

artificial". Existen muchas definiciones posibles, pero, parafraseando a Kate Crawford[1], para empezar, la IA tiene dos problemas: no es ni inteligente ni artificial. No es inteligente si la comparamos con la inteligencia humana que puede comprender casi toda la semántica y el contexto de una situación dada, incluso cuando la situación es nueva. De hecho, donde la IA es superior al ser humano es en otros aspectos tales como la velocidad de cálculo y la capacidad de memoria, pero no aún en el consumo de energía o su capacidad cognitiva completa.

Tampoco es completamente artificial pues para que la IA funcione, se necesita de gran cantidad de ser humanos altamente calificados para construirla. Asimismo, se consumen enormes cantidades de recursos naturales para desarrollarla y utilizarla, incluyendo electricidad y todos los componentes materiales de un computador.

Otro de los problemas del concepto es que la mayoría de sus definiciones no consideran el contexto social y tecnológico donde está inmersa la IA. Es decir, un sistema o innovación de este tipo puede ser creada o utilizada para un propósito en un lugar determinado y luego para otro fin en otro lugar, implicando diferentes riesgos e impactos. Por ello, toda definición de IA debe abarcar la naturaleza sociotécnica de su concepción y función en la sociedad.

Todo ello conduce al marco ético de la IA. El sesgo tecnológico de los seres humanos nos lleva a humanizar la IA, atribuyéndole nociones de "justicia" a "ética" a los algoritmos, cuando ambas características debieran ser solamente usadas para los seres humanos detrás de esos algoritmos. Lo apropiado es hablar del uso ético de la IA, el cual debe anclarse en uno o varios principios éticos, planificando un diseño, implementación y prueba cautelosa del sistema, antes de su uso masivo. Esto es crucial, por ejemplo, en servicios públicos que buscan dirigir el uso de la IA hacia el bienestar social, pero al mismo tiempo alinear los incentivos privados con el interés general.

Una de las funciones de la ética en este campo es no agravar la situación de vulnerabilidad de algunas personas ni las condiciones desfavorables o de desigualdad en las que ya de por sí viven y de las cuales no es causante la IA por sí misma. En este contexto, un sistema de IA puede mejorar la vida de la mayoría empeorando al mismo tiempo la vida de unos pocos. Debe ser, entonces, la sociedad en su conjunto la que decida si un sistema debe utilizarse o no, y en qué casos se puede utilizar. Este aspecto de evaluación ética de los riesgos e impactos, junto al aspecto legal, son componentes esenciales del marco orientador que proponemos.

[1] Localizable en: https://www.theguardian.com/technology/2021/jun/06/microsofts-kate-crawford-ai-is-neither-artificial-nor-intelligent

2. Bases Fundamentales para un Marco Ético de la IA

En este capítulo se estudian alguno de los principales marcos éticos de adhesión voluntaria en materia de IA que han sido elaborado por relevantes organizaciones internacionales, alguna de las cuales tienen relación directa o vínculos con América Latina. Estos principios tienen en común que son de carácter universal y que su objetivo principal es orientar y guiar la construcción de políticas éticas y regulatorias para un uso responsable de la IA a nivel estatal, siendo plenamente aplicables a la adopción de políticas y normas internas nivel privado.

En esta sección, se hace un análisis exploratorio de los Principios de la OCDE sobre Inteligencia Artificial y del Marco de Clasificación de Sistemas de IA propuesto por dicha organización, la Recomendación sobre la Ética de la IA elaborada por UNESCO, las Recomendaciones de la OMS para el uso de inteligencia artificial en el sector salud y la Iniciativa global del IEEE sobre la ética de los sistemas autónomos e inteligentes, y su publicación "*Ethically Align Design"* para el diseño responsable y ético de dichos sistemas.

2.1 Organización para la Cooperación y el Desarrollo Económico

América Latina ya cuenta con cuatro países (México, Chile, Colombia y Costa Rica) que forman parte de la Organización para la Cooperación y el Desarrollo Económico (OCDE). Además, Argentina, Brasil y Perú ya iniciaron su proceso con miras a adherirse. La adhesión de estos países a esta importante organización de buenas prácticas, así como la influencia del organismo en las políticas de la región, hace prudente la revisión de los distintos insumos, instrumentos y lineamientos emitidos por ésta para informar la elaboración de políticas públicas y regulaciones en materia de Inteligencia Artificial (IA).

La OCDE publicó sus *Principios de la OCDE sobre Inteligencia Artificial*[2] en mayo del 2019, convirtiéndose en una de las primeras organizaciones a nivel mundial en emitir lineamientos éticos relacionados con el uso de IA. En el año 2022, para dar continuidad a los esfuerzos puestos en marcha, la red de expertos de IA de la OCDE ("ONE AI") publicó un documento denominado *Marco para la clasificación de los sistemas de Inteligencia Artificial*[3] (el "Marco"), una metodología para asistir a los responsables políticos y a otros actores relevantes a clasificar los sistemas de IA con base en una política de riesgos en función de su posible impacto en los ámbitos cubiertos por los Principios de IA de la OCDE para una adecuada implementación normativa de estos.

Por su parte, en junio de 2021, la OCDE emitió un primer informe sobre el estado de la implementación de los Principios de la OCDE, denominado *State of Implementation of the OECD AI Principles: Insights from National AI Policies.*[4] El informe analiza las diferentes estrategias y políticas de IA elaboradas por países miembros para la materialización de sus principios.

---

[2] OECD (2021), "*Recommendation of the Council on Artificial Intelligence*", May 29, 2021, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#mainText

[3] OECD (2022), "OECD Framework for the Classification of AI systems", OECD Digital Economy Papers, No. 323, OECD Publishing, Paris, https://doi.org/10.1787/cb6d9eca-en.

[4] OECD (2021), "State of implementation of the OECD AI Principles: Insights from national AI policies", *OECD Digital Economy Papers*, No. 311, OECD Publishing, Paris,https://doi.org/10.1787/1cd40c44-en.

A continuación, se efectuará un breve análisis de cada uno de los documentos mencionados, con el objeto de identificar aspectos que puedan contribuir a la elaboración de un marco de gobernanza adecuado para la IA y una estrategia integral de Inteligencia Artificial.

a. ***Principios de la OCDE sobre Inteligencia Artificial***

El 21 de mayo de 2019, los 36 países miembros de la OCDE, junto con Argentina, Brasil, Colombia, Costa Rica, Perú y Rumanía, suscribieron en París, los Principios de la OCDE sobre Inteligencia Artificial (de ahora en adelante los "Principios" o la "Recomendación del Consejo"). Dichos principios fueron elaborados por un grupo de expertos conformado por más de 50 personas pertenecientes a distintos sectores como gobiernos, instituciones académicas, empresas y más. Hoy en día, los Principios se encuentran vigentes y además de los países miembros de la OCDE, ocho países figuran como adherentes a los mismos.

La Recomendación del Consejo está dividida en un preámbulo y dos secciones sustantivas. El preámbulo ofrece un resumen del contexto en el cual se adoptan los Principios y, además, incluye una serie de definiciones en relación con el concepto de Inteligencia Artificial, su ciclo de vida y su funcionamiento. A su vez, la primera sección sustantiva propone una serie de principios para un "enfoque responsable en apoyo a una IA digna de confianza". Mientras que, la segunda parte sustantiva consiste en "políticas nacionales y de cooperación internacional para el apoyo de una Inteligencia Artificial digna de confianza".

El preámbulo establece una serie de elementos importantes para entender las dos secciones sustantivas que le siguen. Por un lado, el preámbulo hace un resumen general de los diferentes instrumentos tanto a lo interno de la OCDE, como los instrumentos de Derecho Internacional Público e incluso, las motivaciones que inspiraron a la adopción de los Principios. La relevancia de este Capítulo es poder situar la Recomendación dentro una jerarquía de normas, lo cual, a su vez, resalta el carácter normativo-legal de los Principios.

Por otro lado, se ofrece una serie de definiciones relevantes para entender el campo de aplicación de los principios. Por ejemplo, se explica la definición de sistema de inteligencia artificial, lo cual resulta útil para zanjar debates con respecto a una definición que ha evolucionado sobre el tiempo. Asimismo, se establece lo que se entiende por el "ciclo de vida" de un sistema de IA, conocimiento de IA, actores de IA y partes interesadas, definiciones que se utilizan en el desarrollo de los Principios.

La primera sección sustantiva informa sobre los cinco **principios gobernantes** para un enfoque responsable en apoyo a una IA confiable. Vale rescatar que los llamados a aplicar estos principios son los actores de IA, es decir, aquellos que juegan un rol activo en el ciclo de vida de un sistema de IA, incluyendo organizaciones e individuos. A continuación, un resumen de los principios:

a) *Crecimiento inclusivo, desarrollo sostenible y bienestar*: las partes interesadas deberían adoptar de manera proactiva un enfoque responsable de IA que promueva resultados beneficiosos para los individuos y el planeta. Por ejemplo, el mejoramiento de las capacidades humanas, la creatividad humana, la inclusión de grupos subrepresentados, reducción de la inequidad social o económica.

b) *Valores centrados en el humano y la igualdad*: los actores de IA deben respetar el Estado de Derecho, los derechos humanos y los valores democráticos durante el

ciclo de vida de un sistema de IA. Asimismo, deberían instaurar garantías y mecanismos, como la intervención de un humano en la toma de decisiones finales, según el contexto y los desarrollos más recientes.

c) *Transparencia y que sea explicable ("explicabilidad"):* los actores de IA deben asegurar la transparencia, de manera que las personas conozcan cuándo están interactuando con un sistema de IA, comprendan las decisiones adoptadas sobre ellas con base en esta tecnología y puedan impugnar dichas decisiones. Notar que hay técnicas de aprendizaje automático, como el aprendizaje profundo, que son cajas negras y por lo tanto difíciles de explicar. Por otro lado, es un mito de que se pierda calidad si no se usan estas técnicas, todo depende del caso de uso.

d) *Robustez, seguridad y salvaguarda:* los sistemas de IA deben ser seguros de tal manera que ante cualquier situación de uso normal o incluso un mal uso previsible no generen un riesgo desproporcionado a la seguridad. Esto incluye la trazabilidad de los datos durante todo el ciclo de vida del sistema de IA. También, los actores de IA deberían tener sus roles delimitados y con mecanismos para mitigar los riesgos, especialmente con aspectos relacionados con la vida privada, la seguridad digital y parcialidad.

e) *Responsabilidad*: los actores de IA son responsables por su buen funcionamiento.

La segunda sección sustantiva propone **guías para la elaboración de políticas nacionales** y de cooperación internacional. Esta sección refuerza que los países deben adoptar los principios consagrados en la primera sección sustantiva dentro de sus políticas nacionales y de cooperación internacional, con particular énfasis en las pequeñas y medianas empresas. A continuación, un resumen de los cinco principios enunciados:

a) **Inversión en la investigación y desarrollo en materia de IA**: los poderes públicos deberían prever inversiones a largo plazo e inversiones privadas en la investigación y desarrollo interdisciplinario de la IA. Asimismo, los poderes públicos deberían considerar inversiones públicas y privadas que propicien el desarrollo responsable de la IA a través de bases de datos abiertas.

b) **Favorecer la instauración de un ecosistema digital para la IA**: los poderes públicos deberían favorecer el desarrollo y la accesibilidad a un ecosistema digital para elaborar una IA digna de confianza. Por ejemplo, mecanismos como los fideicomisos de datos (o "*data trusts*").

c) **Modelar un campo de acción favorable a la IA**: los poderes públicos deberían propiciar un ambiente que apoye una transición ágil entre la investigación, desarrollo y lanzamiento de sistemas IA. Conviene crear espacios controlados para la experimentación, prueba y determinar la escala adecuada. La regulación debe ser revisada y adaptada para fomentar la innovación y competencia.

d) **Reforzar la capacidad humana y preparar la transformación del mercado laboral**: los poderes públicos deberían trabajar en estricta colaboración con las partes interesadas para preparar la transformación del mercado laboral y de la sociedad. Esto significa dotar a las personas de los medios para utilizar e interactuar con los sistemas de IA. Asimismo, los poderes públicos deberían entrar en un diálogo con la sociedad para asegurar una transición adecuada en el uso de la IA.

También, los poderes públicos deberían trabajar con las partes interesadas para asegurarse que los sistemas de IA que se utilicen de manera responsable en el medio laboral.

e) **Favorecer la cooperación internacional para una IA digna de confianza**: los poderes públicos, incluyendo aquellos países en vías de desarrollo, junto con las partes interesadas deberían cooperar activamente para garantizar la aplicación de los Principios. En este sentido, se deberían aprovechar los espacios como la OCDE y otros organismos internacionales para compartir los conocimientos en materia de IA. Además, se debería promover la creación de normas técnicas internacionales basadas en investigaciones de consenso. Incluso, los poderes públicos deberían fomentar métricas comparables internacionalmente para medir el avance de aplicación de los principios en la investigación, desarrollo y lanzamiento de IA.

b. El marco de la OCDE para la clasificación de los sistemas de IA

El marco publicado en 2022 clasifica los sistemas de IA en cuatro dimensiones: 1) Contexto, 2) Datos o "*Input*", 3) Modelo de IA y 4) Resultado o "*Outcome*". Estas dimensiones se basan en la visión conceptual de un sistema de IA genérico establecida en trabajos anteriores de la OCDE y las detallamos a continuación.

El **contexto** describe el entorno socioeconómico en el que se despliega y utiliza el sistema de IA. Las características principales de esta dimensión incluyen el sector en el que se despliega el sistema (por ejemplo, sector salud, sector financiero, policía y seguridad, etc.); su función empresarial; su carácter crítico; el nivel de despliegue; las partes interesadas afectadas; los usuarios; y las repercusiones en los derechos humanos y en el bienestar.

Los **datos o "*input*"** se refieren a los datos y/o a los inputs que utiliza el modelo de IA para construir una representación del entorno. Las características principales de esta dimensión son: la procedencia u origen de los datos, el método de recopilación (por máquinas o por humanos), la estructura y el formato, y las propiedades de los datos (por ejemplo, el tipo o el acceso, si incorpora datos personales o no, etc.)

La tercera dimensión, **Modelo de IA,** es una representación computacional de los procesos, objetos, ideas, personas y/o interacciones del mundo real que incluye suposiciones sobre la realidad. En palabras simples, es el algoritmo y el motor del sistema de IA que procesa la información y los datos para, a partir de éstos, generar un resultado. Las características principales de esta dimensión incluyen las características del modelo; cómo se construye el sistema (por ejemplo, utilizando conocimiento experto, aprendizaje automático, redes neuronales, etc.); y cómo se utiliza (por ejemplo, para qué objetivos y utilizando qué medidas de rendimiento).

Por último, la **tarea y el resultado** se refieren a las tareas que realiza el sistema, por ejemplo, la personalización de contenido, el reconocimiento de una persona, la asignación de un crédito, entre otros. Se refiere a la acción resultante de la puesta en marcha del modelo, que influye en el contexto. Las características principales de esta dimensión incluyen la(s) tarea(s) del sistema; la autonomía de acción; los sistemas que combinan tareas y acciones, como los vehículos autónomos; y las áreas de aplicación principales, como la visión por ordenador.

Tal y como se observa, cada una de las dimensiones del Marco está compuesta, a su vez, por "sub-dimensiones". Estas "sub-dimensiones" son las características o parámetros que deben evaluarse en una aplicación específica de IA para luego cuantificar su riesgo frente a los principios éticos y valores democráticos concebidos por la OCDE. A modo de ejemplo, en la valoración del contexto de un sistema de IA, la OCDE propone que se evalúen los siguientes elementos, pues de éstos dependerá el nivel de riesgo (bajo o alto) del sistema:

a) **Sector industrial**: por ejemplo, el nivel de impacto de la IA será diferente en el sector salud que en el sector construcción.

b) **Áreas funcionales de negocio**: por ejemplo, un algoritmo para selección de personal tendrá diferentes implicaciones que un algoritmo utilizado para logística.
c) **Actividades críticas**: en este parámetro se evalúa si el sistema de IA se implementará o no en actividades o infraestructuras de interés general cuya interrupción o disrupción pueda afectar la salud, la seguridad o la economía. Este factor repercute en el Principio de Seguridad y Robustez.

d) **Escala de implementación y madurez tecnológica**: se valora, entre otros, la cantidad de personas afectadas, lo cual dependerá de si se trata de un proyecto piloto o un despliegue a nivel de sector, país o región en general.

e) **Usuarios del Sistem***a*: se analiza si los que interactuarán con el sistema son expertos en IA (desarrolladores), o amateurs. Este parámetro es importante frente al Principio de Transparencia y explicabilidad.

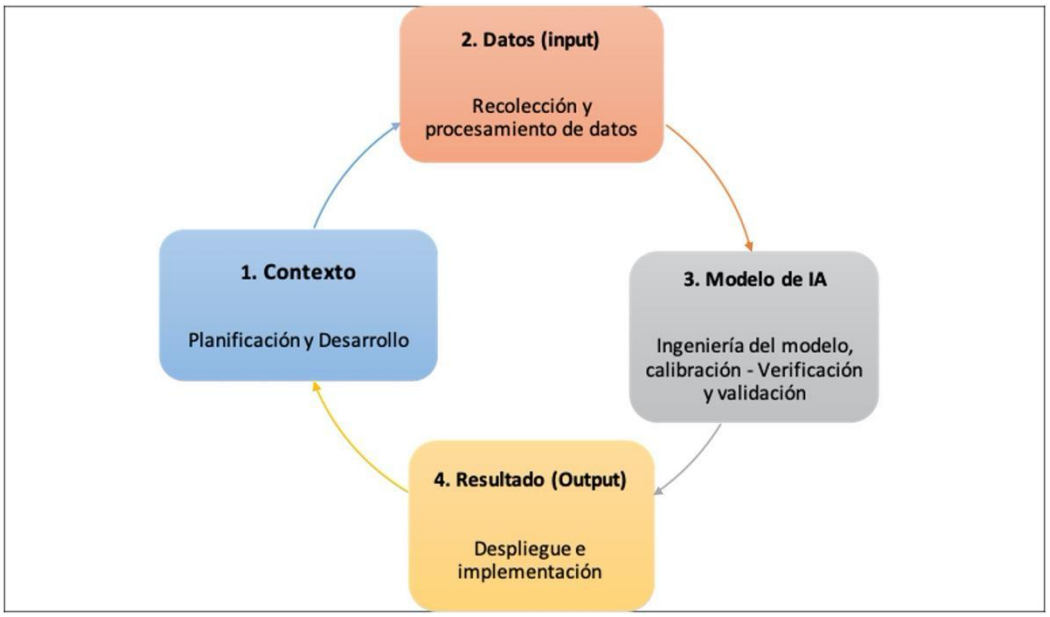
f) **Partes interesadas afectadas por el sistema y opcionalidad**: se evalúa la naturaleza de la población afectada. Por ejemplo, menores, personas en situación de vulnerabilidad, consumidores, trabajadores, entre otros. Además, se analiza si el usuario final puede optar por ser excluido de los efectos del sistema o cambiarse a otro.

g) **Beneficios y riesgos para los derechos humanos y valores democráticos**: se valora si el sistema tiene impacto sobre los derechos humanos de las personas (por ejemplo, derecho a la salud, derecho a la privacidad, debido proceso, igualdad y no discriminación, entre otros).

