



Gobernanza Ética de la Inteligencia Artificial:

Grandes Empresas Tecnológicas,
Inteligencia Artificial y la Gobernanza
Global Tecnológica



EIIS | EUROPEAN INSTITUTE
OF INTERNATIONAL STUDIES

Gobernanza Ética de la Inteligencia Artificial:

Grandes Empresas Tecnológicas,
Inteligencia Artificial y la Gobernanza
Global Tecnológica

Gobernanza Ética de la Inteligencia Artificial: Grandes Empresas Tecnológicas, Inteligencia Artificial y la Gobernanza Global Tecnológica

Autor:

Sistema Económico Latinoamericano y del Caribe (SELA)

Director Editorial:

Clarems Endara, Secretario Permanente del SELA

Coordinación de Publicaciones:

Yeimy Ramirez Ávila, Jefe de Gabinete

Supervisión Editorial:

Yeimy Ramirez Ávila, Jefe de Gabinete

Gustavo Herrera, Coordinador de Desarrollo Social

Klibis Marín, Oficial de Comunicaciones

Maquetación y Diseño de portada:

Mauricio Gaitán

Impreso en:

MACRO SRL. La Paz - Bolivia

ISBN – Impreso: 978-980-6458-40-6

ISBN – Digital: 978-980-6458-41-3

Depósito Legal – Impreso: DC2025001195

Depósito Legal – Digital: DC2025001194

Copyright © SELA, Julio de 2025. www.sela.org

© Sistema Económico Latinoamericano y del Caribe (SELA), 2025 Torre Europa, pisos 4 y 5 Avenida Francisco de Miranda, Urbanización Campo Alegre Caracas, 1060, República Bolivariana de Venezuela Apartado 17035, Caracas 1010-A.

Todos los derechos reservados. Prohibida su venta. No se permite la reproducción total o parcial de este documento, ni su almacenamiento en un sistema informático, ni su transmisión en cualquier forma o por cualquier medio, electrónico, mecánico, fotocopia u otros métodos, sin el permiso previo del SELA.

ÍNDICE

Servicio Exterior, Diplomacia Digital e Inteligencia Artificial: gobernanza, actores emergentes y dilemas éticos en América Latina y el Caribe <i>Alejandro Garofali</i>	8
La regulación de la Inteligencia Artificial en la Unión Europea: innovación, riesgos y gobernanza <i>Aline Beltrame de Moura</i>	16
Inteligencia Artificial centrada en el ser humano en los países nórdicos <i>Anna Virkama</i>	28
La regulación y la gobernanza global de la Inteligencia Artificial <i>Antonio Núñez y García-Saúco</i>	38
Iniciativas gubernamentales sobre la gobernanza ética de la IA en América y Europa: Canadá, Estados Unidos, Brasil, Reino Unido y Francia. <i>Arturo Flores López</i>	47
Gobernanza ética de las iniciativas de IA en Asia: China, Japón, Singapur, Corea del Sur, Taiwán <i>Cátia Miriam Costa</i>	58
El rol de Naciones Unidas en la construcción de la Gobernanza ética de la IA y la coyuntura latinoamericana <i>Cecilia Celeste Danesi</i>	69
La Inteligencia Artificial y la Geopolítica <i>Jesús Rafael Argumosa Pila</i>	79

Diplomacia en la era de la inteligencia artificial: Nuevos desafíos y oportunidades <i>José Beraún Aranibar</i>	88
La Inteligencia Artificial como desafío regional: Propuesta para una arquitectura de Gobernanza Ética en América Latina y el Caribe <i>Lourdes Daza Aramayo Ph.D</i>	98
La gobernanza ética de la inteligencia artificial: Big Tech Companies & Techplomacy <i>Mario Torres Jarrín</i>	122
Inteligencia artificial y protección de datos personales: el caso del RGPD de la Unión Europea y sus implicaciones para América Latina y el Caribe <i>Mónica Velasco Pufleau</i>	128
Inteligencia Artificial (IA) y Sistemas Judiciarios: el desafío de la Justicia Numérica <i>Naiara Posenato</i>	139
La Soberanía Digital como pilar estratégico para el desarrollo de América Latina y el Caribe en la era de la Gobernanza Tecnológica Global <i>Gustavo Herrera</i> <i>Luis Emilio Martínez</i>	151

PRÓLOGO

El Sistema Económico Latinoamericano y del Caribe (SELA), en continuidad con la publicación *Inteligencia Artificial y Diplomacia: las relaciones internacionales en la era de las tecnologías disruptivas* (2024), presenta esta nueva entrega con el objetivo de continuar fortaleciendo las capacidades negociadoras y diplomáticas en la región de América Latina y el Caribe, en consonancia con los principios y prioridades definidos en el Programa de Trabajo 2022-2026 del organismo.

Esta segunda edición, titulada *Gobernanza Ética de la Inteligencia Artificial: Grandes Empresas Tecnológicas, Inteligencia Artificial y la Gobernanza Global Tecnológica*, forma parte de un esfuerzo sostenido por analizar las implicaciones éticas, geopolíticas y estratégicas de la Inteligencia Artificial (IA) en el escenario internacional contemporáneo, así como, los desafíos emergentes que plantea la gobernanza de las tecnologías avanzadas.

Esta publicación responde al compromiso del SELA de orientar, promover y continuar fomentando el desarrollo de capacidades negociadoras para una mejor comprensión de la aplicación de la diplomacia a los retos políticos, geopolíticos y geoeconómicos que surgen en el ciberespacio. En ese sentido, forma parte de un proceso de formación regional más amplio, desarrollado por el SELA en alianza con el European Institute of International Studies (EIIS) y la colaboración de la Universidad Pontificia de Salamanca.

Positivamente, esta publicación busca dotar a diplomáticos, académicos, altos funcionarios y tomadores de decisiones, de herramientas conceptuales y operativas para comprender y anticipar los impactos de la revolución digital, la inteligencia artificial, la Big Data y otras tecnologías emergentes inmersas en la diplomacia, la cooperación internacional y la seguridad global.

Este segundo volumen refleja los contenidos de las principales reflexiones, conceptos y debates abordados durante los cursos internacionales sobre “*Gobernanza ética de la Inteligencia Artificial*” y “*La Gobernanza Global Tecnológica: Big Tech Companies e Inteligencia Artificial*”, celebrados en mayo de 2025 en el marco de la conmemoración del 50° aniversario del SELA.

Con base en este enfoque, reunimos a expertos internacionales en ética digital, gobernanza, algorítmica, relaciones internacionales, regulación tecnológica y derecho internacional, con el fin de examinar tanto los retos como las oportunidades que plantean estas tecnologías en constante evolución en un mundo cada vez más digitalizado, donde la IA desempeña un papel central en diversos sectores.

En esta entrega se abordan temáticas sobre la necesidad de establecer marcos de gobernanza ética para el desarrollo y uso responsable de la inteligencia artificial, reconociendo que esta tecnología no es neutral y que sus impactos pueden amplificar desigualdades, impactar derechos fundamentales y moldear decisiones de alto impacto. Se reflexiona sobre principios esenciales como la equidad, la privacidad, la transparencia, la seguridad, la no discriminación, la supervisión humana; y se destacan herramientas como auditorías algorítmicas, evaluaciones de impacto ético, marcos regulatorios adaptativos y mecanismos de cooperación internacional.

De la misma forma, los tópicos abordados se centran en la gobernanza tecnológica en un contexto global dominado por grandes empresas tecnológicas transnacionales. Se analiza como el poder de las *Big Tech companies* influye en la geopolítica, en la configuración de los marcos regulatorios y en la soberanía de los Estados. Se examinan los dilemas sobre el control de los flujos de información, la concentración del poder económico y tecnológico, la fragmentación regulatoria y el rol de los organismos multilaterales frente a estos nuevos actores. Asimismo, se plantea la necesidad de avanzar hacia un sistema de gobernanza global tecnológica que combine estándares éticos comunes, cooperación multilateral efectiva, desarrollo inclusivo de capacidades y participación ciudadana.

Para el SELA, es claro que el desarrollo acelerado de estas tecnologías debe analizarse multidisciplinariamente al margen de la política internacional. El poder de las grandes empresas tecnológicas —las llamadas *Big Tech*— ha alcanzado un nivel sin precedentes, con capacidad de influir sobre decisiones gubernamentales, elecciones, mercados financieros y marcos regulatorios a escala planetaria. En este contexto, la gobernanza global de la tecnología ha dejado de ser una aspiración teórica para convertirse en una necesidad apremiante.

La Inteligencia Artificial no es una herramienta neutral. Los algoritmos que la sustentan reflejan valores e intereses de sus diseñadores. Por ello, avanzar hacia una gobernanza ética y responsable de la IA no es un lujo, sino una exigencia para resguardar los derechos fundamentales, la equidad, la transparencia y la confianza pública.

Desde el SELA, esperamos que esta compilación de distintos estudios y perspectivas sea de utilidad para la membresía, así como, para todos aquellos actores que trabajan en el diseño de políticas públicas, marcos normativos e iniciativas internacionales vinculadas al desarrollo ético, justo y sostenible de la inteligencia artificial y las tecnologías emergentes.

En un momento de aceleración digital y creciente interdependencia, promover una gobernanza centrada en las personas, orientada al bien común y basada en valores compartidos es no solo una responsabilidad, sino también una oportunidad estratégica para nuestra región.

Dr. Clarems Endara
Embajador

Secretario Permanente
Sistema Económico Latinoamericano
Y del Caribe (SELA)

Servicio Exterior, Diplomacia Digital e Inteligencia Artificial: gobernanza, actores emergentes y dilemas éticos en América Latina y el Caribe

Emb. Dr. Alejandro Garofali
*Presidente del European Institute
of International Studies (Suecia)*

La Inteligencia Artificial (IA) y la diplomacia digital están cambiando la manera en que se hace diplomacia hoy en día, por ende, la forma en que se ejerce la profesión diplomática. Esa combinación virtuosa tiene gran impacto en todo, como la forma en que el mundo se administra, los principios éticos de la tecnología, e incluso cómo participan los actores importantes, tanto si son del sector público como si no lo son (especialmente si reconocemos en el accionar de las Big Tech, elementos asimilables a actores diplomáticos no estatales).

Mirando a América Latina y el Caribe (ALC), pese a todos los problemas y las diferencias en tecnología que se encuentran, también surgen oportunidades de desarrollar nuestras propias maneras de hacer diplomacia digital y el rol del diplomático de la región en ese entorno. Y por qué no, también en diálogo con las Big Tech y sus correlatos en los ecosistemas digitales regionales.

En ese sentido, es menester analizar desde varias perspectivas, cómo la IA y la diplomacia digital se juntan, desde teorías importantes hasta los nuevos jugadores, y también los problemas éticos y políticos, con un ojo puesto en lo que pasa en ALC.

1. Algunas nociones preliminares:

La diplomacia, que solemos ver como actividad de vinculación entre países y en foros de la comunidad internacional organizada, usando canales y reglas formales, ahora está sufriendo

cambios sustanciales dada la digitalización y la expansión de campos de aplicación de la inteligencia artificial.

Esta mezcla de elementos está cambiando la manera en que entendemos el poder, como se nos representa y negociamos, cambian los actores y los espacios y tiempos en que se dan tales procesos. Además, aparecen nuevos problemas éticos y técnicos en este contexto (Torres, 2022).

En ALC, región que nos ocupa, esta onda de digitalización en las relaciones internacionales se está desarrollando en un ambiente con persistencia de desigualdades en acceso a la tecnología, con un tanto de dependencia de afuera y muchas veces participando en dichos procesos a través de instituciones fragmentadas. No obstante, también hay espacio para ideas renovadoras e integradoras en la diplomacia, y oportunidades de forjar lazos en la región e impulsar políticas que busquen mayor independencia en la tecnología y hacer que las cosas sean un poco más justas para todos (De Souza, 2021; Mariño, 2022). Una forma de transformación digital priorizando la equidad y la sostenibilidad, integrando IA en diplomacia para beneficio de todos los sectores de la sociedad, establecer alianzas multilaterales y regionales para maximizar el potencial de aplicación de la tecnología y su uso ético a todo nivel.

Nos proponemos hacer un análisis lo más sintético posible y crítico, y con varias perspectivas sobre la diplomacia digital en la época de la inteligencia artificial, con un enfoque especial en América Latina y el Caribe (ALC).

Recurramos a los marcos teóricos, los nuevos personajes involucrados, y los problemas morales y políticos que aparecen con esta nueva forma de poder global (blando y algo más).

2. Bases Teóricas y Conceptos

La diplomacia digital, se trata de cómo determinados actores de ámbito internacional gestionan el poder en el mundo mediante el uso de herramientas digitales para sus estrategias: así, promueven lo que les conviene, cuentan sus historias y manejan sus relaciones (Bjola & Holmes, 2015).

La inteligencia artificial, a su manera, implica sistemas capaces de tomar decisiones por sí mismos o de aprender de información de manera progresiva, sin la necesidad de la intervención directa de un humano. Esto está cambiando la forma en que analizamos, predecimos e incluso cómo nos comunicamos, todos, todo (Russell & Norvig, 2020).

Ahora bien, al juntar estos dos conceptos, surge algo muy interesante: la diplomacia digital, a veces llamada diplomacia algorítmica. Aquí es donde los algoritmos entran en juego, influyendo en la interacción diplomática. Pueden, por ejemplo, ayudar a tomar decisiones sobre política exterior, analizar riesgos, incluso hacer simulaciones de escenarios. Aunque también plantea dilemas éticos complicados, mucho de los cuales aún no tienen respuesta (Bjola, 2019; Zhang et al., 2021) y nos deben mantener alerta.

Y si nos fijamos un poco más, es claro que la diplomacia digital tiene que ser estudiada teniendo en cuenta el grado en que las jerarquías de poder impactan su desarrollo, sobre todo en lugares como Latinoamérica (Torres, 2022). Y un aspecto que no nos cansamos de subrayar, la inteligencia artificial no es precisamente imparcial, refleja los prejuicios y las maneras de control que existen en los sistemas que la crean (Crawford, 2021; Floridi et al., 2018).

Así, resulta fundamental conectar las ideas que abordan la equidad de los algoritmos, la inclusión social y -por supuesto-, la autonomía tecnológica, tanto como posible. Estas ideas nos dan una base para pensar una mejor diplomacia digital.

Una diplomacia que evite reproducir y perpetuar las desigualdades existentes entre naciones, sino que al contrario realmente las enfrente, utilizando un análisis profundo y un enfoque humano basado en la ética de la tecnología y sus aplicaciones, con sentido de trascendencia de los servicios diplomáticos como bien público al momento de su incorporación como herramienta o su consideración como objeto de negociación y gestión como proceso internacional.

3. Nuevos protagonistas en escena

El mundo digital ha cambiado la diplomacia por completo. Ha sumado una complejidad increíble, además de un montón de nuevos personajes en juego. Ya no son solo los gobiernos los que hablan y regulan. Ahora tenemos organizaciones internacionales, grandes empresas tecnológicas, ONG's y hasta movimientos sociales involucrados, que opinan y hasta generan mecanismos de autorregulación.

Y las mayores empresas tecnológicas, las "Big Tech", tienen un poder enorme sobre cómo fluye la información y sobre toda la infraestructura digital, algo que, aunque no lo pretendan, afecta las políticas de un sinnúmero de países y altera ecuaciones de poder global (Suzor, 2019; Zuboff, 2020).

En América Latina y el Caribe, la situación es un tanto agri dulce. Por un lado, dependemos bastante de la tecnología de otras regiones, lo que nos pone un límite en nuestra independencia en la diplomacia digital. Pero, por otro lado, están surgiendo ideas autónomas, muy valiosas, desde los gobiernos locales, las universidades y grupos de ciudadanos que están usando tecnología abierta, redes que no tienen un centro y nuevas formas de participación para tener influencia en lo que pasa en todo el mundo.

Un ejemplo que salta a la vista es la diplomacia digital feminista, como iniciativa impulsada por la Red de Politólogas, junto con la Diplomacia Ciudadana Andina, que les da voz a grupos indígenas y en la Comunidad Andina. Estas propuestas ofrecen una perspectiva diferente. Van más allá de las formas usuales de cómo un país se representa, y lo más importante, ayudan a poner de relieve el valor del conocimiento que nace desde la tierra, de las bases de la sociedad (Gómez-Barris, 2018).

Estos actores están construyendo un “ecosistema diplomático más grande”, un espacio donde los límites entre el gobierno y las organizaciones, digamos, no gubernamentales, se desdibujan, en un fenómeno que termina siendo más glocal que local. Todo esto abre camino a nuevas maneras de influir en la geopolítica mundial (Torres, 2022).

4. El gobierno mundial y las diferencias regionales

Un reto gigante que nos ocupa actualmente en las relaciones internacionales es comprender cómo dominar la inteligencia artificial. Estados Unidos y China, por ejemplo, están invirtiendo fuertemente para desarrollarla y sacar provecho, sino dominarla. Por su parte, otras zonas, como América Latina y el Caribe, están planteando y evaluando cómo enfrentar la falta legislación, capital e infraestructuras, así como la dependencia de la tecnología (UNESCO, 2021; Filgueira, 2023).

En este escenario, la diplomacia digital es casi un campo de marte, un terreno de batalla, un lugar donde se pelea por el poder mundial. Autores como Deibert (2020) y Torres (2022) insisten en que es urgente establecer reglas que rijan la IA y que involucren a todos los países, sobre todo a los del Sur Global, para garantizar que haya transparencia, responsabilidad y que todos tengan la misma oportunidad de acceder a la tecnología.

En América Latina y Caribe, la Comunidad de Estados Latinoamericanos y Caribeños (CELAC), así como, la Organización de los Estados Americanos (OEA) y el Sistema Económico Latinoamericano y del Caribe (SELA), intentan promover la transformación digital y la ciencia abierta, conscientes de que es un largo camino que requerirá amplia inversión en tiempo y paciencia. Crear conciencia y movilizar tanto las maquinarias públicas y privadas, como la sociedad civil, a incorporar estas temáticas, considerar opciones, juntar actores

y producir soluciones comunes adaptadas a nuestras realidades, es una tarea de grandes proporciones e inevitable, tarde o temprano. Todo es empezar a tomar conciencia y actuar o quedar relegado.

En esta región, al asunto de la fragmentación, sumado a la insuficiente inversión en investigación y desarrollo, se superpone la forma en que las grandes corporaciones tecnológicas actúan o interactúan con y en la política global y local en cada caso. Todo esto hace que resulte difícil alcanzar una diplomacia digital que de verdad podamos controlar, como apuntan Ceballos, Maisonnave, Britto Londoño (2020), así como, De Souza (2021).

Ahora refiramos a los aspectos éticos y otras maneras de ver estos procesos. El tema es qué abarca la IA en la diplomacia. Ello abre un montón de preguntas éticas relevantes. Usar algoritmos para abordar discursos, adivinar cómo se comportan las personas o planificar cosas y ahí radica el problema. Porque ello fácilmente podría aumentar la vigilancia, la discriminación o la manipulación. Y comprometer seriamente la información personal y vulnerabilidad de los servicios públicos. Y ello afecta los derechos humanos básicos como la privacidad, la libertad de expresión y hasta nuestra cultura, como explican Whittaker y colegas (2018) y también Crawford (2021).

Por eso, la ética en los algoritmos necesita ir más allá y no sólo mirar la tecnología y el individuo. Necesitamos una visión contextualizada que entienda de historia, geopolítica y cultura. En este sentido, entendemos, Torres (2022), alude a que tenemos que buscar una IA que incluya a todos y respete las diferentes culturas.

Juntar la diplomacia digital con la inteligencia artificial está cambiando por completo la forma en que entendemos el poder y la soberanía en este siglo XXI. Sabemos de las problemáticas de fondo respecto de transparencia, sesgos y desigualdades regionales, así como, de los asimétricos y lentos avances normativos, algunas veces muy fragmentados que no terminan de incorporar en sistemas nacionales lineamientos claros de gestión de la IA. No obstante, empezamos a ver esfuerzos ambiciosos en otras regiones (como órdenes ejecutivas en EE. UU. y normas sobre IA en la UE), que inspiran desarrollos autónomos en nuestros países. La pregunta de fondo es si los países con menos capacidades en IA pueden beneficiarse del conocimiento regional compartido que se viene dando en la materia.

Ante los riesgos que amenazan con marginarnos, concentrar el poder y dominarnos digitalmente, la región de ALC necesita urgentemente construir sus propias opciones. Es decir, alternativas sólidas que se basen en la justicia social, la soberanía tecnológica y por supuesto, la participación democrática.

Sintetizando, para desarrollar los grandes pilares de una estrategia digital efectiva (gobernanza, infraestructuras y tecnología), es imperativo profundizar el análisis con visión crítica sobre estos procesos, entrelazando de manera práctica ideas teóricas, dilemas morales y ejemplos concretos de nuestra región. Y es que la diplomacia digital no puede ignorar las estructuras de poder mundial, ni tampoco las diferencias históricas que se verifican en grandes regiones del planeta.

Así pues, es importantísimo fortalecer las capacidades diplomáticas dentro de la región. Los gobiernos deben dialogar, alinear sus estrategias y acordar bases comunes. Al mismo tiempo, estos procesos deben legitimarse, entendiendo necesario impulsar una diplomacia que tenga inspiración en las bases de la sociedad, que escuche la voz de la gente, de las comunidades locales y de género, intergeneracional, aunándolas con las de los demás actores de estas sociedades. Y que esas voces no solo sean escuchadas, sino que también todos los actores sientan que tienen un papel importante en este cambio digital a nivel global.

Apostar por la democratización de la tecnología y estar alertas a las sensibilidades que genera la digitalización de procesos (alertando en forma temprana sobre riesgos y marcando oportunidades que la tecnología presenta para la sociedad toda), debieran ser ejes de trabajo futuro en cualquier foro y a todo nivel, en tónica glocal.

La proposición de fondo es que si usamos la IA de una forma que incluya a todos, promueva la equidad y el bienestar y nos dé independencia tecnológica, podremos tener una diplomacia digital que trascienda la región y de contribuya a un mundo más justo. Y viceversa, una diplomacia alerta, presta a incorporar la IA para una transformación de procesos hacia la eficiencia y eficacia de la actividad diplomática, debe acompañar la evolución exponencial de la tecnología, usarla como herramienta de cambio, complementar decisiones complejas con su uso, potenciar la gobernanza global digital y asegurar que la diplomacia digital destaca como bien público regional y global.

Ello sin duda supondrá una adaptación estratégica de nuestros servicios exteriores, que deberán necesariamente cooperar, compartir sus estrategias (emular buenas prácticas, como el uso de IA en ciberseguridad, así como, en procesos consulares digitales, caso de Chile), ampliar la complementación entre los institutos diplomáticos (con aplicación de herramientas de simulación en base a modelos digitales, por ejemplo), para ir delineando un perfil de diplomático tipo latinoamericano, que incorpore habilidades tecnológicas propias del ejercicio del poder blando para ser un actor relevante en el ámbito digital.

El proceso ha iniciado y no debiera tener pausa, los cursos de capacitación para altos funcionarios de nuestros servicios exteriores como las ofrecidas por el SELA y estrategias

afines de construcción y fortalecimiento de capacidades para la gobernanza regional digital, así como, la vinculación más estrecha de las academias diplomáticas y sus planes de formación, son parte de un camino que se viene pavimentando y está dando sus frutos.

De igual manera, debe promoverse la participación de nuestros países en procesos de construcción colaborativa multilateral en la temática (como la Cumbre de Futuro de la ONU y su primer esbozo de compendio, el Global Digital Compact).

Se deberán redoblar esfuerzos, la velocidad de avance de la IA es exponencial, como decíamos antes, a la velocidad de la luz, cuando las posibilidades de su incorporación y dominio en el campo diplomático la siguen a paso humano.

Referencias

Bjola, C. (2019). “La Diplomacia en la Era de la Inteligencia Artificial”. Real Instituto Elcano, ARI 98/2019.

Bjola, C., & Holmes, M. (2015). “Diplomacia Digital: Teoría y Práctica”. Routledge.

Ceballos, L., Maisonnave, M., Britto Londoño, C. (2020). “Soberanía tecnológica digital en Latinoamérica”. Revista Propuestas para el Desarrollo, año IV, número IV, octubre 2020, páginas 151-167.

Crawford, K. (2021). “Atlas de la IA: Poder, Política y los Costos Planetarios de la Inteligencia Artificial”. Yale University Press.

De Souza, J. (2021). “Colonialismo digital”. Editora UBU.

Deibert, R. (2020). “Reiniciar: Recuperando Internet para la Sociedad Civil”. House of Anansi.

Floridi, L., Cowls, J., Beltrametti, M.,... (2018). “AI4People-Un Marco Ético para una Buena Sociedad de la IA: Oportunidades, Riesgos, Principios y Recomendaciones”. Minds & Machines 28, 689–707.

Filgueira, F. (2023). Retos de gobernanza de la inteligencia artificial en América Latina. Infraestructura, descolonización y nueva dependencia. Revista Del CLAD Reforma Y Democracia, 87, 44-70.

Gómez-Barris, M. (2018). "Más allá de la Marea Rosa: Arte y Corrientes Políticas Subyacentes en las Américas". University of California Press.

Russell, S., Norvig, P. (2020). "Inteligencia Artificial: Un Enfoque Moderno" (4ª ed.). Pearson.

Suzor, N. (2019). "Sin Ley: Las Reglas Secretas que Gobiernan Nuestras Vidas Digitales". Cambridge University Press.

Torres, M. (2022). "Diplomacia algorítmica y justicia digital: repensar la gobernanza desde el Sur." Revista Latinoamericana de Política Internacional, 15(2), 101-128.

UNESCO. (2021). "Recomendación sobre la Ética de la Inteligencia Artificial". UNESCO.

Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V. & Schwartz, O. (2018). "Informe AI Now 2018". AI Now Institute.

Zhang, Y., Wu, M., Tian, G. Y., Zhang, G. & Lu, J. (2021), "Ética y privacidad de la inteligencia artificial: Comprensiones desde la bibliometría", Knowledge-based Systems.

Zuboff, S. (2020). "La era del capitalismo de la vigilancia". Paidós.

La regulación de la Inteligencia Artificial en la Unión Europea: innovación, riesgos y gobernanza

Aline Beltrame de Moura¹

*Directora de la Cátedra de Estudios sobre la Unión Europea
del European Institute of International Studies (Suecia)*

La regulación de la Inteligencia Artificial (IA) en la Unión Europea (UE) es un área en rápida evolución, destacándose por el compromiso de la UE en liderar un modelo de gobernanza de IA que sea ético, seguro, confiable y centrado en el ser humano. A diferencia de otros modelos, la UE intenta promover un enfoque de la gobernanza digital centrado en el ser humano, que haga hincapié en las normas y los valores democráticos². Este modelo tiene como objetivo mantener altos niveles de protección de la salud, seguridad y derechos fundamentales conforme a lo establecido en la Carta de los Derechos Fundamentales de la Unión Europea (CDFUE).

Considerando este escenario, la IA deberá servir como instrumento para las personas con el objetivo de aumentar el bienestar humano y proteger valores como la democracia, el estado de derecho y el medio ambiente sin crear riesgos ni perjudicar intereses públicos

1 Profesor de Derecho en la Universidad Federal de Santa Catarina (Brasil). Titular de la Cátedra Jean Monnet de Derecho de la Unión Europea. Coordinadora del Módulo Jean Monnet (2018-2021), de la Red Jean Monnet – Proyecto BRIDGE (2020-2023) y del Debate de Políticas Jean Monnet – BRIDGE Watch (2023-2026). Todos los proyectos son cofinanciados por el Programa Erasmus+ de la Comisión Europea. Directora de la Cátedra de Estudios de la Unión Europea en el Instituto Europeo de Estudios Internacionales (Suecia). Coordinadora del Centro Latinoamericano de Estudios Europeos (LACES) y Editora en Jefe de la *Latin American Journal of European Studies*. Doctora en Derecho Internacional por la Università degli Studi di Milano (Italia).

2 Van T’Klooster, K., & Hosli, M. O. (2024). Dilemmas and tradeoffs of digital governance models: The European Union, China and United States compared. In M. Torres Jarrín (Ed.), *Diplomacy and digital age: Science, technology and global digital governance* (p. 117). Peter Lang.

y derechos fundamentales. El objetivo de la regulación es mejorar el funcionamiento del mercado interno mediante la creación de un marco jurídico uniforme para el desarrollo, comercialización, implementación y uso de sistemas de IA en la UE.

1. Década Digital de Europa: el camino hasta la Ley de IA

En un mundo dinámico y en constante transformación, queda claro que impulsar la digitalización trae muchos beneficios para la sociedad. La UE entonces desea reforzar su soberanía digital y definir sus propias normas para estar preparada para la era digital. Una de las principales iniciativas políticas de la Unión Europea fue la implementación del Mercado Único Digital en 2015³, basado en tres pilares fundamentales: acceso, entorno y crecimiento. Este marco sentó las bases para el desarrollo de la estrategia digital europea, configurando su visión a largo plazo para la transformación digital e integración en todos los sectores.

Para guiar esta transformación, la Comisión Europea lanzó el programa de políticas de la Década Digital de Europa⁴ con metas concretas para 2030. Estos objetivos incluyen 4 ejes principales: competencias digitales (capacitar a la población con habilidades digitales avanzadas), infraestructuras seguras y sostenibles (desarrollar infraestructuras tecnológicas seguras y ecológicas), transformación digital de las empresas (fomentar la digitalización en el sector empresarial) y digitalización de los servicios públicos (asegurar que los servicios públicos sean accesibles y eficientes digitalmente).

A partir del reciente Informe sobre el estado de la Década Digital 2024⁵ se desprenden cinco recomendaciones clave dirigidas a los Estados miembros para avanzar en la transformación digital de la UE de manera inclusiva, sostenible y centrada en las personas. En primer lugar, se destaca la importancia de una coordinación e implementación eficiente de la normativa digital existente. La Comisión insta a los Estados miembros a aplicar de forma coherente marcos regulatorios fundamentales como el Reglamento de Servicios Digitales (DSA)⁶, el Reglamento

3 European Commission. (2015). A Digital Single Market strategy for Europe [Communication]. Retrieved from https://ec.europa.eu/commission/presscorner/api/files/attachment/8210/DSM_communication.pdf.

4 Unión Europea. (2022). *Decisión (UE) 2022/2481 del Parlamento Europeo y del Consejo de 14 de diciembre de 2022 relativa al programa Europa Digital y por la que se deroga la Decisión (UE) 2021/1075* [Decisión publicada en el Diario Oficial de la Unión Europea, L 323, 19.12.2022, p. 4–33]. EUR-Lex. <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32022D2481>.

5 Comisión Europea. (2024). *Comunicación de la Comisión: Estado de la Década Digital 2024 – Informe principal* [Informe]. Recuperado de [<https://digital-strategy.ec.europa.eu/en/library/report-state-digital-decade-2024>].

6 Comisión Europea. (2022). *Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022 relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE (Ley de Servicios Digitales)*, DO L 277, 27.10.2022, p. 1–102, <https://eur-lex.europa.eu/eli/reg/2022/2065/oj>; <https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses>; <https://digital-strategy.ec.europa.eu/en/policies/dsa-enforcement>.

de Mercados Digitales (DMA)⁷, la Ley de Inteligencia Artificial (Ley AI)⁸, el Reglamento sobre la identidad digital europea (EUDI)⁹, y legislación sobre ciberseguridad y datos¹⁰. Para ello, se requiere una reducción de cargas administrativas, sinergias entre estructuras de gobernanza y un enfoque gubernamental integral que abarque todos los niveles de la administración pública.

En segundo lugar, se recomienda reforzar la competitividad y la soberanía digital europea. Esto implica aumentar las inversiones públicas y privadas en tecnologías clave, superar barreras del mercado único, fomentar redes colaborativas de conectividad, computación y microelectrónica, y modernizar los sistemas educativos y de formación con una visión estratégica de largo plazo en torno a las competencias digitales.

Una tercera recomendación consiste en promover un despliegue equitativo de las tecnologías digitales en todos los territorios de la UE. La Comisión subraya la necesidad de ampliar la cooperación entre actores locales, como pymes, universidades, autoridades regionales y municipales, mediante la creación de consorcios y *hubs* digitales, de modo que los beneficios de la digitalización alcancen a todas las regiones y sectores sociales, reduciendo las brechas de desarrollo.

En cuarto lugar, se enfatiza el aprovechamiento de la transformación digital para impulsar una transición ecológica inteligente. Los Estados miembros deben alinear sus estrategias digitales con los objetivos climáticos, estableciendo metas cuantificables de sostenibilidad digital y garantizando la interoperabilidad entre las agendas digital y climática, en especial a través de los Planes Nacionales de Energía y Clima (NECPs)¹¹, pasando de proyectos piloto a soluciones a gran escala.

7 Comisión Europea. (2022). *Reglamento (UE) 2022/1925 del Parlamento Europeo y del Consejo de 14 de septiembre de 2022 sobre mercados disputables y equitativos en el sector digital y por el que se modifican las Directivas (UE) 2019/1937 y (UE) 2020/1828 (Ley de Mercados Digitales)*, DO L 265, 12.10.2022, p. 1–66, <http://data.europa.eu/eli/reg/2022/1925/oj>; DMA designated Gatekeepers (europa.eu).

8 Comisión Europea. (2024). *Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo relativo al Espacio Europeo de Datos Sanitarios*. DO L 168, 12.07.2024, p. 1–84.

9 Comisión Europea. (2024). *Reglamento (UE) 2024/1183 del Parlamento Europeo y del Consejo relativo al Espacio Europeo de Datos Sanitarios y por el que se modifican los Reglamentos (UE) n.º 2018/1724 y (UE) 2019/1021*, DO L 183, 30.04.2024, p. 1–84.

10 Como con la actualización de la Directiva de Seguridad de las Redes y los Sistemas de Información, la Ley de Ciberseguridad y la Ley de Ciberresiliencia, el Reglamento Europeo de Identidad Digital (EUDI) y el Reglamento General de Protección de Datos (GDPR).

11 Los planes nacionales de energía y clima (PNEC) fueron introducidos por el Reglamento sobre la gobernanza de la unión de la energía y la acción por el clima (UE)2018/1999, acordado como parte del paquete Energía limpia para todos los europeos que se adoptó en 2019.

Por último, se hace un llamado a colocar a las personas y sus derechos en el centro del proceso de transformación digital. Esto incluye reforzar los derechos digitales, garantizar la protección de los menores en línea, cerrar las brechas sociales y digitales, y combatir riesgos como la desinformación, la exclusión digital y las amenazas a la salud pública y la democracia.

Estas cinco líneas de acción buscan acelerar el cumplimiento de los objetivos establecidos en el programa de la Década Digital, consolidando una Europa digital más resiliente, justa y competitiva, promoviendo la inclusión digital, mejorando las infraestructuras tecnológicas y asegurando que todos los ciudadanos europeos puedan participar plenamente en la era digital.

En este sentido, la regulación de la IA en la Unión Europea, representada por la Ley de Inteligencia Artificial, es un componente crucial de este esfuerzo de transformación digital. La Ley IA busca establecer un marco jurídico uniforme para el desarrollo, comercialización, implementación y uso de sistemas de IA en la UE. Es importante destacar que este reglamento no es un acto aislado, una vez que forma parte de una estrategia legislativa integral que pretende responder a un conjunto de desafíos en el ámbito digital, los cuales analizaremos con detalle más adelante.

2. De la propuesta inicial al texto final

La regulación europea sobre IA tiene su origen en un proceso iniciado en 2017, cuando el Parlamento Europeo comenzó a debatir el tema bajo términos como “robótica”¹². En 2018, la Comisión Europea creó un grupo de expertos en IA¹³ y, en 2019, publicó las primeras Directrices Éticas para una IA confiable¹⁴, un hito que destacó la importancia de respetar los derechos humanos. En 2020, la Comisión presentó el Libro Blanco sobre IA¹⁵, lo que permitió lanzar una amplia consulta pública y realizar una evaluación de impacto. Con base en ello, en abril de 2021, la Comisión, único órgano con iniciativa legislativa en la UE, presentó oficialmente la propuesta de Reglamento de IA.

12 European Parliament. (2017, February 16). European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics. (2015/2103(INL)) (P8_TA(2017)0051). *Official Journal of the European Union*, C 252, 239-257. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52017IP0051>.

13 European Commission. (2018, March 7). *The Commission presents a European approach to artificial intelligence*. European Commission – Press Corner. https://ec.europa.eu/commission/presscorner/detail/en/ip_18_1381.

14 High-Level Expert Group on Artificial Intelligence. (2019). *Ethics guidelines for trustworthy AI*. European Commission. <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>.

15 European Commission. (2020). *White paper on artificial intelligence: A European approach to excellence and trust*. https://commission.europa.eu/document/download/d2ec4039-c5be-423a-81ef-b9e44e79825b_en?filename=commission-white-paper-artificial-intelligence-feb2020_en.pdf.

A lo largo de 2022 y 2023, el Consejo y el Parlamento Europeo discutieron el texto, logrando un acuerdo político en diciembre de 2023. Este proceso duró dos años y medio, considerado ágil para los estándares legislativos europeos. El resultado final de la Ley de la IA difiere significativamente de la propuesta original de 2021, que no contemplaba la IA generativa como ChatGPT. Fue solo tras el auge de esta tecnología a finales de 2022 que el Parlamento Europeo introdujo cambios importantes, incorporando obligaciones específicas para desarrolladores de modelos de IA.

Entre las nuevas disposiciones, se destacan la exigencia de respeto a los derechos fundamentales, evaluaciones de riesgo, y medidas de transparencia para informar a los usuarios sobre contenidos generados por IA. También se incluyó la protección del estado de derecho y del medio ambiente, considerando el alto consumo energético de los centros de datos. Asimismo, el reglamento pasó a abarcar tanto los modelos de propósito general como los generativos, imponiendo obligaciones diferenciadas según sus riesgos. El proceso refleja una evolución legislativa ágil y adaptativa ante el ritmo acelerado de la innovación tecnológica.

En este proceso de elaboración de la Ley IA, uno de los aspectos que más debate suscitó en el Parlamento Europeo y en el Consejo durante la tramitación del Reglamento fue la definición del término “sistema de inteligencia artificial”. La formulación finalmente adoptada por la Ley de IA busca ser tecnológicamente neutra y lo suficientemente flexible para mantenerse vigente frente a la rápida evolución tecnológica y del mercado.

En ese sentido, la definición establecida en el artículo 3.1 es la siguiente: “Un sistema de inteligencia artificial es un sistema basado en máquinas, diseñado para operar con distintos niveles de autonomía, que puede demostrar capacidad de adaptación tras su puesta en funcionamiento y que, con fines explícitos o implícitos, infiere a partir de los datos de entrada cómo generar resultados como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales.” Esta definición abarca una variedad de técnicas que conforman la inteligencia artificial, tales como el aprendizaje automático, el procesamiento del lenguaje natural y los sistemas de apoyo a la toma de decisiones, buscando ofrecer una perspectiva práctica sobre su funcionamiento. En todo momento, se ha destacado como elementos clave la autonomía y la capacidad de adaptación que caracterizan a estos sistemas.¹⁶

Ella es amplia y dinámica puesto que pretende abarcar tanto las tecnologías actuales como aquellas que puedan desarrollarse en el futuro, garantizando así una regulación eficaz sin

¹⁶ Cotino Hueso, L. (2024). ¿Qué es «inteligencia artificial» para el Reglamento? Análisis, delimitación y aplicaciones prácticas. En L. Cotino Hueso & P. Simón Castellano (Dir.), *Tratado sobre el Reglamento de Inteligencia Artificial de la Unión Europea* (p. 121). Aranzadi.

limitar el potencial innovador del sector. Justamente con el objetivo de proporcionar la seguridad jurídica y las adaptaciones necesarias a los nuevos avances tecnológicos, la Comisión evaluará anualmente la necesidad de modificar la lista del anexo III, que establece cuales son los sistemas de IA considerados de alto riesgo, y la lista de prácticas de IA prohibidas previstas en el artículo 5 del reglamento.

En definitiva, el Reglamento (UE) 2024/1689 fue aprobado por el Parlamento en marzo de 2024 y publicado en el Diario Oficial en julio de 2024. El texto incluye 180 considerandos, 113 artículos y 13 anexos, consolidando el primer marco legal integral sobre inteligencia artificial en el mundo. Sin embargo, el calendario de implementación de la Ley de IA prevé una aplicación escalonada a partir del 1 de agosto de 2024, según su artículo 113. Las prohibiciones sobre prácticas de riesgo inaceptable entrarán en vigor el 2 de febrero de 2025. A partir del 2 de agosto de 2025, se aplicarán obligaciones para modelos de IA de propósito general, así como normas de gobernanza y sanciones. El 2 de agosto de 2026, el reglamento será plenamente aplicable a la mayoría de los sistemas, especialmente los de alto riesgo. Finalmente, el 2 de agosto de 2027 comenzarán a regir las obligaciones específicas para sistemas de alto riesgo integrados en productos regulados. Este enfoque gradual busca facilitar la adaptación y el cumplimiento efectivo por parte de todos los actores involucrados.

3. Categorización de los sistemas de IA con un enfoque basado en los riesgos

En relación con la forma de categorización y regulación de los sistemas de IA, la Europa ha adoptado un enfoque basado en el riesgo, evaluando los sistemas de IA de manera diferenciada. De acuerdo con el considerando 26 de la Ley de IA, este enfoque asegura que la regulación sea proporcional al riesgo que presentan los sistemas, evitando regulaciones excesivamente restrictivas para sistemas de bajo riesgo y garantizando una protección robusta para aquellos que representan mayores peligros.

Por lo tanto, el Reglamento clasifica los sistemas de inteligencia artificial en cuatro niveles de riesgo: inaceptable, alto, limitado (relacionado con la transparencia) y mínimo. Esta categorización establece el grado de supervisión necesario y define los requisitos específicos que deben cumplir tanto los desarrolladores como los operadores de estos sistemas. De este modo, el marco normativo ajusta proporcionalmente el tipo y el alcance de las obligaciones conforme al nivel de riesgo que cada sistema de IA pueda implicar para los derechos fundamentales, la seguridad y el bienestar de las personas, garantizando así una regulación equilibrada que protege a los individuos sin obstaculizar el desarrollo tecnológico.

3.1 Riesgos Inaceptables

En primer lugar, es importante aclarar que los sistemas de IA con riesgos “inaceptables” están prohibidos, y el Capítulo II, artículo 5, establece una lista de prácticas vetadas. Así, el reglamento prohíbe ciertos usos de la IA considerados una amenaza clara para los derechos y la seguridad de los ciudadanos europeos, ya que contravienen los valores de la Unión.

La regulación de la UE sobre inteligencia artificial prohíbe determinados usos considerados de riesgo inaceptable por vulnerar derechos fundamentales como la privacidad, la dignidad humana y la no discriminación. Entre estas prácticas se destaca la puntuación social, como la que existe en China¹⁷, donde los ciudadanos son monitoreados de forma permanente y evaluados según su comportamiento. Este sistema asigna puntuaciones que inciden directamente en el acceso al empleo, a servicios públicos o al derecho a viajar, siendo una forma de control social incompatible con los valores democráticos europeos.

Otra práctica prohibida es la manipulación cognitivo-conductual de personas o grupos vulnerables, como niños o personas con discapacidades físicas o mentales. Los sistemas de IA no pueden explotar estas vulnerabilidades para influir negativamente en el comportamiento de los individuos, como ocurre en la publicidad abusiva dirigida a menores, en la venta de productos a precios inflados a personas dependientes, o en plataformas de juego que estimulan adicciones mediante técnicas manipulativas.

También se prohíbe el uso de sistemas de identificación biométrica remota en tiempo real en espacios públicos, debido a su carácter altamente intrusivo y al potencial efecto intimidatorio sobre las libertades fundamentales. Solo se permite en circunstancias excepcionales de interés público, como la búsqueda de personas desaparecidas o amenazas terroristas.

Asimismo, están prohibidos los sistemas de IA que emplean técnicas subliminales o engañosas para distorsionar las decisiones informadas de las personas, como influir en votos, fomentar compras irresponsables o inducir a conductas perjudiciales sin el conocimiento del usuario. Igualmente, se excluyen los sistemas de evaluación automatizada del riesgo penal que se basan únicamente en perfiles de comportamiento o rasgos de personalidad, sin una evaluación humana fundamentada en hechos objetivos, como en decisiones sobre libertad condicional o vigilancia predictiva.

Estas restricciones reflejan el compromiso de la UE con la transparencia, la ética y la supervisión humana en el uso de la IA. La prohibición de estas prácticas es esencial para proteger

17 Başer, T. (2022). *Artificial Intelligence and Social Credit System In China* [M.S. - Master Of Science Without Thesis]. Middle East Technical University.

a las personas y a la sociedad frente a los efectos nocivos de la manipulación digital, fortalecer la confianza pública en las nuevas tecnologías y asegurar que el desarrollo de la IA sea seguro, beneficioso y respetuoso de los derechos fundamentales.

3.2 Alto Riesgo

El Capítulo III de la Ley de la IA europea establece normas específicas para los sistemas de IA considerados de alto riesgo debido a su potencial impacto sobre la salud, la seguridad y los derechos fundamentales de las personas. Estos sistemas solo pueden comercializarse en el mercado europeo si cumplen requisitos obligatorios y se someten a una evaluación de conformidad previa. La clasificación como “alto riesgo” no se basa únicamente en la función del sistema, sino también en su propósito y forma de uso. Deben cumplir con estándares estrictos en gestión de riesgos, solidez técnica, gobernanza de datos, transparencia, supervisión humana y ciberseguridad.

El reglamento distingue dos grandes categorías de IA de alto riesgo: sistemas integrados como componentes de seguridad de productos ya regulados, y sistemas independientes que afectan derechos fundamentales, enumerados en el Anexo III. La Comisión Europea puede actualizar esta lista mediante actos delegados para incorporar nuevos usos.

De este modo, entre los ejemplos de prácticas consideradas de alto riesgo se destacan los sistemas utilizados en migración, asilo y control fronterizo, los cuales requieren especial atención por su impacto directo en personas en situación de vulnerabilidad. Estos sistemas deben operar con la máxima precisión, evitar cualquier forma de discriminación y cumplir estrictamente con normas internacionales, como la Convención de Ginebra sobre el Estatuto de los Refugiados¹⁸ y el principio de no devolución (*non-refoulement*)¹⁹, garantizando así la protección efectiva de los derechos humanos fundamentales.

También se consideran de alto riesgo los sistemas que seleccionan currículos, pues pueden reproducir sesgos discriminatorios. Lo mismo ocurre con los sistemas biométricos, como la categorización basada en atributos sensibles o el reconocimiento de emociones, que están estrictamente regulados. Se excluyen aquellos utilizados únicamente para confirmar la identidad.

La IA aplicada a infraestructuras críticas como transporte o suministro de agua y energía también entra en esta categoría, por su posible afectación a la vida cotidiana. En el ámbito educativo, los sistemas que determinan acceso, evaluación o monitoreo también son

18 ACNUR. (1951). *Convención sobre el Estatuto de los Refugiados*. https://www.acnur.org/sites/default/files/2023-05/Convencion_1951.pdf.

19 Hathaway, J. C. (2005). *The Rights of Refugees under International Law*. Cambridge University Press.

regulados. Lo mismo aplica para IA usada en servicios sociales y asistencia pública, cuyo mal uso puede vulnerar derechos fundamentales como la dignidad y la protección social. Asimismo, los sistemas utilizados para evaluar la solvencia o riesgo crediticio están incluidos, a excepción de aquellos destinados a la detección de fraude.

En el ámbito judicial, la IA es de alto riesgo por la necesidad de respetar la independencia judicial, evitar sesgos y garantizar la supervisión humana constante. Los programadores de estos sistemas no pueden estar vinculados a los poderes Ejecutivo o Legislativo, para asegurar la separación de poderes. Igualmente, se consideran de alto riesgo los sistemas diseñados para influir en procesos democráticos, como elecciones y referendos, según el Reglamento (UE) 2024/900 sobre propaganda política.²⁰

Todos estos sistemas deben superar evaluaciones de conformidad antes y durante su vida útil, y los ciudadanos tienen derecho a presentar quejas ante las autoridades competentes. Si un proveedor considera que su sistema no es de alto riesgo, debe justificarlo documentadamente. En resumen, esta clasificación busca garantizar un desarrollo y uso ético y responsable de la IA, protegiendo los derechos y minimizando restricciones al comercio internacional, en un delicado equilibrio entre innovación y protección fundamental.

3.3 Riesgo Limitado (o de transparencia)

Los sistemas de inteligencia artificial clasificados como de riesgo limitado, o riesgo de transparencia, están regulados en el Capítulo IV de la Ley de la IA, que establece obligaciones específicas para los proveedores y responsables del despliegue de sistemas diseñados para interactuar directamente con las personas, como los *chatbots* o los sistemas de reconocimiento emocional. Aunque estos sistemas no representan un riesgo elevado, pueden inducir a errores o malentendidos, por lo que deben cumplir estrictamente con normas de transparencia e información al usuario.

Uno de los requisitos fundamentales es que los usuarios sean informados de manera clara y comprensible cuando están interactuando con un sistema de IA o cuando sus emociones o características personales son reconocidas de forma automatizada. Además, los proveedores tienen la obligación de diseñar sus sistemas de modo que eviten la generación de contenidos ilegales, como ultrafalsificaciones (*deepfakes*), y deben implementar medidas de mitigación efectivas. También están obligados a publicar resúmenes que detallen los datos protegidos por derechos de autor utilizados para entrenar los modelos de IA.

20 European Union. (2024). *Regulation (EU) 2024/900 of the European Parliament and of the Council of 11 April 2024 on the transparency and targeting of political advertising*. Official Journal of the European Union, L 90, 1–40. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202400900.

En el caso particular de los modelos de IA generativa, como ChatGPT y DeepSeek, aunque no se los clasifica dentro de la categoría de alto riesgo, deben igualmente observar estas obligaciones de transparencia y cumplir con la normativa de la UE en materia de propiedad intelectual. Esto implica, por ejemplo, que los contenidos generados o modificados mediante IA, como imágenes, audios o videos, deben estar claramente etiquetados como tales, garantizando que los usuarios sean conscientes de su origen y naturaleza.

En resumen, los sistemas de IA de riesgo limitado deben respetar rigurosos requisitos de etiquetado y transparencia para proteger los derechos de los usuarios, fomentar la confianza en la tecnología y garantizar un uso ético y responsable de la inteligencia artificial. Esta regulación contribuye de manera decisiva a crear un entorno digital más seguro, justo y respetuoso tanto para los consumidores como para los desarrolladores.

3.4 Riesgo Mínimo

Los sistemas de inteligencia artificial clasificados como de “riesgo mínimo” son aquellos que no representan una amenaza significativa para la salud, la seguridad o los derechos fundamentales de las personas. Generalmente se utilizan en tareas no críticas y, por ello, están sujetos a una regulación más ligera en comparación con los sistemas de alto riesgo. Sin embargo, aun cuando no se les imponen obligaciones específicas adicionales en el marco de la Ley de la IA, estos sistemas deben cumplir con principios generales de transparencia y buenas prácticas, especialmente en lo que respecta a la protección de datos personales.

Entre los ejemplos más comunes se encuentran los filtros de spam, los videojuegos en los que los personajes son controlados por IA, los sistemas de recomendación de contenidos (como música o películas) y los asistentes virtuales como Siri o Alexa. Aunque la Ley de la IA no introduce nuevas exigencias regulatorias para estos sistemas, siguen estando plenamente sometidos a normativas generales de la Unión Europea, como el Reglamento General de Protección de Datos (GDPR)²¹, que garantiza la protección de los derechos de los usuarios y regula el uso de datos personales.

En definitiva, aunque los sistemas de riesgo mínimo no requieren medidas regulatorias tan estrictas como los de mayor riesgo, su funcionamiento debe alinearse con los estándares generales vigentes, asegurando así un entorno tecnológico seguro, transparente y respetuoso con los derechos fundamentales.

21 Parlamento Europeo & Consejo de la Unión Europea. (2016). *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*. Diario Oficial de la Unión Europea, L 119, 1–88. <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A32016R0679>.

4. Innovación, Gobernanza y Proyección Global de la Ley de la IA

Uno de los principales desafíos al regular tecnologías emergentes como la IA es equilibrar la innovación con la protección de derechos fundamentales. Aunque algunas grandes empresas tecnológicas han criticado la regulación por su potencial para frenar el desarrollo, una normativa bien diseñada puede, en realidad, fortalecer la seguridad jurídica, fomentar la competencia justa y orientar el progreso tecnológico de manera ética y responsable. La Ley de la IA responde a este desafío mediante diversas herramientas previstas en los artículos 57 al 63, entre las que se destacan los espacios controlados de prueba (*regulatory sandboxes*), mecanismos de gobernanza multinivel, un régimen sancionador riguroso y un enfoque extraterritorial que proyecta los valores europeos a escala global.

Los *regulatory sandboxes* permiten que desarrolladores, empresas y autoridades colaboren en entornos supervisados donde pueden experimentar, evaluar y adaptar sistemas de IA antes de su entrada al mercado. Estos espacios ofrecen un entorno seguro para identificar riesgos y garantizar el cumplimiento de estándares éticos y legales, reduciendo así los costes regulatorios iniciales, especialmente para pymes y startups. Cada Estado miembro debe disponer al menos de uno de estos entornos, y se promueven medidas para facilitar su acceso, como exenciones o ayudas específicas.

En cuanto a su estructura de implementación, la Ley de la IA adopta un modelo de gobernanza multinivel en el Capítulo VII. A nivel europeo, la Comisión Europea actúa como órgano coordinador con facultades para investigar, sancionar e introducir actualizaciones normativas, como la ampliación del Anexo III sobre usos de alto riesgo. La Oficina de Inteligencia Artificial de la UE (*IA Office*), con sede en Bruselas, lidera los esfuerzos de implementación armonizada y promueve la cooperación internacional. Además, el Consejo Europeo de Inteligencia Artificial (*European AI Board*), integrado por representantes de los 27 Estados miembros y presidido por la Comisión, emite directrices y recomendaciones para asegurar la coherencia en la aplicación del reglamento. Se ha también establecido un Foro Consultivo (*Advisory Forum*), que reúne a representantes de la industria, las empresas emergentes, las pymes, la sociedad civil y el mundo académico. Este foro proporciona asesoramiento técnico tanto al Comité Europeo de IA como a la Comisión, garantizando una representación plural y equilibrada en el proceso de implementación del Reglamento.

A nivel nacional, cada Estado miembro debe establecer o designar al menos una autoridad notificante y al menos una autoridad de vigilancia del mercado como autoridades nacionales competentes. Esta autoridad es encargada de supervisar el cumplimiento de la Ley de IA, realizar auditorías e investigaciones locales, así como recibir y tramitar las reclamaciones y denuncias en virtud de los artículos 85 a 87 del Reglamento. Según el calendario

establecido por la Ley de la IA, los Estados miembros deben designar a sus autoridades competentes antes del 2 de agosto de 2025, pero hay instituciones non-profit que mantienen una lista consolidada de las autoridades designadas, que se actualiza de manera continua.²²

Además de la estructura gubernativa, para asegurar el cumplimiento efectivo del reglamento, la Ley de la IA establece un régimen de sanciones robusto. Las infracciones más graves, como el uso de sistemas de IA prohibidos, pueden acarrear multas de hasta 35 millones de euros o el 7 % del volumen de negocios anual global. Otras violaciones, como el incumplimiento de obligaciones generales o la provisión de información falsa, pueden ser sancionadas con hasta 15 millones o el 3 % y 7,5 millones o el 1 %, respectivamente. Aunque se aplican los mismos porcentajes a pymes y startups, se establece un umbral inferior para proteger su viabilidad económica. Además, los ciudadanos y entidades jurídicas tienen el derecho de presentar quejas ante las autoridades competentes y de obtener explicaciones sobre decisiones automatizadas.

Finalmente, la Ley de la IA posee una marcada dimensión extraterritorial. Inspirada en el conocido “efecto Bruselas”²³ del GDPR, cuyas disposiciones sobre transferencias internacionales se han convertido en un vehículo clave para proyectar los derechos fundamentales de la UE a nivel global e influir directamente en la evolución de la legislación internacional sobre protección de datos²⁴, la Ley de la IA europea también extiende su alcance más allá de las fronteras europeas. Así, sus normas se aplican a sistemas desarrollados fuera del territorio de la Unión, como en China, Brasil o Estados Unidos, siempre que estén destinados al mercado europeo o impliquen el procesamiento de datos de ciudadanos europeos.

En conjunto, estas disposiciones de la Ley de la IA europea sugieren que la regulación no tiene por qué ser necesariamente enemiga de la innovación. Más bien, la normativa europea busca ofrecer un marco que proteja a los ciudadanos, impulse el desarrollo tecnológico y refleje los valores de la UE en la gobernanza global de la inteligencia artificial.

22 Artificial Intelligence Act. (n.d.). *National implementation plans*. Retrieved May 20, 2025, from <https://artificialintelligenceact.eu/national-implementation-plans/>.

23 Bradford, A. (2012). *The Brussels effect*. *Northwestern University Law Review*, 107, 1–68.

24 Beltrame de Moura, A., & Henrique Ceron Trevisol, G. (2024). The protection of fundamental rights in international data transfers from the Schrems II Case. In M. Torres Jarrín (Ed.), *Diplomacy and digital age: Science, technology and global digital governance* (p. 105). Peter Lang.

Inteligencia Artificial centrada en el ser humano en los países nórdicos

Anna Virkama²⁵

European Institute of International Studies (Suecia)

1. Introducción

La gobernanza de la Inteligencia Artificial (IA) está cobrando impulso mundial a medida que países y gobiernos compiten por el liderazgo en el establecimiento de normas, la ética y la regulación de la IA. Organizaciones internacionales y multilaterales como la UNESCO, la OCDE y el Banco Mundial se han interesado por la forma en que la Inteligencia Artificial (IA) está transformando profundamente las sociedades y han formulado sus propios marcos normativos y directrices.

Los gobiernos de todo el mundo están buscando diferentes formas de impulsar sus capacidades de IA para el uso del sector público y privado, pero al mismo tiempo surgen voces de preocupación sobre cómo gestionar todas las transformaciones económicas y socioculturales que la IA está generando. Es evidente la necesidad de una estrecha coordinación y colaboración de todos los sectores, públicos y privados, así como la necesidad de debatir y reflexionar sobre la gobernanza ética de la IA.

Como señaló el filósofo sueco Nick Bostrom (2014) hace más de una década, la IA puede ser simultáneamente el mayor riesgo pero también la mejor opción para que la humanidad alcance su pleno potencial. En su libro *Superinteligencia*, sostenía que la mayoría de los grandes retos de nuestra era pueden abordarse con una inteligencia que supere las capacidades cognitivas humanas. Sin embargo, el uso de la IA en todo su potencial para la humanidad conlleva un riesgo existencial que debe gestionarse mediante la gobernanza y la

²⁵ Anna Virkama, MSc, Doctoranda Investigadora del European Institute of International Studies (Suecia), Investigadora Asociada, Instituto de Estudios Europeos y Derechos Humanos, Centro Europeo de Estudios Internacionales, Universidad Pontificia de Salamanca.

ingeniería de seguridad (Bostrom, 2014). Para los gobiernos, la IA encierra la gran promesa de mejorar la eficiencia, la falta de recursos y la falta de competencia en el sector público (Toll & al., 2020).

Las reglas de la IA no nacen en el vacío, sino que tienden a reflejar los valores generales, las normas, las prioridades éticas y el entorno cultural de las sociedades donde se formulan. Con esta premisa en mente, este documento explora la IA centrada en el ser humano en los países nórdicos (Dinamarca, Finlandia, Islandia, Noruega y Suecia) examinando sus Estrategias Nacionales de IA y sus iniciativas públicas para el uso de la IA, y se pregunta cómo el enfoque nórdico de la IA centrada en el ser humano beneficia y desafía a los estados de bienestar nórdicos. ¿Cómo se reflejan los valores nórdicos en la forma en que estas sociedades articulan el uso ético de la IA? El artículo aplica un punto de vista cultural y sociológico a la gobernanza de la IA, y señala que las estrategias de gobernanza de la IA reproducen y cuestionan los valores sociales existentes.

En este artículo se analiza el denominado “enfoque centrado en el ser humano” de la IA en el contexto de las sociedades nórdicas, que suelen considerarse pioneras en la gobernanza de la IA (Silo AI, 2021). Los valores y normas que comparten los países nórdicos, como la democracia, la igualdad, el bienestar social y la educación, se analizan en el contexto de la creciente importancia que se concede al uso de la IA.

2. Búsqueda de una IA centrada en el ser humano en las sociedades nórdicas

Antes de analizar el enfoque nórdico de la gobernanza de la IA, echemos un vistazo a algunas de las características de las sociedades nórdicas, porque la forma en que las personas y las sociedades utilizan las nuevas tecnologías y la digitalización, y las normas culturales, valores, instituciones y creencias influyen en la adopción de la IA. Aunque cada país nórdico es independiente y tiene su propia historia y legislación, existen algunos valores compartidos que pueden describirse como “valores nórdicos”. Después de todo, cinco países nórdicos -Dinamarca, Finlandia, Islandia, Noruega y Suecia- comparten una base de valores similar en lo que respecta a la relación entre el Estado y la ciudadanía, y su enfoque también es similar al de otros países miembros de la Unión Europea en general (a pesar de que Islandia y Noruega no son Estados miembros de la UE), que hacen hincapié en la democracia y en los marcos jurídicos centrados en el ser humano y basados en los derechos.

Los países nórdicos son conocidos por ser sociedades de bienestar con servicios que el Estado ofrece a los ciudadanos a cambio del dinero de sus impuestos. El sector público atiende las necesidades de los ciudadanos en sectores como la sanidad, la educación, la atención a la tercera edad, las bajas por jubilación y enfermedad, etcétera. Culturalmente se consideran

“países de baja jerarquía”, con una relación relativamente directa entre los ciudadanos y las entidades estatales. La confianza, la igualdad, la innovación, la competitividad, la economía basada en el conocimiento y la sostenibilidad son denominadores comunes de los países nórdicos. Los países nórdicos son relativamente homogéneos culturalmente, y las lenguas que se hablan en ellos pertenecen a pequeños grupos lingüísticos.

Los ciudadanos nórdicos suelen expresar un alto grado de confianza en sus instituciones y manifiestan niveles de satisfacción muy elevados. Año tras año, Finlandia se sitúa a la cabeza del índice de felicidad de las Naciones Unidas en el Informe Mundial sobre la Felicidad, destacando como el país más feliz del mundo, seguido de otros países nórdicos. La palabra “felicidad” puede resultar un poco engañosa si se interpreta como una emoción o un sentimiento, pero los resultados de la encuesta de este índice sirven para demostrar que los finlandeses están globalmente contentos con la calidad de vida que consiguen gracias a los distintos servicios que presta la sociedad.

Las sociedades nórdicas también son conocidas por su búsqueda de la igualdad y la cohesión social en todos los ámbitos de la sociedad. La educación pública de alta calidad y los sistemas sanitarios universalistas pretenden garantizar la igualdad de oportunidades para todos los niños, de todos los orígenes sociales. Las guarderías públicas y los largos permisos parentales han contribuido a sostener a las familias y han permitido a las mujeres labrarse una carrera profesional. En cuanto a los resultados educativos, Finlandia ocupa desde hace años el primer puesto en muchas clasificaciones internacionales. Lo mismo ocurre con los sistemas de asistencia sanitaria y seguridad social.

Aunque las sociedades nórdicas tienen en común una amplia participación cívica centrada en el bienestar de los ciudadanos, no son entidades estáticas, como señalan Huby et al. (2024): han sido moldeadas y transformadas por influencias globales. Mencionan la década de 1980 como punto de inflexión para que las sociedades nórdicas se unieran a otras economías europeas en una orientación hacia el liberalismo de mercado, seguida de una reforma conocida como Nueva Gestión Pública (NGP), que combinada con las estructuras de bienestar existentes “produjo acuerdos complejos que implicaban a ciudadanos, agentes públicos, privados y del tercer sector en la coproducción del Estado” (Huby et al., 2024, véase también Osborne, 2018; Pedersen y Kuhnle, 2017). El papel de la digitalización en este contexto puede ser doble, ya que puede simplificar o aumentar la complejidad.

La competitividad y la innovación también caracterizan a estas pequeñas naciones, que a pesar de su población relativamente pequeña, tienen industrias muy competitivas y son mundialmente conocidas por sus empresas innovadoras y a menudo digitalmente avanzadas. Dinamarca, Suecia y Finlandia son los países más innovadores de la Unión Europea, según

el Cuadro de Indicadores de Innovación de la Comisión Europea (2024). Las sociedades nórdicas también son conocidas por su fuerte inversión en la construcción de economías del conocimiento, un factor que, sin duda, ha contribuido al crecimiento del sector digital.

Por último, los países nórdicos han sido durante mucho tiempo pioneros en conciencia medioambiental y pioneros en sostenibilidad. El creciente consumo energético de las supercomputadoras y la necesidad de minerales de tierras raras y otros recursos escasos constituyen una grave preocupación, especialmente para los países del Sur Global que proporcionan esas materias primas (Galofari, Riordan y Torres Jarrín, 2023:5-6).

Por tanto, no es de extrañar que, en el caso de los Estados nórdicos, las estrategias nacionales para el uso de la IA reflejen estos valores comunes, que se analizarán, país por país, más adelante en este documento. Pero antes, se examinan las diferencias y similitudes de las nociones de IA ética y centrada en el ser humano.

3. El enfoque nórdico de la gobernanza digital

La IA centrada en el ser humano refleja la idea de Bostrom (2014: 254) de que la superinteligencia solo debe desarrollarse si puede utilizarse en beneficio de toda la humanidad. Toll et al. (2020) han distinguido cuatro valores de gobernanza digital, basados en el modelo desarrollado por Rose et al. (2015, p. 542), que describen como “...*de origen escandinavo y se corresponde bien con la cultura sueca y los sistemas de bienestar del sector público. El modelo también representa las expectativas y responsabilidades de las organizaciones gubernamentales escandinavas.*” Según este modelo, existen cuatro ideales de valor para la gestión de la gobernanza electrónica: *eficiencia, servicio, profesionalidad y compromiso.*

Al igual que en muchos otros países del mundo, los gobiernos nórdicos han comprendido el potencial de la IA para crear cambios positivos en la sociedad. Pero, al mismo tiempo, existe cierto “revuelo” en torno a la IA y, por lo tanto, es necesario ver más allá de ese “revuelo” y considerar las repercusiones en la vida real y entablar un debate sobre las consecuencias éticas (Leikas, 2019). Sin embargo, la IA centrada en el ser humano no solo tiene que ver con las normativas y restricciones, sino también con cómo utilizar la IA de forma proactiva en beneficio de los ciudadanos.

4. Estrategias nacionales de IA y ejemplos de avances en IA en los países nórdicos

En 2024, todas las sociedades nórdicas, excepto Islandia, habían publicado sus Estrategias Nacionales de Inteligencia Artificial (NAIS). Para este trabajo, se analizaron las versiones

oficiales en inglés de las NAIS nórdicas a partir de la comparativa publicada por NordForsk. Las NAIS no son documentos jurídicamente vinculantes, sino directrices de política.

Finlandia aspira a convertirse en un centro de datos de IA en los países nórdicos, y líder en el desarrollo de IA. Finlandia ocupa el segundo lugar en el Índice de Calidad de Vida Digital y cuenta con una Estrategia Nacional de IA desde 2017 (actualizada en 2018 y 2020). El clima frío del norte de Finlandia se considera una ventaja natural para albergar una supercomputadora de pre-exaescala para el procesamiento de datos, LUMI (“lumi” significa “nieve” en finlandés). Lo que hace interesante a esta supercomputadora es su integración en el entorno. LUMI se encuentra en un centro de datos de la ciudad de Kajaani, donde el clima es frío. El centro de datos de CSC se enorgullece de ser uno de los más ecoeficientes del mundo, ya que el calor extra de LUMI sirve para calentar cientos de hogares al año (Lumi Supercomputer, 2025).

Finlandia también ha adoptado un sistema de IA centrado en el ser humano denominado AuroraAI, lanzado por el Ministerio de Hacienda en 2020 y disponible para personas y organizaciones a partir de 2022 (Ministerio de Hacienda finlandés, 2020). Ayuda a los ciudadanos a navegar por el sistema de los diferentes servicios que el Estado ofrece a los ciudadanos en función de los diferentes acontecimientos y necesidades de la vida registrados automáticamente. Finlandia también ha expresado su ambición de integrar la IA en los servicios de seguridad ciudadana.