Cada dimensión posee sus propias características, las cuales deben ser evaluadas una a una, para obtener una concepción integral del riesgo involucrado en cada aplicación de IA. Ahora bien, de acuerdo con el Marco, las cuatro dimensiones pueden ser asociadas a diferentes etapas del ciclo de vida del sistema de IA. Esto, a su vez, permite identificar los actores relevantes en cada etapa de ese ciclo; identificación que resulta relevante frente al Principio de Rendición de Cuentas (*Accountability*). Por ejemplo, en la dimensión de "Contexto" la OCDE identifica como actores relevantes a los operadores del sistema, es decir, a quienes lo diseñan, planifican y evalúan.

En términos generales, el Marco de la OCDE provee una metodología útil para identificar las características más importantes de un sistema de IA, lo que, a su vez, permite entender de mejor manera su potencial impacto y medir su riesgo integral frente a los principios y valores que la OCDE ha dispuesto y tomar acciones ya sea en el plano de políticas públicas o normativas vinculantes.

El Ciclo de Vida de los sistemas de IA y su relación con la dimensión correspondiente se ilustran en la siguiente figura.



Fuente: Ciclo de vida basado en el "*OECD FRAMEWORK FOR THE CLASSIFICATION OF AI SYSTEMS – PUBLIC CONSULTATION ON PRELIMINARY FINDINGS*", 2021.

c. Informe sobre el Estado de Implementación de los Principios de la OCDE

En junio de 2021, se publicó el informe sobre el estado de implementación de los principios de la OCDE sobre la IA. En dicho documento se establece una medición inicial del progreso alcanzado en cuanto a la implementación de los Principios en una serie de países miembros de la OCDE. El estudio se centra en los esfuerzos implementados por los Gobiernos para llevar a cabo las recomendaciones de la OCDE tendientes a materializar los principios, todo ello a través de las diferentes fases del ciclo de vida de las estrategias de IA.

En síntesis, el informe destaca que los países se encuentran en distintas fases del desarrollo e implementación de sus políticas de IA. Además, también varían los enfoques de gobernanza para la construcción de sus estrategias de IA. Algunos países optan por asignar la función a Ministerios o Departamentos ya establecidos mientras otros se inclinan por asignar la labor a un nuevo órgano de coordinación intergubernamental.

Para invertir en I+D en IA, los países están financiando institutos y proyectos nacionales de investigación relacionados con la IA mediante subvenciones; consolidando redes de investigación y plataformas de colaboración en IA; dando prioridad a las inversiones en IA en sectores económicos específicos; aplicando políticas de innovación orientadas a misiones relacionadas con la IA; y adquiriendo sistemas de IA para el sector público.

En cuanto a infraestructura tecnológica necesaria para avanzar en la materia, el informe destaca que el acceso abierto a los datos del sector público sigue siendo una prioridad, ya que las estrategias nacionales de datos se centran cada vez más en la IA para fomentar un ecosistema digital sólido para la IA e impulsar la I+D en IA. El informe indica que las políticas para promover el acceso a los datos públicos y las iniciativas que permiten compartir datos del sector privado incluyen fideicomisos de datos, presas de datos y espacios de datos.

Como parte de su estrategia de IA, varios países han desarrollado o están desarrollando repositorios centralizados y accesibles de conjuntos de datos públicos abiertos, como registros sanitarios gubernamentales anonimizados y datos de satélites (por ejemplo, Chile, España, Estados Unidos, Noruega y Portugal). Otros están buscando formas de incentivar el intercambio de datos en el sector privado (por ejemplo, el Reino Unido y la Unión Europea).

2.2 UNESCO: Acuerdo Mundial Sobre Ética de la Inteligencia Artificial

En marzo del 2020, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), convocó a un grupo especial de expertos encargados de elaborar un proyecto de *Recomendación sobre la Ética de la Inteligencia Artificial*[5], grupo que en mayo de ese mismo año entregó un borrador que fue sometido a una consulta pública entre junio y agosto de 2020 entre múltiples partes interesadas. El texto definitivo fue publicado en noviembre de 2021.

El acuerdo parte por un preámbulo, en que lo más significativo está representado por el reconocimiento que la tecnología tiene en la vida de las personas y su impacto en el pensamiento, interacciones y en la adopción de decisiones, por la contribución de UNESCO a la colaboración entre las naciones para asegurar el respeto universal de los derechos humanos y por el convencimiento de la *necesidad de proveer un instrumento de derecho internacional que permita entregar una orientación responsable del uso de las tecnologías de inteligencia artificial*.

El preámbulo también resalta la idea que la tecnología presenta desafíos diferentes para los países de ingresos medios y bajos por lo que requieren promoción y protección de su cultura a fin de desarrollar economías sostenibles, al mismo tiempo que reconoce que el uso de esta tecnología puede ser beneficiosa para el medio ambiente pero que igualmente puede generar daños y repercusiones en el medio ambiente y los ecosistemas.

Finalmente, el instrumento deja claro que la preocupación por la ética no debe ser un freno a la innovación, sino que, por el contrario, ésta debe ser una oportunidad para estimular prácticas nuevas y responsables de investigación e innovación que además pueden influir poderosamente en la elaboración de medidas políticas y normas jurídicas.

En cuanto a su ámbito de aplicación, la recomendación “trata de las cuestiones éticas relacionadas con la IA. Aborda la ética de la IA como una reflexión normativa sistemática, basada en un marco integral y evolutivo de valores, principios y acciones interdependientes, que puede guiar a las sociedades a la hora de afrontar de manera responsable los efectos conocidos y desconocidos de las tecnologías de la IA en los seres humanos, las sociedades y el medio ambiente y los ecosistemas, y les ofrece una base para aceptar o rechazar las tecnologías de la IA.”

La recomendación se hace cargo de las repercusiones éticas en las principales esferas de competencia de la UNESCO, esto es, educación, ciencia, identidad y diversidad cultural, comunicación e información, con la idea de proporcionar un marco universal de valores, principios y acciones que ayuden a los Estados a orientar sus leyes y cualquier otro instrumento relativo a la IA, para orientar las acciones para incorporar la ética en las

[5] UNESCO (2020), Outcome document: first draft of the Recommendation on the Ethics of Artificial Intelligence, https://unesdoc.unesco.org/ark:/48223/pf0000373434

personas, grupos, comunidades y empresas en todo el ciclo de vida de la IA, para promover el respeto por la dignidad humana, el entendimiento multidisciplinario y promover el acceso equitativo a los progresos de la ciencia.

Entre los valores que aspira promover se encuentran:

a) **Respeto, protección y promoción de la dignidad humana, los derechos humanos y las libertades fundamentales.** La dignidad de cada ser humano constituye la base del sistema indivisible de derechos humanos y libertades fundamentales y es esencial a lo largo del ciclo de vida de los sistemas de IA.

   Ningún ser humano debería sufrir daños físicos, económicos, sociales, políticos o mentales durante ninguna etapa del ciclo de vida de los sistemas de IA.

   En el marco de esas interacciones, las personas nunca deberían ser cosificadas, no debería socavarse su dignidad, y sus derechos humanos nunca deberían ser objeto de violación o abusos.

   La promoción y respeto por los derechos humanos deben ser una prioridad de los gobiernos, el sector privado, la sociedad civil, las organizaciones internacionales, las comunidades técnicas y las universidades.

b) **Prosperidad del medio ambiente y los ecosistemas.** Todos los actores que participan en el ciclo de vida de los sistemas de IA deben respetar el derecho internacional y las leyes, normas y prácticas nacionales pertinentes, como la precaución, concebidas para la protección y la restauración del medio ambiente y los ecosistemas y para el desarrollo sostenible. Deberían reducir el impacto ambiental de los sistemas de IA, incluida, entre otras cosas, su huella de carbono, para asegurar la minimización del cambio climático y los factores de riesgo ambiental, y prevenir la explotación, la utilización y la transformación no sostenibles de los recursos naturales que contribuyen al deterioro del medio ambiente y a la degradación de los ecosistemas.

c) **Diversidad e inclusión.** El respeto, la protección y la promoción de la diversidad y la inclusión deberían garantizarse a lo largo del ciclo de vida de los sistemas de IA, como mínimo de conformidad con el derecho, las normas y los principios internacionales de derechos humanos, así como con la diversidad y la inclusión demográfica, cultural, social y de género. Toda tendencia a la homogeneización debería ser vigilada y corregida.

   La diversidad de las elecciones de estilo de vida, creencias, opiniones, expresiones o experiencias personales, incluida la utilización opcional de sistemas de IA y la concepción conjunta de estas arquitecturas, no debería restringirse en modo alguno durante ninguna etapa del ciclo de vida de dichos sistemas.

d) **Vivir en armonía y paz.** Los actores de la IA deberían desempeñar una función propicia para la vida armoniosa y pacífica, garantizando un futuro interconectado en beneficio de todos. El valor de vivir en armonía y paz apunta al potencial de los sistemas de IA para contribuir a lo largo de su ciclo de vida a la interconexión de todas las criaturas vivas entre sí y con el medio natural.

Los principios que promueve este instrumento son:

a) **Proporcionalidad e inocuidad.** En su virtud, la elección de un método de IA debería justificarse de las siguientes maneras: a) el método de IA elegido debería ser conveniente y proporcional para lograr un objetivo legítimo determinado; b) el método de IA elegido no debería repercutir negativamente en los valores fundamentales enunciados en el presente documento; c) el método de IA debería ser adecuado al contexto y basarse en fundamentos científicos rigurosos. En los casos que implican decisiones de vida o muerte, la decisión final debería ser adoptada por un ser humano.

b) **Seguridad y protección.** En su virtud, los daños no deseados (riesgos de seguridad) y las vulnerabilidades a los ataques (riesgos de protección) deberían evitarse a lo largo del ciclo de vida de los sistemas de IA para garantizar la seguridad y la protección de los seres humanos y del medio ambiente y los ecosistemas. La seguridad y la protección de la IA será posible gracias al desarrollo de marcos de acceso a los datos que sean sostenibles, respeten la privacidad y fomenten un mejor entrenamiento de los modelos de IA que utilicen datos de calidad.

c) **Equidad y no discriminación.** En su virtud, **l**os actores de la IA deberían promover la justicia social, respetando la equidad. La equidad supone compartir los beneficios de las tecnologías de la IA en los planos local, nacional e internacional, teniendo en cuenta las necesidades específicas de los diferentes grupos de edad, sistemas culturales, diferentes grupos lingüísticos, personas con discapacidad, niñas y mujeres, y poblaciones desfavorecidas, marginadas y vulnerables.

   Los actores de la IA deberían hacer todo lo posible por reducir los sesgos socio técnicos inadecuados basados en prejuicios a fin de garantizar la equidad de dichos sistemas. Debería existir la posibilidad de poder recurrir contra la determinación y la discriminación algorítmicas injustas.

   La equidad supone abordar la brecha digital y de conocimientos y las desigualdades mundiales en lo que respecta al acceso a la tecnología, los datos, la conectividad, los conocimientos y las competencias, así como a la participación de las comunidades afectadas como parte de la etapa de concepción, de manera que todas las personas sean tratadas equitativamente.

d) **Sostenibilidad.** En su virtud, debe llevarse a cabo una evaluación continua de los efectos sociales, culturales, económicos y ambientales de las tecnologías de la IA con pleno conocimiento de las repercusiones de dichas tecnologías en la sostenibilidad como un conjunto de metas en constante evolución en toda una serie de dimensiones, como las que se identifican actualmente en los Objetivos de Desarrollo Sostenible (ODS) de las Naciones Unidas.

e) **Intimidad y Protección de Datos.** Se trata de un derecho esencial para la protección de la dignidad, la autonomía y la capacidad de actuar de los seres humanos, debe ser respetada, protegida y promovida a lo largo del ciclo de vida de los sistemas de IA, tanto a nivel personal como colectivo. Es fundamental que los datos de la IA se recopilen, utilicen, compartan, archiven y supriman de forma coherente con los valores y principios enunciados.

Los sistemas algorítmicos requieren evaluaciones exhaustivas del impacto en la privacidad que incluyan también consideraciones sociales y éticas de su utilización y un empleo innovador del enfoque de privacidad desde la etapa de concepción.

f) **Supervisión y decisión humanas.** En su virtud, siempre debe ser posible atribuir la responsabilidad ética y jurídica, en cualquier etapa del ciclo de vida de los sistemas de IA, a personas físicas o a entidades jurídicas existentes. La supervisión humana se refiere, por tanto, no solo a la supervisión humana individual, sino también a la supervisión pública, según corresponda. Un sistema de IA nunca podrá reemplazar la responsabilidad y la rendición de cuentas final por parte de un ser humano.

g) **Transparencia y explicabilidad.** Es una condición previa fundamental para garantizar el respeto, la protección y la promoción de los derechos humanos fundamentales y los principios éticos. La transparencia es necesaria para que la legislación nacional e internacional pertinente en materia de responsabilidad funcione eficazmente.

El grado de transparencia y explicabilidad debería ser siempre adecuado al contexto, ya que es preciso encontrar el equilibrio justo entre la transparencia y la explicabilidad y otros principios como la seguridad y la protección. Las personas tienen derecho a saber cuándo se toma una decisión sobre la base de algoritmos de IA y, en esas circunstancias, exigir o solicitar explicaciones e información a empresas del sector privado o instituciones del sector público.

La explicabilidad supone hacer inteligibles los resultados de los sistemas de IA y facilitar información sobre ellos, pero también se refiere a la inteligibilidad de la entrada, salida y comportamiento de cada componente algorítmico y la forma en que contribuye a los resultados de los sistemas.

h) **Responsabilidad y rendición de cuentas.** En su virtud, los actores de la IA deberían asumir la responsabilidad ética y jurídica de conformidad con el derecho nacional e internacional vigente, en particular el derecho, los principios y las normas internacionales de derechos humanos, y las directrices éticas establecidas durante todo el ciclo de vida de los sistemas de IA respetando, protegiendo y promoviendo los derechos humanos y al tiempo que se fomenta la protección del medio ambiente y los ecosistemas. Para esto deberían elaborarse mecanismos adecuados de supervisión, evaluación del impacto y diligencia debida para garantizar la rendición de cuentas respecto de los sistemas de IA y de su impacto a lo largo de su ciclo de vida, garantizando la auditabilidad y la trazabilidad (del funcionamiento) de los sistemas de IA, en particular para intentar solucionar cualquier conflicto con los derechos humanos y las amenazas al bienestar del medio ambiente y los ecosistemas.

i) **Sensibilización y educación.** En su virtud, debería promoverse mediante una educación abierta y accesible, la participación cívica, las competencias digitales y la capacitación en materia de ética de la IA, la alfabetización mediática e informacional y la capacitación dirigida conjuntamente por los gobiernos, las organizaciones intergubernamentales, la sociedad civil, las universidades, los medios de comunicación, los dirigentes comunitarios y el sector privado, y teniendo en cuenta la diversidad lingüística, social y cultural existente, a fin de garantizar una participación pública efectiva, de modo que todos los miembros de la sociedad puedan adoptar decisiones informadas sobre su utilización de los sistemas de IA y estén protegidos de influencias indebidas.

j) **Gobernanza y colaboración adaptativas y de múltiples partes interesadas**. En su virtud, los datos deben ser utilizados de manera lícita. La soberanía de los datos significa que los Estados, en cumplimiento del derecho internacional, regulan los datos generados dentro de sus territorios o que pasan por ellos y adoptan medidas para la regulación efectiva de los datos sobre la base del respeto del derecho a la privacidad y otros derechos humanos.

La participación de las diferentes partes interesadas a lo largo del ciclo de vida de los sistemas de IA es necesaria para la gobernanza inclusiva de la IA, el aprovechamiento compartido de los beneficios de la IA y el progreso tecnológico equitativo y su contribución a los objetivos de desarrollo. Entre las partes interesadas figuran, entre otros, los gobiernos, las organizaciones intergubernamentales, la comunidad técnica, la sociedad civil, los investigadores y los círculos universitarios, los medios de comunicación, los responsables de la educación, los encargados de formular políticas, las empresas del sector privado, las instituciones de derechos humanos y los organismos de fomento de la igualdad, los órganos de vigilancia de la lucha contra la discriminación y los grupos de jóvenes y niños.

La recomendación también entrega directrices para los ámbitos de acción política donde enfatiza sobre la necesidad de hacer evaluaciones de impacto ético, sobre la creación de mecanismos de gobernanza y de administración ética, de políticas de datos, en desarrollo y cooperación internacional, en medio ambiente y ecosistemas, en asunto de género, cultura, educación e investigación, comunicación e información, economía y trabajo y en salud y bienestar social. Para todo lo anterior, es necesario tener políticas claras, flexibles y transparentes acerca del tratamiento de los datos y en particular de su calidad.

2.3 Organización Mundial de la Salud: Ética y Gobierno de la IA para la Salud

El 26 de junio de 2021, la Organización Mundial de la Salud (OMS), luego de dos años de consultas realizadas por a través de expertos (en salud, medicina, derechos humanos, tecnología y ética) ha emitido una recomendación[6] cuyo objetivo general es visibilizar las oportunidades que una tecnología como la inteligencia artificial ofrece pero también alertar que si se usa indebidamente ésta puede causar daños, por lo que la ética y los derechos humanos deben tener un lugar central en su concepción, de manera que su texto es una guía que busca aumentar sus potenciales beneficios, reduciendo los riesgos allí donde sea posible.

La OMS reconoce que esta tecnología es prometedora para la práctica de la salud pública y de la medicina en general y que para aprovechar sus beneficios se deben abordar los desafíos éticos para los sistemas de atención de la salud, los profesionales y los beneficiarios de los servicios médicos y de salud pública, por lo que “oportunidades y desafíos están indisolublemente ligados”.