Suecia tiene su primera estrategia de IA desde 2018. Hace hincapié en la educación, la innovación y la investigación. La última estrategia data de marzo de 2024 y menciona los retos a los que se enfrenta la sociedad sueca, como el cambio climático, el aumento de las necesidades energéticas, el envejecimiento de la población y la amenaza a la seguridad y la democracia del país. AI Sweden es un centro nacional de IA aplicada que pretende, con sus más de 160 socios, acelerar el uso de la IA en beneficio de la sociedad y para la competitividad (AI Sweden, 2025).

Suecia también tiene su supercomputadora MIMER, que, al igual que LUMI en Finlandia, también está cofinanciada por la Unión Europea. En cuanto a la innovación en IA social, los robots Furhat son un ejemplo interesante de Suecia. Los Furhat son robots humanoides conversacionales. Estos robots se han utilizado, por ejemplo, en sociología médica, donde ayudan a recoger las opiniones de los pacientes. También se han utilizado para la terapia del autismo en programas de entrenamiento asistidos por robots y adaptados a niños autistas, y para el aprendizaje personalizado de idiomas para inmigrantes (Furhat Robotics, 2025).

Noruega

Noruega presentó por primera vez su Estrategia Nacional de IA en 2020. Esta estrategia hacía hincapié en la ética y la IA centrada en el ser humano, la prevención de daños y el uso de la IA únicamente para los fines previstos. También hacía hincapié en la protección de la privacidad y los datos, la transparencia, la inclusión y la igualdad, la responsabilidad y los beneficios para el medio ambiente. También se mencionan el crecimiento económico y los beneficios para la sociedad.

En el caso de Noruega, se hace hincapié en el poder de liderar el potencial de innovación en la aplicación de la IA. Noruega pretende reforzar el intercambio entre los sectores público y privado para impulsar la innovación y crear valor. Noruega busca reforzar el uso de la IA especialmente en las áreas en las que ya se ha posicionado como líder en negocios e investigación, como la salud, el petróleo y el gas, la energía, las industrias marina y marítima y el sector público.

A diferencia de Suecia, Noruega no cuenta con iniciativas públicas para financiar, por ejemplo, nuevas empresas de IA, pero sí con un plan más amplio llamado SkatteFUNN que promueve la reducción de impuestos para las empresas en sus costos de investigación y desarrollo. A través de este fondo, las empresas pueden solicitar recibir hasta 50 horas de asistencia de un investigador para resolver un problema, lo que puede ser una ayuda importante para las pequeñas empresas.

El uso de la IA en el sector público también es importante para los noruegos y el gobierno cree en el potencial de la digitalización y el uso de la IA para racionalizar y crear mejores servicios para los ciudadanos utilizando la IA. Al mismo tiempo, les preocupa la privacidad de los datos, la parcialidad y la transparencia, por ejemplo, el uso de algoritmos para identificar a los ciudadanos que podrían considerarse delincuentes potenciales. En el sector público, Noruega ha utilizado la IA con fines de verificación de residencia. En 2018 hubo un proyecto en el que se utilizó IA para verificar la residencia real de estudiantes que solicitaban préstamos estudiantiles. Noruega reconoce la necesidad de competencias digitales, innovación, infraestructura y ciberseguridad. También insta a equilibrar los avances tecnológicos con consideraciones éticas, y estudia el uso de un “sandbox” regulatorio para las nuevas innovaciones.

Dinamarca

Dinamarca cuenta con una estrategia nacional de IA a partir de 2019, que sirve de hoja de ruta para el uso sostenible de la IA con énfasis en los principios éticos y las competencias

de la IA. Su objetivo es beneficiar a los ciudadanos, las empresas y la sociedad en general. Los principios éticos y las competencias digitales también ocupan un lugar destacado en la agenda danesa. Existe una preocupación por el uso de la IA, especialmente cuando están surgiendo nuevas formas de toma de decisiones. La estrategia también reconoce el problema de construir soluciones de IA basadas en el reconocimiento de lenguas, especialmente cuando se trata de una lengua perteneciente a un grupo lingüístico pequeño. Para ello, una iniciativa clave de esta estrategia es un recurso lingüístico danés compartido. Se ha desarrollado y es de uso gratuito.

Cabe mencionar igualmente algunas iniciativas interesantes para impulsar el sector público en los ámbitos de la salud, el bienestar social y el empleo. Según la Agencia Danesa para la Gobernanza Digital, el sector público digital danés fue nombrado el mejor del mundo por cuarto año consecutivo en 2024 en la encuesta sobre administración electrónica de la ONU. Un ejemplo interesante para mencionar es el uso de la IA en el sector asistencial en Dinamarca; se trata de la “enfermera de guardia”, un sistema que se ha utilizado en el cuidado de ancianos para observar y alertar al personal en caso de emergencia.

Islandia

Islandia creó su primera Estrategia Nacional de IA en 2021. Tiene una importante base ética: la IA debe aportar beneficios a toda la sociedad. La estrategia define un marco para evaluar las implicaciones éticas y morales en las decisiones promovidas por la IA.

Como país pequeño, es importante que Islandia promueva la cooperación internacional en el uso de la IA y participe activamente en el debate sobre regulación y normas, y que siga implicada en la cooperación nórdico-báltica, así como en los proyectos del Consejo de Ministros para promover los servicios digitales en los servicios transfronterizos y transversales. La estrategia reconoce los retos de seguridad. Islandia quiere promover la educación y la experiencia en IA, basándose en los conocimientos locales, y adaptar las tecnologías de IA al servicio de la industria islandesa. Como país con infraestructuras avanzadas, Islandia ha invertido en energía verde y tiene un bajo nivel de complejidad para diseñar soluciones.

El gobierno islandés ha designado una importante financiación para ampliar la infraestructura y los servicios digitales, incluido el desarrollo de la computación en nube, la mejora de los servicios públicos y la mejora de la competitividad. Como en el caso de Dinamarca, la lengua islandesa también pertenece a la familia de las lenguas pequeñas y fue la segunda lengua utilizada en ChatGPT después del inglés.

5. Conclusiones

A pesar de ser países relativamente pequeños en términos de población, los países nórdicos tienen planes ambiciosos para convertirse en líderes mundiales y centros neurálgicos del desarrollo de la IA, aunque todavía están por detrás de países como Canadá, Reino Unido, Singapur y Estados Unidos en cuanto a su adopción. Este artículo ha esbozado algunos de los principios rectores de las sociedades nórdicas en términos de valores, normas y desarrollo social, que son la columna vertebral de la innovación en IA centrada en el ser humano.

La Inteligencia Artificial puede ser tanto una oportunidad como una amenaza para el modelo nórdico de sociedad del bienestar, tal y como se expresa en las estrategias nacionales de IA y en las iniciativas clave para el desarrollo de la IA de los países nórdicos. Mientras que los países nórdicos en general reconocen los riesgos de la IA para la democracia, la igualdad de derechos, la seguridad nacional, el medio ambiente, etc., también han expresado su interés en buscar activamente la integración de la IA en el sector público en beneficio de los ciudadanos.

No cabe duda de que es necesario un control ético constante, como señalan Leikas et al (2022). Sin embargo, la IA centrada en el ser humano en el contexto nórdico significa que los países no solo regulan y dirigen, sino que también educan y capacitan a los ciudadanos sobre sus derechos y, al mismo tiempo, refuerzan la agencia humana. El uso de la IA centrada en el ser humano también puede significar, por ejemplo, la prestación de servicios sociales hechos a la medida que consideren la situación integral del individuo, como en el ejemplo del servicio AuroraAI en Finlandia.

Aunque cada país nórdico tiene su propia estrategia nacional de IA, los países nórdicos comparten valores similares en sus iniciativas de gobernanza de la IA, y cooperan no sólo a nivel estatal sino también a través de diferentes iniciativas de investigación, innovación y negocio. El futuro mostrará cuántas innovaciones nórdicas de IA centrada en el ser humano serán escalables a otros países del mundo.

Referencias

Agencia para el Gobierno Digital (2024) 'Denmark Tops UN E-Government Survey'. <https://en.digst.dk/news/news-archive/2024/oktober/denmark-tops-un-e-government-survey/>

AI Sweden (2025) Centro Nacional de IA Aplicada. <https://www.ai.se/en>

Bostrom, N. (2014) *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, Oxford.

Índice de Calidad de Vida Digital (2024). Surfshark. <https://surfshark.com/research/dqi>

Furhat Robotics (2025) <https://www.furhatrobotics.com/>

Cuadro de Indicadores de la Unión por la Innovación 2024 (EIS) de la Comisión Europea. Bruselas: Unión Europea 2024. Web

Garofali, A., Riordan, S. & Torres Jarrín, M. (2023) 'The Environmental and Ethical Challenges of Artificial Intelligence'. Towards Reformed Multilateralism: Transforming Global Institutions and Frameworks. Task Force 7: T20 Policy Brief.

Huby, G.; Roste, R. & Reutter, L. M. (2024) 'Digitalization -Saving or Undermining the Nordic Welfare State?'. Nordisk välfärdsforskning. 120-126. 2024. DOI: <https://doi.org/10.18261/nwr.9.2.1>

Leikas, J., Johri, A., Latvanen, M., Wessberg, N. & Hahto, A. (2022) 'Governing Ethical AI Transformation: A Case Study of AuroraAI'. AI For Human Learning and Behavior Change. Volume 5 -2022. DOI: <https://doi.org/10.3389/frai.2022.836557>

Lumi Supercomputer (2025) 'About LUMI' <https://www.lumi-supercomputer.eu/about-lumi/>

Ministry of Finance (2020) 'The AuroraAI national artificial intelligence programme begins – with the aim of using artificial intelligence to bring people and services together in a better way'. Valtioneuvosto, Finnish Government. <http://tiny.cc/1dil001>

Osborne, S. P. (2018). From public service-dominant logic to public service logic: are public service organizations capable of co-production and value co-creation? *Public Management Review*, 20(2), 225–231. <https://doi.org/10.1080/14719037.2017.1350461>

Pedersen, A. W., & Kuhnle, S. (2017). The Nordic welfare state model. In O.P. Knutsen (Ed.), *The Nordic models in political science: Challenged, but still viable?* (pp. 249–272). Fagbokforlaget.

RISE (2025) 'MIMER: Sweden's AI Factory Powers Up - Supercomputing Race to Tomorrow's AI Solutions.' The Research Institute of Sweden. <https://www.ri.se/en/mimer-swedens-ai-factory-powers-up-supercomputing-race-to-tomorrows-ai-solutions>

Robinson, C. (2020) 'Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI)'. *Technology in Society*, Vol. 63, Article 101421. DOI: <https://doi.org/10.1016/j.techsoc.2020.101421>

Silo AI. (2021). The Nordic state of AI, the 2021 report. Silo AI. <https://silo.ai/nordic-state-of-ai/> (Accessed September 2022).

Strange, M. & Tucker, J. (2024). A paradigm shift in plain sight? AI and the future of health-care in the Nordic states. *Nordisk välfärdsforskning | Nordic Welfare Research*, 9(2), 168–179. DOI: <https://doi.org/10.18261/nwr.9.2.5>

Toll, D., Lindgren, I., Melin, U. & Madsen, C. (2020) Values, Benefits, Considerations and Risks of AI in Government: A Study of AI Policy Documents in Sweden. *Journal of Democracy, JeDEM* 12(1), 40-60. DOI:10.29379/jedem.v12i1.593

Informe Mundial de la Felicidad (2025) <https://worldhappiness.report/>

La regulación y la gobernanza global de la Inteligencia Artificial

Emb. Dr. Antonio Núñez y García-Saúco
*Miembro del European Institute
of International Studies (Suecia).*

Resumen: *Ante la urgente necesidad de regular la Inteligencia Artificial (IA) son constatables actitudes reticentes y negativas que cubren un amplio espectro de resistencia u oposición alegando argumentos de distinta naturaleza, lo que dificulta la tarea regulatoria y la gobernanza de aquella. Una precisión conceptual podría quizás contribuir a aclarar en parte este debate. Cuatro serían los conceptos para definir y aclarar: globalidad, inteligencia artificial, regulación y gobernanza. De sus definiciones y funciones pueden extraerse algunas conclusiones útiles respecto a la interrelación práctica entre regulación y gobernanza global de la IA.*

I

La globalidad. - El concepto de *globalidad* suele asociarse a los de *globalización* y *globalismo*.

La *globalidad* es el resultado de la *globalización*, siendo ésta un proceso de continua y creciente interdependencia, a nivel mundial, en las actividades que llevan a cabo los distintos actores que intervienen en los diferentes ámbitos de la esfera internacional. La *globalidad* sería, por así decir, el estado de lo global en un momento dado y, por tanto, una parte fundamental y configurante de un período histórico como el nuestro.

Tres características acompañan la *globalidad* actual.

Primero, es un conjunto integrado y complejo de profunda interdependencia que sobrepasa la mera yuxtaposición de acciones aisladas para constituir un sistema interactivo de influjos mutuos y recíprocos.

Segundo, la *globalidad* puede afectar a todos los ámbitos y aspectos de las actividades, las relaciones y las interrelaciones humanas.

Tercero, en este sistema global se integran lo nacional y lo internacional sin solución de continuidad, desconfigurando la línea divisoria entre lo local y lo global para configurar el orden planetario definitorio de nuestro tiempo (*la aldea global*).

Por su parte, el *globalismo* es la ideología o doctrina que propugna y defiende la *globalización* y la *globalidad*, como objetivos deseables.

En resumen: la *globalización* es el proceso evolutivo que conduce a la *globalidad*, mientras el *globalismo* es la ideología que lo defiende y propugna.

Para comprender en su totalidad el fenómeno de la *globalidad* de nuestro resulta imprescindible conocer cómo y cuándo comenzó el proceso de *globalización* y cuáles fueron sus factores determinantes.

La actual globalización comenzó en la Modernidad con el desplazamiento del método cognitivo. El concepto filosófico de *verdad* quedó postergado por el científico de *hipótesis*, especie de *verdad provisional*, pero útil, mientras la experiencia no anulara su validez y utilidad (empirismo).

El método filosófico, en cuanto analítico-deductivo, no ampliaba el conocimiento: sólo concretizaba verdades previas generales que ya contenían lo concretizado. Si todo hombre es mortal y Sócrates es hombre, es, en consecuencia, mortal. El moderno método científico, por el contrario, será sintético-inductivo: formulará principios o leyes generales a partir de situaciones empíricas concretas, posibilitando ampliar el conocimiento. Había nacido la *Sociedad del Conocimiento*.

Si el empirismo pronto tendería a la ciencia aplicada, el utilitarismo buscaría sus beneficios prácticos. El ya pujante capitalismo industrial no tardaría en forjar una alianza indestructible con la nueva ciencia aplicada dotándola de ingentes recursos. Así aparecerían las revoluciones industriales en las que se apoyaría el proceso de *globalización* hasta nuestros días.

Muchos fueron los descubrimientos e innovaciones que aportaron todas estas revoluciones: la primera (1760-1850) trajo el telégrafo, el ferrocarril y la navegación a vapor; la segunda (1870-1914), junto a la electricidad, el motor de combustión interna (automovilismo) y la producción en masa; la tercera (1950-1980), llamada también primera digital, la informática, los ordenadores, la automatización, la electrónica y las telecomunicaciones, incluido internet; la cuarta y actual: la IA, los big data, el internet de las cosas (*IoT*), la robótica avanzada, la biotecnología, la computación cuántica...

A la *Sociedad del Conocimiento* se superpondría la *Sociedad de la Comunicación y la Información*.

II

La inteligencia artificial. - En sentido formal, la IA puede definirse como el ámbito de la informática orientado a la creación de sistemas capaces de programar en máquinas tareas inteligentes propias de la inteligencia humana. O, si se prefiere, desde el lado inverso: la potencialidad de las máquinas para realizar tareas hasta ahora exclusivas de la inteligencia humana: percibir, aprender, razonar, recordar o decidir.

Si la *globalización* se forjó en las revoluciones científico-industriales, la IA, que también surgió en este marco, como hemos visto, se fraguó a partir de las neurociencias, en concreto, de la neurobiología y de la neuropsicología o psicología cognitiva.

La aspiración a unir el mundo humano con el mundo inanimado ha sido atávica en la historia de los hombres. No es difícil encontrar intentos de ello en la Antigüedad clásica y durante la Edad Media, así como reflexiones filosóficas, ya en la Modernidad, en Descartes, Locke o Leibnitz.

Fue, sin embargo, en nuestro tiempo, cuando se logró proyectar sobre las máquinas la capacidad intelectual humana. Partiendo de los descubrimientos del Premio Nobel español, Ramón y Cajal (1852-1934) sobre la estructura del sistema nervioso y los circuitos de realimentación, McCulloch y Pitts formalizaron (1943) el concepto de *neurona artificial M-P* (de sus iniciales) o unidad lógica de umbral (TLU: *Threthold Logic Unit*). Pocos años después, Norbert Wiener consolidó la cibernética: ciencia que estudia la analogía de comportamiento entre organismos biológicos y máquinas (1948).

Como consecuencia, las redes neurológicas artificiales (RNA) se diseñaron como las neuronas biológicas humanas, con nodos interconectados (*artificial neurons*) que operan, como aquellas, transmitiendo señales entre sí y aprendiendo de forma similar, ajustando sus interconexiones a los datos proporcionados. En su estructura, las RNA son más simples

que las cerebrales y menos intrincadas. En su funcionamiento, la IA copia al cerebro. Éste fortalece las conexiones neuronales en base a las sensaciones que recibe de vivencias y experiencias externas. Las RNA se entrenan con los enormes conjuntos de datos proporcionados, ajustando parámetros internos para mejorar el rendimiento.

Si la neurobiología inspiró el diseño de las RNA, la psicología ofreció el modelo de los procesos de aprendizaje: atención, comprensión, reflexión, pensamiento, memoria... aplicable a las máquinas.

Por ejemplo, el diseño de algoritmos en el aprendizaje automático (*machine learning*) replica el procesamiento de la información en la psicología cognitiva. Estos algoritmos aprenden de los datos suministrados, pudiendo identificarlos, ordenarlos y clasificarlos, elaborar en base a ellos patrones, modelos o pautas, configurar tendencias, hacer predicciones y tomar decisiones. En el *deep learning*, una potente subcategoría dentro del *machine learning*, los algoritmos pueden reconocer imágenes, procesar el lenguaje natural y hacer traducción automática.

En conclusión, la inteligencia artificial se ha consolidado como un elemento central y un factor clave en el entramado tecnológico que impulsa y acelera los procesos de globalización contemporáneos. En estrecha articulación con otras tecnologías, como internet, constituye uno de los vectores más influyentes en la configuración de la globalidad propia de la actual era digital.

III-IV

Regulación y gobernanza. - Aunque similares y utilizados indistintamente, los conceptos de *regulación* y *gobernanza*, a los que habría que añadir un tercero, el de *gobernabilidad*, difieren en significado y extensión y todos presentan implicaciones relevantes para la IA y consecuencias prácticas significativas.

Empecemos por la *regulación*. Puede entenderse por *regulación* el marco compendiador de todas las leyes, normas, reglas y preceptos vinculantes que los órganos competentes, nacionales o supranacionales, dictan para el diseño, desarrollo y uso de la IA.

Se trata, pues, de un marco regulatorio estricto y formal, esto es, de un marco competencial de normas o leyes obligatorias y específicas con incumplimiento sancionable y punible. El objetivo global último de la *regulación* es proteger los derechos, garantizar el uso y disfrute de los beneficios de la IA y evitar y/o minimizar sus riesgos.

La *gobernanza*, por su parte, abarca todo el conjunto de principios, criterios, procesos y mecanismos establecidos o que se puedan establecer para que el diseño, el desarrollo, la ejecución y el uso de sistemas de IA proporcionen seguridad y responsabilidades y se respeten los valores éticos y el bien común. La orientación moral que conlleva el significado hace que el término sea con frecuencia referido como *buena gobernanza*.

La *gobernanza* es claramente, pues, un concepto más amplio y extenso que el de *regulación* que, como acabamos de ver, es más restringido y preciso. La *gobernanza* va más allá de leyes y normas formalmente obligatorias, para incorporar principios o directrices, incluso informales, en orden a reforzar la garantía de los procesos en favor de la ética, la transparencia, la seguridad o la responsabilidad. Al ir más allá de las leyes, también va más allá de los órganos legislativos o formalmente reguladores, pudiéndose incorporar al marco de la *gobernanza*, además de los Estados, múltiples instituciones no gubernamentales de diferente carácter: social, empresarial, científico etc..., bien de índole internacional o nacional, incluso expertos individuales, para recoger e incorporar criterios de orientación, referencias de buenas prácticas o directrices de utilidad. Por último, la orientación finalista de la *gobernanza*, además de la referencia a marcos y estándares, la impulsa a conceder gran importancia a los procesos de supervisión y evaluación para prevenir violaciones de derechos fundamentales y mitigar riesgos.

Conviene dejar bien sentado que *regulación* y *gobernanza* no son conceptos ni contrapuestos ni excluyentes, sino complementarios. La primera se inspira en la segunda y la concretiza, y ambas tienden a ordenar o normatizar la IA: la *regulación*, de forma estricta, mediante leyes; la *gobernanza*, de manera flexible, a través de principios. En muchos casos, la *gobernanza*, como norma ética, precede y conduce a la *regulación* formal.

Finalmente, el concepto de *gobernabilidad* hace referencia a la capacidad de gestionar los sistemas de IA conforme a lo establecido en los marcos de *regulación* y de *gobernanza*. La *gobernabilidad* presentaría, así, una doble funcionalidad: como criterio de orientación y como índice de resultados. Conforme al primer criterio, de índole cualitativa, la *gobernabilidad* podría entenderse como el proceso de adhesión al conjunto de leyes y normas, así como de principios, criterios o pautas con los que regir el desarrollo equitativo y sostenible de la IA en favor del bien común humanitario. Según el segundo criterio, de naturaleza cuantitativa, la *gobernabilidad* se apreciaría por el nivel de efectividad logrado en la aplicación y puesta en práctica de las normas y de los criterios establecidos para alcanzar los objetivos previstos en la *regulación* y la *gobernanza*. En ambos casos, se trataría de un enfoque valorativo, modular o de grado respecto del cómo y del cuánto en el cumplimiento de todo lo establecido en la *regulación* y en la *gobernanza*.



La IA, como el resto de las modernas tecnologías, en las que se integra, ofrece un elenco enorme de beneficios, pero conlleva también un considerable número de riesgos que, por conocidos, no parece necesario enumerar ni detallar.

Ante ello, *la regulación y la gobernanza*, como se acaba de indicar, se orientan a impedir o minimizar riesgos y a facilitar y maximizar beneficios. Esta bifrontalidad ha planteado el dilema de qué opción priorizar.

Los beneficios, debido a las ventajas que ofrecen, tienden a ser socialmente aceptados. La regulación, en esta situación, apenas tendría otro papel que tratar de posibilitar el disfrute natural y pacífico de estas ventajas. En consecuencia, la complaciente recepción de los beneficios por la sociedad, unida al alto nivel de especialización de las nuevas tecnologías, propiciaron la inhibición de los poderes públicos, relegando la necesidad de regulación.

Sólo cuando se evidenciaron los riesgos (violación de la intimidad personal, ciberataques...) apareció la urgencia regulatoria y, como reacción, las posiciones contrarias o reticentes a ella.

Este asunto adquirió relevancia cuando las máquinas alcanzaron capacidad autodidacta o de aprendizaje automático (*Machine Self-Teaching*), esto es, cuando se demostró que aquellas podían aprender solas, sin la tutela o supervisión de los hombres, exclusivamente en función de la información procurada por los datos proporcionados.

Frente a esta situación y ante la eventualidad de ser reguladas, algunas grandes empresas tecnológicas emprendieron un proceso de autoregulación como parte de su propio código de conducta empresarial, mostrándose dispuestas a corregir desviaciones en sus algoritmos (*algorithmic biases*). Pero, no sólo esto. Se propuso una función relevante a los algoritmos en la regulación. Ajenos a emociones, los algoritmos, como es sabido, no sólo pueden analizar, de forma objetiva y rápida, enormes cantidades de datos también pueden detectar riesgos propios y autocorregirse. Fue esta capacidad la que movió a considerar atribuirles capacidad regulatoria.

Pero, los algoritmos, además de que pueden ser ciegos y opacos, son jurídicamente irresponsables y los empresarios dueños de los algoritmos no pueden garantizar que en un conflicto de intereses, prevalecerá la objetividad y la equidad sobre el propio perjuicio. Nadie puede ser juez y parte. Los empresarios y los algoritmos pueden contribuir grandemente a la regulación, pero ésta no puede depositarse ni quedar solo en sus manos.

La oposición a la autoregulación y a los algoritmos reguladores no cerró el debate sobre la regulación.

Uno de los argumentos más repetidos frente a la regulación apeló a su inoportunidad. La IA de nuestros días, a pesar de sus grandes logros, está todavía a enorme distancia de haber alcanzado la plenitud de sus capacidades y potencialidades tanto en términos de beneficios como de riesgos. En realidad, la IA se encontraría a niveles de desarrollo todavía muy básicos y claramente insuficientes para ser regulada con eficacia, corriendo el peligro de que cualquier intento de regulación sea repetida e incesantemente anulado por los hechos. Sería excesivamente arriesgado y peligroso tratar de regular fenómenos con efectos todavía desconocidos.

Un razonamiento similar acudió a los riesgos implícitos que comporta toda regulación y a las consecuencias indeseadas o efectos imprevistos de cualquier diseño regulatorio de los algoritmos que se basara en presupuestos erróneos o equivocados, donde los perjuicios serían desproporcionados y los beneficios nulos.

Otro de los alegatos más comunes alude a la inutilidad. La IA, como una de las tecnologías modernas, se desarrolla a un ritmo vertiginoso, claramente inalcanzable por ningún regulador en tiempo útil. Es, por ende, inútil cualquier regulación de la IA porque nunca podrá acompasarse al ritmo del progreso tecnológico y resultará inservible por tardía. La argumentación recuerda la clásica aporía del filósofo griego Parménides sobre Aquiles, “el de los alados pies”, y la tortuga.

Vinculado a este último alegato, está el que considera que toda regulación, además de tardía, sería retardataria. Toda regulación impone condiciones y requisitos que, de forma inevitable, tienden a limitar la libertad y la capacidad innovadora. Por añadidura, la imposición de nuevas condiciones y requisitos suele obligar a gastos adicionales y a la prolongación de los procesos, haciendo a las empresas menos competitivas y más costosa y lenta la investigación.

Por último, no podía faltar la invocación técnica. La IA está tan íntimamente vinculada a otras tecnologías y es tan transversal en su funcionamiento y operatividad que no puede ser regulada aisladamente sin que se regulen simultáneamente todas las otras tecnologías, lo que convierte la pretendida regulación del IA en una tarea ciclópea, imposible de abordar, tanto por su dimensión, como por la constante e incesante innovación que experimentan la propia IA y las múltiples tecnologías con las se imbrica. En resumen, ¿es posible regular la IA?

La regulación de la IA debe ser posible porque es necesaria. Regular solo la IA puede que no baste. Pero, no regularla no es una opción.

Comencemos por distribuir la regulación en las causas y la regulación de los efectos.

La regulación en las causas puede entenderse como regulación en origen, es decir, la regulación de los algoritmos. La denominamos regulación en origen porque los algoritmos son el cerebro de donde emanan las capacidades, las instrucciones y el potencial operativo de cada sistema de IA. Se trata, en consecuencia, de una regulación de carácter técnico sobre la que la informática ha ido acumulando amplia experiencia y un conocimiento valioso y útil. Esta regulación ya se ha demostrado técnicamente posible y se ha articulado en una serie de mecanismos de inspección, intervención y control, pruebas, auditorías, supervisiones y seguimiento que actúan sobre el diseño, los datos, los componentes, la composición, la estructura y el funcionamiento de los algoritmos así como de dispositivos y sistemas rápidos de corrección, desactivación y reversión al estado anterior. Su finalidad es evitar sesgos y desvíos en los algoritmos que puedan lesionar derechos o intereses legítimos.

La regulación de los efectos plantea cuestiones más complicadas y diversas.

La primera de ella está relacionada con la dimensión *global* como concepto. Ya quedó expuesto que la *globalidad* posee la capacidad de influenciar todos los ámbitos y aspectos relativos a las actividades, relaciones e interrelaciones humanas, tanto a nivel nacional como internacional. La pregunta inmediata, apuntada con anterioridad, es cómo regular en su totalidad los innumerables efectos que en todos los sectores (social, cultural, formativo, informativo, político, legal, económico, laboral, recreativo etc.) pueden derivarse de la IA, directamente y a través de las numerosas tecnologías con las que se entrelaza. En términos más directos: ¿es posible regular la *globalidad* de los efectos asociados a la IA y a las modernas tecnologías con las que opera?

La segunda cuestión está relacionada con la posible respuesta a esta cuestión y a cómo implementarla. Ante la envergadura de la tarea, se ha sugerido un orden de prelación en la regulación. El tema radica en cómo establecerlo y con qué criterios.

Para empezar, como ya quedó indicado, de los dos objetivos básicos que pretende la regulación, el orientado a prevenir y mitigar los riesgos se presenta como más urgente e importante que el de maximizar los beneficios. Sin un umbral mínimo de seguridad es imposible disfrutar los beneficios.

Adicionalmente, es necesario plantear la relación alternativa o sincrónica entre *regulación* y *gobernanza* y, más en concreto, la forma de ordenar normativamente los diferentes

sectores o ámbitos donde pueden llegar los efectos de la IA. Es decir, cuáles de ellos, por su importancia, exigen una *regulación* y cuáles pueden ser sometidos a *gobernanza*. Ambos aspectos se han intentado clarificar en función de un doble criterio: el valor de lo amenazado y el nivel del riesgo. En función de ello, parece haberse logrado un cierto consenso. Qué valores exigen *regulación* estricta:

- los derechos fundamentales de las personas,
- las infraestructuras críticas (ciberseguridad) y,
- aquellos derechos o intereses importantes y legítimos susceptibles de sufrir daño irreversible.

A esto se asimilarían aquellas decisiones automatizadas sin posibilidad de revisión o apelación y aquellas otras que impidan la transparencia (cajas negras). Por el contrario, parece más apropiado acudir a la *gobernanza* o regulación flexible, es decir, no por leyes, sino por principios, en situaciones o casos de:

- bajo riesgo inmediato para derechos fundamentales o infraestructuras básicas,
- gran diversidad cultural, multiplicidad étnica o pluralidad de valores éticos,
- tecnologías emergentes o experimentación innovativa.

En resumen, la *gobernanza* vendría a ordenar y normatizar aquellas situaciones donde una *regulación* estricta no es imperativa, urgente, posible o conveniente. En estos supuestos, la *gobernanza* aportaría los principios, criterios, estándares, pautas y buenas prácticas de orientación ética, equitativa y sostenible, que deben, en cualquier caso, inspirar toda *regulación* y aplicarse en su ausencia.

Finalmente, es preciso tener en cuenta que la *globalidad* requiere y exige un marco global con una doble exigencia de participación y negociación.

En este sentido, tanto la *regulación* como la *gobernanza* exigen la participación representativa y colaboradora de todos los actores clave implicados en las interacciones de la IA, a saber: Estados nacionales, Organizaciones internacionales, gubernamentales y no gubernamentales, en cuanto sociedad civil, empresas multinacionales, especialmente las tecnológicas, así como expertos y asociaciones de expertos.

En cuanto al marco de negociación, además de la responsabilidad natural de los Estados y del loable y meritorio esfuerzo de numerosas Organizaciones internacionales de carácter regional, el marco de las Naciones Unidas se presenta como el más adecuado para el nivel global que exigen tanto el consenso sobre los valores de la *buena gobernanza* como el progresivo avance desde ella hacia la cada vez más necesaria *regulación*.

Iniciativas gubernamentales sobre la gobernanza ética de la IA en América y Europa: Canadá, Estados Unidos, Brasil, Reino Unido y Francia.

Dr. Arturo Flores López

*Investigador Asociado del European Institute
of International Studies (Suecia)*

Introducción

La Inteligencia Artificial (IA) ha surgido como un dominio crítico en la política y la gobernanza contemporáneas, planteando urgentes desafíos éticos, legales y sociotécnicos a través de las fronteras nacionales. Los gobiernos de todo el mundo tienen la tarea de diseñar mecanismos normativos e institucionales que faciliten la innovación y mitiguen riesgos como la parcialidad, la discriminación, las violaciones de la privacidad y las lagunas en la rendición de cuentas (Calo, 2017; Jobin et al., 2019). A medida que los sistemas de IA influyen cada vez más en la toma de decisiones en áreas sensibles -que van desde la atención médica y el empleo hasta la vigilancia policial y las finanzas-, la gobernanza ética se ha convertido en una piedra angular de las estrategias nacionales e internacionales (Binns, 2018; Burrell, 2016; Bryson, 2018).

El presente documento ofrece un análisis comparativo de cinco iniciativas nacionales -Canadá, Estados Unidos, Brasil, Reino Unido y Francia-, cada una de las cuales presenta un modelo distintivo de gobernanza ética de la IA configurado por sus instituciones políticas, culturas reguladoras y prioridades estratégicas. A pesar del amplio consenso sobre principios fundamentales como la transparencia, la equidad, la responsabilidad y el diseño centrado en el ser humano (Comisión Europea, 2021; The White House, 2023), la aplicación de estos principios difiere sustancialmente entre jurisdicciones (Taeihagh, 2021). Estas diferencias

reflejan una tensión más amplia entre la adaptabilidad tecnológica, la aplicabilidad de la normativa y la inclusión de las partes interesadas (Zarsky, 2016; Scherer, 2016).

Los casos seleccionados representan diversos enfoques de diseño institucional, desde estrategias centralizadas integradas en marcos legislativos más amplios -como se observa en Francia y Canadá- hasta modelos más descentralizados e impulsados por la innovación ejemplificados por Estados Unidos y el Reino Unido (Gobierno de Canadá, 2023b; DSIT, 2024; Francia - Estrategia Nacional para la IA, 2023; NIST, 2025). Brasil, por el contrario, ofrece un caso de regulación participativa adaptada a la dinámica sociotécnica local y a los objetivos de desarrollo regional (Lobato, 2024; Ministerio de Ciencia, Tecnología e Innovaciones, 2021). Al examinar estos cinco marcos nacionales, este documento pretende identificar convergencias y divergencias en la gobernanza ética de la IA y evaluar cómo cada modelo equilibra imperativos contrapuestos: fomentar la innovación, proteger los derechos fundamentales y garantizar la responsabilidad democrática. El análisis también sitúa las iniciativas nacionales dentro de tendencias y movimientos reguladores mundiales más amplios, como la Ley de Inteligencia Artificial de la Unión Europea y las normas mundiales emergentes sobre gestión de riesgos de la IA (Comisión Europea, s.f.; AI Standards Hub, 2024).