En particular, el documento señala que, por un lado, “La inteligencia artificial puede utilizarse, lo que ya se está haciendo en algunos países ricos, para mejorar la velocidad y la precisión del diagnóstico y la detección de enfermedades; facilitar la atención clínica; reforzar la investigación en el ámbito de la salud y el desarrollo de medicamentos, y apoyar

[6] OMS (2021), Ética y Gobernanza para la Inteligencia Artificial en Salud, https://www.who.int/publications/i/item/9789240029200

diversas intervenciones de salud pública, como la vigilancia de la morbilidad, la respuesta a los brotes y la gestión de los sistemas de salud".

Otro de los beneficios que advierte este informe dice relación con que "la IA también podría permitir que los pacientes tuvieran un mayor control de su propia atención de salud y comprendieran mejor la evolución de sus necesidades". También, "podría facilitar el acceso a los servicios de salud en los países con escasos recursos y las comunidades rurales, donde los pacientes a menudo tienen dificultades para acceder a los agentes de salud o al personal médico".

Por otro lado, el informe advierte "del peligro de sobreestimar las ventajas de la IA en el ámbito de la salud, sobre todo cuando esto se hace en detrimento de inversiones y estrategias básicas que son necesarias para lograr la cobertura sanitaria universal (…) las oportunidades conllevan desafíos y riesgos, como la recopilación y utilización poco éticas de los datos sobre salud; los sesgos codificados en los algoritmos, y los riesgos que presenta la IA para la seguridad del paciente, la ciberseguridad y el medio ambiente".

La guía es extensa pero precisa en cuanto a las temáticas abordadas, así, por ejemplo, cuando se refiere a la aplicación de IA en salud, circunscribe su recomendación a materias relacionadas con cuidado de la salud, investigación salud y desarrollo de fármacos, gestión y planificación en sistemas de salud, salud pública y vigilancia sanitaria.

En cuanto a las leyes, políticas y principios que se aplican al uso de la IA para la salud, la guía entrega recomendaciones en materia de IA y derechos humanos, leyes y políticas en materia de protección de datos, leyes y políticas vigentes relacionadas con datos de salud, principios para el uso de la IA en salud y leyes y políticas de bioética.

En el capítulo que da nombre a la guía se ofrecen 6 principios éticos que deberían regir la IA en sus aplicaciones en salud:

a) **Preservar la autonomía del ser humano**: En su virtud, significa que los seres humanos deberían seguir siendo dueños de los sistemas de atención de salud y las decisiones médicas; se debería preservar la privacidad y la confidencialidad, y los pacientes deben dar su consentimiento informado y válido por medio de marcos jurídicos adecuados para la protección de datos.

b) **Promover el bienestar y la seguridad de las personas y el interés público.** En su virtud, los diseñadores de tecnologías de IA deberían cumplir los requisitos normativos en materia de seguridad, precisión y eficacia para indicaciones o usos bien definidos. Se deben instaurar medidas de control de la calidad en la práctica y de mejora de la calidad en la utilización de la IA.

c) **Garantizar la transparencia, la claridad y la inteligibilidad.** En su virtud, se exige que se publique o documente información suficiente antes de la concepción o el despliegue de una tecnología de IA. Esa información debe ser fácilmente accesible y facilitar consultas y debates provechosos sobre la concepción de la tecnología y sobre el uso que se debería hacer o no de esta.

d) **Promover la responsabilidad y la rendición de cuentas.** En su virtud, incumbe a las partes interesadas velar por que estas sean utilizadas en condiciones apropiadas y por personas debidamente formadas. Se deberían instaurar mecanismos eficaces para

que las personas y los grupos que se vean perjudicados por decisiones basadas en algoritmos puedan cuestionarlas y obtener reparación.

e) **Garantizar la inclusión y la equidad.** En su virtud, requiere que la IA aplicada a la salud sea concebida de manera que aliente la utilización y el acceso equitativos en la mayor medida posible, con independencia de la edad, el sexo, el género, el ingreso, el origen étnico, la orientación sexual, la capacidad u otras características amparadas por los códigos de derechos humanos.

f) **Promover una IA con capacidad de respuesta y sostenible.** En su virtud, los diseñadores, desarrolladores y usuarios deberían evaluar de forma continua y transparente las aplicaciones de la IA en situación real a fin de determinar si esta responde de manera adecuada y apropiada a las expectativas y las necesidades.

Luego de la enumeración de estos principios, el documento aborda los principales desafíos éticos en materia de IA para el cuidado de la salud, poniendo el énfasis en la evaluación de si se debe o no utilizar IA en salud, como implementar IA cuando existe una brecha digital, como recoger datos, cómo rendir cuentas responsablemente por decisiones que se adopten por medio de IA, la toma de decisiones autónomas, los sesgos de los programas de IA, entre otros desafíos.

El documento también innova en la aportación de reflexiones vinculadas con los *Regímenes de responsabilidad por inteligencia artificial para la salud*, en particular respecto de la atención clínica, compensación de errores, el papel que les incumbe a las agencias reguladoras y el papel que les compete a los países de ingreso medio y bajo.

Por último, esta guía ofrece una serie de elementos para la construcción de un *Marco de Gobernanza para la IA en el ámbito de la Salud*.

2.4 La Ética de los Sistemas Autónomos e Inteligentes del IEEE

El Instituto de Ingenieros Eléctricos y Electrónicos ("IEEE", por sus siglas en inglés) es una de las organizaciones técnico-profesionales más grandes y reconocidas del mundo en materia de tecnología. Desde hace algunos años, el IEEE lanzó su "*Iniciativa global del IEEE sobre la ética de los sistemas autónomos e inteligentes*"[7], con la misión de garantizar que todas las partes implicadas en el diseño y el desarrollo de sistemas de IA reciban educación, formación y capacitación para dar prioridad a las consideraciones éticas, de modo que estas tecnologías avancen en beneficio de la humanidad.

Esta concepción fue materializada en el documento "*Ethically Align Design*"[8] ("Diseño éticamente alineado"). El propósito es que este informe proporcione ideas y recomendaciones pragmáticas y orientativas, que sirvan de referencia para el trabajo de tecnólogos, educadores y actores responsables de políticas públicas.

[7] IEEE, The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, https://standards.ieee.org/industry-connections/ec/autonomous-systems.html

[8] The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, First Edition. IEEE, 2019. https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/ autonomous-systems.html

El contenido del informe abarca análisis científico, recursos, principios generales, y recomendaciones. Ofrece, además, guías para la elaboración de estándares, certificaciones, regulaciones o legislación para el diseño, producción y utilización de sistemas inteligentes u autónomos que se alineen de forma demostrada al bienestar social.

Los tres pilares sobre los que descansa el informe son:

a) **Valores humanos universales**: Para el IEEE, los sistemas autónomos/inteligentes ("SA/I") pueden ser una fuerza para el bien de la sociedad siempre que se diseñen para respetar los derechos humanos, se alineen con los valores humanos y aumenten de forma holística el bienestar, al tiempo que empoderen al mayor número de personas posible. También deben diseñarse para salvaguardar el medio ambiente y los recursos naturales. Estos valores deben guiar a los responsables políticos, así como a los ingenieros, diseñadores y desarrolladores. Los avances en inteligencia artificial deben estar al servicio de todas las personas, en lugar de beneficiar únicamente a pequeños grupos, a una sola nación o a una corporación.

b) **Autodeterminación política y control sobre los datos**: Los SA/I -si se diseñan y aplican correctamente- tienen un gran potencial para fomentar la libertad política y la democracia, de acuerdo con los preceptos culturales de cada sociedad, cuando las personas tienen acceso y control sobre los datos que constituyen y representan su identidad. Estos sistemas pueden mejorar la eficacia y la responsabilidad de los gobiernos, fomentar la confianza y proteger la intimidad y privacidad, pero sólo cuando las personas tienen agencia y control sobre su identidad digital, y sus datos están protegidos de forma demostrable.

c) **Fiabilidad técnica**: En última instancia, los SA/I deben prestar servicios en los que se pueda confiar. Esta confianza significa que los SA/I cumplirán de forma fiable, segura y activa los objetivos para los que fueron diseñados, al tiempo que promueven los valores humanos que pretenden reflejar. Las tecnologías deben ser supervisadas para garantizar que su funcionamiento cumple con los objetivos éticos predeterminados que se ajustan a los valores humanos y respetan los derechos codificados. Además, deberían desarrollarse procesos de validación y verificación, incluyendo aspectos de explicabilidad, que podrían conducir a una mejor auditabilidad y a la certificación de los SA/I.

El documento contiene múltiples guías e insumos valiosos para el diseño ético de IA, según sus áreas de impacto y según los principales riesgos que se han logrado identificar asociados a esta tecnología.

Para efectos de este trabajo, sin embargo, conviene destacar los **Principios Éticos** que, según la IEEE, descansan sobre los pilares antes mencionados y que deben ser considerados al diseñar, desarrollar e implementar sistemas de inteligencia artificial:

a) **Derechos Humanos.** El sistema autónomo/inteligente se creará y operará para respetar, promover y proteger los derechos humanos internacionalmente reconocidos.

b) **Bienestar.** Los creadores de SA/I adoptarán el aumento del bienestar humano como principal criterio de éxito para el desarrollo. El IEEE ha creado un estándar (IEEE 7010-2020) para evaluar este criterio en el despliegue de soluciones de IA.

c) **Agencia de datos.** Los creadores de SA/I deberán empoderar a los individuos con la capacidad de acceder y compartir de forma segura sus datos, para mantener la capacidad de las personas de tener control sobre su identidad.

d) **Eficacia.** Los creadores y operadores de SA/I deberán aportar pruebas de la eficacia y adecuación a los fines de los SA/I.

e) **Transparencia.** La base de una decisión basada en un SA/I debe ser siempre descubrible.

f) **Rendición de cuentas.** El SA/I debe ser creado y operado para proporcionar una justificación inequívoca de todas las decisiones tomadas.

g) **Conciencia del mal uso.** Los creadores de los SA/I deberán protegerse de todos los posibles usos indebidos y riesgos de los SA/I en funcionamiento.

h) **Competencia.** Los creadores de los SA/I deberán especificar y los operadores deberán cumplir con los conocimientos y habilidades necesarios para un funcionamiento seguro y eficaz.

Un elemento adicional que merece destacarse del trabajo del IEEE es la elaboración de un estándar para la incorporación de criterios éticos en el diseño de sistemas inteligentes. Se trata de la norma **IEEE 7000-2021** "*IEEE Standard Model Process for Addressing Ethical Concerns during System Design*" (Modelo de proceso de la norma IEEE para abordar las cuestiones éticas durante el diseño del sistema).

Esta norma establece un conjunto de procesos mediante los cuales las organizaciones pueden incluir la consideración de los valores éticos a lo largo de las etapas de exploración y desarrollo del concepto. La gestión y la ingeniería en transparente comunicación con las partes interesadas seleccionadas para la difusión y priorización de los valores éticos es apoyada por esta norma. Ello implica la trazabilidad de los valores éticos a través de un concepto operacional, las propuestas de valor y las disposiciones de valor en el diseño del sistema.

La norma describe, además, los procesos que proporcionan trazabilidad de los valores éticos en el concepto de operaciones, los requisitos éticos y el diseño basado en el riesgo ético. Todos los tamaños y tipos de organizaciones que utilizan sus propios modelos de ciclo de vida son relevantes para esta norma.

Sin duda, los Principios enunciados, los estándares desarrollados por el IEEE y el informe *Ethically Align Design* en general, son un insumo valioso para la construcción de marcos éticos verificables y regulaciones relativas a sistemas autónomos e inteligentes.

2.5 Principios para Sistemas Algorítmicos Responsables de ACM

La Asociación para la Maquinaria de Computación, ACM, por su nombre en inglés (*Association for Computing Machinery*), es la asociación más grande de profesionales de la computación y la informática, con más de 100 mil miembros en todo el mundo. Ya en enero de 2007, el comité de políticas tecnológicas de la ACM publicó su primer conjunto de 7

principios para algoritmos transparentes y con rendición de cuentas (*accountability*).[9] En octubre del 2022 el mismo comité publicó una nueva versión con 9 principios instrumentales para los sistemas algorítmicos responsables que fue anunciada en noviembre.[10] Este documento fue traducido más tarde al castellano,[11] e incluye una motivación de la necesidad de estos principios, del uso de ellos en la gobernanza de la IA y los equilibrios de coste-beneficio que implican.

Es muy importante recalcar que se usa el término *sistemas algorítmicos* y no solamente sistemas de IA para no crear un resquicio legal para sistemas que no usan IA (o dicen no usar IA). Por la misma razón, el borrador de principios de IA del gobierno estadounidense publicado el mismo mes[12] usa el término *sistemas automatizados*, aunque se refiera principalmente a la IA.

Los nueve principios son los siguientes:

1. **Legitimidad y competencia:** Se deben realizar evaluaciones legales y éticas para confirmar que cualquier riesgo introducido por los sistemas será proporcional a los problemas que se aborden, y que todas las partes interesadas relevantes entienden los compromisos entre beneficio y daños. En otras palabras, el sistema debe ser legal, ética y socialmente legítimo. Los creadores del sistema deberían tener tanto las competencias de gestión como la autorización administrativa explicita para construir y desplegar dichos sistemas. Asimismo, deberían tener amplia experiencia en el dominio de la aplicación, una base científica para el uso previsto de dichos sistemas y ser socialmente legitimados por las partes interesadas que pueden ser afectadas por el sistema. En particular, no se deben implementar proyectos sin una base científica clara (por ejemplo, inferir rasgos de personalidad a partir de imágenes faciales).

2. **Minimización del daño:** Administradores, diseñadores, desarrolladores, usuarios y otras partes interesadas en un sistema algorítmico deberían ser conscientes de los posibles errores y sesgos involucrados en su diseño, implementación y uso, así como el potencial daño que un sistema puede causar a los individuos y la sociedad. Las organizaciones deberían realizar evaluaciones de impacto periódicas sobre los sistemas que utilizan, para determinar si el sistema podría causar daños, especialmente discriminación de personas, así como para mitigar posibles problemas.

3. **Seguridad y privacidad:** El riesgo de actores maliciosos puede ser mitigado mediante la introducción de buenas prácticas en seguridad y privacidad en cada fase del ciclo de vida del sistema, incluyendo controles sólidos para mitigar nuevas vulnerabilidades que surjan en el futuro.

4. **Transparencia:** Se recomienda a los desarrolladores de sistemas a documentar claramente la forma en que se seleccionaron los conjuntos de datos, las variables y los modelos específicos utilizados para el desarrollo, entrenamiento, validación y testeo,

[9] Localizable en: https://www.acm.org/articles/bulletins/2017/january/usacm-statement-algorithmic-accountability
[10] Localizable en: https://www.acm.org/articles/bulletins/2022/november/tpc-statement-responsible-algorithmic-systems
[11] Baeza-Yates, R., Matthews, J., et al. Declaración de Principios para Sistemas Algorítmicos Responsables, ACM, 2022.
[12] Blueprint for an AI Bill of Rights, OSTP, White House, EE.UU., 2022.

así como las medidas específicas que se usaron para garantizar la calidad de los datos y los resultados. Los sistemas deberían también informar su nivel de confianza en cada resultado, y un ser humano debería intervenir cuando la confianza sea baja. Los desarrolladores deberían asimismo documentar los enfoques que se utilizaron para explorar posibles sesgos. Para sistemas con un impacto crítico en la vida y el bienestar de las personas, se deberían exigir procedimientos de verificación y validación independientes.

5. **Interpretabilidad y explicabilidad:** Se recomienda a los administradores de sistemas algorítmicos a informar tanto sobre los procesos que siguen los algoritmos empleados (interpretabilidad) como sobre las decisiones concretas que toman (explicabilidad). La explicabilidad puede ser tan importante como la exactitud, especialmente en contextos de políticas públicas o en cualquier entorno en el que haya preocupaciones sobre cómo los algoritmos podrían beneficiar a un grupo en perjuicio de otro sin ser identificado (e.g., sesgos sociales).

6. **Mantenibilidad:** Se recomienda recopilar evidencia de la solidez de todos los sistemas algorítmicos a lo largo de sus ciclos de vida, incluida la documentación de los requisitos de este, el diseño o la implementación de cambios, los casos de prueba y los resultados, y un registro de los errores encontrados y corregidos (e.g., para evitar mermas en el desempeño si los datos de entrada van cambiando). El mantenimiento adecuado puede requerir tener que volver a entrenar a los sistemas con nuevos datos de entrenamiento y/o reemplazar los modelos empleados.

7. **Auditabilidad e Impugnación:** Los reguladores deberían fomentar la adopción de mecanismos que permitan a personas y comunidades cuestionar los resultados y buscar reparación por efectos adversos resultantes de decisiones del sistema. Los administradores deberían asegurarse de que datos, modelos, algoritmos y decisiones se registran para que puedan ser auditados y los resultados replicados en casos en los que se sospecha o se alega daño. Las estrategias de auditoria deberían hacerse públicas para permitir que personas, organizaciones de interés público e investigadores, puedan revisarlas y recomendar mejoras. Es importante recalcar que auditar un sistema no legítimo lo legitima. Por lo tanto, en caso de dudas, debe comenzarse con un análisis legal y ético de este, siguiendo con la auditoría técnica sólo si es legítimo (primer principio).

8. **Rendición de cuentas y responsabilidad:** Las instituciones públicas y privadas deberían rendir cuentas por las decisiones tomadas por los algoritmos que utilizan, incluso si no es factible explicar en detalle cómo esos algoritmos produjeron los resultados. Dichas instituciones deberían ser responsables de los sistemas completos tal como se implementan en sus contextos específicos, no sólo de las partes individuales que componen un sistema determinado (tal como sería, por ejemplo, en un avión). Cuando se detecten problemas en los sistemas automatizados, las organizaciones responsables de implementar dichos sistemas deberían documentar las acciones concretas que tomen para resolver el problema y bajo qué circunstancias el uso de estas tecnologías debería ser suspendido o terminado.

9. **Limitación de los impactos medioambientales:** Los sistemas algorítmicos deben diseñarse para informar estimaciones de los impactos medioambientales, incluidas las emisiones de carbono tanto de la fase de entrenamiento como en su uso. Los sistemas

de IA deben diseñarse para garantizar que su impacto ambiental sea razonable dado el grado de exactitud requerido por el contexto en el que se implementan.

Estos principios fueron luego complementados con un documento con recomendaciones para tener sistemas más seguros[13] y son actualmente los más completos desde el punto de vista técnico. En base a ellos la ACM los ha extendido a la IA generativa,[14] agregando límites y recomendaciones para su despliegue y uso, manejo de propiedad intelectual, control de datos personales y transparencia en los errores encontrados y las correcciones realizadas.

2.6 Conclusiones

De la investigación efectuada, destaca la convergencia que existe entre las distintas organizaciones internacionales en cuanto a los principios éticos que deben guiar el diseño, implementación y uso de IA durante todo el ciclo de vida del sistema. La mayoría de los organismos coinciden en la importancia de que los sistemas de IA respeten valores humanos como la dignidad, la libertad y la inclusión social, así como los Derechos Humanos en general, la transparencia y la explicabilidad, la seguridad, la privacidad, la responsabilidad y la rendición de cuentas.

Se evidencia, además, la importancia de que los principios éticos que se opte por implementar tengan arraigo en los Derechos Humanos reconocidos en los distintos instrumentos internacionales suscritos por los países de la región. Así, tanto el origen como el contenido de los principios tendrán una justificación objetiva y contextualizada que favorezca su implementación voluntaria por parte de los actores involucrados en el uso de IA.

[13] Localizable en: https://dl.acm.org/doi/pdf/10.1145/3582277
[14] Localizable en: https://www.acm.org/articles/bulletins/2023/july/tpc-principles-generative-ai

3. Propuestas Normativas y Marcos Regulatorios Relacionadas con la IA

En la presente sección se realiza un análisis exploratorio sobre marcos legales vigentes y proyectos de leyes en discusión a nivel comparado, relacionados con IA, con el propósito de extraer mejores prácticas regulatorias.

Se analizan tanto normas cuyo ámbito de aplicación especifico es la regulación de la IA, así como aquellas cuyo ámbito regulatorio es parte de otras áreas pero que contienen provisiones aplicables a desarrollos de IA.

3.1 Reglamento General de Protección de Datos de la Unión Europea

En abril del año 2016, la Unión Europea adoptó el Reglamento General de Protección de Datos[15] ("GDPR", como se le conoce por sus siglas en inglés). La nueva regulación pretende fortalecer los derechos fundamentales de las personas en la esfera digital concediéndoles más control sobre sus datos, y construir confianza en los distintos actores para facilitar la consolidación de la economía digital y el flujo de datos necesario para dicha consolidación.

Ámbito de aplicación

En los últimos años, la norma se ha convertido en el referente mundial para el desarrollo legislativo domestico en esta materia. Países de todo el mundo, incluso de la región latinoamericana (por ejemplo, Brasil o Panamá), han efectuado reformas a sus normas de protección de datos, basados en el estándar de protección establecido por la Unión Europea. Una de las particularidades de esta directiva es su ámbito de aplicación extraterritorial, es decir, es aplicable a empresas u organizaciones radicadas fuera de la UE que traten datos personales de ciudadanos europeos.

Derechos y garantías reconocidos

Uno de los aspectos más innovadores de la regulación, es la instauración del derecho a no ser objeto de decisiones automatizadas. La rápida incorporación de nuevas tecnologías como la inteligencia artificial y en particular el área del aprendizaje automático (*machine learning*) por parte de empresas y gobiernos han revelado la importancia de proteger a las personas frente a decisiones automatizadas basadas en perfiles creados a partir del análisis de sus datos personales, lo que puede resultar en discriminaciones veladas.

Se considera que se elaboran perfiles cuando los aspectos personales de las personas son evaluados para elaborar predicciones sobre aquellas, incluso si no se toman decisiones. Por ejemplo, si una empresa u organización evalúa características como la edad, el sexo, o la altura, o incluye a la persona en una categoría, significa que se está elaborando un perfil sobre ésta.

[15] Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE.

El Reglamento concede el derecho a no ser objeto de una decisión basada únicamente en medios automatizados, si la decisión produce efectos jurídicos en la persona o le afecta significativamente de modo similar; por ejemplo, el tratamiento puede afectar significativamente a una persona si ejerce una influencia en sus circunstancias, comportamiento o preferencias. Por ejemplo, el tratamiento automatizado puede dar lugar a la denegación de una solicitud de un crédito por internet.

La regulación europea permite excepcionalmente el uso de algoritmos para decisiones automatizadas cuando la persona haya dado su consentimiento, cuando la decisión sea estrictamente necesaria para celebrar un contrato con la persona, o cuando la ley así lo establezca. Sin embargo, en esos casos, la ley establece que el responsable debe adoptar las medidas adecuadas para salvaguardar los derechos y libertades del interesado y, como mínimo, concederle el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión.

El Reglamento también contiene otras medidas con incidencia en el uso de inteligencia artificial. Por ejemplo, contempla el principio de privacidad por diseño y por defecto, bajo el cual los creadores de productos o aplicaciones de inteligencia artificial deben asegurarse, con la debida atención al estado de la técnica, de tomar en cuenta la protección de datos personales desde el diseño del sistema mismo. Estas medidas son idóneas para reducir algunos de los riesgos de los sistemas de IA en materia de privacidad.

Medidas de cumplimiento

El Reglamento establece el principio de responsabilidad proactiva, de manera que cada compañía está obligada a demostrar objetivamente que cumple con la regulación. Lo anterior lo puede hacer mediante el Registro de Actividades de Tratamiento, en el cual se registran todos los tratamientos de datos que realiza la empresa, así como por la adopción de Protocolos y procedimientos internos. Además, los responsables de los datos pueden acudir a mecanismos de certificación para demostrar su cumplimiento.

El incumplimiento de los derechos de los titulares de datos, incluidos los derechos que concede el Reglamento frente a la toma de decisiones automatizadas, está sancionado con multas administrativas de 20 000 000 EUR como máximo o, tratándose de una empresa, de una cuantía equivalente al 4 % como máximo del volumen de negocio total anual global del ejercicio financiero anterior, optándose por la de mayor cuantía.

3.2 Propuesta de Regulación del Uso de la IA de la Unión Europea

La propuesta de *Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de La Unión*[16], en adelante "Ley de Inteligencia Artificial o "AI Act") elaborada por la Comisión Europea, forma parte de un conjunto de propuestas, entre las cuales se encuentran, el proyecto de Ley de Servicios Digitales, el proyecto de Ley de Mercados Digitales y el proyecto de Ley de Gobernanza de Datos.

[16] Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de La Unión, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

La propuesta se estructura en los marcos jurídicos existentes, y sigue un enfoque basado en los riesgos, imponiendo obligaciones normativas cuando un sistema de IA es clasificado como de alto riesgo, mientras a los restantes sistemas que no son de alto riesgo se les impone una carga regulatoria menos exigente. Sobre los sistemas de IA de alto riesgo se imponen requisitos relativos a los datos de alta calidad, documentación, transparencia, vigilancia humana, precisión y la solidez.

Cabe destacar que, a la fecha de finalización de este documento (Enero 2023), la propuesta de Ley ha sido objeto de modificaciones por parte del Consejo de la Unión Europea y se encuentra actualmente en discusión en el Parlamento Europeo, donde se espera sufra modificaciones adicionales.

Entre las principales modificaciones destacan las siguientes:

a) la definición de "sistema de inteligencia artificial" la cual se reduce en alcance para evitar someter a la regulación aplicaciones y tecnologías que no son parte del objetivo político;
b) se aclaran los objetivos específicos permitidos para utilizar sistemas de reconocimiento facial para fines policiales y criminales;
c) se brinda discrecionalidad a la Comisión para modificar la lista de aplicaciones de alto riesgo, ya sea para incorporar usos de alto riesgo o eliminarlos;
d) se flexibilizan las evaluaciones de conformidad para garantizar su proporcionalidad, especialmente aquellas que deben realizar las pequeñas y medianas empresas que desarrollen o implementen IA de alto riesgo;
e) se introduce regulación especial para los llamados sistemas de uso o propósito general; entre otras.

Se espera que el Consejo y el Parlamento Europeo lleguen a un acuerdo sobre el texto definitivo durante el año 2023.

Ámbito de aplicación

El título I define el objeto del Reglamento y el ámbito de aplicación, indicando al respecto que es aplicables a:

a) Los proveedores que introduzcan en el mercado o pongan en servicio sistemas de IA en la Unión, con independencia de si dichos proveedores están establecidos en la Unión o en un tercer país.
b) Los usuarios de sistemas de IA que se encuentren en la Unión.
c) Los proveedores y usuarios de sistemas de IA que se encuentren en un tercer país, cuando la información de salida generada por el sistema se utilice en la Unión.

La propuesta distingue diferentes niveles de riesgo en relación con sistemas de IA: i) riesgos inaceptables, ii) alto riesgo, iii) riesgos limitados y iv) riesgos mínimos (Título IX). Como riesgo es una variable continua, estos niveles se ejemplifican con casos de usos. Por lo tanto, usos nuevos de la IA son más difíciles de clasificar, sin contar que esta clasificación es arbitraria. Precisamente una de las modificaciones introducidas por el Consejo de la Unión a la propuesta inicial, es precisamente en la manera de clasificar el riesgo de los sistemas de IA, pues se reconoce que el riesgo es un factor cambiante dependiente del contexto.

Esta regulación básicamente se centra en obligaciones que los agentes regulados deben cumplir para el desarrollo, uso y comercialización o puesta en comercialización bajo marca propia de sistemas de IA de alto riesgo.

Derechos y garantías reconocidos

Si bien esta es una propuesta eminentemente comercial, hace un gran énfasis a la protección de derechos fundamentales consagrada en la carta de los Derechos Fundamentales de la Unión Europea, entre estos están reconocidos y consagrados:

a) El derecho a la dignidad humana (artículo 1).
b) El respeto de la vida privada y familiar y la protección de datos de carácter personal (artículos 7 y 8).
c) La no discriminación (artículo 21).
d) La igualdad entre hombres y mujeres (artículo 23).
e) La libertad de expresión (artículo 11) y de reunión (artículo 12).
f) El derecho a la tutela judicial efectiva y a un juez imparcial, la presunción de inocencia y los derechos de la defensa (artículos 47 y 48), así como el principio general de buena administración.
g) Protección a grupos especiales, consagrados a través de los derechos de los trabajadores a unas condiciones de trabajo justas y equitativas (artículo 31), un elevado nivel de protección de los consumidores (artículo 28), los derechos del niño (artículo 24) y la integración de las personas discapacitadas (artículo 26).
h) El derecho a un nivel elevado de protección del medio ambiente y la mejora de su calidad (artículo 37) también es pertinente, en particular en lo que respecta a la salud y la seguridad de las personas.
i) Además, las obligaciones relativas a la realización de pruebas ex ante sobre los sistemas de IA, la gestión de riesgos y la vigilancia humana facilitarán el respeto de otros derechos fundamentales, ya que contribuirán a reducir al mínimo el riesgo de adoptar decisiones asistidas por IA erróneas o sesgadas en esferas críticas como la educación y la formación, el empleo, servicios importantes, la aplicación de la ley y el poder judicial.

Institucionalidad y gobernanza

En el título VI se establecen los sistemas de gobernanza nacionales. Los Estados miembros tendrán que designar a una o más autoridades nacionales competentes y, entre ellas, seleccionar a una autoridad nacional de supervisión que se encargará de supervisar la aplicación y ejecución del Reglamento. Por su parte, el Supervisor Europeo de Protección de Datos actuará como la autoridad competente para la supervisión de las instituciones, las agencias y los organismos de la Unión cuando entren en el ámbito de aplicación del presente Reglamento.

Las autoridades de vigilancia del mercado controlarían también el mercado e investigarían el cumplimiento de las obligaciones y los requisitos aplicables a todos los sistemas de IA de alto riesgo que ya se han introducido en el mercado.

La supervisión ex post debe garantizar que, una vez que un sistema de IA esté en el mercado, las autoridades tengan las competencias y los recursos necesarios para intervenir en caso de que este genere riesgos inesperados, lo que justificaría una rápida actuación.

Asimismo, también vigilarán que los operadores cumplan las obligaciones oportunas que les imponga el Reglamento.

La propuesta no contempla la creación automática de organismos o autoridades adicionales en los Estados miembros. En consecuencia, los Estados miembros podrían designar a las autoridades sectoriales existentes, a las que también confiarían las competencias para vigilar y aplicar las disposiciones del Reglamento, aprovechando así sus conocimientos especializados.

Medidas de cumplimiento

Esta propuesta establece obligaciones que se aplicarán a los proveedores y usuarios de sistemas de IA de alto riesgo. Esta propuesta tiene entre sus finalidades fomentar la confianza de los ciudadanos en el uso de la IA y reforzar los mecanismos de aplicación, imponiendo auditorías y evaluaciones de conformidad, con nuevos requisitos de documentación, trazabilidad y transparencia, calidad de datos entre otros. Además, el marco contemplará medidas específicas de apoyo a la innovación, incluidos los "*sandboxes*" reglamentarios y medidas específicas de apoyo a los pequeños usuarios y proveedores de sistemas de IA de alto riesgo para que cumplan las nuevas normas.

3.3 Ley de Servicios Digitales de la Unión Europea

El 4 de octubre de 2022 el Consejo de la Unión Europea anunció que había llegado a un acuerdo sobre el texto de la ley de servicios digitales (“DSA”, por sus siglas en inglés). Esta ley, junto a la de mercados digitales (“DMA”, por sus siglas en inglés) importan el esfuerzo regulatorio más importante en materia de regulación de la tecnología desde la aprobación del Reglamento Europeo de protección de datos personales (GDPR por sus siglas en inglés) en el año 2018.

Ámbito de Aplicación

Esta ley busca obligar a las grandes empresas tecnológicas a asumir mayores niveles de responsabilidad por el contenido por los servicios digitales que se prestan a través de ellas. Para su aplicación define qué es lo que se entenderá por servicios digitales señalando que son una gran categoría de servicios en línea que va desde “simples” sitios web hasta servicios de infraestructura de internet.

Medidas de Cumplimiento

La manera en que se ha propuesto cumplir con ese objetivo es obligar a las plataformas a eliminar contenido que se considere ilegal utilizando marcadores de confianza que los propios usuarios podrán ejecutar para llamar la atención de la plataforma; transparentar y explicar sus algoritmos a investigadores o usuarios que así lo requieran a propósito de la existencia de los algoritmos de recomendación y, a adoptar medidas contra la desinformación derivada de la manipulación electoral, entre otras.

En materia de protección de datos personales no se podrá dirigir publicidad a menores de edad o en función de categorías especiales de datos personales como el origen étnico, las opiniones políticas o la orientación sexual, ni se podrán usar patrones oscuros. Esta prohibición sin duda tendrá un impacto en la operación de las plataformas, sobre todo las

grandes que utilizan sistemas de inteligencia artificial, por cuanto será obligatorio restringir varios ámbitos que son parte del "negocio".

La ley de servicios digitales plantea un marco de sanciones que pueden ser equivalentes al 6% de la facturación total anual en todo el mundo en el caso de "plataformas en línea de muy gran tamaño". Este tipo de sanciones son inéditas (muy superiores al GDPR) y devuelven a los Estados el poder de modelar las conductas de actores que resultan demasiado poderosos en la vida política de la sociedad contemporánea y que hasta la fecha existían sin ninguna obligación de rendir cuentas por su operación y las externalidades de su "innovación".

Institucionalidad y Gobernanza

Esta norma tiene la particularidad de no derogar ni modificar leyes sectoriales, sino que como lo expresa la Comisión Europea, se trata de reglas "horizontales" que cubren todos los servicios y todo tipo de contenido ilegal. En este sentido el legislador europeo ha sido especialmente criterioso para regular sin dañar la institucionalidad vigente.

Todo lo anterior resulta de mucho interés para América latina donde existen varios países que se han comenzado a plantear la posibilidad de regular plataformas a través de proyectos de ley que se han aventurado a proponer regulaciones sobre plataformas digitales de distinto tipo y cuyos efectos resultan dañinos para regulaciones vigentes y que hasta el momento funcionan bien[17].

3.4 Proyecto de Ley de Responsabilidad Algorítmica de EE. UU.

El proyecto de ley de responsabilidad o rendición de cuentas algorítmica ("*Algorithmic Accountability Act*"[18]) es una norma presentada en el Senado estadounidense en abril de 2019, en el comité de comercio, ciencia y transporte.

Ámbito de aplicación

Sus reglas buscan exigir a las entidades que usan, almacenan o comparten información personal que lleven a cabo evaluaciones de impacto y de protección de datos personales. Entrega a su vez una serie de conceptos, como, por ejemplo, el de "decisiones automatizadas", "evaluación de impacto" y los elementos que estas evaluaciones deben contener. Sus reglas, en principio, se encuentran diseñadas para ser aplicadas a todas las organizaciones cubiertas por la jurisdicción de la Comisión Federal de Comercio (FTC).

Derechos y garantías reconocidos

Llevar a cabo evaluaciones de impacto y de protección de datos personales y hacerse públicas en la medida que la organización sujeta a la jurisdicción de la FTC lo estime pertinente.

[17] El paquete de la Ley de servicios digitales. https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package

[18] Localizable en: https://www.congress.gov/bill/117th-congress/house-bill/6580/text

Institucionalidad y gobernanza

La norma ordena que dentro de los 2 años siguientes a la dictación de la norma se dicte un reglamento para regular a las organizaciones sujetas a su jurisdicción.

Medidas de cumplimiento

Los incumplimientos a la normativa propuesta se pueden perseguir a través de una actuación directa de la FTC o a través de una acción civil promovida por el fiscal general cuando estime que los intereses de los residentes en el Estado han sido amenazados o vulnerados.

3.5 Proyecto de Ley de Privacidad de la Información Biométrica de EE. UU.

El proyecto de ley de privacidad de la información biométrica (*Biometric Information Privacy Act*[19]) es una norma presentada en el Congreso estadounidense en agosto de 2020, en el Comité del Poder Judicial. Dispone que dentro de los 60 días siguientes a la dictación de la norma se establezca por parte de los organismos retenedores de datos biométricos una política de retención y destrucción de identificadores.

Ámbito de aplicación

La ley se aplica a establecimientos privados en posesión de identificadores biométricos.

Derechos y garantías reconocidos

La organización privada que recopile datos biométricos sólo podrá hacerlo si el titular del dato consintió expresamente la recopilación, cuando sea necesario para proporcionar un servicio a la persona o cliente y para otro propósito comercial válido.

Institucionalidad y gobernanza

La norma ordena que dentro de los 2 años siguientes a la dictación de la norma se dicte un reglamento para regular a las organizaciones sujetas a su jurisdicción.

Medidas de cumplimiento

Los incumplimientos a la normativa propuesta se pueden perseguir directamente a través de una acción civil o a través de una acción civil promovida por el fiscal general cuando estime que los intereses de los residentes en el Estado han sido amenazados o vulnerados.