Canadá

Canadá ha adoptado un enfoque polifacético y global de la gobernanza de la Inteligencia Artificial (IA), impulsado por una reforma legislativa, inversiones estratégicas y órganos consultivos de expertos. Un elemento central de este marco es la propuesta de *Ley sobre Inteligencia Artificial y Datos (AIDA)*, presentada a través del proyecto de ley C-27, también conocido como *Ley de Aplicación de la Carta Digital de 2022*. Este proyecto de ley pretende modernizar el régimen de privacidad de Canadá e introducir mecanismos de regulación de los sistemas de IA para mitigar los daños y promover la equidad. Como señala el Ministerio de Justicia, la AIDA establece “salvaguardias, medidas de transparencia y mecanismos de supervisión para garantizar que cualquier intrusión en los derechos sea limitada y adecuada en una sociedad libre y democrática” (Gobierno de Canadá, 2023b).

Como complemento de la AIDA está la *Ley de Protección de la Información Personal y los Documentos Electrónicos (PIPEDA)*, que establece diez Principios de Información Leal que exigen a las organizaciones del sector privado un tratamiento responsable de los datos. Según la Oficina del Comisionado de Privacidad de Canadá (2024), las organizaciones deben “identificar los fines de la recogida de datos y obtener el consentimiento”, y deben garantizar la transparencia, exactitud y seguridad en sus prácticas de tratamiento de datos. La LPRPDE también prohíbe prácticas explotadoras como la elaboración de perfiles discriminatorios y la vigilancia no autorizada.

La política canadiense de IA se estructura además en torno a la *Estrategia Pancanadiense de Inteligencia Artificial*, dirigida por el Instituto Canadiense de Investigación Avanzada (CIFAR). Con una financiación renovada hasta el presupuesto de 2021, la estrategia hace hincapié en tres pilares: comercialización, normas y talento. En el marco del primer pilar, se asignaron 60 millones de dólares para ayudar a los institutos nacionales de IA, como Amii, Mila y el Instituto Vector, a traducir la investigación en aplicaciones comerciales. Además, se destinaron 125 millones de dólares a Agrupaciones de Innovación Global para fomentar la adopción de la IA en sectores estratégicos. El pilar 2 invierte 8,6 millones de dólares en el desarrollo de normas nacionales de IA a través del Consejo de Normas de Canadá, mientras que el pilar 3 proporciona 208 millones de dólares para apoyar la investigación y el desarrollo de talentos, junto con 40 millones de dólares para infraestructura computacional (Estrategia Pancanadiense de Inteligencia Artificial, 2024).

Además, el *Consejo Asesor sobre Inteligencia Artificial*, creado en 2019, desempeña un papel fundamental en la orientación de la política federal. A partir de 2025, el Consejo ha ampliado su mandato para ofrecer asesoramiento político sobre la seguridad y la gobernanza de la IA, contribuyendo a la aplicación de las iniciativas del Presupuesto 2024. Reúne a partes interesadas del mundo académico, la industria y la sociedad civil para garantizar que el desarrollo de la IA refleje valores canadienses como la inclusión, la equidad y la transparencia (Gobierno de Canadá, 2025). Estos marcos jurídicos, institucionales y estratégicos posicionan colectivamente a Canadá como líder en gobernanza de la IA responsable e impulsada por la innovación.

Estados Unidos

Estados Unidos ha adoptado un enfoque integral y de múltiples instrumentos para la gobernanza de la inteligencia artificial (IA), equilibrando la innovación con la supervisión ética y la responsabilidad democrática. En el centro de este marco se encuentra la *Iniciativa Nacional de IA*, autorizada por la *Ley de Iniciativa Nacional de IA de 2020*. Esta estrategia federal pretende “avanzar en la investigación y el desarrollo de la IA, promover una fuerza de trabajo sólida en IA, garantizar el uso ético y fiable de la IA y fomentar la colaboración internacional” (Artificial Intelligence for the American People, s.f.). La iniciativa está coordinada por la *Oficina de la Iniciativa Nacional de IA* (NAIIO), que facilita la cooperación interinstitucional e intersectorial, apoya la creación de Institutos Nacionales de Investigación sobre IA y promueve la integración responsable de la IA en los organismos gubernamentales.

Paralelamente, la *Ley de Responsabilidad Algorítmica de 2023* propone importantes medidas regulatorias dirigidas al uso de la IA en contextos de toma de decisiones de alto riesgo, como la sanidad, la vivienda y el empleo. Exige a las empresas que evalúen y divulguen el impacto de los sistemas automatizados, especialmente cuando los datos personales

sensibles puedan influir en los resultados. La ley obliga tanto a los desarrolladores como a los usuarios a “realizar evaluaciones de impacto e informar de los resultados seleccionados a la Comisión Federal de Comercio (FTC)”, con el objetivo de mitigar las prácticas discriminatorias o sesgadas (Algorithmic Accountability Act of 2023, 2023).

Una herramienta clave de apoyo a la gobernanza del riesgo es el *Marco de Gestión del Riesgo de la IA (AI RMF)*, desarrollado por el Instituto Nacional de Estándares y Tecnología (NIST). Publicado en enero de 2023, este marco voluntario ofrece orientación para integrar la “fiabilidad en el diseño, desarrollo, uso y evaluación de los sistemas de IA” (NIST, 2025). Va acompañado de una serie de recursos, entre los que se incluyen un manual y una hoja de ruta y, en 2024, un perfil especial que aborda los riesgos de la IA generativa. El NIST también ha puesto en marcha el *Centro de Recursos de Inteligencia Artificial Fiable y Responsable* para promover la aplicación y fomentar la armonización internacional.

La Oficina de Política Científica y Tecnológica de la Casa Blanca publicó en 2022 el *Plan para una Declaración de Derechos de la IA*, que articula aún más los principios éticos. Este documento esboza cinco protecciones fundamentales: “sistemas seguros y eficaces, protección frente a la discriminación algorítmica, privacidad de los datos, notificación y explicación, y alternativas humanas o mecanismos de emergencia” (The White House, 2023). Reconociendo la creciente influencia de la IA en ámbitos como la sanidad y las finanzas, el plan subraya que los daños algorítmicos pueden evitarse mediante un diseño y una gobernanza basados en principios. Estas iniciativas demuestran el compromiso del gobierno de EE.UU. para fomentar el liderazgo en IA, salvaguardando los derechos civiles y promoviendo la confianza en las tecnologías digitales.

Brasil

Brasil se ha convertido en un líder regional en la gobernanza de la inteligencia artificial (IA), aplicando un enfoque normativo multidimensional y centrado en la ética que refleja tanto las mejores prácticas internacionales como las preocupaciones sociotécnicas locales. Una piedra angular de este marco en evolución es el proyecto de ley 2338/2023, un marco jurídico integral de la IA aprobado por unanimidad por la Comisión Interna Temporal de Inteligencia Artificial (CTIA) del Senado brasileño en diciembre de 2024. El proyecto de ley adopta un “enfoque basado en el riesgo que reconoce nuevos derechos fundamentales frente a los desafíos planteados por los sistemas automatizados” y sustituye la idea de una autoridad central de IA por una estructura reguladora descentralizada compuesta por organismos sectoriales como Anatel, Anvisa y el Banco Central (Lobato, 2024). La legislación fue el producto de una importante deliberación pública y de la participación de la sociedad civil, incluidas las contribuciones de Data Privacy Brasil.

Como complemento de este esfuerzo normativo está la *Estratégia Brasileira de Inteligência Artificial* (EBIA), la Estrategia Nacional de Inteligencia Artificial de Brasil, lanzada en 2021 por el Ministerio de Ciencia, Tecnología e Innovaciones (MCTI). Basada en los Principios de IA de la OCDE, la EBIA hace hincapié en “el crecimiento inclusivo, los valores centrados en el ser humano, la transparencia, la solidez y la responsabilidad” (Ministerio de Ciencia, Tecnología e Innovaciones, 2021). La estrategia se estructura en torno a nueve ejes estratégicos, que incluyen legislación y ética, gobernanza de la IA, cooperación internacional, educación y formación de la mano de obra, I+D, aplicaciones industriales, modernización del sector público, seguridad pública y emprendimiento. A través de 73 acciones estratégicas, la EBIA pretende fomentar la innovación garantizando al mismo tiempo un acceso equitativo y la protección de los derechos humanos.

Además, la *Ley General de Protección de Datos* (LGPD) de Brasil, en vigor desde 2020, proporciona un marco jurídico fundacional para la gobernanza de datos. Con un alcance similar al del Reglamento General de Protección de Datos (RGPD) de la UE, la LGPD se aplica a cualquier operación de tratamiento de datos, independientemente de su ubicación geográfica o medios tecnológicos. Refuerza derechos clave como el consentimiento, el acceso, la rectificación y la minimización de datos, creando una infraestructura sólida para la protección de datos personales (Data Protection Laws of the World, s.f.).

A pesar de los avances institucionales, Brasil sigue enfrentándose a desafíos. Entre ellos se incluyen “la escasa preparación en ciberseguridad, la inversión insuficiente en I+D, el acceso desigual a la tecnología en las distintas regiones y la persistencia de desigualdades raciales y de género en la educación tecnológica” (Lobato, 2024). En particular, se presta poca atención a la inclusión de las culturas y lenguas indígenas en el desarrollo de la IA, lo que puede reforzar las exclusiones sistémicas. Aunque el público confía más en los servicios tecnológicos del sector privado que en las plataformas digitales del gobierno, Brasil sigue puntuando alto en índices como el Inventario de Datos Abiertos y el Índice de Participación Electrónica, lo que indica una base sólida para la futura gobernanza de la IA.

Reino Unido

El Reino Unido ha adoptado un enfoque multifacético y con visión de futuro para la gobernanza de la inteligencia artificial (IA), con el objetivo de equilibrar la innovación con la seguridad y la confianza pública. Un elemento central de la estrategia del Reino Unido es el *Libro Blanco sobre Regulación de la IA* de 2023, que esboza un “marco regulador de la IA favorable a la innovación” basado en cinco principios intersectoriales que los reguladores existentes aplican dentro de sus dominios (DSIT, 2024). En lugar de establecer un único regulador central de la IA, el Reino Unido promueve un modelo flexible y adaptable que

permite a los organismos sectoriales dar forma a sus respuestas al cambio tecnológico. Este marco se complementa con una nueva función central de coordinación para garantizar la coherencia normativa y colmar las lagunas de gobernanza. Organismos reguladores como la Oficina del Comisionado de Información (ICO) y la Autoridad de Competencia y Mercados (CMA) ya han tomado medidas para alinearse con estas propuestas mediante la revisión de los modelos fundacionales y la actualización de las orientaciones sobre IA y protección de datos.

Como complemento de esta arquitectura normativa está el Centro de Normas de IA, una iniciativa lanzada en el marco de la Estrategia Nacional de IA del Reino Unido. Dirigido por el Alan Turing Institute en colaboración con la British Standards Institution (BSI) y el National Physical Laboratory (NPL), y con el apoyo del Departamento de Ciencia, Innovación y Tecnología (DSIT), el Centro “promueve una IA fiable y responsable a través de una normalización internacional e integradora” (AI Standards Hub, 2024). Estructurado en torno a cuatro pilares -un observatorio, una comunidad de colaboración, un centro de formación y un centro de investigación-, el Centro es una plataforma para el desarrollo de capacidades, el intercambio de conocimientos y la participación de las partes interesadas, especialmente de los grupos menos representados. Apoya el objetivo del Reino Unido de dar forma a las normas mundiales de la IA y fomentar la participación interdisciplinaria.

El Instituto de Seguridad de la IA, que también forma parte del DSIT, aborda los riesgos emergentes asociados a los sistemas avanzados de IA. El objetivo del Instituto es probar tecnologías de IA, informar a los responsables de la formulación de políticas y fomentar la colaboración público-privada para garantizar el desarrollo seguro de la IA. Dirigido por expertos de OpenAI, Google DeepMind y la Universidad de Oxford, y respaldado por más de 100 millones de libras en financiación, el Instituto cuenta con el apoyo de una red de más de 20 asociaciones de investigación y acceso a más de 1.500 millones de libras en infraestructura informática (The AI Security Institute, s.f.). Su misión es hacer avanzar “la investigación de beneficio público y garantizar prácticas seguras de IA”, reflejando el énfasis del Reino Unido en la innovación responsable.

Por último, la Ley de Protección de Datos de 2018, alineada con el GDPR del Reino Unido, sigue siendo una piedra angular de la gobernanza de la IA, especialmente en la investigación. Aunque los datos anonimizados están exentos, los investigadores que manejan datos identificables deben aplicar fuertes salvaguardias y seguir cumpliendo la normativa del Reino Unido y, en su caso, de la UE (DSIT, 2024).

Francia

Las iniciativas de Francia en materia de inteligencia artificial (IA) reflejan una estrategia coordinada a varios niveles centrada en la innovación, la regulación y la supervisión ética. Un elemento central de su visión nacional es la estrategia “Francia 2030”, que se basa en la anterior iniciativa *AI for Humanity* y pretende consolidar el liderazgo mundial de Francia en IA. Adoptada en noviembre de 2021, esta estrategia hace hincapié en el desarrollo del talento, la excelencia científica y la integración económica. Dos programas principales -IA-Cluster e IA-Booster- forman su columna vertebral. IA-Cluster pretende crear “centros de excelencia para la investigación y la formación en IA”, mientras que IA-Booster apoya la transformación digital de las pequeñas y medianas empresas (Francia - Estrategia Nacional para la IA, 2023). Entre los principales objetivos figuran duplicar el número de profesionales de la IA de aquí a 2030, situar a las instituciones francesas entre las 50 mejores del mundo y atraer a 15 destacados investigadores internacionales de IA de aquí a 2024.

El programa de IA de Francia se inscribe en el marco más amplio de la Ley de IA de la Unión Europea y el Reglamento General de Protección de Datos (RGPD). La Commission nationale de l’informatique et des libertés (CNIL) ha desempeñado un papel crucial a la hora de aclarar cómo se aplica el RGPD a los sistemas de IA a través de su *Plan de Acción de IA 2023*. Estas orientaciones fomentan la innovación al tiempo que salvaguardan los derechos fundamentales. Destaca, por ejemplo, que los desarrolladores de IA pueden esbozar la finalidad prevista de los sistemas de IA “aunque no esté totalmente definida en la fase de formación”, y afirma que pueden utilizarse grandes conjuntos de datos cuando “los datos personales se traten con cuidado” (Ley de IA de la UE, s.f.). Otras disposiciones permiten la conservación prolongada de los datos en condiciones estrictas de seguridad y la reutilización legalmente justificada de los datos en consonancia con los fines originales del tratamiento.

Además de la regulación a nivel nacional y de la UE, Etalab, organismo francés de política de datos públicos, contribuye a la transparencia algorítmica en el gobierno. El libro blanco de Etalab de 2019 “*With Great Power Comes Great Responsibility*” identifica la necesidad de responsabilidad pública en la toma de decisiones algorítmicas. Hace hincapié en la integración temprana de salvaguardias -como las evaluaciones de impacto, la mitigación del sesgo de los datos y la capacidad de explicación- para garantizar que los algoritmos en la administración pública defiendan los valores democráticos (Chignard, s.f.).

El hecho de que Francia acoja la Cumbre de Acción sobre IA en febrero de 2025 subraya su intención de mantenerse a la vanguardia del discurso mundial sobre IA. Mediante la combinación de sólidos marcos jurídicos, inversiones estratégicas y supervisión ética, Francia pretende “acelerar la integración de la IA en la economía”, garantizando al mismo tiempo

la alineación con los valores y derechos fundamentales europeos. Este enfoque holístico refleja la ambición de Francia de liderar las dimensiones tecnológica y normativa de la gobernanza de la IA.

Conclusión

El presente documento comparativo de los marcos nacionales de gobernanza de la inteligencia artificial (IA) en Canadá, Estados Unidos, Brasil, Reino Unido y Francia revela trayectorias convergentes y divergencias duraderas en la forma en que los Estados tratan de equilibrar la innovación, la regulación y la supervisión ética. Un consenso cada vez mayor en todas las jurisdicciones hace hincapié en la necesidad de promover sistemas de IA fiables que se ajusten a los valores democráticos, los derechos humanos y la responsabilidad pública (The White House, 2023; NIST, 2025; EU Artificial Intelligence Act, s.f.). Sin embargo, los enfoques nacionales divergen significativamente en cuanto a acuerdos institucionales, diseño normativo y prioridades estratégicas.

Canadá y Francia representan modelos de gobernanza anticipatoria, caracterizados por una coordinación centralizada de las políticas y un firme compromiso con la supervisión ética. La Ley de Inteligencia Artificial y Datos (AIDA) de Canadá y la Estrategia Pancanadiense de IA reflejan un énfasis normativo en la equidad, la transparencia y la integración responsable de la IA en todos los sectores (Gobierno de Canadá, 2023b; Estrategia Pancanadiense de Inteligencia Artificial, 2024). Francia avanza de forma similar en un enfoque integrado a través de la alineación con la Ley de IA de la Unión Europea y la autoridad reguladora de la CNIL, haciendo hincapié en la responsabilidad algorítmica y la protección de los derechos fundamentales (Francia - Estrategia Nacional para la IA, 2023).

Por el contrario, Estados Unidos y el Reino Unido favorecen modelos de gobernanza más descentralizados y sectoriales. El marco estadounidense -anclado por la Iniciativa Nacional de IA y la Carta de Derechos de la IA- se basa en la colaboración entre organismos y en normas voluntarias, con el objetivo de aprovechar la innovación y promover al mismo tiempo salvaguardias mediante mecanismos de derecho indicativo (Ley de Responsabilidad Algorítmica de 2023; NIST, 2025). El enfoque del Reino Unido, articulado en su Libro Blanco sobre la Regulación de la IA y apoyado por instituciones como el Instituto de Seguridad de la IA y el Instituto Alan Turing, da prioridad a la agilidad normativa y a la capacidad de establecimiento de normas mundiales (DSIT, 2024).

Brasil ofrece un caso convincente de una economía emergente que adapta los principios mundiales a su contexto nacional. La propuesta de marco jurídico de la IA (proyecto de ley 2338/2023), junto con la Ley General de Protección de Datos (LGPD), señala un com-

promiso con la gobernanza de la IA basada en los derechos. Sin embargo, Brasil sigue enfrentándose a retos estructurales, como la desigualdad digital y la limitada capacidad reguladora, que pueden limitar su aplicación efectiva (Ministerio de Ciencia, Tecnología e Innovaciones, 2021).

De cara al futuro, la durabilidad y legitimidad de las estrategias nacionales de IA dependerán de su capacidad para institucionalizar principios éticos, garantizar salvaguardias aplicables y fomentar la coherencia internacional. La creciente naturaleza transnacional de los sistemas de IA exige mecanismos sólidos de interoperabilidad normativa, especialmente a la luz de iniciativas como la Ley de IA de la UE, los Principios de IA de la OCDE y los Códigos de Conducta del G7. A medida que los Estados se enfrentan a las implicaciones sociales de la IA avanzada, incluidos los modelos generativos y los sistemas autónomos de toma de decisiones, la integración de modelos de gobernanza inclusivos, participativos y adaptativos será fundamental para mantener la confianza pública y la responsabilidad democrática (Ley de Inteligencia Artificial de la UE, s.f.; Leyes de Protección de Datos del Mundo, s.f.).

En conclusión, aunque todos los países han dado pasos significativos hacia una gobernanza ética de la IA, siguen existiendo compensaciones entre la flexibilidad normativa, la fuerza de la aplicación y la inclusión de las partes interesadas. El panorama comparativo sugiere la necesidad de modelos de gobernanza híbridos que combinen la innovación nacional con la cooperación transnacional y la regulación adaptativa.

Referencias

AI Standards Hub. (2024). Acerca de. <https://aistandardshub.org>

Algorithmic Accountability Act of 2023. (2023). Congreso de los Estados Unidos.

Binns, R. (2018). On the idea of fairness in AI. *IEEE Internet Computing*, 22(2), 3–10. <https://doi.org/10.1109/MIC.2018.0222313>

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Bryson, J. J. (2018). The ethics of artificial intelligence. In K. T. F. K. T. H. H. T. (Ed.), *Handbook of ethics in artificial intelligence* (pp. 4–8). Springer.

Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12. <https://doi.org/10.1177/2053951715622512>

Calo, R. (2017). Artificial intelligence policy: A primer and roadmap. *U.C. Davis Law Review*, 51(2), 393–436.

Cave, S., & Dignum, V. (2019). *Responsible AI: A global perspective*. Springer.

Chignard, S. (s.f.). With great power comes great responsibility. Etalab.

Data Protection Laws of the World. (s.f.). <https://www.dlapiperdataprotection.com/>

Departamento de Ciencia, Innovación y Tecnología (DSIT). (2024). Libro blanco sobre la regulación de la IA.

Ley de Inteligencia Artificial de la UE. (s.f.).

Comisión Europea. (2021). Proposal for a regulation laying down harmonized rules on artificial intelligence (Ley sobre Inteligencia Artificial). https://ec.europa.eu/info/sites/default/files/commission-proposal-regulation-ai-act_en.pdf

Francia - Estrategia nacional para la IA (2023).

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerization? *Technological Forecasting and Social Change*, 114, 254–280. <https://doi.org/10.1016/j.techfore.2016.08.019>

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Gobierno de Canadá. (2023b). Artificial Intelligence and Data Act (AIDA).

Gobierno de Canadá. (2025). Advisory Council on AI.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>

Lobato, L. (2024). Brazil's AI law: Participatory regulation in a digital age. *Data Privacy Brasil*.

Ministerio de Ciencia, Tecnología e Innovación (2021). *Estratégia Brasileira de Inteligência Artificial*.

Instituto Nacional de Normas y Tecnología (NIST) (2025). *AI Risk Management Framework*.

Oficina del Comisionado de Privacidad de Canadá (2024).

Estrategia Pancanadiense de Inteligencia Artificial (2024).

Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human supervision of automation. In *Human Factors in Systems and Design* (pp. 77–99). CRC Press.

Scherer, M. (2016). Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. *Harvard Journal of Law & Technology*, 29(2), 351–377.

Taeihagh, A. (2021). Governance of artificial intelligence: Challenges and opportunities for regulatory frameworks. *Journal of Artificial Intelligence Research*, 70, 113–130. <https://doi.org/10.1613/jair.1.12479>

Instituto de Seguridad de la IA (s.f.).

The White House. (2023). Blueprint for an AI Bill of Rights. <https://www.whitehouse.gov>

Zarsky, T. Z. (2016). The trouble with algorithmic decisions: An analytic road map to examine the issues of algorithmic decision-making. *Science, Technology, & Human Values*, 41(1), 118–136. <https://doi.org/10.1177/0162243915605575>

Gobernanza ética de las iniciativas de IA en Asia: China, Japón, Singapur, Corea del Sur, Taiwán

Cátia Miriam Costa

*Cátedra de Interregionalismo y Gobernanza Global del
European Institute of International Studies (Suecia)*

Resumen

Este capítulo analiza las prácticas de gobernanza ética en torno a la inteligencia artificial (IA) en cinco países asiáticos tecnológicamente avanzados: China, Japón, Singapur, Corea del Sur y Taiwán. Estos territorios representan una muestra representativa diversa de modelos de gobernanza. El estudio analiza las tendencias de cada país en la elaboración de políticas de IA, destacando los valores compartidos -como la transparencia, la responsabilidad y la equidad- y los enfoques distintivos conformados por contextos culturales, políticos y económicos. Además de los esfuerzos nacionales, el documento investiga los mecanismos de cooperación regional y los diálogos mundiales, como las iniciativas de la ASEAN, las disposiciones sobre comercio electrónico de la OMC y los acuerdos tecnológicos trilaterales, que promueven el desarrollo responsable de la IA y la interoperabilidad transfronteriza. También aborda la contribución del sistema de las Naciones Unidas a la gobernanza mundial y su influencia sobre las normas reguladoras regionales y nacionales.

Explora cómo los países seleccionados en los estudios de caso interpretan normas mundiales como los principios de la IA de la OCDE e incorporan salvaguardias éticas a la legislación, la práctica institucional y el diseño de políticas. La interacción entre la aplicación local y la influencia mundial surge como un tema central, en particular en la forma en que la gobernanza de la IA se enmarca en torno a la protección de los datos personales, la soberanía de los datos y la responsabilidad algorítmica.

Mediante un análisis político comparativo y una revisión empírica de los documentos estratégicos de 2020 a 2025, el documento revela cómo los experimentos reguladores de Asia contribuyen a la evolución de las normas mundiales sobre IA.

Concluye que los modelos híbridos de la región -que combinan innovación, ética y gobernanza- aportan valiosas ideas sobre las mejores prácticas internacionales. Este trabajo hace avanzar el discurso académico y político al ofrecer una perspectiva comparada de la gobernanza de la IA y su posible contribución a la gobernanza mundial.

1.Introducción

La gobernanza de la Inteligencia Artificial (IA) se refiere al conjunto de políticas, procesos, normas y marcos institucionales que garantizan el desarrollo y la aplicación ética, legal y socialmente responsable de las tecnologías de IA. Estos sistemas están diseñados para garantizar que el desarrollo de la IA sirva al bien público, promueva la equidad y evite daños a las personas y las sociedades. Los mecanismos de gobernanza abordan cuestiones críticas como la discriminación, la vigilancia, la opacidad algorítmica y el potencial de la automatización para desplazar puestos de trabajo y concentrar el poder.

A medida que las aplicaciones de IA se extienden a servicios esenciales y procesos de toma de decisiones -como la calificación crediticia, los diagnósticos médicos, la vigilancia policial predictiva y la gestión laboral-, se ha intensificado la preocupación pública por la parcialidad, la rendición de cuentas y el uso indebido de los datos. La gobernanza de la IA requiere, por tanto, un enfoque multidimensional que incluya regulación legal, principios éticos, supervisión institucional y mecanismos participativos para garantizar la inclusión y la equidad. Además, la transparencia en el diseño algorítmico, el control de calidad de los datos y la capacidad de explicación del sistema son fundamentales para aumentar la confianza de los usuarios y salvaguardar las libertades civiles.

A nivel mundial, marcos como el Reglamento General de Protección de Datos (RGPD) de la Unión Europea y los Principios de la IA de la OCDE ofrecen modelos fundacionales que hacen hincapié en la rendición de cuentas, los valores centrados en el ser humano y la gestión de riesgos. Sin embargo, el éxito de la aplicación de estas normas depende principalmente de cómo las interpretan, adaptan y apliquen los países en sus contextos nacionales. Asia es especialmente significativa a este respecto, destacando los ejemplos de China, Japón, Singapur, Corea del Sur y Taiwán, cada uno de los cuales adopta enfoques distintivos basados en sus tradiciones de gobernanza, prioridades económicas y capacidades tecnológicas.

El presente documento investiga cómo estos cinco países articulan y aplican la gobernanza de la IA a través de marcos normativos, acuerdos institucionales y asociaciones internacionales.

Sitúa sus esfuerzos en el marco de debates más amplios sobre la IA ética, la soberanía digital y la elaboración de normas mundiales, y sostiene que Asia no se limita a adoptar las normas mundiales, sino que les da forma activamente.

2. Metodología

El presente estudio emplea una metodología cualitativa de análisis político comparativo para examinar la gobernanza ética de la IA en cinco economías asiáticas: China, Japón, Singapur, Corea del Sur y Taiwán. Estos países fueron seleccionados por su liderazgo en innovación digital y sus diferentes orientaciones normativas, lo que ofrece una visión general de la diversidad de la gobernanza en la región. El estudio adopta un enfoque de múltiples fuentes, integrando documentos oficiales y publicaciones académicas.

Las fuentes primarias se seleccionaron en función de su relevancia para las iniciativas reguladoras de los gobiernos. Los datos complementarios se extrajeron de la información institucional facilitada por organizaciones internacionales como la OCDE, la ASEAN y la APEC y de iniciativas de colaboración entre gobiernos asiáticos y socios mundiales. Se analizó la coherencia temática, la intención reguladora, los mecanismos de aplicación y la alineación de estos materiales con los principios éticos de la IA reconocidos en todo el mundo.

La atención analítica se dirigió a comparar los componentes estructurales de la gobernanza que impulsan las políticas públicas. Además, el estudio evalúa la integración de los valores éticos -transparencia, equidad, inclusión y responsabilidad- en la gobernanza de la IA en las distintas jurisdicciones.

El enfoque de análisis comparativo permitió identificar patrones, divergencias y prácticas híbridas de gobernanza. También facilitó la exploración de la armonización de las políticas regionales y la posible influencia de los modelos asiáticos en los debates internacionales sobre la gobernanza de la IA. Al presentar cinco estudios de casos asiáticos, el estudio pretende aclarar las prácticas éticas de gobernanza de la IA y estimular el debate sobre cómo las experiencias nacionales y regionales pueden contribuir a la gobernanza mundial.

3. Compromiso regional y mundial

La cooperación regional y el compromiso mundial influyen cada vez más en el enfoque asiático de la gobernanza de la IA. Países como Japón, Singapur, Corea del Sur, China y Taiwán contribuyen activamente al desarrollo de políticas internacionales de IA a través de foros multilaterales, acuerdos bilaterales y la participación en organizaciones intergubernamentales. La publicación por parte de la ASEAN de la Guía sobre Gobernanza y Ética de la IA en 2024 representa un hito en la coordinación regional, ya que ofrece recomendaciones

prácticas para alinear las estrategias nacionales y fomentar la interoperabilidad entre los Estados Miembros.

La Cooperación Económica Asia-Pacífico (APEC) también ha configurado activamente la integración económica de la IA. Su informe de 2022 sobre la IA en la formulación de políticas económicas hizo hincapié en la necesidad de un crecimiento integrador, un despliegue ético y una regulación armonizada. Paralelamente, Japón, Singapur y Australia iniciaron diálogos sobre comercio digital a través de la Declaración Conjunta sobre Comercio Electrónico de la OMC, abordando la localización de datos, la confianza digital y los mecanismos de rendición de cuentas de la IA.

Las asociaciones bilaterales y trilaterales refuerzan aún más la influencia de Asia en la gobernanza. Por ejemplo, Corea del Sur, India y Estados Unidos convocaron un diálogo tecnológico en 2024 para coordinar enfoques sobre sistemas de IA de alto riesgo y flujos de datos transfronterizos. Del mismo modo, China y Estados Unidos iniciaron conversaciones bilaterales sobre los riesgos de la IA, con el objetivo de reducir las tensiones geopolíticas al tiempo que se garantiza la seguridad y el uso ético de las tecnologías de IA. Estas interacciones demuestran un creciente reconocimiento de la naturaleza interconectada de los retos de la IA y de la importancia de los marcos de gobernanza compartidos.

Eventos mundiales como el Foro de Gobernanza de Internet de la ONU y la Cumbre Mundial de IA para el Bien de la Unión Internacional de Telecomunicaciones también cuentan con contribuciones activas de partes interesadas asiáticas. A través de estas plataformas, los gobiernos asiáticos abogan por un modelo equilibrado que incorpore la innovación y la moderación ética. Esta sinergia mundial-regional subraya el liderazgo emergente de Asia en la configuración de la gobernanza internacional de la IA y los contornos normativos.

El enfoque de análisis comparativo permitió identificar patrones, divergencias y prácticas híbridas de gobernanza. También facilitó la exploración de la armonización de las políticas regionales y la posible influencia de los modelos asiáticos en los debates internacionales sobre la gobernanza de la IA. Al situar los esfuerzos nacionales en un contexto político mundial más amplio, el estudio pretende aportar tanto ideas descriptivas como recomendaciones normativas para futuras prácticas de gobernanza ética de la IA.

4. Estudios de casos nacionales

Los estudios de casos reflejan las economías digitales más innovadoras de Asia y las que presentan un bagaje más sofisticado a la hora de abordar cuestiones de tecnología y gobernanza de la IA. China, Japón, Singapur, Corea del Sur y Taiwán articulan la regulación nacional con las interacciones regionales y globales y han experimentado paradigmas avanzados en

la gobernanza de la IA. Curiosamente y aunque con experiencias y sistemas políticos diferentes, todos los países presentan una voluntad similar de tener soberanía nacional sobre los datos resultantes del uso de la IA.

4.1 China

El enfoque chino de la gobernanza de la IA se basa en un modelo de innovación impulsada por el Estado y supervisión reguladora. El énfasis del gobierno en la soberanía digital y la estabilidad social ha dado lugar a una estructura de gobernanza única que combina el avance tecnológico con la orientación. Documentos normativos como las “Disposiciones sobre la Administración de los Servicios de Información de Internet de Síntesis Profunda” de 2022 y la normativa de 2023 sobre IA generativa ilustran un marco integral centrado en la seguridad de los contenidos, la transparencia algorítmica y la responsabilidad de los usuarios.

Estas políticas reflejan una preocupación por la armonización de los flujos de información y la promoción de la IA como motor del crecimiento económico y el control de las posibilidades de perjudicar el uso de las herramientas de IA. La Iniciativa de Gobernanza Global de la IA de China, lanzada a finales de 2023, pretende internacionalizar su modelo de gobernanza proponiendo principios de seguridad, equidad y respeto mutuo. Además de la regulación nacional, China ejerce influencia a través de iniciativas multilaterales, incluida la cooperación con los países BRICS y su papel activo en los esfuerzos de establecimiento de normas de IA de la ONU. El ecosistema digital de exportación de China, ejemplificado por empresas como Baidu y Alibaba, también desempeña un papel crucial en la aplicación de protocolos de gobernanza.

La presencia internacional de las empresas chinas y la sofisticación de las tecnologías de IA del país favorecieron su capacidad para actuar como un actor internacional dinámico en relación con los asuntos de gobernanza de la IA. China enmarca su modelo de gobernanza de la IA como una de las posibles soluciones globales para la gobernanza de la IA y como alternativa a la no gobernanza o a modelos menos avanzados. Normalmente, el argumento de China para mostrar las bondades de su modelo hace referencia al desarrollo económico, el éxito de la economía digital, la soberanía y la armonía social (Xu et al., 2024; Mahapatra, 2025).