3.6 Directiva sobre Decisiones Automatizadas de Canadá

La directiva sobre sistemas de toma de decisiones automatizadas (*Directive on Automated Decision-Making*[20]) canadiense corresponde a una herramienta obligatoria de evaluación de riesgos destinada a apoyar a la Directiva del Consejo del Tesoro de Canadá sobre la toma de decisiones automatizada. Compuesto por 48 riesgos y 33 preguntas de mitigación,

[19] Localizable en: https://www.congress.gov/bill/116th-congress/senate-bill/4400/text
[20] Localizable en: https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592

es básicamente un cuestionario que determina el nivel de impacto de un sistema de toma de decisiones automatizado. La evaluación está organizada de acuerdo con las consideraciones políticas, éticas y de derecho administrativo del gobierno sobre las áreas de riesgo de los sistemas de decisión automatizados. Esta evaluación de impacto debe ser complementada al principio de la fase de diseño de un proyecto, y los resultados deben publicarse en un formato accesible en las dos lenguas oficiales de Canadá.

Ámbito de aplicación

Esta Directiva se aplica a sistemas que brindan servicios públicos externos tal como se define en la Política de Servicios y Digital, se aplica a cualquier sistema, herramienta o modelo estadístico utilizado para recomendar o tomar una decisión administrativa dentro de la administración del estado y a cualquier sistema de decisión automatizado desarrollado o adquirido después 1 de abril de 2020.

Derechos y garantías reconocidos

El objetivo de esta Directiva es garantizar que los sistemas de decisión automatizados se implementen por el Estado de una manera que reduzca los riesgos para los canadienses y las instituciones federales, y conduzca a decisiones más eficientes, precisas, coherentes e interpretables tomadas de conformidad con la ley canadiense.

Institucionalidad y gobernanza

La Secretaría de la Junta del Tesoro de Canadá es responsable de proporcionar orientación a todo el gobierno sobre el uso de sistemas de decisión automatizados. También es la encargada de desarrollar y mantener las Evaluaciones de Impacto Algorítmico y cualquier documentación de respaldo.

Medidas de cumplimiento

Las consecuencias del incumplimiento de esta política pueden incluir cualquier medida permitida por la Ley de Administración Financiera que la Junta del Tesoro determine como apropiada y aceptable en las circunstancias.

3.7 Ley General de Protección de Datos Personales de Brasil

En agosto 2018, el Congreso Nacional de Brasil aprobó la Ley nº 13.709 de 14 de agosto de 2018, Ley General de Protección de Datos Personales[21] (modificada por la Ley nº 13.853 de 8 de julio de 2019) ("LGPD") la cual entró en vigor el 18 de septiembre de 2020. La LGPD es una ley integral de protección de datos que cubre las actividades de los controladores y procesadores de datos y crea requisitos modernos sobre el tratamiento de la información personal de las personas.

La regulación es una de las más recientes en la región latinoamericana y contiene algunas reglas novedosas -basadas en la GDPR- relacionadas con el uso de datos personales en soluciones de inteligencia artificial. Por lo tanto, conviene revisar estas reglas para extraer mejores prácticas potencialmente útiles para el diseño de un marco ético y/o regulatorio para la IA en el continente.

[21] Localizable en: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm

Ámbito de aplicación

La Ley se aplica a cualquier operación de tratamiento realizada por una persona física o jurídica, pública o privada, siempre que:

a) la operación de tratamiento se lleve a cabo en territorio brasileño;
b) la actividad de tratamiento tenga por objeto el ofrecimiento o la prestación de bienes o servicios, o el tratamiento de datos de personas físicas situadas en el territorio brasileño; o
c) los datos personales objeto de tratamiento sean recogidos en el territorio brasileño.

La norma considera como datos recogidos en el territorio nacional aquellos cuyo titular se encuentra en el territorio brasileño en el momento de la recogida.

Derechos y garantías reconocidos

Según la norma brasileña, toda persona sujeta a una decisión automatizada basada en el tratamiento de sus datos personales, tiene derecho a solicitar la revisión de esas decisiones, en el tanto afecten sus intereses. En esta categoría de decisiones se incluyen aquellas destinadas a definir un perfil personal ("*profiling*"), profesional, de consumo y de crédito, o aspectos de la personalidad de la persona.

Asimismo, la persona tiene derecho a que el responsable del tratamiento de los datos le proporcione información **clara** y **adecuada** sobre los criterios y procedimientos utilizados para la decisión automatizada, respetándose, eso sí, el secreto comercial e industrial. Es decir, la norma reconoce el derecho a una explicación frente a la toma de decisiones automatizadas, similar a la protección que otorga la GDPR.

Una garantía adicional que provee la regulación brasileña es que, si el responsable se opone a brindar una explicación fundamentado en que, con ello, comprometería sus secretos comerciales o industriales, entonces la Autoridad competente podrá realizar una auditoría a la empresa para verificar que la explicación verdadera detrás de una decisión automatizada no se base en características personales que den pie a un trato discriminatorio. No obstante, no queda claro cómo sería el funcionamiento en la práctica de estas auditorías y si basta con invocar el secreto comercial para evadir la obligación de brindar una explicación.

Estas protecciones son especialmente relevantes frente a decisiones automatizadas que puedan tener un impacto significativo en los derechos e intereses de las personas. Por ejemplo, decisiones de reclutamiento en las que se deniega un puesto de trabajo, el rechazo de un crédito bancario o la denegación de acceso a educación o a un bono social.

Las garantías mencionadas son, además, consistentes con los principios de la OCDE sobre IA, especialmente los principios 1.2 "valores centrados en el ser humano y justicia", y 1.3 "Transparencia y Explicabilidad", y con los instrumentos internacionales de Derechos Humanos, en los que se garantiza el debido proceso y el derecho a la igualdad y no discriminación.

Por último, es de resaltar el hecho de que la norma regule en capítulo separado (artículo 23 y siguientes) el uso de datos personales por parte del sector público. La Ley define con

suficiente especificidad las condiciones bajo las cuales podrán las autoridades tratar e intercambiar datos personales.

Medidas de cumplimiento

Similar a lo que ocurre con la GDPR, la legislación brasileña favorece la responsabilidad proactiva de las empresas, de manera que cada compañía puede optar por protocolos, procedimientos internos o normas técnicas para garantizar el cumplimiento de la norma.
La regulación prevé la posibilidad de emitir advertencias previo a la imposición de multas, con indicación de un plazo para la adopción de medidas correctoras. También se contemplan multas simples de hasta el 2% de los ingresos por ventas de la persona jurídica, grupo o conglomerado en Brasil en su último ejercicio fiscal, multas diarias, teniendo en cuenta el límite total anterior; divulgación de la infracción después de haber sido debidamente investigada y confirmada su ocurrencia; bloqueo de los datos personales a los que se refiere la infracción; eliminación de los datos personales a los que se refiere la infracción; suspensión parcial del funcionamiento de las bases de datos objeto de la acción infractora durante un máximo de seis meses; y prohibición parcial o total de ejecutar actividades relacionadas con el tratamiento de datos.

3.8 Marco Regulatorio de Protección de Datos Personales, al Consumidor y Derecho a la Información Pública de Argentina

La Constitución argentina reconoce en el artículo 43[22] el derecho que tiene toda persona a interponer un recurso de amparo "para tomar conocimiento de los datos a ella referidos y de su finalidad, que consten en registros o bancos de datos públicos, o los privados destinados a proveer informes, y en caso de falsedad o discriminación, para exigir la supresión, rectificación, confidencialidad o actualización de aquéllos. No podrá afectarse el secreto de las fuentes de información periodística".
Cuenta también desde el año 2000 con la ley Nº 25.326[23] sobre protección de datos personales el que ofrece acceso libre a los datos generados o procesados por sistemas de inteligencia artificial a través de la posibilidad de ejercer el derecho de información consagrado en el artículo 13º de esta misma ley.

Ámbito de aplicación

Sus reglas permiten a los titulares de la información acceder a todas las bases de datos, públicas y privadas, que guarden información sobre su persona.

Derechos y garantías reconocidos

La ley reconoce el derecho a la información, el derecho de acceso, a la rectificación, actualización o supresión, etc.

Institucionalidad y gobernanza

[22] Localizable en: http://servicios.infoleg.gob.ar/infolegInternet/anexos/0-4999/804/norma.htm
[23] Localizable en: https://www.argentina.gob.ar/normativa/nacional/ley-25326-64790/texto

Cuenta con un órgano de control que goza de autonomía funcional y que se encuentra descentralizado en el Ministerio de Justicia y Derechos Humanos.

Medidas de cumplimiento

La ley contempla la posibilidad de establecer sanciones administrativas impuestas por la autoridad de control, pero en casos especiales son igualmente procedentes sanciones civiles derivadas de los daños y perjuicios derivados de la aplicación de esta ley.

En Argentina, además, el sistema de protección de datos personales se ve reforzado por la existencia de otros dos cuerpos normativos: el de protección al consumidor y el de acceso a la información pública.
Por un lado, a través de la ley protección al consumidor es posible perseguir responsabilidad a través de lo prescrito por los artículos 2º y 40º de la Ley de Defensa del Consumidor, por cuanto reconocen responsabilidad al productor, fabricante, proveedor, vendedor, entre otros, por los daños resultantes de la prestación del servicio, como así también del vicio o riesgo de la cosa. De manera que ella resultaría perfectamente aplicable si el vicio o riesgo de la prestación es resultado de procesos derivados del uso de inteligencia artificial.

Por otro, la ley Nº 27.275, del año 2016, reconoce el derecho de acceso a la información pública. Esta norma, como la mayoría de las reglas de esta naturaleza, contempla en el art. 8º un listado de situaciones que hacen excepción al régimen general de acceso, el que, para efectos de este estudio, es el relativo a "Secretos industriales, comerciales, financieros, científicos, técnicos o tecnológicos cuya revelación pudiera perjudicar el nivel de competitividad o lesionar los intereses del sujeto obligado."

3.9 Protección de Datos Personales y Acceso a Información Pública de Uruguay

Uruguay cuenta desde el año 2008 con la ley Nº 18.331 de protección de datos personales[24]. La ley entrega una serie de definiciones y principios que permiten su integración en caso de vacíos.

En el caso de aplicaciones relacionadas con inteligencia artificial, no hay normativa específica que se refiera a ella, sin perjuicio de lo cual han sido ingentes los montos fiscales que se han destinado a la obtención de este tipo de tecnología en el sector público. Un ejemplo de ello es la compra de software de predicción del delito y vigilancia PREDPOL que "elabora mapas predictivos que predice mediante algoritmos, dónde ocurrirán las incidencias delictivas de las próximas 24 horas. Estas predicciones son a partir de un intercambio de información e interoperabilidad directa con el Sistema de Gestión de Seguridad Pública (SGSP)"[25].

Es una práctica frecuente en América latina usar la ley de presupuestos para autorizar ciertas prácticas públicas, camufladas en una glosa presupuestaria sin el debido ejercicio deliberativo sobre los impactos de su implementación[26].

[24] Localizable en: https://www.impo.com.uy/bases/leyes/18331-2008
[25] Scrollini, Fabrizio (Febrero 26, 2020), Zenodo, localizable en: https://zenodo.org/record/4564556#.YUdecS2b5sM
[26] Clark, X. y Zaror, D. 2020. Marcos Legales Estadísticos de América Latina. http://dx.doi.org/10.18235/0002938, pág. 116.

La ley Nº 18.381 de agosto de 2008 sobre acceso a la información pública, por su parte, es aplicable a todos los servicios que forman parte de la organización del Estado; entre sus normas más relevantes se encuentra el artículo 1º que establece que "se considera información pública toda la que emane o esté en posesión de cualquier organismo público sea o no estatal, salvo las excepciones establecidas por la ley, así como las informaciones reservadas o confidenciales." En efecto, el artículo 8º señala que "Las excepciones a la información pública serán de interpretación estricta y comprenderán aquellas definidas como secretas por la ley y las que se definan seguidamente como de carácter reservado y confidencial."

Ámbito de aplicación

Ella se aplica a las personas naturales y también a las jurídicas en cuanto corresponda y respecto de cualquiera que administre una base de datos susceptible de tratamientos, ya sea público o privado.

Derechos y garantías reconocidos

La ley reconoce el derecho a la información, acceso, rectificación, actualización, inclusión o supresión y a la impugnación de valoraciones personales.

Institucionalidad y gobernanza

La ejecución de esta ley se encuentra a cargo de la Agencia para el Desarrollo del Gobierno de Gestión Electrónica y de la Sociedad de la Información y del Conocimiento (AGESIC).

Medidas de cumplimiento

Para perseguir el cumplimiento de las normas establecidas por esta ley, la ley contempla una serie de sanciones administrativas de menor a mayor intensidad.

3.10 Ley de Protección de Datos Personales de Panamá

Panamá aprobó la Ley 81 sobre Protección de Datos Personales[27] el 26 de marzo de 2019, y entró en vigor el 27 de marzo del 2021, luego de 2 años de *vacatio legis*. En términos generales, la normativa se basa en el Reglamento General de Protección de Datos europeo, pero con importantes matices.

Ámbito de aplicación

La normativa panameña se aplica a:

- las bases de datos ubicadas en el territorio de la República de Panamá;
- las bases de datos que almacenen o contengan datos personales de nacionales o extranjeros
- cualquier persona encargada del tratamiento de datos que esté domiciliada en Panamá; y,

[27] Localizable en: https://www.asamblea.gob.pa/APPS/LEGISPAN/PDF_NORMAS/2010/2019/2019_645_3008.pdf

- cualquier actividad comercial en línea de empresas extranjeras que se dirija al mercado panameño.

Derechos y garantías reconocidos

El artículo 19 de la Ley 81 regula lo relativo a las decisiones basadas en tratamientos automatizados de los datos personales. Similar a lo que establece la normativa europea y la brasileña, se les concede a las personas el derecho a no ser sujeto de decisiones de esta naturaleza, salvo que lo hayan consentido expresamente, que la decisión sea necesaria para celebrar o dar cumplimiento a un contrato, ó si lo autorizan leyes especiales.
A diferencia de las salvaguardas adicionales previstas por la GDPR y la LGPD brasileña, en caso de toma de decisiones automatizadas o elaboración de perfiles, la normativa panameña no impone al responsable del tratamiento la obligación de aplicar medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, como lo sería el reconocimiento del derecho a obtener intervención humana, el derecho a expresar su punto de vista y el derecho de impugnar la decisión.
Pareciera que esta omisión de la Ley 81 quiso ser enmendada con el Reglamento a la Ley adoptado en mayo de 2021, cuyo artículo 24, inciso 7), estipula la obligación del responsable de proveer al destinatario de la decisión, "información significativa sobre la lógica aplicada, así como la importancia y consecuencias de la decisión". Sin embargo, se sigue omitiendo por completo el derecho a obtener intervención humana frente a decisiones basadas en inteligencia artificial, cuando estas hayan sido consentidas por el usuario, sean necesarias para la ejecución de un contrato con el usuario, o lo permita expresamente una ley especial.

Mecanismos de cumplimiento

Panamá optó por el principio de responsabilidad proactiva. Por lo tanto, las empresas deben mantener una ficha técnica con sus protocolos y procedimientos para demostrar el cumplimiento de la legislación.

Las sanciones por incumplimiento a la Ley van de 1.000 a 10.000 dólares de multa, según la gravedad de la infracción.

3.11 Tratado Internacional: Acuerdo de Asociación de Economía Digital

En el año 2019, las naciones de Chile, Nueva Zelandia y Singapur iniciaron negociaciones para pactar un acuerdo internacional en materia de economía digital. Estas negociaciones culminaron con la firma del Acuerdo de Asociación de Economía Digital[28] ("DEPA", por sus siglas en inglés) en junio de 2020, el primer acuerdo de su tipo en el mundo.

Su objetivo es promover el comercio digital en países de economías pequeñas y medianas para promocionar a estos países como plataformas de la economía digital. Esto incluye no solo contar con un marco amigable para las empresas, donde puedan exportar sus servicios y productos digitales, sino también proporcionar mayor transparencia y certeza, así como incrementar la confianza en el entorno digital y el desarrollo económico inclusivo.

[28] Localizable en: http://www.sice.oas.org/trade/DEPA/DEPA_Text_e.pdf

Si bien el DEPA es un acuerdo pensado para beneficio de economías más limitadas, potencias mundiales como Canadá y Corea del Sur, están valorando adherirse al Acuerdo o han iniciado el proceso formal para adherirse, lo cual refuerza la importancia que está adoptando esta iniciativa a nivel internacional. A finales de diciembre del 2022, otro país latinoamericano -Costa Rica- anunció que ha iniciado el proceso formal para adherirse al Acuerdo.

Ámbito de aplicación

Por ahora es un tratado internacional suscrito únicamente por Chile, Nueva Zelandia y Singapur y aplica a las medidas adoptadas o mantenidas por dichas naciones que afecten al comercio en la economía digital.

Contenido relacionado con IA

En lo que concierne a inteligencia artificial propiamente, el Acuerdo contiene un artículo específico -el 8.2- que reconoce la importancia económica y social de desarrollar marcos éticos y de gobernanza para el uso confiable, seguro y responsable de las tecnologías de IA. Por ello, dispone que las Partes del Acuerdo se esforzarán por promover la adopción de marcos éticos y de gobernanza que apoyen el uso confiable, seguro y responsable de las tecnologías de IA (Marcos de Gobernanza de IA), que, a su vez, incorporen los principios o directrices reconocidos internacionalmente, incluida la explicabilidad, la transparencia, la equidad y los valores centrados en el ser humano.

Institucionalidad y gobernanza

El Acuerdo también contiene propuestas y recomendaciones en materia de gobernanza de datos y protección de datos personales. Estas provisiones son relevantes para el presente análisis debido a que la disponibilidad de datos es esencial para el diseño, entrenamiento y despliegue de sistemas confiables y responsables de IA, así como para apalancar innovaciones y emprendimientos tecnológicos. Algunas de las recomendaciones destacables derivadas del DEPA son la de garantizar los flujos de datos transfronterizos y el intercambio de datos como facilitadores de la innovación basada en datos y poner en marcha pruebas de concepto y entornos de prueba controlados regulatorios (*regulatory data sandboxes*) que permitan anticipar los beneficios y riesgos de compartir determinados datos para determinadas aplicaciones. También, destaca la obligación que asumen las Partes de facilitar el acceso público y el uso de la información gubernamental para fomentar el desarrollo económico y social, la competitividad y la innovación, y promover la cooperación internacional para fomentar modelos de licencia de datos abiertos en forma de licencias públicas estandarizadas disponibles en línea.

En resumen, este novedoso acuerdo internacional de economía digital contiene provisiones que -unidas a otras medidas- podrían potencialmente contribuir a los objetivos del continente en materia de inteligencia artificial responsable.

3.12 Hallazgos

Sobre este análisis exploratorio, destaca el hecho de que las normas especiales sobre inteligencia artificial son la excepción. Por el contrario, la mayoría de los países han regulado aspectos específicos relacionados con la IA de forma directa o indirecta a través

de sus leyes sectoriales, como, por ejemplo, en las normas de protección de datos, protección al consumidor, o acceso a la información pública.

Algunos de los derechos relacionados con la IA que la mayoría de los países estudiados han adoptado en su legislación sobre protección de datos son, entre otros, el derecho a la explicación, a la supervisión humana y a la revisión de decisiones que impliquen el tratamiento automatizado de datos personales o la generación de perfiles a partir de la evaluación de rasgos de la personalidad del titular de la información, siguiendo el desarrollo legislativo europeo. En ese sentido, es indudable el efecto que ha tenido el Reglamento General de Protección de Datos de la Unión Europea en las recientes legislaciones latinoamericanas sobre esta materia, como es el caso de Brasil, Uruguay y Panamá.

Del estudio sobresale que un número importante de países han optado por mecanismos de derecho blando no vinculantes (*soft law*) y autorregulación para una gestión basada en riesgos en materia de IA. Es decir, se ha recurrido a lineamientos o marcos éticos, guías de conducta, experimentación regulatoria y prototipos de regulación, en lugar de legislación vinculante. Esto, posiblemente motivado en: 1) la poca preparación, capacitación y el desconocimiento que se tiene del fenómeno tecnológico, 2) el acelerado dinamismo que suele caracterizar a la IA, y 3) el temor a desincentivar el uso de la IA.

También se identificaron propuestas normativas dirigidas a regular de manera integral y especial la inteligencia artificial, siendo la más importante la Propuesta de Ley de Inteligencia Artificial de la Comisión Europea (ver Sección 3.2). Esta propuesta tiene entre sus finalidades fomentar la confianza de los ciudadanos en el uso de la IA mediante un enfoque basado en los riesgos, imponiendo cargas normativas cuando un sistema de IA es clasificado como de alto riesgo, mientras los restantes sistemas de menor riesgo se les impone una normativa más laxa. Sobre los sistemas de IA de alto riesgo se establecen requisitos relativos a la calidad de los datos, obligaciones de documentación, pruebas de conformidad, auditorías, transparencia, vigilancia humana, precisión y la solidez. Asimismo, dicho marco contempla medidas específicas de apoyo a la innovación, incluidos los entornos de prueba controlados (*sandboxes* regulatorios) y medidas específicas de apoyo a los pequeños usuarios y proveedores de sistemas de IA de alto riesgo para que cumplan las nuevas normas.

4. Propuesta de Marco Ético y Normativo para una IA Responsable

La IA debe ser implementada en contextos apropiados donde existan beneficios identificables por los ciudadanos, que estén alineados con la cultura y contexto local de cada país latinoamericano. Lo anterior se traduce en que tanto el marco ético como normativo que se implemente debe estar contextualizado para el cumplimiento de dicho fin.

Para ello es altamente recomendable incorporar un enfoque participativo e inclusivo, en el sentido que estos principios y normas reflejen reglas y valores propios de los diferentes sectores, pero especialmente de las localidades y comunidades donde se aplicará la IA en cada país, dando relevancia a su contexto histórico y social. Lo anterior, generará además el efecto positivo que tanto el origen como el contenido de marcos éticos y normativos tendrán una justificación objetiva y cercana que favorecería su adopción.

4.1 Recomendaciones de Principios Éticos para el Uso Responsable de la IA

A continuación, se propone un marco ético de principios contextualizados a la realidad social e histórica latinoamericana que podría ser tomado como base o guía para el desarrollo de políticas públicas, Estrategias o Políticas Nacionales de Inteligencia Artificial.

Para la selección de estos principios se utilizaron como fuentes, esencialmente, las recomendaciones de la OCDE y la UNESCO sobre ética de la inteligencia artificial, desarrolladas en el Capítulo I y la experiencia de los autores.

Además de la importancia del contexto ya mencionado en el título anterior, estos principios deberían estar directamente relacionados con recomendaciones para llevarlos a la práctica de una forma factible, tanto desde el punto de vista técnico como legal. Precisamente, la falta de operativización de los principios éticos, y su aplicación inconsistente entre países y regiones es uno de los principales retos que enfrentan los gobiernos en esta materia.

Los principios éticos que consideramos apropiado para un uso ético y responsable de la IA en Latinoamérica son los siguientes:

a) Respeto a los Derechos Humanos

Los derechos humanos, y en especial la dignidad humana y la no discriminación, han de ser respetados, protegidos y promovidos a lo largo del ciclo de vida de los sistemas de IA. Todas las organizaciones involucradas, tanto públicas como privadas, deben respetar los instrumentos y marcos de Derechos Humanos en los procesos que rodean el ciclo de vida de los sistemas de IA en Latinoamérica.

Justificación

El desarrollo de una IA responsable y ética debe estar guiado por los impactos que pueden producir, principalmente si estos impactos conllevan potenciales vulneraciones a los Derechos Humanos.

Los Derechos Humanos descansan sobre la dignidad humana. La extensión de la dignidad humana se ramifica a la protección de Derechos Humanos que se han visto vulnerados por el uso de ciertos sistemas de IA. Al respecto, está altamente documentado que en algunos casos cuando se utilizan sistemas de IA para resolver problemas sociales profundos, éstos perpetúan y/o amplifican los prejuicios existentes en sociedad, situación que puede ser aún más profunda en el contexto latinoamericano.

Ejemplos de este tipo de tecnologías están presentes en materias como predicción de delitos, puntuación de riesgo de reincidencia penal o de aplicación de determinadas penas. Dentro de estos, se encuentran los sistemas de reconocimiento facial, creados en los años setenta con el objeto de identificar a sospechosos delictuales contrastando fotografías contenidas en bases de datos policiales. Hasta hoy, los sistemas policiales de reconocimiento facial se construyen con bases de datos históricas, sin tomar en consideración que muchos datos son incompletos, sesgados, reflejo de detenciones ilegales o de racismo policial, lo cual explica, además, que la prevención de delitos mediante esta tecnología posea un alto margen de error.

Por su parte, muchos sistemas y usos de la IA, en particular aquellos de vigilancia masiva, tienen un impacto amplio y profundo no solo en el derecho de la privacidad, sino que se extiende al derecho a la integridad personal, la honra y dignidad humana, incluso el derecho a la autonomía, libertad individual y libertad de expresión, lo que se traduce en el hecho que las personas se vean obligadas a actuar de cierta manera por temor a ser calificadas negativamente por sistemas de vigilancia.

Asimismo, existe abundante evidencia de que existe un patrón detectado en el uso de sistemas de IA cuando se emplean en materias de relevancia social. En dichos casos, grupos históricamente excluidos y oprimidos, tales como, pueblos originarios, comunidades LGBTIQ+, minorías religiosas y personas de escasos recursos, entre otras, son comúnmente afectados de forma desproporcionada, reproduciendo y perpetuando injusticias estructurales. Sistemas predictivos de abuso infantil que asocian injustamente pobreza con abusos, sistemas predictivos de delitos que asocian delincuencia a racismo policial histórico, sistemas predictivos que detectan erróneamente fraudes de beneficios sociales, son algunos ejemplos.

b) Transparencia y Rendición de Cuentas

La transparencia y rendición de cuentas de los sistemas de IA son una condición previa para garantizar la protección de Derechos Humanos y valores democráticos.

Justificación

Durante los últimos años se ha discutido sobre el nivel de transparencia que debe existir en el desarrollo y uso de sistemas de IA, en particular, en aquellos que se relacionan con problemáticas sociales complejas y que potencialmente pueden tener un impacto negativo sobre las personas.

En ese sentido, se entiende que la transparencia y exigencia de información estratégica y relevante, es un requisito para la construcción de la confianza entre los ciudadanos y entidades públicas o privadas, con la finalidad que las personas puedan contar con antecedentes necesarios para tomar la decisión de aceptar con cierta confianza el uso de un modelo algorítmico de inteligencia artificial. Pero esto es cierto sólo respecto de una parte de la población, ya que respecto de personas vulnerables o de escasos recursos, el uso de sistemas tecnológicos en temáticas que les impactan no les es consultado y menos explicado. Hasta cierto punto, exigir y obtener transparencia se convierte en una suerte de “privilegio” y un elemento más que puede agravar la desigualdad existente en un país determinado, por lo que se debe procurar un acceso libre y democrático a la información.

De la misma manera, se debe tener presente que la IA entendida como un sistema socio-técnico no solo es conceptualizada desde la técnica, ya que junto con ella se despliegan motivaciones, y decisiones económicas y políticas sobre las personas. Todos estos

elementos técnicos y sociales en su conjunto tienen una relación directa en los impactos negativos que puede generar la IA. Por lo tanto, la transparencia debe ser entendida no solo desde la técnica sino incluyendo también el contexto social.

Por su parte, los desarrolladores deben ser responsables del correcto funcionamiento de los sistemas de IA y del respeto de los restantes principios, considerando sus funciones, el contexto y de conformidad con el estado de la técnica, y tienen por lo tanto la obligación de rendir cuentas.

La rendición de cuentas en su sentido amplio se relaciona con la palabra "*accountability"*, un concepto anglosajón, que ha evolucionado desde la contabilidad hacia conceptos que involucran transparencia, equidad, eficiencia, responsabilidad, entre otros, correspondiendo a un concepto evaluativo en términos positivos o de integridad de un ente.

Como concepto evaluativo no existe un estándar único y determinado y su comprensión depende del contexto histórico y social donde se utilice. En materia de sistemas de IA se utiliza la expresión "*accountability*" en su sentido estricto, como sinónimo de rendición de cuenta que se define como "*una relación entre un actor y un foro, en la que el actor tiene la obligación de explicar y justificar su conducta y el foro puede plantear preguntas y emitir juicios, pudiendo el actor estar expuesto a consecuencias*".

En este caso, la rendición de cuentas es la relación entre el actor, que es quien utiliza, comercializa o diseña un sistema de IA y el foro que puede hacer cumplir normas de conducta, estando el actor sujeto a consecuencias. Esta relación en general es construida a través de legislación vinculante y en menor medida de políticas y recomendaciones.

También esta relación se genera comúnmente desde la sociedad civil principalmente a través de auditorías algorítmicas de carácter externo. En consecuencia, los compromisos voluntarios de rendición de cuentas son menos efectivos y por ende no recomendables, pues no constituyen parte del concepto propiamente tal, ya que la existencia del actor, el foro externo y la obligatoriedad de consecuencias, son esenciales para su efectividad.

c) Privacidad y Protección de Datos Personales

Los sistemas, aplicaciones y soluciones basadas en IA que se implementen en Latinoamérica deben respetar el derecho fundamental de las personas a la vida privada y proteger de manera efectiva los datos personales que puedan estar involucrados en el diseño, entrenamiento, despliegue e implementación de sistemas basados en IA, incluyendo todas las formas de aprendizaje automático.

Se recomienda implementar una legislación moderna sobre protección de datos personales o una reforma a las leyes existentes en la materia, en aquellos países que lo requieran, de manera que se proteja de forma efectiva a las personas frente al tratamiento de sus datos, especialmente ante el tratamiento automatizado basado en IA, a la vez que garantice el flujo libre pero responsable de datos para la innovación tecnológica.

Justificación

La privacidad, en sentido estricto, es el derecho fundamental de toda persona a mantener determinados ámbitos de su vida como privados, lejos de la injerencia de terceros y, sobre todo, del propio Estado. El artículo 11 de la Convención Americana de Derechos Humanos del 22 de noviembre de 1969 o "Pacto de San José", suscrita por la mayoría de países de Latinoamérica, protege este derecho a la vida íntima.

Si bien la protección de datos personales como tal no está expresamente reconocida en el Pacto de San José, la Corte Interamericana de Derechos Humanos (Corte IDH), sí se ha referido a la necesidad de adaptar los alcances del derecho a la vida privada, especialmente ante la creciente influencia de las tecnologías de la información en la vida de las personas. De ahí que el Estado deba asumir un compromiso aún mayor, con el fin de adecuar a los tiempos actuales las formas tradicionales de protección del derecho a la vida privada.

La privacidad y la protección de datos personales son derechos que deben condicionar el uso de la inteligencia artificial en Latinoamérica, no solo por el impacto significativo de la IA sobre dichos derechos, sino porque el uso adecuado y responsable de los datos ayuda a construir la confianza necesaria para el desarrollo de la IA. Ante el estado incipiente de esta tecnología en Latinoamérica y la poca apropiación digital que puede existir en la región, construir esa confianza resulta especialmente relevante.

d) Seguridad

Los efectos no deseados (riesgos de seguridad) y las vulnerabilidades a los ataques (riesgos de protección tanto externos como internos) deberían ser evitados y deberían tenerse en cuenta, prevenirse y eliminarse a lo largo del ciclo de vida de los sistemas de IA para garantizar la seguridad y la protección de los seres humanos, del medio ambiente y de los ecosistemas. La seguridad y la protección de la IA se propiciarán mediante el desarrollo de marcos de acceso a los datos que sean sostenibles, respeten la privacidad y fomenten un mejor entrenamiento y validación de los modelos de IA que utilicen datos de calidad.

Justificación

Este principio se justifica pues pone el foco en la necesidad de crear estructuras que eviten los daños o vulnerabilidades que generan los sistemas de IA, o al menos permitan mitigar sus efectos, cualquiera sea la etapa del ciclo de vida de la IA en relación con las personas, el medio ambiente y los ecosistemas.

Este principio se encuentra fuertemente ligado a la integridad u observancia de ciertas prácticas en materia de inteligencia artificial. En efecto, la OCDE en su Recomendación, lo entiende como una característica que todo sistema de inteligencia artificial debe proveer “durante todo su ciclo de vida, para durante su uso normal, uso previsible o mal uso, o cualquier otra condición adversa, ellos funcionen apropiadamente y no supongan un irracional riesgo de seguridad”.

Por lo anterior, es que su incorporación supone también altos niveles de integración y articulación entre todos los integrantes del ecosistema de la inteligencia artificial de Latinoamérica, pues la operación y control, en muchos casos dependerán de varios entes los que al menos deben compartir ciertos análisis, tareas y acciones para el logro de esos objetivos.

e) Sostenibilidad

El desarrollo de sociedades sostenibles depende del logro de un complejo conjunto de objetivos relacionados con distintas dimensiones humanas, sociales, culturales, económicas y ambientales, dimensiones que deben necesariamente estar integradas, pues las intervenciones en un área impactarán necesariamente en el área de las otras. Es decir, todos los aspectos posibles de sostenibilidad.

La llegada de las tecnologías de la IA puede beneficiar los objetivos de sostenibilidad o dificultar su consecución, dependiendo de la forma en que se apliquen en países con diferentes niveles de desarrollo. Por consiguiente, la evaluación continua de los efectos humanos, sociales, culturales, económicos y ambientales de las tecnologías de la IA debería llevarse a cabo con pleno conocimiento de las repercusiones de dichas tecnologías en la sostenibilidad como un conjunto de metas en constante evolución en toda una serie de dimensiones, como las que se definen actualmente en los Objetivos de Desarrollo Sostenible (ODS) de las Naciones Unidas. Esta también ha sido una preocupación académica en la que se evidencia una brecha de investigación crítica ya que como se ha señalado, la inteligencia artificial tiene la capacidad para influir en todos los objetivos de desarrollo sostenible.

Justificación

Tal como lo señala la UNESCO, contar con un medio ambiente sano es una pre condición esencial para cualquier conversación relativa al progreso social. La protección y restauración del medio ambiente resultan esenciales para el logro de un desarrollo sostenible, es decir, para el logro de un desarrollo que asegure las necesidades del presente sin comprometer las necesidades de futuras generaciones.

Por otro lado, la OCDE también lo concibe como un objetivo deseado para las personas y el planeta, para promover de esta forma entornos naturales que promuevan un crecimiento inclusivo y que genere bienestar.

En efecto, son múltiples los casos donde la aplicación de la inteligencia artificial puede venir a prestar un apoyo indispensable, como en la eficiencia en el uso de riego en la pequeña agricultura o en el monitoreo climático.

f) Gobernanza Colaborativa

La participación de las diferentes partes interesadas a lo largo del ciclo de vida de los sistemas de IA es necesaria para garantizar enfoques inclusivos de la gobernanza de la IA, de modo que los beneficios puedan ser compartidos por todos, y para contribuir al desarrollo sostenible. Entre las partes interesadas figuran, entre otros, los gobiernos, las organizaciones intergubernamentales, la comunidad técnica, la sociedad civil, los investigadores y los círculos universitarios, los medios de comunicación, los responsables de la educación, los encargados de formular políticas, las empresas del sector privado, las instituciones de derechos humanos y los organismos de fomento de la igualdad, los órganos de vigilancia de la lucha contra la discriminación y los grupos de jóvenes y niños.

Deben adoptarse medidas para tener en cuenta los cambios en las tecnologías y la aparición de nuevos grupos de partes interesadas y para permitir una participación significativa de las personas, las comunidades y los grupos marginados.

Justificación

En sentido amplio, la gobernanza refiere a la cultura y el entorno institucional en el que los ciudadanos, las instituciones relevantes y las partes interesadas interactúan entre sí y participan en los asuntos públicos.

La ausencia de una estrategia nacional y de una adecuada gobernanza en términos generales en Latinoamérica ha generado que los esfuerzos por avanzar en materia de IA sean aislados. En ese sentido, se denota una ausencia de coordinación entre las distintas

instituciones en los países de la región, que impacta en la transferencia tecnológica, en la investigación y desarrollo (I+D), en la alfabetización digital, en la visión intersectorial e inclusiva de la IA y, en general, en las posibilidades reales de potenciar los beneficios de la IA en favor de la población.

Latinoamérica requiere de una sincronización colaborativa, interdisciplinaria e inclusiva, donde los diferentes integrantes de la sociedad participen de forma significativa en la construcción del marco de gobernanza de la IA. Esta colaboración permanente contribuiría a legitimar eventuales marcos éticos y regulatorios para la IA que se lleguen a implementar en los países.