4.2 Japón

La estrategia japonesa de gobernanza de la IA se caracteriza por la agilidad y la inclusión. Guiado por su Estrategia de IA 2022, el país se basa en gran medida en la orientación no vinculante, la reciprocidad y el diálogo con las partes interesadas. En lugar de aplicar leyes

rígidas sobre IA, Japón favorece marcos adaptables que permiten una aplicación específica para cada sector. Esta estrategia de gobernanza incluye la coordinación con asociaciones industriales, grupos de defensa de los derechos de los consumidores y socios internacionales para mantener la coherencia normativa y la confianza pública.

Instrumentos jurídicos clave como la Ley de Protección de la Información Personal (PIPA) y la Ley de Responsabilidad de Productos desempeñan un papel indirecto en la gobernanza de la IA, mientras que iniciativas como las Directrices de Gobernanza de la IA elaboradas por la Oficina del Gabinete apoyan el despliegue ético. Estas políticas hacen hincapié en la responsabilidad, la dignidad humana y la capacidad de explicación de los algoritmos, y se ajustan estrechamente a los principios de la OCDE. La Agencia Digital de Japón refuerza aún más la supervisión, trabajando en estrecha colaboración con la industria y el mundo académico para fomentar la innovación y mitigar los riesgos sociales.

La implicación multilateral de Japón incluye la participación en los diálogos sobre gobernanza digital del G7, asociaciones con la ASEAN y la UE, y contribuciones activas a los debates políticos sobre IA de la OCDE. El gobierno también colabora con organizaciones internacionales en programas de capacitación que promueven una IA fiable. La dirección estratégica de Japón incluye esfuerzos para liderar el desarrollo de normas mundiales sobre ética y gobernanza de la IA, abogando por la neutralidad tecnológica, la inclusión y la sostenibilidad.

Además, Japón está integrando la IA en su visión de la Sociedad 5.0, cuyo objetivo es construir una sociedad superinteligente en la que la tecnología esté al servicio de la productividad económica y la inclusión social. Las consideraciones éticas están integradas en esta agenda nacional, garantizando que el progreso tecnológico no se produzca a expensas de las libertades individuales o la cohesión social. El modelo de gobernanza iterativo y basado en el diálogo de Japón refleja un compromiso con los valores democráticos y el multilateralismo frente al rápido cambio tecnológico (Xu et al., 2024).

4.3 Singapur

Singapur es ampliamente reconocido por su marco de gobernanza de IA proactivo y completo. El Marco Modelo de Gobernanza de la IA, publicado por primera vez en 2019 y actualizado en 2020, ofrece orientación detallada sobre cómo las organizaciones pueden poner en funcionamiento la IA responsable. Los principios básicos incluyen la equidad, la transparencia, la orientación al ser humano y la capacidad de explicación. La estrategia de gobernanza de Singapur se refuerza a través de kits de herramientas específicos del sector, *sandboxes* normativos y un marco de implementación diseñado para fomentar la responsabilidad y la confianza. El marco es un documento vivo que se revisa continuamente para abordar los riesgos emergentes y los comentarios de las partes interesadas.

El ecosistema de gobernanza de Singapur está coordinado por la Autoridad de Desarrollo de Medios Infocomm (IMDA) y la Comisión de Protección de Datos Personales (PDPC), que trabajan de forma conjunta para garantizar el cumplimiento, promover la innovación y salvaguardar el interés público. La PDPC también administra la Ley de Protección de Datos Personales (PDPA), que es fundamental en la gestión de la privacidad de los datos y el consentimiento y es esencial para los sistemas de IA entrenados en grandes conjuntos de datos. Estos organismos colaboran estrechamente con las empresas, las instituciones académicas y la sociedad civil para elaborar orientaciones específicas para cada sector y facilitar la reciprocidad.

Singapur se ha posicionado internacionalmente como un poder de convocatoria en el discurso mundial sobre la ética de la IA. Se ha asociado con Australia para probar los principios de gobernanza de la IA, ha contribuido al Centro para la Cuarta Revolución Industrial del Foro Económico Mundial y ha participado en acuerdos de economía digital con países como el Reino Unido y Chile. Estas colaboraciones permiten la interoperabilidad normativa y ayudan a gestionar los retos transfronterizos en materia de gobernanza de datos. El liderazgo de Singapur en la Asociación Mundial sobre Inteligencia Artificial (GPAI) refuerza aún más su influencia normativa en las normas políticas internacionales sobre IA.

A través de estos esfuerzos nacionales e internacionales, Singapur ha construido un modelo de gobernanza que alinea el avance tecnológico con los imperativos éticos. Su enfoque orientado a los resultados y con múltiples partes interesadas garantiza el fomento de la innovación y su dirección responsable. Esto sitúa a Singapur como pionero en materia de regulación y como punto de referencia para los países que pretenden gobernar la IA de forma equilibrada y orientada al futuro (Mahapatra, 2025).

4.4 Corea del Sur

Corea del Sur combina la inversión estratégica en IA con la previsión normativa para posicionarse como líder mundial en IA. La propuesta de Ley para la Promoción de la Industria de la IA y el Establecimiento de un Marco de Confianza en la IA clasifica los sistemas de IA en función del riesgo, establece requisitos para las pruebas y la certificación y crea un marco para la supervisión gubernamental. El modelo de gobernanza de Corea del Sur pretende equilibrar el apoyo a las empresas nacionales de IA con una sólida protección de los consumidores y los trabajadores. El Ministerio de Ciencia y TIC, en colaboración con la Agencia Nacional de la Sociedad de la Información (NIA), dirige la política y la estrategia de aplicación de la IA del país.

Corea del Sur también ha establecido las Normas Éticas de la Inteligencia Artificial, que proporcionan directrices voluntarias centradas en el ser humano, la inclusión y

la sostenibilidad. Estas normas refuerzan las consultas públicas, los comentarios de la industria y la alineación con las normas éticas mundiales. Se han puesto en marcha espacios aislados de regulación para permitir la experimentación en entornos reales bajo estrecha supervisión, lo que permite la innovación y la supervisión responsable. En particular, la IA se está aplicando en las ciudades inteligentes, la educación y la atención sanitaria, donde el despliegue ético es fundamental.

El énfasis de Corea del Sur en la I+D y la educación es evidente en sus planes nacionales de inversión, incluida la financiación para el desarrollo de talentos en IA, la cooperación internacional en investigación sobre IA y la integración de la alfabetización en IA en la educación pública. El gobierno también promueve la gobernanza digital transfronteriza a través de diálogos trilaterales con Estados Unidos e India, haciendo hincapié en la innovación responsable y los valores democráticos. Como parte de los grupos de trabajo de la OCDE y del Observatorio de IA, Corea del Sur contribuye a los esfuerzos de intercambio de conocimientos y establecimiento de normas sobre transparencia algorítmica, equidad y rendición de cuentas.

Mediante estas iniciativas, Corea del Sur aspira a convertirse en un centro tecnológico y un líder normativo en la gobernanza mundial de la IA. Su doble compromiso con la competitividad económica y la responsabilidad ética la sitúan como modelo para las economías digitales emergentes que navegan por las complejidades de la regulación de la IA (Xu et al., 2024).

4.5 Taiwán

La gobernanza de la IA de Taiwán depende del Ministerio Digital, que lidera iniciativas para integrar los valores públicos en el diseño y despliegue de la tecnología. Taiwán promueve un enfoque de la sociedad civil basado en el gobierno abierto y el compromiso cívico a través de *hackatones*, plataformas de consulta en línea y diálogos intersectoriales. Estos mecanismos participativos informan del uso ético de la IA en los servicios públicos, incluidos la sanidad, la educación y el bienestar social. Las aplicaciones de IA en la prestación de servicios públicos deben someterse a evaluaciones de sesgo, capacidad de explicación y derechos de los usuarios antes de ser aprobadas.

En cuanto a la estructura institucional, el Consejo Nacional de Ciencia y Tecnología de Taiwán supervisa la coordinación de la política de IA, mientras que la Oficina de Protección de Datos Personales garantiza que el tratamiento de los datos cumple las normas legales y éticas. El marco de gobernanza de Taiwán incorpora normas internacionales, incluida la alineación con los principios de la OCDE, y participa activamente en la cooperación digital a través de la

APEC, el grupo Digital Nations y otros entornos multilaterales. Estas colaboraciones ayudan a Taiwán a crear capacidad reguladora y promover normas compartidas.

El enfoque de gobernanza de Taiwán se basa en la transparencia. Las políticas de contratación pública exigen que los algoritmos utilizados en los servicios públicos estén abiertos a auditorías, sean comprensibles para los usuarios y estén sujetos a mecanismos de compensación. Además, Taiwán fomenta el desarrollo de herramientas de IA de código abierto y apoya la participación de la sociedad civil en la supervisión ética y la evaluación tecnológica. Su enfoque de los derechos digitales, basado en parte en el GDPR europeo, se considera progresista en Asia.

Al incorporar principios democráticos a sus estrategias de IA, Taiwán ofrece una alternativa a modelos de gobernanza más centralizados o impulsados por intereses comerciales. Demuestra cómo los territorios pequeños pueden liderar iniciativas digitales al fomentar enfoques integradores de las tecnologías emergentes. El ejemplo de Taiwán ilustra que incluso los países o territorios más pequeños pueden ofrecer propuestas interesantes para la gobernanza regional y mundial de la gobernanza de la IA (Xu et al., 2024).

Conclusión

La gobernanza ética de la IA en Asia es dinámica, diversa y estratégicamente significativa. El liderazgo político, los ecosistemas de innovación y la experimentación normativa de la región ofrecen valiosas lecciones para los marcos mundiales. Países como China, Japón, Singapur, Corea del Sur y Taiwán han desarrollado modelos únicos que reflejan sus realidades sociopolíticas y aspiraciones económicas, al tiempo que mantienen firmes compromisos con normas internacionales como la transparencia, la responsabilidad y los derechos humanos. Estos esfuerzos reflejan la capacidad de Asia no sólo para adaptarse a las tendencias éticas mundiales, sino también para contribuir de forma sustantiva a dar forma a las normas internacionales sobre IA.

Una característica notable del panorama de la gobernanza de la IA en Asia es la búsqueda de un modelo híbrido que equilibre la innovación con la supervisión y la soberanía nacional con la colaboración mundial. Organizaciones regionales como la ASEAN y la APEC han sentado las bases de marcos de gobernanza compartidos que promueven la interoperabilidad transfronteriza y la coherencia normativa. Al mismo tiempo, los compromisos bilaterales y trilaterales con los principales actores mundiales subrayan la postura proactiva de la región en la configuración de las futuras normas de IA. Estos avances señalan un movimiento hacia sistemas de gobernanza de la IA más integrados y orientados al futuro.

La investigación futura debería explorar las repercusiones sociales a largo plazo de los marcos de gobernanza de la IA en Asia, especialmente en ámbitos como los derechos digitales, el empleo, la educación y la atención sanitaria. Los estudiosos también podrían examinar cómo influyen los distintos modelos reguladores en la confianza pública y la participación ascendente en los procesos políticos de la IA. Los estudios longitudinales transnacionales y las entrevistas a las partes interesadas pueden revelar cómo experimentan las comunidades locales las implicaciones éticas de los sistemas de IA en la práctica.

Las recomendaciones políticas para los gobiernos y las organizaciones internacionales incluyen el fortalecimiento de la capacidad institucional para la supervisión ética, el fomento de procesos de formulación de políticas inclusivas y la inversión en educación para promover la alfabetización en IA. Los gobiernos deben garantizar que la regulación de la IA siga siendo adaptable, transparente y basada en valores sociales. Organismos internacionales como la ONU, la OMC y la OCDE deben seguir apoyando marcos de cooperación que mitiguen los riesgos éticos respetando las diferencias nacionales. A medida que las tecnologías de IA evolucionan rápidamente, la gobernanza ética debe posicionarse como un pilar fundamental para sostener sociedades inclusivas, democráticas y equitativas en todo el mundo.

Referencias

Aaronson, S. A. (2020). Data Governance, AI, and Trade: Asia as a Case Study (IIEP-WP-2020-6). Instituto de Política Económica Internacional, Universidad George Washington.

ASEAN. (2024). ASEAN Guide on AI Governance and Ethics.

Cooperación Económica Asia-Pacífico. (2022). AI in Economic Policymaking.

Mahapatra, S. (2025). Ethical Governance of AI and the Prevention of Digital Authoritarianism in South and Southeast Asia: Case Studies of India and Singapore. Laboratorio de Transformación Digital de GIGA.

Struett, T., Aaronson, S. A., & Zable, A. (2024). Data Governance Mapping Project, Year 4. Centro de Comercio Digital y Gobernanza de Datos.

Walter, Y. (2024). Managing the race to the moon: Global policy and governance in Artificial Intelligence regulation—A contemporary overview and an analysis of socioeconomic consequences. Discover Artificial Intelligence.

Xu, J., Lee, T., & Goggin, G. (2024). AI governance in Asia: policies, praxis and approaches. *Communication Research and Practice*, 10(3), 275–287.

Yigitcanlar, T., Agdas, D., & Degirmenci, K. (2023). Artificial intelligence in local governments: perceptions of city managers on prospects, constraints and choices. *AI & Society*, 38, 1135–1150.

Yoshija Walter (2023), Managing the race to the moon: Global policy and governance in Artificial Intelligence regulation—A contemporary overview and an analysis of socio-economic consequences, *Discover Artificial Intelligence*, 4, 14 (2024). <https://doi.org/10.1007/s44163-024-00109-4>

El rol de Naciones Unidas en la construcción de la Gobernanza ética de la IA y la coyuntura latinoamericana

Cecilia Celeste Danesi

*Investigadora del Instituto de Estudios Europeos y Derechos Humanos,
Universidad Pontificia de Salamanca*

El desarrollo acelerado de la Inteligencia Artificial (IA) ha generado transformaciones profundas en todos los sectores de la sociedad. Desde la salud y la educación hasta la seguridad y las finanzas, las aplicaciones de IA ofrecen tanto oportunidades de innovación como desafíos significativos para los derechos humanos, la equidad y la sostenibilidad global. En este contexto, el debate sobre la gobernanza ética de la IA se ha vuelto urgente, sobre todo para que el desequilibrio entre el norte y el sur global no se siga incrementando. La necesidad de establecer principios, marcos normativos y estructuras institucionales que regulen su uso de forma responsable ha motivado a actores multilaterales, gobiernos, empresas y sociedad civil a proponer iniciativas que prioricen una inteligencia artificial centrada en el ser humano.

En América Latina y el Caribe, el tema adquiere una relevancia estratégica. El curso “Gobernanza ética de la inteligencia artificial”, organizado por el SELA y el European Institute of International Studies, con colaboración de la Universidad Pontificia de Salamanca, propuso un espacio de reflexión para diplomáticos y funcionarios públicos sobre los desafíos éticos de la IA y las respuestas normativas que están surgiendo a nivel internacional.

Este artículo retoma las principales ideas expuestas durante la clase titulada “Naciones Unidas & Inteligencia Artificial: El Órgano Asesor de Alto Nivel sobre Inteligencia Artificial, la Unión Internacional de las Telecomunicaciones: ‘AI for Good’ y la UNESCO”, articulando

un análisis sobre el papel de las Naciones Unidas y otros organismos internacionales en la construcción de una gobernanza ética global con foco en América Latina.

1. La gobernanza ética de la inteligencia artificial: principios y características

El concepto de gobernanza ética de la IA implica establecer directrices, principios y mecanismos institucionales que orienten el desarrollo y la aplicación de tecnologías inteligentes hacia el bien común. Consiste en un plan estratégico común que gobierna organizaciones, instituciones, países, pueblos y cualquier espacio que genere o consuma inteligencia artificial.

Esta gobernanza se caracteriza por ser transdisciplinaria, es decir, integrar múltiples campos del conocimiento; multisectorial, al involucrar a gobiernos, sector privado, sociedad civil y comunidad científica; periódica, mediante procesos de revisión y actualización continua; transversal, dado su impacto en todos los sectores de la vida social; y global, buscando armonizar marcos normativos más allá de las fronteras nacionales (Danesi & Torres 2025). La mayoría de los planes estratégicos de IA adoptados por los países y organismos internacionales comparten algunos principios fundamentales: equidad, transparencia, responsabilidad, privacidad, explicabilidad, y respeto por los derechos humanos. No obstante, su implementación enfrenta importantes desafíos normativos, institucionales y técnicos. La falta de consensos globales, las asimetrías tecnológicas y la rapidez del cambio científico plantean preguntas sobre cómo garantizar que la IA no profundice desigualdades existentes ni socave las instituciones democráticas.

A continuación exploraremos el enfoque que están tomando los organismos internacionales, con foco en las Naciones Unidas, ya que constituyen el espacio y la guía propicia para cimentar una gobernanza ética de la IA global.

2. El papel de las Naciones Unidas en la construcción de una gobernanza ética global

Las Naciones Unidas han asumido un rol central en el impulso de una IA centrada en el ser humano. Este compromiso se manifiesta a través de múltiples agencias, comisiones y órganos consultivos que han abordado los riesgos y oportunidades de la IA en el desarrollo sostenible, los derechos humanos y la igualdad de género.

2.1. Conclusiones de la CSW67

En marzo de 2023, la Comisión de la Condición Jurídica y Social de la Mujer (CSW67) celebró su 67º período de sesiones con el tema prioritario: “Innovación y cambio tecnológico, y educación en la era digital para lograr la igualdad de género y el empoderamiento de todas

las mujeres y niñas”. La elección de este tema fue sumamente relevante para poner en agenda el impacto de la IA en los derechos de las mujeres.

Las conclusiones acordadas enfatizaron la necesidad de promover una transformación digital inclusiva y equitativa. Entre las recomendaciones destacadas se encuentra la urgencia de cerrar la brecha digital de género, fomentar una educación tecnológica con perspectiva de género, garantizar la participación equitativa de las mujeres en el diseño y desarrollo de tecnologías, y prevenir la violencia de género en entornos digitales (ONU Mujeres, 2023). Estas conclusiones reflejan una creciente preocupación por el sesgo algorítmico (Dwivedi, 2024), la exclusión digital y la reproducción de desigualdades en los sistemas de IA. Asimismo, resaltan la importancia de construir una IA feminista, inclusiva y sensible a las múltiples formas de discriminación.

2.2. El órgano asesor de alto nivel sobre inteligencia artificial

El Secretario General de la ONU creó el Órgano Asesor de Alto Nivel sobre Inteligencia Artificial, integrado por especialistas internacionales provenientes del mundo académico, científico, político y tecnológico. Este grupo publicó un primer informe con recomendaciones orientadas a promover un sistema de gobernanza global para la IA. Entre las propuestas figuran la creación de una oficina de IA dentro de la Secretaría General, la consolidación de un marco mundial de datos para IA, el establecimiento de un fondo global, y el fortalecimiento del desarrollo de capacidades en países del sur global.

Este órgano propone fomentar el intercambio de buenas prácticas regulatorias, desarrollar normas técnicas comunes y habilitar canales de diálogo político entre actores públicos y privados. Su enfoque busca compatibilizar la innovación con la ética, anticipar riesgos y promover una distribución justa de los beneficios tecnológicos (United Nations, 2023).

2.3. La UNESCO y la ética de la IA

La UNESCO ha tenido un rol clave para el desarrollo de una IA ética, no solo como una de las pioneras en la materia sino también por el alcance e impacto que han tenido sus iniciativas. En 2021, adoptó la “Recomendación sobre la Ética de la Inteligencia Artificial”, primer instrumento normativo global sobre el tema que cuenta con la firma de más de ciento noventa países. Este documento establece principios como la proporcionalidad, la seguridad, la no discriminación, la sostenibilidad ambiental, la privacidad, la transparencia, la rendición de cuentas y la supervisión humana. Además de los principios, la Recomendación plantea valores fundamentales: el respeto por los derechos humanos, la protección de los ecosistemas, la inclusión y la paz.

Uno de los aspectos más destacables del documento es que cuenta con herramientas prácticas para su implementación. Entre ellas, se encuentran las evaluaciones de impacto ético (EIA), que permite evaluar los riesgos antes y después de que el sistema se lance al público incluyendo todo el proceso de diseño, desarrollo e implementación de un sistema de IA. Brinda un análisis detallado del mismo en materia de Derechos Humanos así como también recomendaciones para mejorarlo. Por su parte, la RAM (Readiness Assessment Methodology) es una herramienta clave para apoyar a los Estados Miembros en la implementación de la Recomendación ya que proporciona información detallada y completa sobre las diferentes dimensiones de la preparación para la IA. Ayuda a resaltar las brechas institucionales y regulatorias y permite a la UNESCO adaptar el apoyo a los gobiernos para llenar esas brechas, a fin de garantizar un ecosistema de IA ético en línea con la Recomendación.

Finalmente, la plataforma “Women for Ethical AI” promueve la inclusión de mujeres en los procesos de diseño y regulación de tecnologías emergentes.

2.4. La Unión Internacional de Telecomunicaciones y la iniciativa “AI for Good”

La Unión Internacional de Telecomunicaciones (UIT), agencia especializada de la ONU en tecnologías de la información y la comunicación, impulsa desde 2017 la plataforma “AI for Good”. Esta iniciativa promueve el uso de la inteligencia artificial para acelerar el cumplimiento de los Objetivos de Desarrollo Sostenible (ODS), incentivando la colaboración entre gobiernos, organismos internacionales, empresas tecnológicas, academia y sociedad civil. Su objetivo principal es articular soluciones de IA que contribuyan al bien público global, en áreas como salud, educación, cambio climático, seguridad alimentaria y reducción de desigualdades.

“AI for Good” funciona como un espacio de encuentro donde expertos de múltiples sectores presentan casos de uso, comparten investigaciones y debaten regulaciones emergentes. A través de cumbres anuales, publicaciones y talleres técnicos, la UIT busca facilitar una gobernanza multilateral de la IA que sea innovadora, ética y centrada en las personas. Uno de los aportes más relevantes de esta iniciativa ha sido visibilizar el potencial transformador de la IA cuando se alinea con valores universales y principios de justicia social (ITU, 2023).

2.5. La OMS: Ética de la IA en salud y grandes modelos multimodales

La Organización Mundial de la Salud (OMS) también ha expresado preocupación por el impacto de la IA en el ámbito sanitario. En 2021 publicó su primera guía sobre ética y gobernanza de la inteligencia artificial en salud, y en 2023 amplió sus lineamientos para abordar

los desafíos éticos de los grandes modelos multimodales (LMMs), como los sistemas de lenguaje de gran escala.

Entre los usos potenciales de estos modelos en salud, la OMS identificó aplicaciones clínicas (diagnóstico y atención), orientación al paciente, tareas administrativas, formación médica y desarrollo farmacéutico. A su vez, advirtió sobre riesgos como el sesgo en los datos, la desinformación, la falta de transparencia y la pérdida de supervisión humana. Por ello, aboga por marcos éticos y regulatorios robustos que garanticen la seguridad, la equidad y la confidencialidad de los datos médicos (WHO, 2023).

3. Otras iniciativas internacionales y su aporte a la gobernanza ética de la IA

Más allá del sistema de Naciones Unidas, otros organismos internacionales han avanzado en propuestas normativas, observatorios y guías éticas sobre IA, contribuyendo a consolidar un ecosistema global de gobernanza.

3.1. La OCDE y el Observatorio de Políticas de IA

La Organización para la Cooperación y el Desarrollo Económicos (OCDE) lanzó en 2019 el “AI Policy Observatory” (OECD.AI), un repositorio digital que reúne marcos regulatorios, políticas públicas, datos e investigaciones relacionadas con la inteligencia artificial en más de 60 países. Este observatorio tiene como objetivo facilitar el acceso a información comparada, promover buenas prácticas y ofrecer asistencia técnica a los Estados miembros.

La OCDE también fue pionera en la adopción de los “Principios de IA” (2019), reconocidos como una referencia normativa clave a nivel internacional. Estos principios abogan por una IA que sea robusta, segura, centrada en el ser humano, transparente, responsable y que promueva el bienestar inclusivo y sostenible.

3.2. La Unión Europea: del AI Act a las Directrices Éticas

La Unión Europea ha desarrollado uno de los marcos regulatorios más ambiciosos a nivel global. El “AI Act” establece un sistema de clasificación de riesgos que obliga a los desarrolladores de sistemas de IA a cumplir con requisitos específicos según el nivel de riesgo que representen. Las aplicaciones de “alto riesgo” deberán cumplir con medidas de prevención y pasar evaluaciones estrictas en materia de transparencia, supervisión humana, gestión de datos, ciberseguridad y respeto a los derechos fundamentales.

Este enfoque regulatorio se complementa con las “Directrices Éticas para una IA Fiable” publicadas por el Grupo de Expertos de Alto Nivel en Inteligencia Artificial de la Comisión

Europea (2019). Las directrices insisten en tres componentes fundamentales: legalidad, ética y solidez técnica. Además, promueven la participación activa de todos los sectores interesados en el diseño de sistemas de IA (European Commission, 2019).

4. Desafíos para América Latina y el Caribe en la gobernanza de la inteligencia artificial

América Latina y el Caribe enfrentan desafíos particulares en el camino hacia una gobernanza ética de la inteligencia artificial. La región se caracteriza por profundas desigualdades socioeconómicas, limitaciones en infraestructura digital, disparidades educativas, baja inversión en investigación y desarrollo, y una escasa presencia en los foros internacionales donde se definen los estándares tecnológicos globales.

Uno de los principales obstáculos es la brecha digital. Según datos de la CEPAL (2025), aquella forma parte de las desigualdades estructurales que caracterizan los modelos de desarrollo en América Latina y el Caribe. Si bien se han registrado avances en conectividad y acceso a dispositivos tecnológicos, persisten desigualdades significativas en el acceso, el uso efectivo de herramientas digitales y, sobre todo, en las habilidades necesarias para aprovechar plenamente las oportunidades que ofrece la era digital.

A esta brecha tecnológica se suma una brecha de capacidades. Muchos países latinoamericanos carecen de marcos regulatorios adecuados para el uso de IA, así como de recursos humanos especializados en ética, programación, análisis de datos, auditoría algorítmica o derecho digital. En este contexto, se corre el riesgo de importar tecnologías desarrolladas en otras regiones sin capacidad crítica para evaluar sus impactos locales, reproduciendo lógicas de dependencia tecnológica.

Además, en la región predominan sistemas políticos con alta volatilidad institucional, limitada coordinación interministerial y escasa articulación con la sociedad civil y la academia. Esto dificulta la elaboración de políticas públicas consistentes que integren la ética en el diseño e implementación de tecnologías emergentes. También existe una tendencia a priorizar la digitalización sin una adecuada evaluación de riesgos éticos y sociales, lo que puede derivar en decisiones apresuradas o guiadas únicamente por criterios de eficiencia.

5. Propuestas para una agenda latinoamericana de gobernanza ética de la IA

Frente a estos desafíos, es urgente diseñar una agenda regional que articule principios éticos con políticas públicas concretas, adaptadas al contexto latinoamericano y caribeño. Algunas propuestas clave incluyen:

5.1. Construcción de marcos normativos con enfoque de derechos humanos

Los marcos regulatorios sobre IA deben estar alineados con los estándares internacionales de derechos humanos, incluyendo los tratados interamericanos y los compromisos adquiridos ante organismos multilaterales. Esto implica integrar principios como la no discriminación, la privacidad, la participación ciudadana, la rendición de cuentas y la justicia social. Las leyes deben ser claras, inclusivas y aplicables tanto al sector público como al privado, garantizando mecanismos de reparación en caso de vulneraciones.

5.2. Promoción de alfabetización digital y ética tecnológica

La educación juega un rol central en la construcción de una ciudadanía crítica y empoderada ante la inteligencia artificial: Se propone desarrollar programas de alfabetización digital y ética tecnológica desde edades tempranas, así como capacitar a funcionarios públicos, periodistas, jueces, educadores y activistas en temas como sesgo algorítmico, transparencia, privacidad y gobernanza de datos. La educación con perspectiva de género, interculturalidad y derechos humanos debe ser eje transversal de estas iniciativas.

5.3. Fortalecimiento de capacidades institucionales y creación de organismos de supervisión

Es fundamental contar con instituciones capaces de evaluar y fiscalizar el uso de IA, tanto en el ámbito estatal como en el empresarial. Para ello, se sugiere la creación de agencias nacionales de supervisión ética de tecnologías emergentes, que puedan emitir dictámenes, realizar auditorías, acompañar procesos de implementación y garantizar la transparencia algorítmica. Estas agencias deberían operar de forma autónoma, con participación multi-actoral y enfoque preventivo (Danesi & Torres 2024).

5.4. Impulso a la cooperación regional e internacional

La gobernanza ética de la IA requiere colaboración entre países para compartir experiencias, armonizar estándares, desarrollar infraestructura común y reducir asimetrías. En este sentido, América Latina y el Caribe pueden aprovechar espacios de diálogo como el SELA, CEPAL, MERCOSUR, CARICOM o la Alianza del Pacífico para construir posiciones regionales conjuntas. Asimismo, es crucial fortalecer la participación en foros globales como la UNESCO, la UIT, la OCDE o el G20, para incidir en la definición de estándares éticos universales desde una mirada del Sur Global.

5.5. Apoyo a la innovación tecnológica inclusiva y responsable

El desarrollo de soluciones tecnológicas locales con enfoque ético debe ser incentivado mediante políticas de fomento a la investigación, financiamiento público y alianzas entre

academia, sector privado y sociedad civil. Es clave apoyar a startups y emprendimientos sociales que utilicen IA para resolver problemas estructurales de la región, como la desigualdad educativa, el acceso a la salud, la gestión ambiental o la mejora de servicios públicos. Este impulso debe estar orientado por valores de inclusión, sostenibilidad y justicia distributiva.

6. El rol de actores no estatales en la gobernanza ética de la inteligencia artificial

Además de los Estados y organismos multilaterales, otros actores tienen un papel decisivo en la configuración de una inteligencia artificial ética. La gobernanza algorítmica no puede entenderse únicamente como una función del gobierno, sino como un proceso colectivo en el que participan empresas tecnológicas, universidades, centros de investigación, organizaciones de la sociedad civil, medios de comunicación y comunidades afectadas.

6.1. Empresas tecnológicas y responsabilidad social corporativa

Las grandes empresas del sector tecnológico (GAFAM, startups, proveedores de IA) deben asumir compromisos éticos que vayan más allá del cumplimiento legal. Esto incluye implementar evaluaciones de impacto ético, publicar informes de transparencia, establecer mecanismos de gobernanza interna para la IA, y diseñar algoritmos que respeten los principios de equidad y no discriminación. La rendición de cuentas empresarial debe ser una exigencia regulatoria y una práctica estándar.

6.2. Sociedad civil y vigilancia ciudadana

Las organizaciones de la sociedad civil juegan un rol clave en la defensa de derechos ante posibles abusos tecnológicos. Su labor incluye la denuncia de casos de discriminación algorítmica, la elaboración de informes de incidencia, la formación de redes de vigilancia ciudadana, y la promoción del acceso abierto a los datos y a los sistemas de decisión automatizada. Una gobernanza democrática de la IA requiere el reconocimiento y fortalecimiento de estos espacios.

6.3. Academia y producción de conocimiento contextualizado

Las universidades e institutos de investigación deben ser protagonistas en el desarrollo de marcos éticos, metodologías de evaluación, sistemas de auditoría y herramientas pedagógicas adaptadas a los contextos regionales. La producción científica en el campo de la IA debe orientarse a resolver problemas sociales y ambientales, respetando la diversidad cultural y epistemológica de la región.

6.4. Medios de comunicación y construcción de narrativas críticas

Los medios cumplen una función pedagógica fundamental al informar a la ciudadanía sobre los riesgos y oportunidades de la IA. También tienen la capacidad de instalar agendas, visibilizar casos de vulneración de derechos, y promover debates públicos sobre tecnología y ética. Es necesario fomentar un periodismo especializado en ciencia y tecnología, así como estrategias de alfabetización mediática.

Conclusiones: hacia un pacto social global por una IA ética y justa

La gobernanza ética de la inteligencia artificial es uno de los desafíos más complejos y urgentes del siglo XXI. No se trata solo de regular una tecnología, sino de decidir qué tipo de sociedad queremos construir en un contexto de automatización creciente y transformación digital acelerada. Frente a los riesgos de concentración de poder, vigilancia masiva, reproducción de sesgos y exclusión digital, es necesario afirmar con claridad que la IA debe estar al servicio de la dignidad humana, la justicia social y la sostenibilidad del planeta.

Las Naciones Unidas y otros organismos multilaterales han dado pasos importantes en esta dirección, promoviendo principios comunes, recomendaciones normativas y espacios de cooperación. No obstante, la tarea es inmensa y requiere el compromiso activo de todos los sectores sociales. América Latina y el Caribe, con su diversidad cultural, potencial creativo y tradición de lucha por los derechos humanos, tiene mucho que aportar en este proceso. Avanzar hacia una IA ética y centrada en las personas exige construir un nuevo pacto social global que reconozca la tecnología como un bien común, regule su uso en función del interés público, y garantice que nadie quede atrás en la revolución digital. La ética no puede ser una aspiración abstracta ni una estrategia de marketing: debe ser el principio rector de toda innovación tecnológica.

Bibliografía

- Danesi, C. and Torres, M. "Artificial Intelligence & Foreign Affairs: reflections and general overview on AI, human rights, ethics and global governance". Peter Lang (2025).
- Danesi, C. and Torres, M. "Auditing Algorithms as a way to prevent Algorithmic Bias.", policy brief, G20 Brasil (2024) https://t20brasil.org/media/documentos/arquivos/TF05_ST_05_Auditing_algorithms66e080b615311.pdf
- Dwivedi, D. Algorithmic Bias: A Challenge for Ethical Artificial Intelligence (AI). In: Heggde, G.S., Patra, S.K., Panda, R. (eds) Immersive Technology and Experiences. Palgrave Macmillan, Singapore (2024). https://doi.org/10.1007/978-981-99-8834-1_5.

- European Commission. (2019). *Ethics Guidelines for Trustworthy AI*. High-Level Expert Group on AI. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- Floridi, L., Cowls, J., Beltrametti, M. et al. AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds & Machines* 28, 689–707 (2018). <https://doi.org/10.1007/s11023-018-9482-5>.
- ITU. (2023). *AI for Good Global Summit*. <https://aiforgood.itu.int>
- OECD. (2019). *OECD Principles on Artificial Intelligence*. <https://www.oecd.org/going-digital/ai/principles>
- ONU Mujeres. (2023). *Conclusiones acordadas de la CSW67*. <https://www.unwomen.org>
- UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. <https://unesdoc.unesco.org>
- United Nations. (2023). *Interim Report of the High-Level Advisory Body on Artificial Intelligence*. <https://www.un.org/en/ai-advisory-body>
- WHO. (2023). *Ethics and governance of artificial intelligence for health: guidance on large multi-modal models*. <https://www.who.int>

La Inteligencia Artificial y la Geopolítica

GD Jesús Rafael Argumosa Pila

Director de la Cátedra de Geopolítica y

Estudios Estratégicos del

European Institute of International Studies (Suecia)

La Inteligencia Artificial (IA) se ha convertido en el tema crucial de la geopolítica actual. Su capacidad para influir en la opinión pública, desacreditar pruebas reales y manipular procesos electorales ha generado preocupación en todo el mundo. La IA generativa, como el ChatGPT, puede ser utilizado por actores nacionales y extranjeros para difundir desinformación y socavar la confianza en la democracia.

Hoy está fuera de duda de que la Inteligencia Artificial proporciona un conjunto de beneficios y ventajas tecnológicas, especialmente en los campos de la eficiencia, innovación, creatividad, imaginación y seguridad disminuyendo el margen de error humano, optimizando los procesos en diversas operaciones y tareas, agilizando la toma de decisiones que se pueden aplicar y ya se aplican en múltiples sectores de la actividad humana.

Pero también la implantación de la IA puede producir una serie de riesgos y desafíos entre los que podemos destacar la desinformación y la manipulación de la opinión pública a través de la IA generativa; el uso de IA para influir en los procesos electorales y socavar la confianza en la democracia; la falta de regulación y control sobre el desarrollo y el uso de la IA; y la posibilidad de que la IA sea utilizada para fines maliciosos como la creación de armas autónomas o la realización de ataques cibernéticos.

Para que esto no ocurra se están estableciendo iniciativas y normas por diferentes instituciones como la Ley de Inteligencia Artificial de la Unión Europea sobre la regulación de la IA, que entró en vigor en agosto de 2024; el informe de un grupo de expertos de Naciones Unidas, del pasado mes de septiembre de 2024, afirmando que es necesario

regular a escala mundial el floreciente campo de Inteligencia Artificial, añadiendo que el desarrollo y el uso de esta tecnología no puede dejarse al capricho de los mercados, en tanto que algunos países, como Italia, han prohibido el uso del ChatGPT debido a preocupaciones sobre la privacidad y la seguridad.

La IA constituye una herramienta geopolítica importante, con Estados Unidos y China como líderes en su desarrollo. Sin embargo, otros actores geopolíticos también están desarrollando capacidades de Inteligencia Artificial avanzadas, lo que puede originar una mayor competencia y tensión en la escena internacional.

Consideraciones generales

Actualmente ChatGPT es la *app* más descargada del mundo. Durante el primer trimestre de 2025, ChatGPT tuvo un crecimiento explosivo del 148% en descargas respecto al mismo periodo del año anterior. Todo apunta a que la implementación de funciones creativas, como la generación de imágenes, fue clave para atraer a una nueva oleada de usuarios, muchos de los cuales desconocían por completo las capacidades del chatbot.²⁶

Este crecimiento no solo refleja su éxito en descargas, sino también en usuarios activos y suscripciones de pago. Con 151,5 millones de usuarios activos, la aplicación ha entrado de lleno en la vida cotidiana de millones de personas. Y en un panorama saturado de herramientas impulsadas por IA, ChatGPT es, para muchos, sinónimo de IA, dejando en un segundo plano a competidores como Grok o DeepSeek. Si sigue creciendo de esta manera, la IA de OpenAI puede convertirse en un fenómeno cultural.

En la Cumbre del Futuro que celebró Naciones Unidas en el mes de septiembre de 2024, una de las cuestiones más sobresalientes fue la aprobación del Pacto Mundial Digital del Futuro, que es el primer Acuerdo verdaderamente mundial sobre la regulación de la Inteligencia Artificial y se basa en la idea de que la tecnología debe beneficiar a todos.

El Pacto pretende reducir la brecha digital y garantizar que las tecnologías de IA se utilicen de forma responsable, fomentando la cooperación mundial tanto en las capacidades de IA como en las amenazas a la seguridad. Los gobiernos también quedaron comprometidos a formar un Panel Científico mundial imparcial sobre IA y a iniciar una conversación internacional sobre la gobernanza de la IA en el seno de la ONU.²⁷

²⁶ www.hd-tecnologia.com/chatgpt-se-convierte-en-la-app-mas-descargada-del-mundo. Datado 16-04-25

²⁷ Microsoft PowerPoint - Summit of the Future deck_October2024_23Oct. Datado 16-04-25

Es evidente que la IA ha dado un paso significativo después de pasar de algoritmo con reglas de funcionamiento predefinidas a modelos generativos avanzados. Estos algoritmos no tienen conciencia, ni sentido de trascendencia, ni intenciones de tomar decisiones autónomas como las personas.

Hasta ahora, la IA ha estado dominada por Estados Unidos y China, pero tanto este año como en el futuro cercano nuevos actores geopolíticos serán capaces de desarrollar y adquirir capacidades de IA avanzadas, para el uso de la seguridad nacional. Es comprensible que despierte tanto interés como para desatar luchas ya arraigadas, protagonizadas por todos los agentes involucrados: está en juego su dimensión ética y legal en la medida en que implanta la relación persona-máquina que se ha estado viviendo en los últimos años. Por desgracia, no todos los usos son buenos, y los ciberdelincuentes aprovechan las capacidades para cometer ataques y engaños.

El año pasado, grandes instituciones financieras como Goldman Sachs, Sequoia Capital, Moody o Barclays, publicaron informes poniendo en duda la rentabilidad de las enormes infraestructuras que demandan las empresas de IA, con promesas que no se han materializado hasta ahora. De hecho, el pasado 27 de enero, se presentó el modelo de IA chino denominado *DeepSeek* que ha dado la prueba que se necesitaba para demostrar que para entrenar grandes modelos no se requieren 100.000 tarjetas Nvidia H100s ni una arquitectura faraónica en continua expansión. Es decir, se pueden conseguir modelos modernos y eficientes con muchos menos costes.

Existe una estrecha y profunda relación entre Inteligencia Artificial, la geopolítica y el poder. El uso de la IA como herramienta de poder proporciona a los países que lideran esta tecnología disruptiva ventajas estratégicas superiores en espionaje, ciberseguridad y economía basada en datos. Asimismo, la posibilidad de controlar las infraestructuras digitales como chips, redes o plataformas aumenta fuertemente la influencia y poderío de los actores estratégicos.

Los distintos modelos de agentes de IA tendrán implicaciones muy diferentes. Con un enfoque de la IA como asesor, los agentes ofrecerían recomendaciones calibradas individualmente a los responsables humanos del asunto, dejando a estos siempre la toma de decisiones. Pero con el modelo de IA autónoma, los autómatas tomarán la decisión en nombre de los seres humanos. Se trata de una distinción profunda con repercusiones de largo alcance.

En la actual competencia geoestratégica global, Estados Unidos y China compiten por liderar el desarrollo de la IA generativa a escala global mientras Europa avanza en la investigación

de esta tecnología²⁸. En general, la geopolítica de la Inteligencia Artificial sitúa a las principales potencias y economías del planeta en una carrera que redefinirá el equilibrio mundial. Uno de los aspectos de esta guerra tecnológica lo constituye los intentos de acceso a los semiconductores producidos por la Taiwán Semiconductor Manufacturing Company (TSMC).

La citada competencia entre Estados Unidos y China por el liderazgo en Inteligencia Artificial no es solo una carrera tecnológica, sino una lucha por la supremacía geopolítica. Con Estados Unidos centrado en la innovación liderada por el sector privado y China impulsando un enfoque más centralizado, ambos países están definiendo el futuro de la IA, con implicaciones que se extienden más allá de sus fronteras.

Análisis geopolítico

La Universidad de Stanford, ha publicado recientemente que, en 2024, las inversiones estadounidenses en IA ascendieron a 109.100 millones de dólares, casi 12 veces más los 9.300 millones de dólares de China y 24 veces más los 4.500 millones de dólares del Reino Unido. Dicha información señala que la IA generativa experimentó un impulso especialmente fuerte al año pasado atrayendo 39.000 millones de dólares a nivel mundial en inversión privada, un aumento del 18,7% con respecto a 2023.²⁹

En general, resulta sorprendente que desde 2013 hasta el presente año se han recaudado cerca de medio billón de dólares para IA en Estados Unidos siendo esta cantidad mayor que la del resto del mundo combinado - 471.000 millones de dólares frente a 289.000 millones de dólares - de acuerdo con los datos de la Universidad de Stanford. En dicho periodo, hay cuatro países europeos, Reino Unido, Alemania, Suecia y Francia, como los mayores inversores en IA.

Continúa el informe señalando que, en 2024, Estados financió 1.073 empresas de IA, Reino Unido lo hizo con 116 empresas, China financió a 98 e India a 74. Canadá, Israel, Corea del Sur, Singapur o Japón están entre los mayores inversores en IA, pero con enorme diferencia frente a los líderes mundiales - Japón invirtió 7.000 millones en los últimos 11 años -. Brasil es el único país de Iberoamérica que se acerca a la inversión de estos últimos países.

En la Unión Europea hay cerca de 7.000 nuevas empresas trabajando con IA. El problema es que son tan pequeñas que no tienen capacidad. Por eso se quiere reunir las e invertir en

²⁸ La geopolítica de la inteligencia artificial • The OMS José Suárez de Lezo. 03-06-24

²⁹ Los países del mundo que más invierten en IA, ilustrados en un gráfico dominado por EEUU.

nuestros superordenadores. El objetivo es no perder el tren en la IA, un campo en el que Europa se está quedando atrás.³⁰

En relación con el futuro mundo multipolar que se avecina, la IA tiene capacidad para reforzar o crear hegemonías, redistribuir el poder cuando aparecen nuevos actores geopolíticos como los que se acaban de mencionar que adoptan y utilizan estratégicamente esta tecnología.

La IA como infraestructura estratégica global está dejando de ser una herramienta tecnológica para convertirse en una infraestructura crítica, comparable a la energía o a las telecomunicaciones. Esto implica una dependencia estructural de las economías, de los gobiernos y de los sistemas militares que dependen de IA para su funcionamiento, así como el control y dominio de los algoritmos fundamentales - modelos funcionales, redes de chips o datos globales - que pueden ejercer una gran influencia global sin necesidad de presencia física.

Un ejemplo lo tenemos en la distinta forma de operar en su objetivo de conseguir la superioridad en el sistema IA en la competencia entre Estados Unidos y China. Mientras China invierte fuertemente en IA militar y vigilancia social, Estados Unidos domina las plataformas digitales globales y chips avanzados. Esto genera una tensión estructural similar a la Guerra Fría, pero con armas invisibles: datos, algoritmos y procesadores.

Es indudable que el orden mundial se está moviendo hacia la multipolaridad algorítmica tecnológica, en la que varios centros de poder compiten no solo por la hegemonía territorial o económica, sino por definir las reglas de la IA. Estados Unidos establece un ecosistema de innovación privada - Open AI, Google, Nvidia, etc... -; China utiliza un enfoque estatal con integración virtual: datos, aplicaciones, gobernanza; la Unión Europea busca el liderazgo ético y regulatorio de la IA, pero sin paso tecnológico proporcional; y otros actores emergentes como India, Canadá e Israel y potenciales alianzas Sur-Sur caminan por sistemas tecnológicos mixtos.

En la era de la IA, el poder duro - capacidad militar y coercitiva - se transforma en defensa cibernética autónoma, drones y robots armados con autonomía selectiva, vigilancia predictiva en la línea del modelo chino, y en inteligencia estratégica basada en big data + IA. En cuanto al poder blando - influencia cultural, normativa y psicología - se redefine en la manipulación informativa con IA generativa, en plataformas digitales que moldean la opinión pública global y por el colonialismo algorítmico mediante la exportación de sistemas IA como forma de influencia. Todo ello crea un campo de batalla híbrido donde no hay fronteras claras entre guerra y paz, lo civil y lo militar, lo nacional y lo extranjero.

30 Virkhunen H. Vicepresidenta de Soberanía Digital de la Unión Europea. El País (11-04-2025)

Entre los riesgos sistémicos que emergen en este nuevo escenario destacan, principalmente, la concentración de poder ya que un puñado de empresas y gobiernos controlan los modelos más avanzados de IA; en el campo de la dependencia estructural los Estados más débiles no pueden operar sin sistemas de IA externos; dentro de la capacidad algorítmica, las decisiones se vuelven incomprensibles y no auditables, incluso para sus propios creadores; y en el entorno de la desigualdad global, la brecha entre los países que desarrollan IA y los que solo la consumen se amplía, generando nuevos conflictos.

Se puede considerar a la IA como el *pilar central del poder global* ya que, en el futuro cercano, la IA se convertirá en uno de los principales factores determinantes de poder a nivel mundial, de la misma forma que hoy lo son los recursos energéticos, la economía o las Fuerzas Armadas. Los países que lideren el desarrollo y la implementación de la Inteligencia Artificial ostentarán un elevado poder estratégico, especialmente en los tres campos que se consideran a continuación

En la *capacidad de decisión y análisis*, los países con sistemas de IA más avanzados tendrán la posibilidad de tomar decisiones más rápidas y precisas, lo que permitirá dominar sectores clave como la economía, la política exterior y la defensa.

En el *control de la información*, las tecnologías de IA, como el procesamiento de lenguaje natural (PLN) y el aprendizaje automático, jugará un papel fundamental en la gestión de la información y la desinformación. Aquellos países que controlen las plataformas de IA podrán influir en las narrativas globales y moldear la opinión pública de manera más efectiva.

En el *desarrollo tecnológico*, la IA será un motor principal para innovaciones en áreas como biotecnología, la nanotecnología y la energía, lo que podría reconfigurar las relaciones internacionales basadas en la competencia por recursos naturales y tecnología.

En resumen, la IA no solo está redefiniendo la forma en que operan los Estados, sino también quién tiene poder, cómo se ejerce y qué se considera legítimo. Estamos entrando en una era donde la geopolítica se vuelve computacional y donde los futuros conflictos no serán por tierra o recursos, sino por datos, chips, y soberanía algorítmica.

Visión futurista

En relación con posibles escenarios futuros en el entorno 2040-2050, se pueden considerar los señalados en la siguiente tabla:

Escenario	Características principales	Riesgos clave
Hegemonía tecnológica	Una o dos potencias dominan el desarrollo y uso global de IA.	Supresión de alternativos <i>dependencia total</i> .
Multipolaridad regulada	Varias potencias cooperan en normas comunes.	Dificultad de implantación y vigilancia.
Anarquía algorítmica	IA se descentraliza sin control total.	Riesgos sistémicos, conflictos por errores o fallos.
Rebote humanista	Reacción ética limitada al uso de IA en ciertas áreas	Ralentización progresiva fragmentación normativa.

Fuente: Elaboración propia.

En el nuevo mapa geopolítico que se está vislumbrando, las grandes potencias están inmersas en una carrera por dominar las tecnologías de IA, con Estados Unidos, China y la UE como los principales contendientes. No obstante, en una visión futurista, se podría encontrar un panorama más diversificado, en el que China podría consolidarse como líder en IA debido a su modelo de desarrollo de tecnologías basadas en la cooperación entre el gobierno y las empresas tecnológicas, además de su amplio banco de datos proveniente de su población.

Estados Unidos seguirá siendo un líder clave en la innovación debido a su ecosistema de investigación avanzada y empresas tecnológicas de vanguardia, aunque podría enfrentarse a desafíos internos por la regulación y la ética de la IA. En cuanto a la Unión Europea podría establecerse como un regulador global de la IA, con su enfoque en la ética, la privacidad y los derechos humanos, creando marcos regulatorios que podrían influir en cómo se implementa la IA en todo el mundo.

En el mundo de la transformación militar y de la guerra del futuro, la IA tendrá un impacto profundo desde el punto de vista geopolítico. Ya se están viendo avances en la automatización de las armas, en las armas de energía dirigida y, en general, en los sistemas de defensa. En un futuro no muy lejano podemos ser testigos de unos modernos procedimientos y modos de combatir productos de una transformación completa de las dinámicas de la guerra.

Así, no es un tema menor tener en cuenta que con los robots de combate, drones y sistemas de armas autónomas se podría cambiar la naturaleza de los conflictos, permitiendo a las

naciones librar guerras más precisas, pero también más letales. El debate sobre la regulación y la ética de las armas autónomas será crucial para definir los nuevos tipos de guerra que se perciben en el cercano horizonte.

Si tratamos el campo de la ciberseguridad y ciberdefensa, es necesario señalar que con la IA mejorando las capacidades de defensa cibernética, los países con la mejor IA serán capaces de proteger sus infraestructuras críticas de ciberataques. Sin embargo, esto también abre la puerta a nuevas formas de guerra cibernética, donde los ataques invisibles pueden desestabilizar economías enteras.

En cuanto a las estrategias predictivas, las potencias podrían utilizar la IA para predecir y anticipar los movimientos de los adversarios, basándose en grandes volúmenes de datos, lo que cambiaría las estrategias defensivas y ofensivas. Los algoritmos predictivos serán tan importantes como las armas físicas en la geopolítica militar.

Los actores en la geopolítica del futuro no serán solo los Estados-nación. La IA dará lugar a nuevos actores que influirán en el poder global, reconfigurando las relaciones internacionales, con la inclusión de nuevos tipos de actores y nuevas alianzas, donde sobresalen las entidades e instituciones señaladas a continuación:

Las principales *Organizaciones Internacionales (OOII)* existente suponen un prototipo ya muy consolidado. Las OOII como la ONU, la OMC o la OTAN tendrán que adaptarse a esta nueva realidad creando marcos regulatorios y celebraciones internacionales para gestionar el impacto global de la IA, la ética de la IA y las tensiones geopolíticas que surjan a partir de ella.

Las *corporaciones tecnológicas* constituyen otro tipo de nuevos actores. Empresas como Google, Microsoft, Amazon, Alibaba y Tencent tienen la capacidad de acumular tanto poder como los propios gobiernos debido a su control sobre los datos, el desarrollo de la IA y la infraestructura digital global.

Las *Redes descentralizadas* junto a actores no estatales conforman un tercer tipo. Las criptomonedas, la blockchain y las tecnologías de IA descentralizadas pueden permitir a actores no estatales (ONG,s, grupos extremistas, o incluso individuos) desafiar al poder tradicional de los Estados. Esto podría llevar a un mundo más fragmentado y menos predecible.

Conclusiones

La relación entre la IA, la geopolítica y el poder está evolucionando a una velocidad impresionante, y su intersección promete cambiar la forma en que las naciones

interactúan entre sí, tanto a nivel económico como militar, social y cultural. En una visión futurista, podríamos estar ante un mundo donde la IA redefine el equilibrio de poder global, introduce nuevos actores en la escena internacional y provoca disrupciones en las estructuras de poder tradicionales.

En el próximo futuro existirá una gran competencia entre los países que participan estratégicamente en el uso y desarrollo de la IA, con independencia del liderazgo competitivo que actualmente se está desarrollando entre Estados Unidos y China. Otros actores estratégicos como la UE, India, el Reino Unido, Canadá, Israel, Singapur, Japón y Brasil tendrán una especial influencia en la nueva configuración geoestratégica de poder mundial que se avecina.

La IA constituye un instrumento fundamental y decisivo en la actual y previsible geopolítica. Con independencia de los beneficios y ventajas que aporta, su desarrollo y uso deben ser regulados y controlados para prevenir riesgos y desafíos. Es importante que los gobiernos, las organizaciones internacionales y la sociedad civil trabajen juntos para establecer normas y regulaciones que garanticen el uso responsable y ético de la Inteligencia Artificial.

A medida que la IA revoluciona la geopolítica y se avanza en un mundo cada vez más multipolar, las superpotencias, grandes potencias y potencias emergentes, jugarán papeles importantes en la toma de decisiones globales. Este escenario también podría llevar a un mundo más interconectado y cooperativo en áreas específicas, como la regulación de la IA, la lucha contra el cambio climático y el control de la proliferación de armas atómicas.

Desde el punto de vista geopolítico, la IA tendrá un fuerte impacto en la guerra del futuro. Con la disposición de sistemas de armas autónomas, muchas de ellas de carácter mortal, se puede mudar la naturaleza y los procedimientos de los conflictos, posibilitando a los países librar guerras más precisas y con mayor letalidad, a grandes distancias y eliminando bajas propias, de impredecibles consecuencias para la seguridad internacional.

En conclusión, la IA tiene el potencial de transformar, de una forma radical, la geopolítica y la distribución del poder en el futuro, pero también plantea enormes desafíos éticos, económicos y sociales que se debe abordar colectivamente si se quiere evitar un mundo fragmentado y desigual. La clave será encontrar un equilibrio entre el desarrollo tecnológico y la creación de normas globales que promuevan la cooperación internacional y la paz.

Su influencia en geopolítica irá en aumento en los próximos años. La IA está llamada a influir en el equilibrio de poder mundial y en la relación entre los Estados. Constituye un factor estratégico que puede acelerar la concentración de poder y favorecer la desigualdad económica.

Diplomacia en la era de la inteligencia artificial: Nuevos desafíos y oportunidades

Emb. José Beraún Aranibar

European Institute of International Studies (Suecia)

En primer lugar, para comprender la diplomacia en la era de la inteligencia artificial debemos retroceder y mirar que visiones sobresalían de la diplomacia en el siglo 20.

Para ello utilizaré a dos autores/intelectuales que son una referencia fundamental para entender cómo se configuró la diplomacia en el siglo pasado. Me refiero a Sir Harold Nicolson (1886 - 1968) y a Hans Morgenthau (1904-1980).

Ambos sostuvieron visiones muy diferentes de lo que debiera ser la diplomacia, pero su impacto y trascendencia a lo largo del tiempo fue y es innegable. También porque apreciamos en el siglo 21 un resurgir de ambas visiones moduladas esta vez por las tecnologías disruptivas dentro de las que destaca la inteligencia artificial.

Sir Harold Nicolson fue una figura clave en la historia de la diplomacia del siglo XX, tanto por su labor como diplomático británico como por su influyente obra teórica. Su obra más influyente es *"Diplomacy"* (1939), donde distingue tres tipos de diplomacia:

- **La diplomacia antigua:** secreta, aristocrática y personalista.
- **La diplomacia democrática moderna:** más abierta, sometida a control público y parlamentario.
- **La diplomacia idealista:** guiada por principios morales, a veces en tensión con los intereses nacionales.

Su influencia en la diplomacia moderna

Tal vez uno de sus principales legados e importancia radica en que Nicolson fue uno de los primeros en reflexionar sistemáticamente sobre la tensión entre los **intereses nacionales** y los **valores internacionales**, anticipando debates fundamentales que apreciamos hoy con intensidad en las relaciones internacionales contemporáneas. Debemos recordar que los principios y valores que han regido el sistema internacional, así como la convivencia entre las naciones se negoció y acordó en diferentes acuerdos y convenios internacionales entre los que destacamos la Carta de las Naciones Unidas. Hoy nuevamente apreciamos ha resurgido esta tensión entre los intereses nacionales y los valores y principios internacionales, así como debe conducirse la diplomacia, si como sostenía Nicolson o como propugnaba Morgenthau.

Hans Morgenthau fue un influyente teórico de las relaciones internacionales y uno de los principales exponentes del **realismo clásico**. Su obra más conocida, *Politics Among Nations* (1948), estableció los fundamentos del realismo político, enfatizando que la política internacional está gobernada por la lucha por el poder y el interés nacional.

Morgenthau argumentaba que los Estados actúan principalmente en función de su **interés nacional definido en términos de poder**, y que la moralidad abstracta no puede aplicarse directamente a la política internacional. También desarrolló los **Seis Principios del Realismo**, que explican cómo la política internacional se basa en leyes objetivas derivadas de la naturaleza humana. Estos principios, son:

1. **La política internacional se rige por leyes objetivas arraigadas en la naturaleza humana.** La política, como la sociedad en general, está gobernada por leyes objetivas arraigadas en la naturaleza humana, que es invariable; por lo tanto, es posible desarrollar una teoría racional que refleje estas leyes objetivas.
2. **El interés nacional se define en términos de poder.** El rasgo principal del realismo político es el concepto de *interés*, definido en términos de poder que infunde un orden racional al objeto de la política, y de ese modo hace posible la comprensión teórica de la política. El realismo político hace hincapié en lo racional, lo objetivo y lo no emocional.
3. **El interés nacional siempre es dinámico.** El realismo asume que el interés definido como poder es una categoría objetiva universalmente válida, pero no con una definición fijada de una vez y para siempre. El poder es el control del hombre sobre el hombre.

4. Los principios morales abstractos no pueden aplicarse a la política internacional.

El realismo político es consciente del significado moral de la acción política. Es también consciente de la tensión entre el control moral y las exigencias de la acción política eficaz. Por lo tanto, el realismo no es inmoral, sino que su objeto de estudio no es la moral.

5. Diferencia entre la aspiración moral de una nación y el principio moral universal.

El realismo político se rehúsa a identificar las aspiraciones morales de una nación en particular con las leyes que gobiernan el universo. Es el concepto de interés definido en términos de poder lo que nos salva de los excesos morales y la torpeza política.

6. La autonomía de la política internacional. El realista político sostiene la autonomía de la esfera política. Se pregunta: ¿cómo afectará esta política el poder de la nación?”. El realismo político está basado en una concepción pluralista de la naturaleza humana. Un hombre pura y exclusivamente político no sería más que una bestia, pues carecería por completo de límites morales. Sin embargo, para desarrollar una teoría autónoma del comportamiento político, el “hombre político” debe abstraerse de los demás aspectos de la naturaleza humana.

Podemos apreciar que Nicolson representa una visión **ética y humanista** de la diplomacia, basada en la transparencia y la confianza mientras que Morgenthau encarna una perspectiva **realista y pragmática**, en la que el poder y el interés nacional son los ejes centrales. También podríamos decir que ambos aportan perspectivas complementarias: Nicolson nos recuerda lo que la diplomacia *debería ser*, mientras que Morgenthau nos advierte lo que la política internacional *realmente es*.

Ahora bien, ambas visiones y perspectivas se desarrollaron a lo largo del siglo 20 y en un entorno internacional que describiremos a continuación. Ello es importante para comprender la irrupción de la inteligencia artificial en la diplomacia y en las relaciones internacionales en el siglo 21.

Comparación entre el entorno internacional en los siglos 20 y 21

El entorno internacional en los siglos 20 y 21 ha experimentado cambios significativos en diversos aspectos. A continuación, presentamos algunas comparaciones entre ambos períodos:

1. Política, relaciones internacionales y diplomacia. Durante el siglo 20, se produjeron importantes conflictos y eventos políticos, como las dos guerras mundiales, la Guerra Fría

y la descolonización de numerosos países. Estos eventos tuvieron un impacto significativo en la configuración del orden mundial y en las relaciones internacionales.

A pesar de los esfuerzos por promover la paz y la cooperación, el siglo 21 ha sido testigo de varios conflictos y tensiones internacionales. Estos incluyen conflictos regionales, como los conflictos en el medio oriente, el conflicto bélico en Ucrania y en Gaza, y desde el inicio de las 2025 tensiones geopolíticas y económicas provocadas por los Estados Unidos y destinadas a prácticamente todos los países del mundo.

2. Economía y globalización. Durante el siglo 20, se produjo un importante crecimiento económico y una mayor integración económica a nivel mundial. Sin embargo, también hubo crisis económicas, como la Gran Depresión de la década de 1930 y la crisis financiera de 2008, que tuvieron un impacto significativo en la economía global.

En el siglo 21, la globalización se intensificó aún más, con un aumento significativo en el comercio internacional, la inversión extranjera y la interconexión global a través de las tecnologías de la información y la comunicación.

3. La era digital y las innovaciones tecnológicas. Durante el siglo 20, se produjeron avances tecnológicos significativos, como la invención de la televisión, la computadora personal y el teléfono móvil. Estos avances transformaron la forma en que las personas se comunicaban y accedían a la información.

El siglo 21 ha sido testigo de avances tecnológicos rápidos y transformaciones digitales en diversos campos, como la inteligencia artificial, la realidad virtual, el big data y el internet de las cosas. Estos avances están teniendo un impacto significativo en la economía, la política, la diplomacia y en la sociedad como un todo. Estos avances han llevado a una mayor conectividad global y han revolucionado la forma en que las personas se comunican, trabajan y acceden a la información.

4. Desafíos globales. Durante el siglo 20, los desafíos globales incluyeron la proliferación nuclear, la carrera armamentista, la pobreza y los conflictos étnicos y religiosos, entre los más significativos.

En el siglo 21, los desafíos globales han evolucionado y se han ampliado. Además de los desafíos tradicionales, también se han destacado problemas como las crisis humanitarias y la migración, las desigualdades globales y la pandemia de COVID-19.

En resumen, todos estos aspectos han moldeado el panorama global y continúan influyendo en la dinámica internacional en la actualidad y en la necesidad de respuestas de la diplomacia

para abordar los desafíos actuales. En ello, la diplomacia constituye la mejor herramienta de trabajo sino acaso la única que disponemos.

Profundizando en el análisis del siglo 21 podemos apreciar que vivimos un sistema internacional definido principalmente **por una incertidumbre sobre el orden existente y su dirección futura**. No solo es la aparición de nuevos actores globales como son las grandes empresas tecnológicas sino, a nivel de países, destaca la competencia por el poder a nivel mundial que los Estados Unidos ha emprendido. Si las nuevas políticas de los Estados Unidos emprendidas en los inicios del 2025 son llevadas a cabo a plenitud veremos que los cambios en el equilibrio de poder se aceleraran y será cada vez más compleja la comprensión de la dinámica y dirección que tomaran las relaciones internacionales.

Un segundo elemento que apreciamos en el sistema internacional del siglo 21 es que **el orden internacional basado en principios y valores fundamentales** que emergieron y se plasmaron, por ejemplo, en la Carta de las Naciones Unidas y en otros instrumentos internacionales y que sustentaron la estabilidad y la prosperidad de gran parte del mundo en el siglo 20, está siendo desafiado por intentos de cambiar unilateralmente el statu quo mediante el uso de la información y desinformación, la propagación del extremismo político y/o violento y también mediante la fuerza y/o la coerción. Todo ello con la participación de la inteligencia artificial.

Un tercer elemento que vemos es **el papel de las tecnologías disruptivas en la comunicación y la toma de decisiones**. La rápida expansión de las **redes sociales** y el acceso a **Internet** ha permitido a los gobiernos y los organismos internacionales comunicarse de manera más rápida y eficiente. Ello hace que la competencia internacional, para tener superioridad tecnológica sea cada vez más intensa, y que cada vez se utilicen más tecnologías como la inteligencia artificial, como herramientas para obtener más poder.

La inteligencia artificial y las dinámicas cambiantes del poder a nivel internacional

La inteligencia artificial está teniendo un impacto significativo en las dinámicas del poder a nivel internacional. Diversos países y actores internacionales están compitiendo por liderar las capacidades generadas por el uso de la Inteligencia Artificial, lo que se presenta como un indicador del presente y futuro liderazgo internacional.

Este avance tecnológico en Inteligencia Artificial está generando desafíos y oportunidades en la competencia a nivel global por el poder.

El desarrollo de la Inteligencia Artificial está modificando las fronteras y el desarrollo formal de las relaciones entre los Estados, lo que repercute en un aumento de eficiencia y

productividad, generando un crecimiento económico que se traduce en un despertar de las dinámicas de poder blando y duro en las relaciones internacionales

En cuanto a la gobernanza nacional e internacional de la Inteligencia Artificial, se están realizando esfuerzos para contrarrestar los riesgos asociados con su uso, como el riesgo de que las máquinas y algoritmos adopten decisiones, el control social de los ciudadanos, y la competencia geopolítica a nivel internacional, incluidas las guerras de información y desinformación.

En la Diplomacia y Política Exterior en los Análisis de Datos la Inteligencia Artificial puede analizar grandes volúmenes de datos para analizar situaciones, hechos, crisis políticas, sociales y económicas, lo que ayuda a la diplomacia a tomar decisiones informadas.

En Ciberseguridad la Inteligencia Artificial se aplica para detectar y contrarrestar ciberataques, así como, para fortalecer la seguridad de infraestructuras críticas, tema fundamental para la diplomacia.

En la Guerra Cibernética los estados utilizan Inteligencia Artificial para desarrollar capacidades ofensivas y defensivas en este ámbito.

En las Políticas Económicas los gobiernos utilizan la Inteligencia Artificial para diseñar y ajustar políticas económicas basadas en análisis predictivos y datos en tiempo real.

En Salud y Educación, la Inteligencia Artificial se utiliza para mejorar los sistemas de salud y educación en los países, facilitando el acceso a recursos y servicios básicos.

Respecto del Cambio Climático, la Inteligencia Artificial ayuda a modelar y predecir los efectos del cambio climático, apoyando la creación de políticas ambientales y acuerdos internacionales.

En la Innovación y la Competencia Tecnológica, los países compiten por el liderazgo en Inteligencia Artificial, lo que influye en las relaciones de poder y la colaboración internacional en investigación y desarrollo.

Entre el derecho internacional y la inteligencia artificial no solo se plantean desafíos y oportunidades. Creemos que constituye uno de los temas prioritarios que debieran ser abordados por todos los países sin excepción. No solo porque el actual sistema internacional se ha construido, se basa y continua gracias a la vigencia y aceptación de principios y valores fundamentales, así como de las normas del derecho internacional. Pero, mirando cuanto

viene sucediendo desde inicios del 2025 en los Estados Unidos, es válido preguntarse si la inteligencia artificial será usada para redefinir y modificar el actual sistema internacional.

Creemos que es fundamental que los países aborden las oportunidades y desafíos que ofrece la inteligencia artificial de manera integral y en cooperación con otros países, para garantizar que su desarrollo y aplicación estén alineados con los principios y valores fundamentales, del derecho internacional y de los derechos humanos.

En resumen:

La Inteligencia Artificial es definitivamente una tecnología disruptiva que está transformando rápidamente las relaciones internacionales, el sistema internacional y por ende la diplomacia, al permitir análisis más rápidos y precisos de las situaciones internacionales, lo que mejora la toma de decisiones. La Inteligencia Artificial mejora la eficiencia y eficacia de las diplomacias nacionales, ayudando a tener respuestas más adecuadas a los desafíos nacionales y globales.