Operatividad de los Principios

Ejemplos de operatividad del Principio de Protección de los Derechos Humanos:

a) Evaluar las implicaciones a largo plazo antes de desarrollar y desplegar sistemas de IA para que ningún grupo de humanos se vuelva irrelevante y ninguna persona sea perjudicada en su vida, sus posibilidades y sus proyecciones a causa del uso de los sistemas de IA.

b) Lo anterior se traduce en una política de evaluación de riesgos y daños a largo plazo en materia de Derechos Humanos antes de embarcarse en nuevos despliegues de innovaciones tecnológicas.

c) Compartir la prosperidad creada por la IA, implementando mecanismos para redistribuir el aumento de la productividad para todos, asegurándose que la IA no aumente la desigualdad ya existente en los países.

d) En todos los desarrollos de IA implementados por el Estado debe primar el interés público el cual tiene por finalidad contribuir a la realización de la dignidad humana en la escala más amplia posible. Lo anterior importa la protección del ser humano en su calidad de tal, el incentivo a su autorrealización y por consiguiente la debida protección de sus Derechos Fundamentales debe estar garantizada en este ámbito.

Ejemplos de operatividad del Principio de Transparencia y Rendición de Cuentas:

a) En particular, respecto de sistemas de IA utilizados en el sector público, la transparencia va más allá de un requisito técnico, sino que debe ser entendida desde un concepto amplio que además incluya información relevante; desde el punto de vista social se traduce en exigir al Estado, además de información técnica, información política y sobre elecciones de diseñadores y tomadores de decisiones, es decir, información sobre las decisiones políticas detrás de las decisiones técnicas.

b) Para el cumplimiento del estándar anterior, esta transparencia lleva implícita la condición que organismos públicos adquieran sistemas de IA que permitan al ciudadano ejercer su derecho de acceso a la información y al organismo público rendir cuentas y brindar información completa sobre estos sistemas. Por ello, debe valorarse si conviene que los sistemas de IA adquiridos por el sector público estén protegidos por secretos comerciales o acuerdos de confidencialidad.

c) En el mismo sentido, es necesario que exista una transparencia activa del Estado, con mecanismos como registros, repositorios y plataformas públicas que permitan

consultar la totalidad de los proyectos y usos de IA que se realizan en el sector público, además de procesos de licitación abiertos. La colaboración público-privada debe ser totalmente transparente, haciendo públicos los conflictos de intereses, contratos con proveedores y cualquier información relevante, cumpliendo con las más altas exigencias de probidad y rendición de cuentas.

d) Sobre la rendición de cuentas, una de las medidas que más relevancia ha tenido son las evaluaciones de impacto algorítmico ("AIAs" por sus siglas en inglés), siendo consideradas en múltiples propuestas legislativas como la reciente propuesta europea. Es altamente recomendable incorporar la obligatoriedad de estas evaluaciones en el sector público e incentivar su utilización en el sector privado.

   Estas evaluaciones buscan determinar el potencial riesgo de un sistema de IA en su contexto, y tratan de comprender, clasificar y responder mejor a los posibles daños o riesgos que puede conllevar.

e) Incorporación de estándares para auditorías algorítmicas, un concepto que es más específico que las AIAs, en el sentido que su función es establecer, por ejemplo, la existencia de un determinado sesgo que perjudica a un grupo o subgrupo de una población en un sistema de IA definido. Si bien no existe aún un estándar internacional generalizado para estas auditorías, se recomienda dar seguimiento a distintas iniciativas internacionales en la materia y obtener retroalimentación de la academia y de los organismos especializados.

f) Las personas destinatarias de decisiones basadas en IA deben conocer, antes de ser objeto de la decisión, que están interactuando con un sistema de esta naturaleza. Además, debe proveérseles información suficiente y comprensible, de forma previa, sobre el funcionamiento general del sistema y las consecuencias de las decisiones para la persona.

g) Los responsables del uso del sistema deben ofrecer evidencia científica que compruebe que éste funciona para lo que se dice, así como acreditar los niveles de certeza de las predicciones o resultados que arroje.

Ejemplos de operatividad del Principio de Privacidad y Protección de Datos:

a) Asegurar legalidad y la calidad de los datos (Principio de Calidad de los Datos) al momento de captar y utilizar datos personales, especialmente durante el entrenamiento del sistema de IA, para mitigar los potenciales sesgos contenidos en los datos. Esto supone utilizar datos suficientemente representativos del entorno en que se aplique el sistema, datos completos y pertinentes, así como efectuar un proceso de limpieza y anotación de datos riguroso.

b) Efectuar estudios de impacto en la privacidad previo a la implementación de un proyecto de IA que afecte significativamente a las personas o sus derechos, en el cual se identifiquen los riesgos del sistema y las medidas de mitigación de esos riesgos, especialmente en el sector público.

c) Privacidad, Ética y Seguridad desde el Diseño y por Defecto: los proyectos de IA deben diseñarse tomando en cuenta, desde su origen, las configuraciones, medidas de seguridad y procesos más respetuosos de la privacidad, así como los principios reconocidos a nivel mundial en materia de protección de datos: principio de

minimización, tanto de la cantidad de datos como del tiempo de almacenamiento de ellos, proporcionalidad y adecuación al fin, entre otros.

d) Utilizar herramientas para anonimizar datos que disminuyan al mínimo el riesgo de re-identificación de las personas cuyos datos hayan sido utilizados para desarrollar los modelos o entrenar los sistemas.

e) Desarrollar una gobernanza de datos que permita al Estado, a los investigadores y a las empresas aprovechar los datos para la toma de decisiones y para la innovación, pero de manera segura, responsable y protegiendo siempre la privacidad y los datos personales involucrados.

Ejemplos de operatividad del Principio de Seguridad:

a) Una vez implementadas, aplicar un testeo y monitoreo constante de pruebas y de supervisión de los sistemas de IA. Esto es especialmente relevante en el uso de tecnología de aprendizaje automático que probablemente evolucionará después de la implementación a medida que siga recibiendo información nueva.

b) Los desarrolladores de sistemas de IA no siempre pueden predecir con exactitud los riesgos asociados a dichos sistemas *ex ante*. También existen riesgos de seguridad asociados a la aplicación de los sistemas de IA que sus creadores no previeron, por lo que debe buscarse que los sistemas adopten "decisiones relativamente seguras".

c) Incorporar la seguridad de los sistemas de IA en el dictado de reglas actualizadas en materia de tratamiento íntegro y de calidad de datos personales;

d) Actualizar materias de ciberseguridad;

e) Promover estándares técnicos en relación con desarrollos de IA dependiendo del sector donde se apliquen;

f) Fomentar sólo el uso de sistemas que permitan utilizar datos en condiciones seguras para los titulares de esa información; y

g) Establecer sistemas de gobernanza que permitan monitorear y evaluar la observancia de condiciones de uso seguro de la IA.

Ejemplos de operatividad del Principio de Sostenibilidad:

a) Incorporación a través de normas de acceso a la información pública que permiten saber qué hace el Estado en relación a estas materias, por ejemplo, a través de canales que informen al ciudadano sobre la existencia de un sistema, de las decisiones que toma y su impacto en sistemas conexos;

b) Incorporación a través de normas de participación ciudadana en instancias de decisión ambiental;

c) Incorporación a través de normas de evaluación del impacto ambiental de los productos o servicios basados en IA, proporcionales al riesgo ambiental de cada sistema de IA por evaluar;

d) Incluir la sostenibilidad en los programas relativos a educación de calidad y trabajos calificados;

e) Incluir la sostenibilidad en el despliegue de programas de energías limpias; y

f) Incluir la sostenibilidad en planes y programas destinados a asegurar el acceso a agua limpia y ciudades inteligentes sostenibles.

Ejemplos de operatividad del Principio de Gobernanza Colaborativa:

a) Fomentar espacios colaborativos, como los denominados “clúster” o ecosistemas de innovación.

b) En esa línea, fAIr LAC iniciativa del Banco Interamericano de Desarrollo es un espacio ideal para integrar a todos los actores relevantes del ecosistema, construir una agenda común e instaurar un sistema de gobernanza propicio para el desarrollo sostenible y responsable de la IA en los países;

c) Para proyectos de IA en el sector público es recomendable instaurar mecanismos de participación y consulta ciudadana durante todo el ciclo de vida del proyecto, pero especialmente en la fase de planificación, previo a su implementación práctica. En estos mecanismos deberían participar, en la medida de lo posible, representantes de las poblaciones destinatarias del sistema de IA, así como personas expertas que contribuyan a identificar de manera prematura los potenciales riesgos y a proponer medidas para mitigar el impacto en los derechos de las personas.

4.2 Recomendaciones Generales para la Estructuración de un Marco Normativo para el Uso Ético y Responsable de la IA

En esta dimensión regulatoria el objetivo debe la protección de las personas, los derechos humanos y a la vez incentivar el desarrollo de este tipo de tecnologías a nivel local buscando una incidencia en el crecimiento económico de un país y en el beneficio concreto de los ciudadanos.

Desde el punto de vista de normas vinculantes, el marco legal debe estar compuesto, en primer lugar, por normas tecnológicamente neutras cuyo objetivo sea compatible con un desarrollo responsable de la IA. Dentro de estas, como se analizó en el capítulo II, cabe destacar leyes contra la discriminación, de protección de los derechos de los consumidores, de ciberseguridad, de propiedad intelectual, transparencia y leyes de protección de datos. Sobre este grupo normativo se debe hacer una revisión exhaustiva y proponer mejoras para reforzar el cumplimiento de sus objetivos y aspectos que estén relacionados con un uso responsable, seguro de sistemas de IA, incorporando los aspectos operativos del marco ético descrito en el acápite anterior.

Dentro de la regulación tecnológica propiamente tal, las normas sobre Investigación y Desarrollo (I+D), fomento a la innovación y Fintech permiten dar certeza jurídica y fomentar la innovación en materia de IA, generalmente en ambientes de prueba controlados por reguladores, por lo que al momento de diseñar estas normas se deben tener en consideración el fomento de desarrollos de IA.

De la misma forma, es posible valorar la conveniencia de introducir una ley o regulación especial de Inteligencia Artificial, como se ya está haciendo en la Unión Europea y otros países, pudiendo contribuir al objetivo de encaminar el uso de la IA hacia el bien común, solventando, eso sí, algunas fallas de mercado que no se resuelven con la legislación vigente y poniendo en el centro de cualquier innovación al ser humano, siempre teniendo en consideración el contexto histórico, social y políticos de los países latinoamericanos.

De acuerdo con lo anterior, proponemos las siguientes recomendaciones generales:

a) Una política de regulación que dé certeza jurídica y permita el crecimiento económico. El marco regulatorio debe ser un elemento habilitante y no obstaculizador para cumplir los fines de una futura estrategia de IA.
b) La importancia de fortalecer las normas sobre derechos humanos, competencia, protección al consumidor, propiedad intelectual, responsabilidad civil, entre otras.
c) La importancia de fortalecer y mitigar las brechas normativas como la falta de capacidad regulatoria, la falta de compromisos y debilidades en contratos públicos-privados y fortalecer la rendición de cuenta de reguladores de protección de datos.
d) No adoptar normativas de carácter general sin antes contextualizarlas. Por ejemplo, el Reglamento Europeo de Protección de Datos (GDPR) y las normas sobre IA implican complejos mecanismos de rendición de cuentas con múltiples actores especializados, con presupuestos de democracias y marcos regulatorios robustos. No se debe intentar diseñar un marco regulador a menos que se comprenda el contexto institucional del país y sus implicaciones para la regulación.
e) Valorar la incorporación de las denominadas "*sunset clauses*" o "cláusulas derogatorias." Este tipo de cláusulas dentro de la legislación sujetan una determinada regla a un periodo prestablecido de vigencia que, cuando se agota, obliga a evaluar la efectividad de la medida y demostrar la necesidad de mantener la regla; de lo contrario se elimina la medida del ordenamiento jurídico. Este tipo de cláusulas podrían ser útiles para probar el efecto de una regla o regulación determinada sobre la IA, de manera que se pueda derogar fácilmente en caso de no cumplir el objetivo deseado.
f) Las sanciones son importantes y su aplicación debe ser lo suficientemente equilibrada para incentivar el cumplimiento de la regulación. Sin embargo, la regulación debe concentrarse en medidas ex ante que procuren alinear el interés público y el privado, con la sostenibilidad y el respeto a los derechos fundamentales.

4.3 Recomendaciones para Regular el Uso de IA en el Sector Público

El Estado tiene algunas particularidades como comprador de soluciones de IA y proveedor de servicios públicos basados en esas soluciones. Estas son:

1. El Estado está sujeto al principio de legalidad y transparencia administrativa;
2. Sus contratos en esencia son públicos-privados, donde un actor privado diseña e implementa un sistema con todos los riesgos que conlleva y los pone a disposición de las entidades públicas;
3. El impacto de las implementaciones de IA es mayor que en el sector privado pues recae sobre gran parte de la población e impacta en servicios que son esenciales para el ciudadano y la ciudadana; y
4. Las personas no pueden elegir el proveedor.

Por ende, es proporcionalmente lógico que el Estado sea el primero en estar sujeto a una regulación mayor. En consecuencia, se recomienda que la regulación de IA para su uso en el sector público tome en cuenta lo siguiente:

1. Que el Estado se someta como comprador y usuario de IA no solo a evaluaciones ex ante de impacto algorítmico, sino a auditorías ex post, tomando en cuenta sus limitaciones, como el hecho de que no existe -hasta hoy- una metodología única o uniforme para su realización.
2. Establecer reglas especiales de contratación pública para una IA confiable.
3. Exigir al proveedor privado suficiente información y pruebas sobre la confiabilidad, seguridad, transparencia y certeza de los sistemas de IA.
4. Para el cumplimiento del estándar anterior, esta transparencia podría incorporar la condición que organismos públicos no adquieran sistemas de IA que estén protegidos por secretos comerciales o acuerdos de confidencialidad. Esto deberá ser valorado por el organismo encargado de diseñar la Estrategia de IA y de compras públicas, pues la imposición tácita a los proveedores de renunciar a la confidencialidad que les garantiza la ley sobre sus secretos comerciales u otra propiedad intelectual puede tener consecuencias no deseadas, como obligar a contratar soluciones menos seguras o de inferior calidad, y desincentivar el mercado.
5. En el mismo sentido, es necesario que exista una transparencia activa del Estado, con mecanismos como registros y plataformas públicas, además de procesos de licitación abiertos.
6. La colaboración público-privada debe ser totalmente transparente, haciendo público conflictos de intereses, contratos con proveedores y cualquier información relevante, cumpliendo con las más altas exigencias de probidad y rendición de cuentas.
7. Asimismo, en el caso de software de uso público, los gobiernos tienen la oportunidad de establecer requisitos técnicos adicionales tanto para su propio desarrollo como para la compra de software desarrollados por terceros. Así, por ejemplo, en la fase de diseño o adquisición se podrían establecer requerimientos de factores pro-transparencia, como disponer de software de código abierto seguro, acceso a artefactos de ingeniería de software, incluidos documentos de requisitos y diseño, seguimiento de errores y bitácoras de cambios en el código, planes de prueba y resultados.
8. Tener listas confiables y precalificadas de proveedores con estándares específicos.
9. Tener calificaciones públicas que estén disponibles en una fuente centralizada e inspeccionables por la población.

5. Conclusiones

Latinoamérica debe tener voz y voto en la mesa de discusión mundial sobre el futuro de la inteligencia artificial. Este documento es un esfuerzo en esa dirección. Los impactos éticos de la inteligencia artificial, si bien no discriminan entre regiones y países, sí tienen diferentes grados de incidencia según el país o la región en donde se desplieguen estos sistemas, siendo Latinoamérica una de esas regiones donde los efectos pueden ser mayores o más perjudiciales. Es por ello que los gobiernos, empresas, academia y la sociedad civil deben construir sólidos fundamentos éticos y regulaciones que sepan equilibrar las bondades de la tecnología y la innovación, con los derechos y la dignidad humana de quienes serán tarde o temprano impactados por la tecnología.

Los instrumentos éticos y normativos estudiados reflejan una corriente de creciente concientización sobre los riesgos de la IA. Es necesario, por ello, que se establezcan principios éticos claros, adaptados al contexto de la región y basados en derechos humanos, que fomenten la transparencia en la toma de decisiones, así como la participación de todos los actores en la creación y aplicación de la inteligencia artificial. Asimismo, se deben materializar dichos principios en acciones concretas que tengan un impacto real y positivo en la vida de todas las personas.

El marco propuesto es una base sobre la cual cada país puede partir para construir su propia estrategia de inteligencia artificial responsable y para crear consciencia entre sus ciudadanos sobre los ingredientes necesarios para hacer de esta tecnología una aliada de la humanidad. Recomendaciones sobre estrategias de IA se encuentran en un documento hermano a este[29] y una introducción a la IA responsable se presenta en este artículo.[30]

[29] López, O., Baeza-Yates, R. Mayorga, R. Estrategias Nacionales de IA: Comparación y Recomendaciones, OptIA, 2023.
[30] Baeza-Yates, R. Introduction to Responsible AI. European Review 31 (4), Cambridge, pp. 406-421, 2023.

Anexo 1: Marco Ético y Normativo del Consejo de Europa

Hemos recogido la metodología y trabajo realizado por el Consejo de Europa a través de su Comité Ad hoc sobre Inteligencia Artificial (CAHAI), en particular, su análisis que relaciona principios rectores de la IA con instrumentos internacionales jurídicamente vinculantes, dándoles a dichos principios contexto y soporte.

El desarrollo de este marco que a la vez es ético y regulatorio en el campo de la IA, se estructura a partir de los siguientes pasos:

a) En primer lugar, se identifican e individualizan los principales instrumentos internacionales jurídicamente vinculantes en materias de derechos humanos ratificados por la mayoría de los países latinoamericanos y que son pertinentes con el contexto de desarrollos e impactos de la IA.

b) Luego, se realizó un análisis exploratorio en torno a las potenciales vulneraciones sobre derechos humanos que se relacionan con el impacto de determinados sistemas de IA.

c) Se proponen principios rectores para el uso responsable y ético de la IA que busquen enfrentar y mitigar los descritos impactos sobre derechos humanos.

d) Los tres componentes anteriores, esto es, (1) normas internacionales vinculantes de derechos humanos, (2) impacto sobre las personas y (3) principios éticos se conectan entre sí, de manera que los principios éticos rectores son contextualizados dentro de los referidos tratados internacionales en materia de derechos humanos que se relacionan con los impactos que producen determinados desarrollos de IA.

Las ventajas de relacionar normas jurídicas, en este caso, instrumentos internacionales jurídicamente vinculantes en materias de derechos humanos y principios rectores para el uso ético y responsable de la IA, desde el contexto latinoamericano, son las siguientes:

a) Situar a los principios éticos rectores de la IA dentro de instrumentos internacionales jurídicamente vinculantes en materias de derechos humanos, en primer lugar, tiene como finalidad alejarlos de su carácter abstracto con diversas interpretaciones, dándoles un contexto y funcionalidad concreta.

b) Se centra en el mayor problema que presentan los impactos de desarrollos y usos de la IA y que se traduce principalmente en la vulneración de derechos humanos.

c) Dicha conexión, además, permite mapear y verificar la existencia de aquellas normas particulares que deben configurar un marco regulatorio apropiado para el desarrollo de la IA de forma ética y responsable.