Asistimos a una nueva etapa de la historia de las relaciones internacionales, en donde el rol de los países tiene que ser replanteado, tanto en competencias como en su acción exterior, abriendo un nuevo e inédito capítulo en el sistema internacional. El desafío para los países es dejar de ser observadores y transformarse en actores en este nuevo sistema internacional. La herramienta que disponen los países siempre ha sido la diplomacia. Desde los tiempos de Nicolson y Morgenthau, no importando la visión y los objetivos, la diplomacia ha sido el instrumento, único y fundamental a disposición de los países.

La Inteligencia Artificial está transformando las dinámicas del poder a nivel internacional, generando tanto desafíos como oportunidades en la competencia internacional.

La colaboración y la competencia en el ámbito de la Inteligencia Artificial tendrán un impacto significativo en el futuro no solo de las relaciones internacionales sino del tipo de sistema internacional en el que viviremos.

Las asimetrías entre los países en el campo de la inteligencia artificial son evidentes y representan un desafío tanto para lograr obtener estos avances tecnológicos, así como para competir a nivel internacional en este ámbito.

La creciente desigualdad en el acceso y desarrollo de la inteligencia artificial también puede limitar su aprovechamiento así como la prevención y resolución de los conflictos, al restringir

la inclusividad de los procesos de negociación y la capacidad de la diplomacia de gestionar los flujos de información de forma eficaz y constructiva.

La gobernanza ética de la inteligencia artificial es un tema complejo, urgente y necesario que requiere la colaboración de actores a nivel nacional e internacional para establecer normativas éticas y para abordar los desafíos teóricos y prácticos que surgen en este contexto. Los riesgos inherentes a la inexistencia de una gobernanza ética de la inteligencia artificial son demasiado grandes.

La gobernanza ética de la Inteligencia Artificial va más allá del cumplimiento de normas legales y comprende la necesidad de adherirse a criterios éticos. La gobernanza ética de la Inteligencia Artificial implica establecer marcos y principios que no solo guíen el desarrollo y la implementación de las tecnologías de Inteligencia Artificial, sino que también den prioridad a las consideraciones éticas, la protección de los derechos y la conducta responsable.

Los países de menores recursos y de acceso a esta tecnología, deben buscar participar activamente en la regulación y gobernanza ética de la Inteligencia Artificial, como medio para recibir sus beneficios y mediatizar sus amenazas.

Los países de América Latina y el Caribe aún no han logrado desarrollar ni conectar plenamente sus recursos académicos con el sector público y privado para establecer una sinergia, un trabajo conjunto y coordinado en temas de inteligencia artificial, lo que crea una brecha significativa en comparación con los países líderes mundiales en este campo.

Necesitamos una diplomacia que dé respuesta a los conflictos que la tecnología está amplificando y diversificando, pues de no disponer de una diplomacia efectiva, el nuevo sistema internacional carecerá de una gobernanza efectiva y será cada vez más difícil encontrar espacio para el desarrollo de las naciones menos desarrolladas.

El uso de la Inteligencia Artificial impulsa a los países a repensar sus relaciones internacionales, sus prioridades en materia de política exterior y su diplomacia. Conocer y entender lo que ocurre en el mundo, es una responsabilidad del diplomático y de todos quienes contribuyen al desarrollo políticas adecuadas al mundo actual y futuro, así como la implementación de la política exterior con terceros países, tanto a nivel bilateral, regional, interregional y multilateral. Ellos deben conocer y capacitarse en las tecnologías disruptivas y en particular en las facetas que adopta la Inteligencia Artificial. El desarrollo de estos conocimientos y habilidades profesionales es hoy primordial e inevitable porque su impacto e influencia trasciende sustantivamente.

Finalmente, vemos a la diplomacia como la herramienta clave para gestionar la incertidumbre que apreciamos en el sistema internacional. También vemos a la diplomacia como mediadora entre valores, principios, intereses y poder. Todo ello en un contexto en el que apreciamos la necesidad de adaptar los principios clásicos de la diplomacia que construimos en el siglo 20 a un nuevo concepto – aún por definir - en el que la inteligencia artificial participa activamente.

Lecturas recomendadas

Sobre teoría diplomática y relaciones internacionales clásicas

1. Nicolson, Harold. *Diplomacy*. Institute for the Study of Diplomacy. 1988. ISBN-100934742529 ISBN-13 978-0934742528
2. Morgenthau, Hans J. *Politics Among Nations: The Struggle for Power and Peace*. Alfred A. Knopf, 1948.
3. Bull, Hedley. *The Anarchical Society: A Study of Order in World Politics*. Columbia University press, 2012. ISBN-10 0231161298 ISBN-13 978-0231161299

Sobre principios y valores del sistema internacional

4. United Nations. *Charter of the United Nations*, 1945.
5. Franck, Thomas M. *The Power of Legitimacy Among Nations*. Oxford University Press, 1990. USBN-10 0195061780 USBN-13 978-0195061789
6. Falk, Richard. *Human Rights Horizons: The Pursuit of Justice in a Globalizing World*. Routledge, 2000. ISBN-10 0415925134 ISBN-13 978-0415925136

Sobre inteligencia artificial y diplomacia

7. Bjola, Corneliu, and Markus Kornprobst (eds.) *Digital Diplomacy: Theory and Practice*. Routledge, 2015. ISBN-10 1138843822 ISBN-13 978-1138843820
8. Allison, Graham. *Destined for War: Can America and China Escape Thucydides's Trap?* Houghton Mifflin Harcourt, 2017. USBN-10 0544935276 USBN-13 978-0544935273

9. West, Darrell M. *The Future of Work: Robots, AI, and Automation*. Brookings Institution Press, 2018. USBN-10 0815732937 USBN-13 978-0815732938
10. United Nations Interagency Working Group on Artificial Intelligence. *AI for Good Global Summit Reports*. ITU.int

La Inteligencia Artificial como Desafío Regional: Propuesta para una Arquitectura de Gobernanza Ética en América Latina y el Caribe

Ing. Lourdes Daza Aramayo Ph.D.

European Institute of International Studies (Suecia)

Anglo American University, Praga

La Inteligencia Artificial (IA) representa una oportunidad histórica para transformar los modelos de desarrollo en América Latina y el Caribe (ALC). Su implementación puede contribuir a mejorar la eficiencia de los servicios públicos, fortalecer la transparencia institucional y ampliar el acceso al conocimiento y la innovación. Sin embargo, sin una gobernanza adecuada, la IA también puede profundizar las desigualdades estructurales, vulnerar derechos fundamentales y consolidar nuevas formas de dependencia tecnológica.

Este capítulo propone una estrategia regional de gobernanza de la IA centrada en los principios de equidad, derechos humanos, justicia distributiva y soberanía digital. Desde un enfoque de economía política, se sostiene que la IA debe estar al servicio de la ampliación de capacidades humanas, el fortalecimiento de la democracia y la inclusión social.

A partir de un análisis comparado de los marcos normativos globales (OCDE, UNESCO, Consejo de Europa, China y Unión Europea), se identifica la necesidad de una arquitectura normativa propia para ALC, construida sobre la propuesta de cuatro pilares estratégicos:

1. Declaración Latinoamericana sobre Derechos Digitales y Gobernanza de la IA, que garantice transparencia, no discriminación, explicabilidad y protección de datos.
2. Infraestructura pública digital soberana, financiada mediante mecanismos de fiscalidad digital progresiva.

3. Observatorios ciudadanos de IA y evaluación de impacto ético y social, con participación de la sociedad civil, academia y gobiernos.
4. Cooperación regional institucionalizada, con articulación entre el SELA, la CELAC y organismos multilaterales especializados.

El documento concluye con un llamado a construir una agenda común de gobernanza digital regional, que posicione a ALC como un actor proactivo, ético y soberano en la economía del conocimiento del siglo XXI.

Palabras clave

Inteligencia artificial, gobernanza, inclusión digital, América Latina, inclusión digital, políticas públicas.

1. Introducción

La Inteligencia Artificial (IA) representa una tecnología general transformadora cuya influencia se extiende a casi todos los dominios de la vida contemporánea: desde la producción industrial y la gestión de servicios públicos, hasta la organización del conocimiento y los regímenes de subjetividad.³¹ En este contexto de profunda disrupción digital, América Latina y el Caribe (ALC) enfrenta un desafío histórico: diseñar un modelo de gobernanza de la IA que no solo mitigue riesgos éticos y sociales, sino que también promueva un desarrollo humano sustentable, equitativo y soberano.

Lejos de ser una herramienta neutral, la IA constituye una tecnología estructurante que puede reproducir —o subvertir— las desigualdades sociales, económicas y epistemológicas que configuran la inserción periférica de la región en la economía global. En este sentido, la gobernanza ética de la inteligencia artificial no es simplemente una cuestión técnica, sino una disputa profundamente política por el diseño del futuro. Como ha sostenido Amartya Sen (1999), el desarrollo debe concebirse como una expansión de las libertades sustantivas de las personas. En línea con esta concepción, este capítulo plantea que la gobernanza de la IA debe orientarse a la ampliación de capacidades humanas, la justicia social y el fortalecimiento institucional, como condiciones necesarias para la libertad real y el bienestar colectivo.

Para abordar esta problemática, se adopta una perspectiva crítica de economía política, articulada en torno a tres ejes conceptuales: (i) el enfoque de capacidades humanas (Sen, 1999;

³¹ Por ejemplo, la IA, al clasificar personas, asignar puntuaciones, filtrar contenidos o tomar decisiones automatizadas, participa en la producción de nuevos regímenes de subjetividad.

Nussbaum, 2011), (ii) la teoría de la justicia como equidad (Rawls, 1971), y (iii) los principios normativos de libertad individual, responsabilidad institucional y pluralismo epistémico.³² Esta perspectiva permite analizar la IA no sólo como infraestructura digital, sino como fenómeno normativo que reconfigura derechos, agencias y formas de distribución social.

La pregunta central que guía este capítulo es: **¿cómo puede ALC construir un modelo de gobernanza de la inteligencia artificial que expanda capacidades humanas, fortalezca derechos y contribuya a un desarrollo inclusivo, ético y soberano?** A partir de esta pregunta, se avanza una hipótesis estructurante: un marco de gobernanza regional, híbrido, con base en principios liberales adaptados al contexto latinoamericano, tiene el potencial de articular innovación tecnológica con justicia social y libertad sustantiva.

El texto se estructura en ocho secciones. Luego de esta introducción, se presentan la tesis e hipótesis de trabajo. La tercera sección expone el marco teórico desde la economía política liberal. La cuarta aborda la metodología utilizada. En la quinta, se analiza comparativamente el panorama global de marcos regulatorios. En la sexta, se discute el potencial transformador de la IA en clave regional. La séptima propone una arquitectura normativa latinoamericana para una gobernanza algorítmica ética, soberana y redistributiva. Finalmente, **la octava sección ofrece una síntesis conclusiva y una propuesta para la región.**

2. Fundamentos normativos y postulados analíticos

La tesis central que sostiene este capítulo es que América Latina y el Caribe (ALC) necesita construir un modelo de gobernanza ética de la inteligencia artificial (IA) enfocado en la equidad, la protección de los derechos fundamentales y la expansión de las capacidades humanas. Tal gobernanza debe asumir la IA como una cuestión de política pública distributiva, y no como una externalidad tecnológica, subordinando su desarrollo y aplicación a los valores democráticos, a la justicia social y a la libertad sustantiva.

A diferencia de los enfoques tecnocráticos o securitarios que dominan muchas agendas globales, esta propuesta se fundamenta en tres supuestos normativos clave:

La IA como infraestructura moral y política: siguiendo a Floridi (2021), los sistemas algorítmicos no son meramente técnicos, sino que configuran formas de poder, representación y toma de decisiones que afectan la agencia individual y colectiva.

La justicia como equidad: en la tradición de John Rawls (1971), se sostiene que las desigualdades tecnológicas solo pueden ser legítimas si benefician a los más desfavorecidos,

³² Una gobernanza de la IA con pluralismo epistémico implica no solo respetar derechos digitales, sino también incluir múltiples visiones del mundo en el diseño de los algoritmos: desde principios éticos indígenas hasta cosmovisiones de género, territoriales o comunitarias.

y si las oportunidades para participar en el desarrollo y control de la IA están abiertas para todos en condiciones de equidad.

Las capacidades como eje del desarrollo: según Sen (1999) y Nussbaum (2011), el desarrollo debe evaluarse en términos de la libertad real de las personas para elegir y realizar vidas que valoran. La IA debe contribuir a ampliar esas capacidades, y no restringirlas a través de sesgos, vigilancia o exclusión.

Desde esta premisa teórica se desprenden tres hipótesis fundamentales que estructuran el análisis propuesto:

- La gobernanza de la IA basada en capacidades humanas contribuye a reducir brechas estructurales y a mejorar la agencia individual y colectiva en contextos de desigualdad persistente, como los que caracterizan a ALC.
- Los marcos internacionales de referencia (OCDE, UNESCO, Consejo de Europa, China) ofrecen elementos valiosos, pero requieren una reinterpretación crítica desde el Sur Global, que tome en cuenta las condiciones institucionales, culturales y epistémicas de la región.
- Una arquitectura normativa regional, sustentada en cooperación intergubernamental, fiscalidad digital progresiva, participación social activa y control público de los datos, es indispensable para evitar una nueva dependencia tecnológica y para ejercer una soberanía digital democrática.

3. Marco teórico

El presente capítulo se apoya en un marco teórico que articula la tradición liberal igualitaria con los debates contemporáneos sobre tecnología, justicia y desarrollo. Esta aproximación reconoce que la inteligencia artificial (IA) no es una tecnología neutra, sino una infraestructura sociotécnica atravesada por disputas normativas, intereses económicos y sesgos institucionales. Por ello, su gobernanza debe responder no solo a criterios de eficiencia o innovación, sino a principios de justicia, equidad y libertad sustantiva.

3.1 El enfoque de capacidades humanas

El principal eje teórico se basa en el enfoque de capacidades desarrollado por Amartya Sen (1999) y Martha Nussbaum (2011). Este paradigma sostiene que el desarrollo no debe medirse por el ingreso o la productividad, sino por la libertad real de las personas para llevar

adelante vidas que consideran valiosas. Las capacidades son, en este sentido, las condiciones sociales, materiales y políticas que permiten ejercer esa libertad.

Aplicado a la IA, este enfoque implica que los sistemas algorítmicos deben diseñarse, implementarse y regularse de forma tal que amplíen —y no restrinjan— las oportunidades humanas. Una IA que refuerza sesgos de clase, género o etnia, o que excluye a comunidades de las decisiones que les afectan, constituye una amenaza a la equidad de capacidades. Como señala Nussbaum (2011), “la dignidad humana exige que los marcos institucionales garanticen mínimamente la igualdad de acceso a los medios para vivir con agencia”.

3.2 La justicia como equidad

El segundo fundamento proviene de la teoría de la justicia como equidad de John Rawls (1971). En este marco, una sociedad justa es aquella en la cual las instituciones distribuyen bienes primarios —como derechos, libertades, oportunidades y recursos— de manera tal que beneficien especialmente a los más desfavorecidos. La IA, en tanto bien público con efectos estructurales, debe someterse a esta prueba de justicia.

En el contexto de ALC, caracterizado por desigualdades históricas y exclusiones sistemáticas, aplicar el principio de la “posición original” de Rawls conduce a imaginar un marco de gobernanza en el que nadie acepte reglas algorítmicas que perjudiquen a los sectores más vulnerables (Rawls, 1971, pp. 136–142). Esto exige diseñar políticas públicas que aseguren transparencia, participación, protección de derechos y mecanismos redistributivos frente a los beneficios concentrados del capitalismo de datos.

3.3 Soberanía individual, pluralismo epistémico y responsabilidad institucional

Finalmente, se incorpora un tercer componente teórico: la noción de soberanía individual y responsabilidad institucional frente al avance tecnológico. Inspirándose en autores como Dworkin (2000), Gutmann y Thompson (2004), y Floridi (2021), se sostiene que una gobernanza democrática de la IA debe reconocer la pluralidad de valores y saberes presentes en las sociedades modernas.

Esto significa, por un lado, evitar visiones monoculturales o tecnocráticas que impongan marcos universales sin atender a las especificidades culturales de las comunidades afectadas (Milan & Treré, 2019). Por otro lado, implica reconocer el papel de las instituciones públicas como garantes de equidad, deliberación informada y protección de derechos frente a dinámicas de mercado opacas.

En suma, el marco teórico aquí adoptado articula tres niveles:

- Normativo: la justicia como equidad (Rawls, 1971).
- Evaluativo: las capacidades como criterio de desarrollo (Sen, 1999; Nussbaum, 2011).
- Institucional: la responsabilidad democrática y el pluralismo en la regulación de la IA (Floridi, 2021; Gutmann & Thompson, 2004).

Este andamiaje conceptual orienta el análisis y la propuesta normativa del capítulo.

4. Metodología

Este capítulo adopta un enfoque metodológico de tipo cualitativo, interpretativo y normativo, basado en el análisis documental y comparativo de marcos regulatorios, principios éticos, políticas públicas y discursos institucionales sobre inteligencia artificial (IA) a nivel internacional y regional. El objetivo no es describir exhaustivamente todos los modelos existentes, sino identificar las tensiones normativas y oportunidades estratégicas que permitan construir una propuesta de gobernanza propia para América Latina y el Caribe (ALC), sustentada en valores democráticos y justicia social.

La elección de este enfoque responde a dos razones fundamentales. Primero, porque la gobernanza de la IA no puede reducirse a cuestiones técnicas o instrumentales, sino que requiere ser comprendida como una práctica normativa: esto es, una actividad institucional que distribuye poder, derechos, riesgos y beneficios en contextos sociales desiguales. Segundo, porque los marcos éticos y legales en torno a la IA se encuentran en rápida evolución y alta disputa, lo que hace necesario un análisis interpretativo que combine teoría normativa con estudio empírico de casos y modelos regulatorios.

4.1 Técnicas y fuentes

Se utilizaron las siguientes técnicas y fuentes:

- Análisis documental de textos normativos de organismos multilaterales (UNESCO, OCDE, Consejo de Europa, entre otros)
- Revisión de literatura académica en economía política, teoría de la justicia, ética de la tecnología y estudios críticos sobre IA.
- Estudio comparativo de políticas nacionales de IA de diferentes regiones (Unión Europea, Estados Unidos, China, India, países de ALC).
Lectura normativa de los principios propuestos, en diálogo con el enfoque de capacidades y la justicia como equidad.

La selección de documentos se basó en su relevancia institucional, su influencia normativa en los debates globales y su aplicabilidad potencial al contexto latinoamericano.

4.2 Limitaciones

Este estudio reconoce algunas limitaciones: (i) la velocidad con la que se desarrollan nuevos marcos regulatorios de IA implica que parte de los documentos analizados podrían quedar obsoletos en el corto plazo; (ii) existe una escasa sistematización de las experiencias latinoamericanas en esta materia, lo que requiere una lectura exploratoria y crítica más que conclusiva.

Pese a estas limitaciones, la metodología empleada permite identificar principios orientadores robustos, marcos comparables y caminos institucionales plausibles para una gobernanza algorítmica con equidad.

5. Análisis comparado de marcos internacionales de gobernanza de la IA

A medida que la inteligencia artificial (IA) redefine estructuras económicas, institucionales y culturales, diversas regiones del mundo han desarrollado marcos normativos para orientar su gobernanza. Estos modelos reflejan visiones políticas y filosóficas distintas sobre el rol del Estado, la libertad individual, los derechos colectivos y la ética pública. Este capítulo presenta un análisis comparado de los principales marcos internacionales, los marcos más influyentes —OCDE, Consejo de Europa, UNESCO, China y Unión Europea— con el objetivo de identificar elementos normativos relevantes para una estrategia latinoamericana.

5.1 OCDE: Principios voluntarios para una IA confiable

La Organización para la Cooperación y el Desarrollo Económicos (OCDE) propuso en 2019³³ cinco principios para una IA confiable: (i) crecimiento inclusivo y bienestar humano; (ii) respeto a los derechos humanos y valores democráticos; (iii) transparencia y explicabilidad; (iv) robustez y seguridad técnica; y (v) responsabilidad. Aunque ampliamente influyentes, estos principios tienen carácter no vinculante y están enfocados en facilitar la innovación privada más que en garantizar justicia redistributiva (OCDE, 2019).

Desde una perspectiva liberal crítica, estos principios ofrecen un punto de partida útil pero insuficiente: carecen de mecanismos coercitivos, no abordan la fiscalidad digital y minimizan los desequilibrios de poder entre grandes corporaciones tecnológicas y Estados periféricos.

³³ Esta organización fue una de las primeras en establecer principios sobre la IA confiable, incluyendo el respeto a los derechos humanos y el estado de derecho, la transparencia, la explicabilidad, la solidez técnica y la seguridad.

5.2 Consejo de Europa: Derechos humanos y marco vinculante

El Consejo de Europa sobre IA³⁴ ha avanzado hacia un marco jurídico vinculante sobre IA, basado en la protección de los derechos humanos, el Estado de derecho y la democracia. La propuesta de un Tratado internacional de gobernanza algorítmica contempla la obligación de realizar evaluaciones de impacto sobre derechos humanos, y exige transparencia, control humano significativo y rendición de cuentas (Consejo de Europa, 2022).

Este enfoque es normativamente robusto y compatible con el paradigma liberal, aunque enfrenta desafíos de implementación práctica y escasa adopción fuera del continente europeo.

5.3 UNESCO: Ética global, diversidad cultural y sostenibilidad

La Recomendación sobre la Ética de la IA de la UNESCO (2021) representa el primer instrumento normativo global adoptado por consenso entre más de 190 Estados miembros. El documento enfatiza la inclusión, la diversidad cultural, la sostenibilidad ambiental y la justicia intergeneracional.³⁵ Introduce conceptos como la “evaluación del impacto ético”, la protección de ecosistemas digitales y la necesidad de marcos legales nacionales adaptados a contextos locales (UNESCO, 2021).

Este modelo resulta especialmente relevante para ALC por su énfasis en la equidad y la pluralidad epistémica, aunque —como en el caso de la OCDE— no es jurídicamente vinculante.

5.4 China: Soberanía tecnológica y planificación estatal

El enfoque chino se basa en un modelo de soberanía digital centralizada, en el que el Estado desempeña un rol hegemónico tanto en el desarrollo de infraestructura como en la regulación algorítmica. China ha integrado la IA a su planificación quinquenal, promoviendo el liderazgo en tecnologías estratégicas y el control estatal sobre los datos. Si bien este modelo ha demostrado eficacia en innovación, plantea serias tensiones con los derechos individuales, la libertad de expresión y el control democrático (Laskai & Webster, 2020).

34 El Consejo de Europa es otra organización internacional, distinta de la UE. Incluye 46 países (más que los 27 de la UE). No tiene competencias legislativas vinculantes como la UE, pero emite estándares de derechos humanos, democracia y Estado de derecho. Su posición se reflejó en el Convenio marco sobre la inteligencia artificial y los derechos humanos, la democracia y el Estado de derecho, adoptado en mayo de 2024. Se trata de un tratado internacional opcional, no una ley directa como la AI Act. Su aplicación no es directa como en AI Act y depende a la ratificación de cada país. Su enfoque es en Derechos Humanos, Democracia y Estado de derecho.

35 la UNESCO enfatiza una IA centrada en el ser humano, promoviendo principios de justicia, sostenibilidad, diversidad cultural, inclusión y solidaridad. Además, introduce la idea de ‘impacto algorítmico’ como dimensión transversal.

Su estrategia, expresada en el plan “Next Generation Artificial Intelligence Development Plan” (2017), busca convertir a China en líder global en IA para 2030 (Laskai & Webster, 2020).

A diferencia de los enfoques occidentales centrados en marcos voluntarios, transparencia y derechos fundamentales, China ha promovido un ecosistema de IA con regulación flexible en el corto plazo, pero alta vigilancia estatal. Esto incluye el uso de tecnologías algorítmicas para monitorear poblaciones, gestionar infraestructuras urbanas, y fortalecer capacidades militares. Sin embargo, en años recientes ha comenzado a formular principios éticos específicos a través de entidades como el Ministerio de Ciencia y Tecnología y la Alianza Nacional de IA, enfatizando conceptos como el “control humano sobre la IA”, la seguridad, la equidad algorítmica y la responsabilidad corporativa.

El modelo chino desafía los paradigmas tradicionales sobre regulación tecnológica al priorizar la soberanía digital y el liderazgo geopolítico. Al mismo tiempo, ha sido criticado por organizaciones internacionales y académicos por sus riesgos en materia de derechos digitales, privacidad y gobernanza algorítmica opaca. Para América Latina y el Caribe, el modelo chino ofrece lecciones tanto sobre capacidad estatal de coordinación, como advertencias sobre la necesidad de salvaguardas democráticas y mecanismos de rendición de cuentas. La región puede observar este modelo no como un ejemplo a replicar, sino como una referencia crítica para reflexionar sobre el equilibrio entre autonomía tecnológica, innovación y protección de libertades fundamentales.

Este enfoque es incompatible con los marcos liberales de gobernanza democrática, aunque ofrece lecciones relevantes sobre inversión pública, autonomía tecnológica y planificación estratégica.

5.5 Unión Europea: Marco regulatorio con enfoque en riesgos

La Unión Europea ha elaborado un Reglamento sobre IA (AI Act)³⁶ que propone un enfoque basado en la clasificación de riesgos. Los sistemas de alto riesgo —como los utilizados en justicia, educación o servicios públicos— estarán sujetos a regulaciones estrictas en cuanto a transparencia, trazabilidad, y derechos de los usuarios. Además, el reglamento prohíbe ciertas prácticas algorítmicas intrusivas, como la vigilancia masiva o el scoring social.

³⁶ AI Act es una ley de la Unión Europea (UE), formalmente conocida como Reglamento de Inteligencia Artificial. Propuesta por la Comisión Europea y adoptada por el Parlamento Europeo y el Consejo de la UE. Su objetivo principal es regular el uso de sistemas de IA dentro del mercado único europeo, con un enfoque basado en riesgos (prohibiciones, requisitos según nivel de riesgo, etc.) y tiene efectos jurídicos obligatorios para todos los Estados miembros de la UE. Su enfoque es regulatorio, orientado a mercado y riesgos.

Este modelo destaca por su equilibrio entre innovación y protección de derechos, y constituye un referente importante para una gobernanza regional en ALC, siempre que se adapte a las condiciones institucionales de la región.

Estos modelos, lejos de ser excluyentes, ofrecen a América Latina y el Caribe una gama de referencias que pueden ser reinterpretadas en clave regional. En lugar de adoptar un marco único, ALC puede construir una estrategia híbrida que combine estándares internacionales con prioridades locales como la equidad, la sostenibilidad, la diversidad cultural y la participación democrática.

Comparación de Enfoques Internacionales sobre la Inteligencia Artificial

Aspecto	AI Act (UE)	Consejo de Europa	UNESCO	OCDE	China
Tipo de documento	Reglamento vinculante	Convenio internacional	Recomendación (no vinculante)	Principios (no vinculantes)	Directrices y regulaciones nacionales
Carácter jurídico	Obligatorio en la UE	Obligatorio si se ratifica	No obligatorio	No obligatorio	Obligatorio dentro del país
Enfoque principal	Riesgos en el mercado	Derechos humanos y democracia	Ética, inclusión, sostenibilidad	Innovación responsable	Seguridad, control social, desarrollo económico
Ámbito geográfico	UE (27 países)	46 países (Europa)	194 países	Miembros de la OCDE	China
Grado de detalle técnico	Alto	Medio	Bajo	Medio	Alto
Mecanismos de implementación	Multas, evaluaciones, obligaciones legales	Obligaciones generales	Promoción voluntaria	Buenas prácticas	Leyes, guías estatales, sanciones
Enfoque sobre DD.HH.	Implícito (Carta de la UE)	Central	Central	Implícito	Explícito en términos de seguridad y orden
Año de adopción	2024	2024	2021	2019	Desde 2017 (regulación creciente)

Fuente: Elaboración propia con base a fuentes oficiales.

En resumen, el análisis comparado muestra que no existe un modelo único o universal de gobernanza algorítmica. Cada región enfatiza distintos valores: libertad individual (OCDE), derechos colectivos (Consejo de Europa), diversidad y sostenibilidad (UNESCO), soberanía tecnológica (China), y enfoque precautorio (UE). América Latina debe construir su propio marco híbrido, anclado en su realidad social y aspiraciones normativas.

6. Implicaciones para América Latina y el Caribe

América Latina y el Caribe (ALC) enfrenta desafíos estructurales que condicionan su inserción en la economía digital global: alta desigualdad social, debilidad institucional, dependencia tecnológica, fragmentación política y bajos niveles de inversión en ciencia

y tecnología. Sin embargo, estos mismos desafíos constituyen la base sobre la cual se puede construir una gobernanza de la inteligencia artificial (IA) con características propias, orientada a la justicia social, la expansión de capacidades humanas y la defensa del bien común.

Desde una perspectiva de economía política crítica, la región debe rechazar la adopción pasiva de marcos regulatorios diseñados en otros contextos, e impulsar una arquitectura normativa que responda a sus realidades históricas y necesidades colectivas. Ello implica repensar el papel del Estado, fortalecer la soberanía digital y construir capacidades institucionales para una regulación ética y efectiva.

6.1. La IA como dispositivo de redistribución y reconocimiento

En un contexto de concentración extrema de datos, infraestructura tecnológica y capacidades computacionales en manos de actores globales, la IA corre el riesgo de profundizar las desigualdades existentes. ALC debe posicionar la IA no solo como vector de innovación, sino como un instrumento de redistribución de oportunidades y reconocimiento de diversidades.

Esto requiere políticas públicas que garanticen:

- Acceso universal y equitativo a tecnologías digitales;
- Protección de comunidades históricamente excluidas frente a sesgos algorítmicos;
- Transparencia y explicabilidad en decisiones automatizadas que afectan derechos fundamentales;
- Participación plural en el diseño, implementación y monitoreo de sistemas algorítmicos.

Como plantea Sen (1999), el desarrollo debe evaluarse por la capacidad efectiva de las personas para elegir entre formas de vida valiosas. Desde esta perspectiva, una IA justa no es aquella que optimiza eficiencia, sino la que amplía la agencia y la libertad real.

6.2. Soberanía tecnológica y fiscalidad digital

La dependencia estructural de ALC respecto a plataformas tecnológicas extranjeras plantea la necesidad urgente de construir soberanía digital. Esto no implica replicar modelos autoritarios como el chino, sino consolidar infraestructuras públicas, marcos regulatorios y alianzas regionales que permitan a los Estados recuperar capacidad de gobernar el espacio digital.

Hoy, casi todo el ecosistema digital en ALC depende de empresas extranjeras (Google, Meta, Amazon, Huawei, etc.). Esto genera vulnerabilidad estratégica: datos personales, decisiones

automatizadas y servicios esenciales están fuera del control regulador local. Con la fiscalidad digital se busca que las grandes empresas tecnológicas (como GAFAM) paguen impuestos en los países donde generan valor y ganancias, incluso si no tienen sede física allí.

Estas empresas obtienen ingresos multimillonarios por publicidad, comercio electrónico, servicios en la nube, etc., sin tributar proporcionalmente en la región.

Esto reduce los recursos fiscales disponibles para invertir en salud, educación digital, innovación pública o conectividad inclusiva.

Un eje central en esta estrategia es la fiscalidad digital progresiva: la tributación justa a grandes empresas tecnológicas (GAFAM)³⁷ que operan en la región sin contribuir proporcionalmente al financiamiento de bienes públicos digitales. Estos recursos permitirían crear fondos regionales para infraestructura, innovación abierta y formación digital inclusiva.

Se trataría de establecer impuestos digitales comunes a nivel regional (como ha hecho la UE). Crear un fondo latinoamericano para soberanía digital, financiado por estos impuestos, para invertir en: investigación pública en IA, infraestructura tecnológica comunitaria, formación crítica en alfabetización digital, protección de derechos en entornos digitales.

6.3. Democracia algorítmica y pluralismo epistémico³⁸

La inclusión epistémica y la justicia cognitiva constituyen dimensiones esenciales para una gobernanza de la Inteligencia Artificial (IA) verdaderamente democrática en América Latina y el Caribe (ALC).

Tradicionalmente, los procesos tecnológicos han estado dominados por formas de conocimiento técnico, corporativo y eurocéntrico que marginan otras epistemologías —como los saberes indígenas, campesinos, populares, feministas y afrodescendientes— que han sido históricamente desvalorizadas.

La inclusión epistémica implica reconocer que existen múltiples formas legítimas de producir conocimiento y comprender el mundo. En el contexto de la IA, esto significa que los algoritmos, los modelos de datos y las decisiones automatizadas no deben ser diseñados

37 Es un acrónimo que designa a las cinco grandes empresas tecnológicas de origen estadounidense que dominan buena parte del ecosistema digital global: G – Google (ahora Alphabet Inc.), A – Apple, F – Facebook (ahora Meta Platforms Inc.), A – Amazon, M – Microsoft.

38 Es un concepto que se refiere a la coexistencia y reconocimiento de múltiples formas de conocimiento, métodos de entender el mundo y criterios de verdad. Es una postura que defiende que no hay una única manera “correcta” o universal de conocer, sino que distintas culturas, disciplinas, lenguajes o cosmovisiones pueden aportar conocimientos válidos y complementarios.

exclusivamente desde centros de poder tecnológico del Norte Global, sino que deben incorporar visiones, valores y prioridades de las comunidades locales. Una IA que se desarrolla sin participación de estas voces perpetúa relaciones coloniales del saber y reproduce exclusión social mediante sistemas invisibilizados, pero altamente efectivos.

La justicia cognitiva, por su parte, va más allá de la representación formal: exige que los marcos de interpretación, las metodologías y los parámetros técnicos respeten la pluralidad cultural, lingüística, territorial y de género. Esto requiere institucionalizar mecanismos de participación efectiva, desarrollar metodologías de evaluación de impacto cultural de tecnologías y promover alianzas con comunidades epistémicas diversas.

También implica revisar críticamente los datos utilizados en la región: ¿quién los produce? ¿con qué sesgos? ¿desde qué lenguas y marcos ontológicos?