El fundamento teórico de este enfoque se basa en el supuesto de que los principios generales proporcionados por los instrumentos internacionales de derechos humanos son aquellos en donde se sustentan todas las sociedades y se desarrollan todas las actividades humanas, incluido el desarrollo de tecnologías como la IA. Adicionalmente, cabe tener presente que los tratados de derechos humanos, a diferencia de otros tipos de tratados no confieren derechos a los Estados, sino a los individuos frente a un Estado, el cual tiene obligaciones para con ellos.

En términos generales, las obligaciones principales de un Estado en materia de derechos humanos son la (a) función de garante, mediante la cual los Estados partes deben asegurar que toda su organización y estructura permitan jurídicamente el libre y pleno ejercicio de los derechos humanos, (b) la obligación de respeto que se traduce en que los Estados deben abstenerse de interferir en el disfrute de los derechos humanos, o de limitarlos y (c) la obligación de protección que exige que los Estados impidan los abusos de los derechos humanos contra individuos y grupos.

Impactos de los Sistemas de Inteligencia Artificial en los Derechos Humanos

El desarrollo normativo de una IA responsable debe estar guiado por los impactos, tanto deseables como no deseables, que pueden eventualmente producir, y con mayor razón si de estos impactos derivan potencialmente vulneraciones sobre los derechos humanos. Los impactos como se ha señalado permiten, a su vez, identificar aquellas normas pertinentes para un marco jurídico eficaz.

Con propósitos prácticos e ilustrativos reproducimos la clasificación de los impactos de la IA según lo ha desarrollado el Consejo de Europa, determinando 4 grandes grupos de vulneraciones contra la dignidad, libertad, igualdad y no discriminación y derechos económicos y sociales. Asimismo, agregamos aquellos impactos sobre la democracia y el estado de derecho.

a) Dignidad humana

Los derechos humanos descansan sobre la dignidad humana. Los Estados deben contribuir a la realización de la dignidad humana en la escala más amplia posible.
La extensión de la dignidad humana se ramifica a la protección de los siguientes derechos que se han visto vulnerados por el uso de ciertos sistemas de IA, como los siguientes:

a) *Seguridad, debido proceso, principio de legalidad*: Está altamente documentado que en algunos casos cuando se utilizan sistemas de IA para resolver problemas sociales profundos perpetúan y amplifican los prejuicios existentes en nuestra sociedad. Ejemplos de estos tipos de sistemas están presentes en materias como, predicción de delitos, puntuación de riesgo de reincidencia penal o de aplicación de determinadas penas.

   Dentro de estos, se encuentran los sistemas de reconocimiento facial, creados en los años 70 con el objeto de identificar a sospechosos delictuales contrastando fotografías contenidas en bases de datos policiales. Hasta hoy, los sistemas policiales de reconocimiento facial se construyen con bases de datos históricas, sin tomar en consideración que muchos datos son incompletos, sesgados, reflejo de detenciones ilegales o de racismo policial, lo cual explica, además, que la prevención de delitos mediante esta tecnología posea un alto margen de error.

b) *Vida privada y familiar, integridad física, psicológica y moral:* Muchos sistemas y usos de la IA, en particular aquellos de vigilancia masiva, tienen un impacto amplio y profundo en el derecho a la intimidad. Este tipo sistemas tiene impactos no solo en la privacidad o datos personales, sino que se extiende al derecho a la integridad personal, la honra y dignidad humana, incluso el derecho a la autonomía, lo que se traduce en

el hecho que las personas se vean obligadas a actuar de cierta manera por temor a ser calificadas negativamente por sistemas de vigilancia.

b) Libertad individual

Se consideran dos libertades individuales imprescindibles:

a) *Libertad de expresión:* Sistemas de IA elaboran perfiles, pudiendo etiquetar o clasificar a personas con base en sus actos. Lo anterior, puede tener un impacto tanto en la privacidad como en su libertad de expresión, ya que las personas no pueden expresarse de forma libre por temor. Por su parte, el uso del reconocimiento facial en espacios públicos puede eliminar el anonimato sin consentimiento cuando individualizan a personas, inhibiendo a las personas de su derecho a protestar o emitir su opinión en público.

b) *Libertad de reunión y asociación:* Sistemas de IA pueden rastrear e identificar automáticamente grupos de personas, inhibiendo sus intenciones de reunión y asociación. Por su parte, el uso del reconocimiento facial en espacios públicos puede inhibir a las personas de reunirse o asociarse en determinadas instancias.

c) Igualdad y no discriminación

Existe abundante evidencia de que existe un patrón detectado en el uso de sistemas de IA cuando se emplean en materias de relevancia social. En dichos casos, grupos históricamente excluidos y oprimidos, tales como, personas afroamericanas, latinas, pueblos originarios, comunidades LGBTIQ+, minorías religiosas, personas de escasos recursos, entre otras, son comúnmente afectados de forma desproporcionada, reproduciendo y perpetuando injusticias sociales.

Sistemas predictivos de abuso infantil que asocian injustamente pobreza con abusos, sistemas predictivos de delitos que asocian delincuencia a racismo policial histórico, sistemas predictivos que detectan erróneamente fraudes de beneficios sociales, son algunos ejemplos.

d) Derechos económicos y sociales

Algunos Sistemas de IA pueden vulnerar derechos de trabajadores, usados como mecanismos de control, seguimiento, evaluación de rendimiento, entre otros. Estas aplicaciones de la IA podrían poner en peligro el derecho a unas condiciones de trabajo dignas, seguras y saludables, como el derecho de sindicación.

e) Impactos sobre valores democráticos y estado de derecho

La IA puede tener impactos negativos sobre determinadas democracias, influyendo aún más en democracias menos robustas, esto puede reflejarse en el impacto que provocan sistemas de IA que son utilizados para evitar o limitar el acceso a información o influenciar a la población de cierta manera, en particular, en períodos eleccionarios.

Por su parte, si bien algunos sistemas de IA pueden aumentar la eficacia en la administración institucional, pueden erosionar la legitimidad procesal, el debido proceso y

el estado de derecho que proporciona el contexto jurídico para que los derechos humanos puedan ser desarrollados en plenitud.

Principios para el Uso Responsable y Ético de la IA

A continuación, enunciamos los ocho principios principales relacionados con el uso responsable y ético de la inteligencia artificial que están directamente relacionados con los impactos anteriormente descritos sobre los derechos humanos:

a) **Desarrollos centrados en el ser humano:** los productores de sistemas de IA deben centrar sus desarrollos en las personas y en el respeto por los derechos humanos, el estado de derecho y valores democráticos.

b) **Transparencia y explicabilidad:** los actores de IA deben proporcionar información significativa, adecuada al contexto y coherente con el estado de la técnica, permitiendo que los afectados por un sistema puedan comprender y cuestionar dichos sistemas.

c) **Robustez, seguridad y protección:** los sistemas de IA deben ser robustos y seguros durante todo su ciclo de vida para que, en condiciones de uso normal, uso previsible o mal uso, u otras condiciones adversas, funcionen adecuadamente y no presenten riesgos de seguridad irrazonables.

d) **Responsabilidad:** los actores de la IA deben ser responsables del correcto funcionamiento de los sistemas de IA y del respeto de los restantes principios, en función de sus funciones, el contexto y de conformidad con el estado de la técnica.

e) **Crecimiento inclusivo, desarrollo sostenible y bienestar:** se debe compartir la prosperidad creada por la IA, implementando mecanismos para redistribuir el aumento de la productividad para todos, asegurándose de que la IA no aumenta la desigualdad y que nadie se quede atrás, procurando que ningún grupo de humanos se vuelva irrelevante a causa de los sistemas de IA.

f) **Prevención de daños:** en su virtud, se elegirá un método de inteligencia artificial cuando esté justificado a través de evidencia científica robusta y sus resultados sean convenientes para los fines perseguidos una vez aplicadas evaluaciones de riesgo y de impacto sobre las personas, en particular, aquellos históricamente excluidos o marginados por la sociedad y sobre el medio ambiente. Será inocua cuando su aplicación no genere daños a los seres humanos, al medio ambiente y ecosistemas.

g) **Igualdad y no-discriminación:** la Inteligencia Artificial debe ser un mecanismo que genere justicia social de manera que sus beneficios deben buscarse procurando alcanzar al mayor número de personas posible sin distinción de raza, edad, situación migratoria, de la identidad de género o nivel socioeconómico. Para ello se debe procurar que en el diseño e implementación de sistemas de IA estén involucradas las visiones de personas pertenecientes a grupos excluidos o marginados, a quienes estos sistemas les impactan más de forma negativa, y dotándolas de la autonomía para auditarlos por su cuenta y oponerse a su implementación.

h) **Autonomía y supervisión humana:** El ser humano siempre debe poder auto determinarse, de manera que conserve el poder de decidir qué decisión tomar sobre

sí mismos, en lugar que lo haga un sistema de inteligencia artificial. Siempre debe ser posible atribuir la responsabilidad ética y jurídica, en cualquier etapa del ciclo de vida de los sistemas de IA, a personas físicas o entidades jurídicas existentes. Esta supervisión humana no es sólo individual, sino que también se refiere a la supervisión pública.

Anexo 2: Tratados Internacionales en Materia de Derechos Humanos

El siguiente cuadro relaciona los derechos humanos específicos que son impactados por la inteligencia artificial, con los principios éticos necesarios para mitigar dicho impacto. Se indican también las obligaciones que tienen los Estados relativas a esos derechos, y los tipos de normativa interna que guardan relación con los impactos mencionados, de manera que los Estados puedan adaptar dicha normativa a los retos de la inteligencia artificial.

Derechos humanos en riesgo de vulneración	**Respeto por la dignidad humana: seguridad, juicio justo, ninguna pena sin ley**
Norma vinculante internacional de derechos humanos ratificado por la Mayoría de países de América Latina.	Convención Americana Sobre Derechos Humanos Pacto Internacional De Derechos Civiles Y Políticos Pacto Internacional De Derechos Económicos, Sociales Y Culturales Protocolo Adicional A La Convención Americana Sobre Derechos Humanos En Materia De Derechos Económicos, Sociales Y Culturales "Protocolo De San Salvador" Convención Sobre Los Derechos Del Niño
Detalle normativo (ejemplos)	Convención americana de derechos humanos. Artículo 7. Derecho a la Libertad Personal Artículo 8. Garantías Judiciales Artículo 9. Principio de Legalidad y de Retroactividad Pacto Internacional de Derechos Civiles y Políticos Artículo 9 1. Todo individuo tiene derecho a la libertad y a la seguridad personales. Nadie podrá ser sometido a detención o prisión arbitrarias. Nadie podrá ser privado de su libertad, salvo por las causas fijadas por ley y con arreglo al procedimiento establecido en ésta. Artículo 15. Principio de Legalidad**.**
Obligación del Estado	Los estados se obligan a: Asegurar que la autoridad competente prevista por el sistema legal decidirá sobre los derechos de toda persona que interponga en sus recursos, desarrollar las posibilidades de recurso judicial, y garantizar el cumplimiento. Asegurar el principio de legalidad, la presunción de inocencia y el debido proceso.
Principios Éticos Asociados	Desarrollos centrados en el ser humano. Transparencia y explicabilidad. Robustez, seguridad y protección. Responsabilidad. Prevención de daños. **Autonomía y supervisión humana.**

Normas vinculantes internas	Normas sobre transparencia y acceso a la información. Normas sobre procesos judiciales.

Derechos humanos en riesgo de vulneración	**Respeto por la dignidad humana** **Vida privada y familiar, integridad física, psicológica y moral**
Norma vinculante internacional de derechos humanos.	Convención Americana Sobre Derechos Humanos Pacto Internacional De Derechos Civiles Y Políticos Pacto Internacional De Derechos Económicos, Sociales Y Culturales Protocolo Adicional A La Convención Americana Sobre Derechos Humanos En Materia De Derechos Económicos, Sociales Y Culturales "Protocolo De San Salvador" Convención Sobre Los Derechos Del Niño
Detalles normativos (ejemplos)	Convención Americana de Derechos Humanos Artículo 5. Derecho a la Integridad Personal. Artículo 11. Protección de la Honra y de la Dignidad.
Obligación del Estado	Los Estados deben asegurar que nadie será objeto de injerencias arbitrarias o ilegales en su vida privada, su familia, su domicilio o su correspondencia, ni de ataques ilegales a su honra y a su reputación. Toda persona tiene derecho a la protección de la ley contra tales injerencias o ataques.
Principios Éticos Asociados	Desarrollos centrados en el ser humano. Robustez, seguridad y protección. Responsabilidad. Prevención de daños. **Autonomía y supervisión humana.**
Normas vinculantes internas	Normas relativas a la protección de privacidad y protección de datos Normas relativas a la protección de autonomía

Derechos humanos en riesgo de vulneración	**Libertad individual** **Libertad de expresión** **Libertad de reunión y asociación**
Norma vinculante internacional de derechos humanos	Convención Americana Sobre Derechos Humanos Pacto Internacional De Derechos Civiles Y Políticos Pacto Internacional De Derechos Económicos, Sociales Y Culturales Protocolo Adicional A La Convención Americana Sobre Derechos Humanos En Materia De Derechos Económicos, Sociales Y Culturales "Protocolo De San Salvador" Convención Sobre Los Derechos Del Niño
Detalles normativos (ejemplos)	Convención Americana de Derechos Humanos Artículo 12. Libertad de Conciencia y de Religión Artículo 13. Libertad de Pensamiento y de Expresión Artículo 15. Derecho de Reunión. Artículo 16. Libertad de Asociación. Pacto Internacional de Derechos Civiles y Políticos.

	Artículo 19. Derecho a la Libertad de expresión. Artículo 21. Derecho de reunión pacífica. Artículo 22. Derecho a asociarse libremente.
Obligación del Estado	Todo individuo tiene derecho a la libertad de opinión y de expresión; este derecho incluye el no ser molestado a causa de sus opiniones, el de investigar y recibir informaciones y opiniones, y el de difundirlas, sin limitación de fronteras, por cualquier medio de expresión. El ejercicio de la libertad de expresión no puede estar sujeto a previa censura sino a responsabilidades ulteriores, las que deben estar expresamente fijadas por la ley y ser necesarias para asegurar el respeto a los derechos o a la reputación de los demás o la protección de la seguridad nacional, el orden público o la salud o la moral públicas.
Principios Éticos Asociados	Desarrollos centrados en el ser humano Robustez, seguridad y protección Igualdad y no-discriminación Autonomía y supervisión humana
Normas vinculantes internas	Normas relativas a regulación de libertad de expresión, libertad personal, de reunión y asociación. Normas relativas a acceso a información y transparencia

Derechos humanos en riesgo de vulneración	**Igualdad y no discriminación**
Norma vinculante internacional de derechos humanos	Convención Americana Sobre Derechos Humanos Pacto Internacional De Derechos Civiles Y Políticos Pacto Internacional De Derechos Económicos, Sociales Y Culturales Protocolo Adicional A La Convención Americana Sobre Derechos Humanos En Materia De Derechos Económicos, Sociales Y Culturales "Protocolo De San Salvador" Convención Sobre Los Derechos Del Niño, Observación general núm. 25 relativa a los derechos de los niños en relación con el entorno digital.
Detalles normativos (ejemplos)	Convención Americana de Derechos Humanos Artículo 1. Obligación de respetar los derechos fundamentales. Artículo 24. Igualdad ante la ley. Observación general núm. 25 relativa a los derechos de los niños en relación con el entorno digital Capitulo III. letra a) Principio general de No discriminación
Obligación del Estado	Los Estados deben abstenerse de realizar acciones que de cualquier manera vayan dirigidas, directa o indirectamente, a crear situaciones de discriminación de jure o de facto. los Estados están obligados a adoptar medidas positivas para revertir o cambiar situaciones discriminatorias existentes en sus sociedades, en perjuicio de determinado grupo de personas. Esto implica el deber especial de protección que el Estado debe ejercer con respecto a actuaciones y prácticas de terceros que,

	bajo su tolerancia o aquiescencia, creen, mantengan o favorezcan las situaciones discriminatorias. El derecho a la no discriminación exige que los Estados parte se aseguren de que todos los niños tengan acceso equitativo y efectivo al entorno digital de manera beneficiosa para ellos. Los Estados parte deben adoptar todas las medidas necesarias para evitar la exclusión digital. Esto incluye proporcionar acceso gratuito y seguro a los niños en lugares públicos específicos e invertir en políticas y programas que apoyen el acceso asequible de todos los niños a las tecnologías digitales y su utilización informada en los entornos educativos, las comunidades y los hogares.
Principios Éticos Asociados	**Igualdad y no-discriminación.** Desarrollos centrados en el ser humano. Transparencia y explicabilidad. Responsabilidad. Prevención de daños. **Autonomía y supervisión humana.**
Normas vinculantes internas	**Normas sobre no discriminación**

Derechos humanos en riesgo de vulneración	**Derechos económicos y sociales.**
Norma vinculante internacional de derechos humanos	Convención Americana Sobre Derechos Humanos Pacto Internacional De Derechos Civiles Y Políticos Pacto Internacional De Derechos Económicos, Sociales Y Culturales Protocolo Adicional A La Convención Americana Sobre Derechos Humanos En Materia De Derechos Económicos, Sociales Y Culturales "Protocolo De San Salvador" Convención Sobre Los Derechos Del Niño
Detalle normativo (ejemplos)	Convención Americana de Derechos Humanos Artículo 26. Desarrollo Progresivo. Protocolo Adicional a La Convención Americana sobre Derechos Humanos en materia de Derechos Económicos, Sociales y Culturales "Protocolo de San Salvador". Artículo 6 Derecho al Trabajo. Artículo 11 Derecho a un Medio Ambiente Sano. Artículo 13 Derecho a la Educación.
Obligación del Estado	Los Estados deben adoptar medidas hasta el máximo de los recursos de que dispongan, para lograr progresivamente la plena efectividad de los derechos económicos y sociales.
Principios Éticos Asociados	Crecimiento inclusivo, desarrollo sostenible y bienestar. Igualdad y no discriminación.
Normas vinculantes internas	Normas relativas a derechos laborales Normas relativas a derechos educativos Normas relacionadas con protección al medio ambiente Normas sobre no discriminación