Aplicar estos principios en ALC permitiría construir una IA arraigada en los contextos históricos, culturales y sociales de la región. Una inteligencia artificial realmente inclusiva no solo mejora la precisión técnica, sino que fortalece la equidad, la autodeterminación y la justicia social. Desde esta perspectiva, la gobernanza de la IA debe ser también una política de conocimiento, donde lo que se reconoce como válido, confiable o útil responda a los principios del pluralismo epistémico y la equidad cognitiva en el Sur Global.

La región tiene la oportunidad de liderar una gobernanza pluralista, que reconozca saberes indígenas, epistemologías del Sur y cosmovisiones alternativas frente al extractivismo digital. Como afirman Milan y Treré (2019), el pluralismo epistémico es clave para una justicia digital intercultural y decolonial.

6.4. Crecimiento con igualdad y sostenibilidad: una agenda estructural para la IA

La adopción estratégica de la inteligencia artificial (IA) en América Latina y el Caribe (ALC) no puede dissociarse de una agenda de transformación productiva, justicia social y sostenibilidad ambiental. En este contexto, la IA debe articularse con un cambio estructural progresivo que supere el modelo extractivista de bajo valor agregado y promueva una economía basada en el conocimiento, la innovación y la equidad (CEPAL, 2020).

En primer lugar, se requiere una transformación productiva que diversifique la matriz económica regional. La IA puede ser un catalizador para el desarrollo de sectores como la agroindustria sostenible, las energías renovables, la biotecnología y las industrias creativas. Estas oportunidades deben ser canalizadas a través de políticas industriales activas que fomenten la sofisticación tecnológica y la integración de cadenas de valor regionales (Rodrik,

2022). La implementación de IA en sectores productivos estratégicos también puede favorecer una transición hacia una economía circular y de bajas emisiones.

Asimismo, es fundamental abordar la dimensión de la desigualdad estructural. La IA puede amplificar brechas existentes si no se acompaña de políticas redistributivas y de inclusión activa. Por ejemplo, los sesgos algorítmicos pueden reproducir formas de exclusión de género, clase y etnia (Eubanks, 2018; Benjamin, 2019). Una agenda estructural debe contemplar mecanismos de auditoría algorítmica, transparencia y participación ciudadana en el diseño y aplicación de estas tecnologías.

En cuanto a la educación, se plantea la necesidad de reformar los sistemas educativos para formar talento humano con competencias digitales, pensamiento crítico y capacidades colaborativas. Esta transformación no solo debe orientarse a las élites tecnológicas, sino democratizarse desde la infancia hasta la educación superior, incluyendo la educación técnica y comunitaria (UNESCO, 2021). La infraestructura digital y la conectividad deben expandirse particularmente en zonas rurales e indígenas, como condición básica para una IA inclusiva (Hilbert, 2020).

Otro eje clave es la fiscalidad. La transición hacia una economía digital e inteligente requiere sistemas fiscales progresivos capaces de financiar políticas sociales universales, infraestructura y transición ecológica. Se deben establecer mecanismos para gravar adecuadamente las grandes plataformas tecnológicas y redistribuir sus beneficios a través de inversión pública estratégica (Piketty, 2020).

La reducción de desigualdades estructurales también implica fortalecer el acceso a salud, cuidados, pensiones y el derecho al territorio para pueblos indígenas y campesinos, en tanto actores fundamentales en la gobernanza territorial y de datos. La IA puede apoyar la planificación y gestión de estos servicios, pero su diseño debe construirse desde marcos de justicia cognitiva y autonomía cultural (D'Ignazio & Klein, 2020).

Por último, un Estado eficaz y transparente es una condición para fomentar el desarrollo digital hacia el bien común. La digitalización de los servicios públicos puede aumentar la eficiencia, pero también la vigilancia. Por tanto, se requiere una arquitectura institucional que garantice la rendición de cuentas, el control ciudadano y la protección de derechos fundamentales (Latonero, 2020).

Este nuevo pacto digital debe anclarse en la integración regional, la cooperación sur-sur y la articulación entre gobiernos, sociedad civil, academia y sector privado. Solo así ALC podrá construir una IA soberana, ética y centrada en el desarrollo humano sostenible.

6.5. El potencial transformador de la inteligencia artificial (IA) en América Latina y el Caribe (ALC)

El potencial se manifiesta en su capacidad para dinamizar sectores productivos, mejorar servicios públicos, reducir desigualdades históricas y enfrentar desafíos estructurales como el cambio climático, la informalidad laboral y la fragmentación institucional. No obstante, este potencial no es automático, sino que depende de las políticas públicas, el diseño institucional y los marcos normativos que guíen su adopción y uso (Crespo & Villalobos, 2022).

En el ámbito productivo, la IA puede facilitar la transformación de sectores clave como la agricultura, la logística, la manufactura y la minería mediante sistemas de predicción climática, mantenimiento predictivo, trazabilidad de insumos y automatización de procesos (Trigo, 2020). Estos avances pueden mejorar la eficiencia y competitividad de las micro, pequeñas y medianas empresas (MIPYMES), siempre que se acompañen de políticas de financiamiento, formación y acceso a infraestructura digital (Rodrik, 2022).

En el campo de la salud, la IA ofrece oportunidades para mejorar el diagnóstico, la prevención y la gestión hospitalaria, particularmente en contextos de escasez de recursos y profesionales. La Organización Mundial de la Salud (WHO, 2021) reconoce que la IA puede contribuir a reducir la mortalidad por enfermedades crónicas y mejorar el acceso a servicios en zonas rurales, pero también advierte sobre los riesgos de exclusión si no se garantiza la equidad en la disponibilidad de datos y tecnologías.

En la educación, los sistemas inteligentes pueden apoyar la personalización del aprendizaje, la detección temprana de dificultades y la creación de plataformas adaptativas para contextos multiculturales (UNESCO, 2021). Sin embargo, Noble (2018) y Eubanks (2018) advierten que las tecnologías educativas no son neutras y pueden reproducir sesgos de clase, raza y género si no se diseñan con criterios de justicia cognitiva.

Respecto a la acción climática, la IA puede optimizar la gestión del agua, prever desastres naturales, modelar escenarios de cambio climático y mejorar la eficiencia energética en ciudades (Hilbert, 2020). Estos usos tecnológicos deben insertarse en estrategias de sostenibilidad ambiental que prioricen la resiliencia de los ecosistemas y las comunidades más vulnerables.

La inclusión financiera y social también puede verse beneficiada mediante sistemas algorítmicos que amplíen el acceso al crédito, seguros, subsidios y servicios públicos para poblaciones tradicionalmente excluidas, como mujeres rurales, pueblos indígenas y trabajadores informales (D'Ignazio & Klein, 2020). No obstante, si estas herramientas no incorporan principios de equidad, corren el riesgo de reforzar mecanismos de vigilancia o control.

Asimismo, la IA puede fortalecer la gobernanza digital, reduciendo la burocracia y mejorando la transparencia en los servicios públicos. Iniciativas de automatización en justicia, seguridad ciudadana y administración tributaria pueden aumentar la eficiencia del Estado, pero también demandan límites claros para evitar abusos, sesgos o decisiones no apelables (Latonero, 2020).

Desde una perspectiva geopolítica, el aprovechamiento transformador de la IA exige que la región no se limite a ser consumidora de tecnologías importadas, sino que construya capacidades locales para su diseño, regulación y evaluación. Esto incluye desde el fortalecimiento de ecosistemas de innovación hasta el establecimiento de políticas de soberanía tecnológica, cooperación científica y producción de datos desde el Sur Global (Floridi, 2021).

Además, en un escenario global caracterizado por crecientes tensiones geopolíticas y una marcada asimetría en la distribución del poder tecnológico, la diplomacia digital emerge como una herramienta clave para que América Latina y el Caribe articule una voz común en los foros multilaterales. La región debe construir alianzas estratégicas con actores que compartan una visión ética, humanista y regulada de la inteligencia artificial, como la Unión Europea.

Tal como argumenta Torres Jarrín (2023), la inteligencia artificial ha emergido como un nuevo campo estratégico de la diplomacia internacional, donde los marcos éticos y los valores compartidos son tan importantes como la infraestructura tecnológica. En este sentido, el fortalecimiento de las relaciones entre América Latina y la Unión Europea en torno a la gobernanza digital representa una oportunidad para avanzar hacia una cooperación basada en la equidad, la reciprocidad y la defensa conjunta de los derechos humanos en entornos algorítmicos.

En síntesis, la IA tiene un potencial transformador significativo en ALC, pero este dependerá de cómo se estructuren los marcos institucionales, éticos y distributivos que la acompañen. Sin una gobernanza deliberada, existe el riesgo de que se convierta en un factor más de concentración de poder y profundización de desigualdades. Con una orientación centrada en derechos, inclusión y sostenibilidad, puede en cambio ser una palanca clave para el desarrollo justo y resiliente de la región.

6.6. La implementación de Inteligencia Artificial (IA) en América Latina y el Caribe (ALC)

ALC enfrenta importantes barreras estructurales y emergentes que deben ser abordadas con un enfoque sistémico, crítico y contextualizado. Estos desafíos no solo responden a déficits tecnológicos o regulatorios, sino que reflejan históricas asimetrías económicas, sociales y

epistémicas. Para construir una IA justa, inclusiva y soberana, la región debe atender las siguientes dimensiones clave:

Brecha digital y desigualdades de acceso: La desigualdad en el acceso a infraestructura digital, conectividad, dispositivos y alfabetización tecnológica sigue siendo uno de los principales obstáculos. Como ha documentado Hilbert (2020), estas brechas se expresan no solo entre países, sino también dentro de los mismos territorios, con efectos más agudos en zonas rurales, poblaciones indígenas y comunidades empobrecidas. La expansión de la IA puede amplificar estas desigualdades si no se garantiza una política de inclusión digital sólida y redistributiva.

Déficit de capacidades locales: Existe una escasez crítica de talento humano en áreas como ciencia de datos, ética de la IA, ciberseguridad, regulación tecnológica y gobernanza algorítmica. Esta carencia limita la autonomía regional para diseñar, auditar e implementar tecnologías de IA adaptadas a contextos locales (Crespo & Villalobos, 2022). ALC requiere una estrategia educativa y de formación continua que democratice el conocimiento técnico y fortalezca la soberanía cognitiva de los países.

Dependencia tecnológica y concentración del poder: La dominancia de grandes corporaciones tecnológicas globales (principalmente de EE. UU. y China) en la provisión de plataformas, software y hardware genera una dependencia crítica. Esta asimetría se traduce en pérdida de control sobre los flujos de datos, la infraestructura digital y los estándares tecnológicos (Couldry & Mejias, 2019). Además, consolida un colonialismo de datos que impide a los países del Sur Global beneficiarse plenamente del valor generado por sus propios datos.

Sesgos algorítmicos y reproducción de desigualdades: Los sistemas de IA entrenados con datos sesgados —que reflejan estructuras históricas de racismo, sexismo y exclusión— pueden amplificar la discriminación en áreas como justicia penal, servicios financieros, salud y empleo (Benjamin, 2019; Eubanks, 2018). Una IA justa en ALC requiere mecanismos obligatorios de auditoría algorítmica, explicabilidad, participación comunitaria en el diseño y protocolos éticos que reconozcan la diversidad cultural y epistémica de la región.

Soberanía de datos y derechos digitales: Uno de los retos emergentes más urgentes es la necesidad de desarrollar marcos jurídicos y políticos que garanticen el control soberano de los datos. Esto incluye no solo protección de la privacidad, sino también derechos colectivos sobre datos generados por comunidades, territorios y ecosistemas (Purtova, 2018). La soberanía de datos debe ser parte integral de cualquier estrategia de IA en ALC, para evitar que las decisiones automatizadas escapen al escrutinio democrático y los intereses locales.

Débil cooperación regional y fragmentación normativa: A pesar de esfuerzos incipientes como la Agenda Digital de la CEPAL y la coalición IA2030 LAC, la región aún carece de un marco común de gobernanza digital. La falta de interoperabilidad normativa, de plataformas compartidas y de mecanismos de evaluación regional limita la posibilidad de crear sinergias y reducir asimetrías entre países. Se necesita una arquitectura de cooperación horizontal, pluralista y basada en principios comunes como los derechos humanos, la inclusión social y la sostenibilidad.

Exclusión de comunidades locales: La mayoría de los marcos tecnológicos vigentes no consideran la participación activa de pueblos indígenas, comunidades rurales o movimientos sociales en el diseño y supervisión de la IA. Esto perpetúa una exclusión epistémica que impide el desarrollo de tecnologías culturalmente pertinentes y éticamente legítimas (Milan & Treré, 2019). Una IA justa debe incorporar la pluralidad de conocimientos y visiones de mundo presentes en ALC, avanzando hacia una democratización del desarrollo tecnológico.

Superar estos desafíos implica una visión integral que articule políticas públicas transformadoras, regulación anticipatoria, inversión en capacidades endógenas y mecanismos de justicia social. Solo así será posible que la IA en ALC contribuya efectivamente a una agenda de desarrollo con igualdad, autonomía y dignidad.

7. Propuesta normativa regional

Frente a los desafíos éticos, económicos y políticos que plantea la inteligencia artificial (IA), América Latina y el Caribe (ALC) necesita construir un modelo normativo propio que combine estándares internacionales con valores regionales de justicia, equidad y pluralismo democrático. Esta propuesta se enmarca en una economía política liberal crítica, que busca ampliar las capacidades humanas, redistribuir oportunidades y democratizar el control sobre las infraestructuras tecnológicas.

A partir del análisis de los marcos internacionales (OCDE, Consejo de Europa, UNESCO, China), de las necesidades estructurales de América Latina y el Caribe (ALC) y del potencial transformador de la inteligencia artificial (IA), se propone un modelo híbrido de gobernanza regional para la IA. Este modelo integra elementos normativos robustos con flexibilidad adaptativa, y prioriza una perspectiva ética centrada en los derechos humanos, la justicia social, la sostenibilidad ambiental y la inclusión digital.

A continuación, se presentan los ejes centrales de una arquitectura normativa regional para la gobernanza de la inteligencia artificial:

7.1. Principios para una Gobernanza Ética de la IA en América Latina y el Caribe

Como destaca Torres Jarrín (2023), la inteligencia artificial no solo plantea desafíos técnicos, sino también diplomáticos, geopolíticos y éticos que requieren respuestas multilaterales y concertadas. Desde esta perspectiva, una declaración latinoamericana que articule derechos digitales y principios éticos puede fortalecer la posición de la región en el diálogo interregional con Europa, promoviendo una gobernanza digital basada en valores compartidos como la justicia, la transparencia y la soberanía tecnológica.

Inspirada en la tradición de los derechos humanos y en documentos como el Reglamento Europeo de Protección de Datos (GDPR) y la Recomendación de la UNESCO (2021), se propone una propuesta de **“Principios para una Gobernanza Ética de la IA en América Latina y el Caribe”**, que garantice:

- Derecho a la explicabilidad y transparencia de decisiones automatizadas;
- Derecho a la no discriminación algorítmica, con criterios de equidad e inclusión;
- Derecho a la autodeterminación informativa y al control sobre los propios datos;
- Derecho a la revisión humana significativa en decisiones que afecten derechos fundamentales;
- Derecho a una reparación efectiva frente a daños ocasionados por decisiones algorítmicas injustas.

Este instrumento debe tener rango supralegal, ser adoptado por los países del MERCOSUR, la CELAC y otros organismos regionales, y constituir el núcleo de una legislación más amplia sobre IA y derechos digitales.

7.2. Infraestructura pública y fondos regionales para soberanía digital

Una gobernanza democrática requiere no solo regulación, sino inversión. Se propone la creación de un Fondo Regional de Infraestructura Digital Soberana, financiado mediante mecanismos de fiscalidad digital progresiva aplicados a las grandes plataformas tecnológicas.

Este fondo permitiría:

- Desarrollar centros de datos regionales bajo control público;
- Impulsar laboratorios de innovación abierta e inclusiva;
- Financiar programas de alfabetización algorítmica y formación crítica en IA;
- Garantizar conectividad universal, segura y libre de discriminación.

De este modo, la región podría reducir su dependencia tecnológica y fortalecer su capacidad estratégica para participar en la economía digital global.

7.3. Observatorios ciudadanos y evaluación de impacto ético

Se propone crear Observatorios Ciudadanos de IA en cada país miembro, articulados en una red regional coordinada. Su misión sería:

- Monitorear el uso de algoritmos en servicios públicos y privados;
- Realizar auditorías sociales de sistemas de IA;
- Generar evidencia sobre impactos diferenciados por género, raza, clase y territorio;
- Promover deliberación democrática sobre estándares éticos y regulatorios.

Junto con ello, se propone instituir la Evaluación de Impacto Ético y Social (EIES)³⁹ como requisito obligatorio para todos los sistemas algorítmicos utilizados en áreas sensibles (educación, salud, justicia, seguridad, finanzas públicas).

7.4. Marco normativo híbrido, abierto y adaptativo

La propuesta normativa debe asumir un carácter híbrido y multinivel: incluir leyes nacionales, tratados regionales, estándares técnicos y principios éticos, con mecanismos de gobernanza colaborativa entre Estados, sociedad civil, academia y sector privado. Asimismo, debe ser abierto y adaptativo, capaz de evolucionar conforme a los cambios tecnológicos y sociales, sin perder su anclaje en los principios de justicia, libertad y equidad.

En conjunto, estas propuestas configuran una arquitectura institucional orientada a democratizar el diseño, uso y supervisión de la IA, desde y para América Latina.

Componentes clave del modelo híbrido:

- **Base normativa garantista:** Inspirado en el enfoque del Consejo de Europa, el modelo propone un marco legal regional vinculante que asegure la protección de derechos fundamentales, evite los abusos algorítmicos y garantice el debido proceso. Esto incluiría la creación de mecanismos de rendición de cuentas, auditoría algorítmica y explicabilidad obligatoria en sistemas públicos de IA (Benjamin, 2019; Latonero, 2020).
- **Flexibilidad de implementación y adaptabilidad local:** Siguiendo la orientación de la OCDE y la UNESCO, el modelo reconoce la diversidad institucional de los países de ALC

³⁹ La Evaluación de Impacto Ético y Social (EIES) es un proceso que evalúa los posibles efectos éticos y sociales de un proyecto o iniciativa, antes de que se implemente. Su objetivo es identificar, predecir, evaluar y mitigar los impactos negativos, promoviendo la sostenibilidad a largo plazo y la equidad social.

y propone una implementación escalonada y adaptativa de los principios, con guías prácticas, marcos éticos sectoriales y procesos participativos de adaptación normativa (OECD, 2019; UNESCO, 2021).

- **Soberanía tecnológica y cooperación regional:** Retomando lecciones del modelo chino, se propone una agenda estratégica regional que fomente la inversión pública en capacidades tecnológicas endógenas, el fortalecimiento de infraestructuras digitales, y la producción de datos desde el Sur Global (Floridi, 2021). Esto debe combinarse con una gobernanza democrática para evitar los riesgos de concentración de poder estatal y violaciones a los derechos digitales.
- **Participación social y justicia epistémica:** Una dimensión clave del modelo es la integración activa de comunidades, pueblos indígenas, mujeres y juventudes en el diseño y supervisión de las tecnologías de IA. Se propone institucionalizar mecanismos de consulta, laboratorios ciudadanos y protocolos de consentimiento comunitario sobre datos e IA (D'Ignazio & Klein, 2020; Milan & Treré, 2019).
- **Interoperabilidad normativa y fiscalidad digital regional:** El modelo aboga por un marco común de fiscalidad digital para gravar las grandes plataformas tecnológicas y redistribuir recursos hacia infraestructura digital, servicios públicos y ciencia abierta. A su vez, se requiere armonizar normas sobre protección de datos, contratación pública de IA y estándares éticos (Piketty, 2020; Susskind, 2022).

Síntesis integradora: Este modelo híbrido retoma lo mejor de los marcos existentes -la normatividad protectora del Consejo de Europa, la adaptabilidad de la OCDE, la visión universalista de la UNESCO, la estrategia soberana de China y las aspiraciones democráticas de los movimientos sociales regionales. Su objetivo es articular una inteligencia artificial que, lejos de reproducir las desigualdades del pasado, se convierta en una herramienta para un futuro más justo, resiliente y sostenible en ALC.

8. Conclusiones

La inteligencia artificial (IA) plantea uno de los desafíos normativos y políticos más relevantes del siglo XXI. Su potencial transformador es inmenso: puede contribuir a resolver problemas estructurales de acceso, desigualdad y exclusión; pero también puede amplificar las brechas existentes, consolidar asimetrías de poder y erosionar derechos fundamentales si no es regulada con criterios de justicia.

Este capítulo ha sostenido que América Latina y el Caribe (ALC) requiere una estrategia regional de gobernanza de IA fundada en los principios de la economía política: justicia

como equidad, libertad sustantiva⁴⁰ y expansión de capacidades humanas. Esta propuesta se distancia de enfoques tecnocráticos, extractivistas o autoritarios, y asume que la IA debe estar al servicio del desarrollo humano, de la democracia y del bien común.

Se ha argumentado que los marcos internacionales existentes —OCDE, UNESCO, Consejo de Europa, Unión Europea y China— ofrecen valiosos insumos normativos, pero resultan insuficientes o inaplicables si no se los adapta críticamente a las realidades institucionales, culturales y epistémicas de la región. La solución no radica en importar modelos, sino en construir uno propio, híbrido, inclusivo y sostenible.

La propuesta normativa presentada incluye una propuesta de Principios para una Gobernanza Ética de la IA en América Latina y el Caribe, la creación de fondos regionales para infraestructura digital soberana, observatorios ciudadanos de IA y una legislación que garantice transparencia, participación y justicia redistributiva. Todos estos componentes deben articularse dentro de una arquitectura abierta y cooperativa, capaz de fortalecer la soberanía tecnológica sin caer en nacionalismos digitales ni subordinaciones tecnocráticas.

En definitiva, gobernar la IA en ALC no es solo un reto técnico: es un proyecto ético y político. Requiere imaginar el futuro como un espacio de libertad ampliada, donde las tecnologías emergentes no refuercen la desigualdad, sino que abran caminos para la dignidad humana, la autonomía colectiva y el desarrollo con justicia.

“Tal como lo plantea Sen (1999), ‘la libertad es tanto el principal objetivo del desarrollo como su medio más eficaz’. Bajo este principio, América Latina y el Caribe tienen hoy la posibilidad histórica de diseñar una inteligencia artificial orientada a la expansión de libertades sustantivas, derechos colectivos y soberanía tecnológica, en el marco de un desarrollo ético y equitativo.”

Referencias

- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim Code*. Polity Press.
- CEPAL. (2020). *Cambio estructural para la igualdad: Una visión integrada del desarrollo*. Comisión Económica para América Latina y el Caribe. <https://www.cepal.org/es/publicaciones>
- CEPAL. (2021). *Agenda digital para América Latina y el Caribe (eLAC2022)*. <https://www.cepal.org/es/temas/sociedad-la-informacion>

40 El concepto de libertad sustantiva se refiere a una libertad real y efectiva, no solo formal o legal.

- Consejo de Europa. (2022). Recommendation on the impact of algorithmic systems on human rights. <https://www.coe.int>
- Couldry, N., & Mejias, U. A. (2019). *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press.
- Crespo, F., & Villalobos, C. (2022). *Gobernanza algorítmica en América Latina: Hacia una agenda democrática*. Banco Interamericano de Desarrollo.
- D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.
- Dworkin, R. (2000). *Sovereign virtue: The theory and practice of equality*. Harvard University Press.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Floridi, L. (2021). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.
- Gutmann, A., & Thompson, D. (2004). *Why deliberative democracy?* Princeton University Press.
- Hilbert, M. (2020). *Digital inequality in Latin America and the Caribbean: Measurement and policy recommendations*. Inter-American Development Bank.
- IA2030 LAC. (2022). *Declaración por una IA ética y justa en América Latina y el Caribe*. <https://ia2030lac.org>
- Laskai, L., & Webster, G. (2020). Translation: China's New Generation Artificial Intelligence Development Plan (2017). DigiChina Project, Stanford University. <https://digichina.stanford.edu>
- Latonero, M. (2020). *Governing artificial intelligence: Upholding human rights and dignity*. Data & Society Research Institute.
- Milan, S., & Treré, E. (2019). Big data from the South(s): Beyond data universalism. *Television & New Media*, 20(4), 319–335.
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
- Nussbaum, M. C. (2011). *Creating capabilities: The human development approach*. Harvard University Press.
- OECD. (2019). *OECD Principles on Artificial Intelligence*. Organisation for Economic Co-operation and Development. <https://www.oecd.org/going-digital/ai/principles/>
- Piketty, T. (2020). *Capital and ideology*. Harvard University Press.
- Purtova, N. (2018). The law of everything: Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology*, 10(1), 40–81.
- Rawls, J. (1971). *A theory of justice*. Harvard University Press.
- Rodrik, D. (2022). *Straight talk on trade: Ideas for a sane world economy*. Princeton University Press.
- Sen, A. (1999). *Development as freedom*. Oxford University Press.

- Susskind, D. (2022). A world without work: Technology, automation, and how we should respond. Metropolitan Books.
- Torres Jarrín, M. (2023). Artificial Intelligence and Foreign Affairs: Diplomacy, Geopolitics and Ethics in Latin America-Europe Relations.
- Trigo, E. (2020). Inteligencia artificial en la agricultura latinoamericana: Oportunidades y retos. Instituto Interamericano de Cooperación para la Agricultura (IICA).
- UNESCO. (2021). Recommendation on the ethics of artificial intelligence. <https://unesdoc.unesco.org>
- WHO. (2021). Ethics and governance of artificial intelligence for health. World Health Organization. <https://www.who.int/publications/i/item/9789240029200>

La gobernanza ética de la inteligencia artificial: *Big Tech Companies & Techplomacy*

Dr. Mario Torres Jarrín
*Director del Instituto de
Estudios Europeos y Derechos Humanos,
Facultad de Ciencias Humanas y Sociales de la
Universidad Pontificia de Salamanca, España.*

Las tecnologías disruptivas siempre han transformado el mundo en el que vivimos, sus surgimientos siempre han representado cambios, incertidumbre y desafíos, pero también oportunidades. Se entiende por tecnología disruptiva aquella tecnología que consigue provocar una transformación en los procesos y mecanismos anteriores a su aparición, así como un cambio en el comportamiento de los usuarios.

Dentro de este contexto, la inteligencia artificial, se ha convertido en una de las innovaciones tecnológicas, y término, que está en boga, tanto en los sectores empresariales, en los medios de comunicación, así como en el mundo político e incluso en el académico. Sin embargo, desde una perspectiva histórica, la inteligencia artificial es una tecnología disruptiva más, tal como lo fue la imprenta, la máquina de vapor, el automóvil, el teléfono, la televisión, el avión, las computadoras, el internet o los teléfonos inteligentes. Cada una de las tecnologías antes descritas desafiaron el orden establecido, generaron cambios en todos los ámbitos de una sociedad: jurídico, económico, político, medioambiental y cultural.

Lo novedoso del período en el que vivimos es que nunca antes en la historia de la humanidad, se ha tenido que hacer frente a tantas tecnologías disruptivas a la vez, y eso sí que es una complejidad y todo un reto, tanto para los gobiernos, como para las empresas y los

ciudadanos. Según Klaus Schwab la nueva era tecnológica podría catalizar un nuevo renacimiento cultural que permitirá sentirnos parte de algo mucho más grande que nosotros mismos: una verdadera civilización global, impulsada por una Cuarta Revolución Industrial⁴¹, la cual se caracteriza por una fusión de tecnologías que está desdibujando las líneas entre las esferas físicas, digital y biológica⁴². Esta nueva revolución no se basa en una sola tecnología, sino en la convergencia de múltiples avances tecnológicos disruptivos como la inteligencia artificial, la robótica, biotecnología, *blockchain*, computación cuántica, realidad aumentada y virtual, entre otros. Estos avances tecnológicos son impulsados por las llamadas *Big Tech Companies*, las cuales, debido a su poder de capitalización y capacidad de movilizar recursos financieros y humanos, son de facto los nuevos actores geopolíticos en el escenario internacional⁴³. Para que nos hagamos una idea de su poder, sólo si consideramos el volumen de capitalización de las tres primeras *Big Tech Companies*: Microsoft, Apple y Nvidia (juntas suman un total de 7,771 U.S. dólares, cifra que es superior al PBI de Alemania y Francia juntas 6,681 U.S. dólares).

La inteligencia artificial es una de las tecnologías más disruptivas de la Cuarta Revolución Industrial, su capacidad para aprender, crear contenidos y tomar decisiones ha transformado múltiples sectores, desde la educación, el transporte, la industria, la medicina, el sector financiero, la vida política o la seguridad y defensa nacional de los países, etc.

El mal uso de la inteligencia artificial puede vulnerar ciertos derechos humanos y fundamentales como el derecho a la privacidad (vigilancia masiva, recolección de datos biométricos sin consentimiento, rastreo de ubicación o reconocimiento facial en espacios públicos), derecho a la no discriminación (generación de sesgos algorítmicos que en algunos casos reproducen o amplifican prejuicios por raza, género, religión, sexo, etc.), derecho a la protección de datos personales o el derecho a la dignidad humana.

En 2017, Dinamarca se convirtió en el primer país del mundo en elevar la tecnología y la digitalización a una prioridad transversal de política exterior y de seguridad⁴⁴. En 2017, el gobierno danés presentó su Estrategia de Política Exterior y de Seguridad 2017-2018, durante su presentación Jeppe Kofod, ministro de asuntos exteriores de Dinamarca señaló:

41 Schwab, K. (2016). La Cuarta Revolución Industrial (1ª ed.), Debate.

42 Schwab, K. (2015). The Fourth Industrial Revolution. What it means and how to respond, Foreign Affairs.

43 Torres Jarrín, M. & Riordan, S. (2023). Science Diplomacy, Cyberdiplomacy and Techplomacy in EU-LAC relations, Springer.

44 Torres Jarrín, M. (2024). La Alianza Digital Unión Europea-América Latina y el caribe: La geopolítica en la era de la gobernanza global tecnológica, en Vera Delzo, P. E. y Ortiza, R. D. América Latina en la Nueva Geopolítica Global. Centro de Estudios Estratégicos del Ejército del Perú y Strategic Studies Institute US Army War College.pp.139-157.

*“Hemos sido demasiado ingenuos durante mucho tiempo sobre la revolución tecnológica. Necesitamos asegurarnos de que los gobiernos democráticos establezcan los límites para la industria tecnológica, y no al revés. Ahí es donde entra en juego la iniciativa danesa Techplomacy”.*⁴⁵

A la vez, la estrategia introducía tres nuevos conceptos en las disciplinas de las relaciones internacionales y la diplomacia: *Techplomacy*, *Tech Embassies* & *Tech Ambassador*.

El primer concepto, *Techplomacy*, se refiere a desarrollar una política exterior y de seguridad focalizada a las *Big Tech Companies*. Por consiguiente, busca establecer relaciones formales con estas empresas con el fin de conocer sus planes, sus avances tecnológicos y los nuevos bienes o servicios que comercializarán, con el objetivo de abordar los desafíos, oportunidades y consecuencias jurídicas, económicas, geopolíticas, sociales, culturales y medioambientales asociadas con dichas tecnologías. La *Tech Diplomacy* pretende ser un instrumento diplomático mediante el cual los gobiernos, organismos internacionales, organismos de integración regional, empresas tecnológicas y otros actores vinculados con el sector tech trabajen juntos para influir en el desarrollo y uso ético de tecnologías avanzadas como la inteligencia artificial.

El segundo concepto, *Tech Embassies*, es la puesta en marcha de la *Techplomacy*, a través de la apertura de representaciones diplomáticas permanentes en las ciudades donde están establecidas y concentradas las llamadas *Big Tech Companies*, como son los casos de Silicon Valley (*hub* tecnológico donde tienen sus sedes empresas como Apple, Google, Meta, Intel, Nvidia, Tesla, etc.) o Beijing (Huawei, Tencent, Alibaba, etc). Dinamarca abrió dos Tech Embassies en estas dos ciudades. Aunque es probable que se amplie el número de estas nuevas *Tech Embassies*, considerando que existen otros “*Silicon Valley*” en el mundo, por ejemplo, en Asia (Shenzhen, Bangalore y Singapur), en Europa (Londres, Berlín y Estocolmo), Medio Oriente (Tel Aviv), en América (Toronto, Sao Paulo, Ciudad de México y Buenos Aires).

En lo concerniente al tercer concepto, el primer *Tech Ambassador* que nombró Dinamarca fue Casper Klynge (2017-2020), en 2020 se nombró una segunda *Tech Ambassador*, Anne Marie Engtoft Larsen. Desde 2017 hasta la actualidad, otros países se han sumado a la iniciativa danesa, como Francia, Bulgaria, Estados Unidos, Estonia, Suiza, Australia, Alemania, India, Brasil, cuentan con una estrategia de política exterior orientada al sector tech, y utilizan la *Techplomacy*, también han comenzado a nombrar sus primeros *Tech Ambassadors*, inclusive la Unión Europea y las Naciones Unidas han nombrado “Enviados Especiales”, con rango de embajador, para que trabajen de cerca con las *Big Tech Companies*.

⁴⁵ Office of the Tech Ambassador of Denmark: <https://techamb.um.dk/>

El tema de la gobernanza ética de la inteligencia artificial, es un área de trabajo contemplada dentro de la *Techplomacy*, dado que la inteligencia artificial es una de las áreas desarrolladas, gestionadas y lideradas por las *Big Tech Companies*. Por consiguiente, podríamos indicar que el origen de buscar una gobernanza ética de la inteligencia artificial nace en el momento que Dinamarca nombra al primer *Tech Ambassador*. Posteriormente, en 2018, Google publica sus “*AI Principles*” tras la controversia del *Project Maven* con el Pentágono. En 2019, la OCDE publica sus Principios sobre IA, el cual representó el primer documento emitido por un organismo internacional que abogaba por un uso ético y una gobernanza ética de la IA.

Ese mismo año, la Unión Europea lanza su Ética para una IA confiable. Desde entonces y hasta la fecha observamos una carrera entre varios países organismos internacionales y organismos de integración regional por liderar la gobernanza ética de la IA a nivel mundial, tales son los casos de la UNESCO que, en 2021, adopta el primer marco global ético sobre IA. Entre 2022-2023 la Unión Europea inició sus trabajos en lo que hoy es el reglamento sobre IA: EU AI Act.

Por otro lado, el G7, G20 y la ONU comienzan a trabajar sobre la necesidad de crear marcos regulatorios y definir estándares a nivel global sobre el desarrollo de una IA segura y ética. Entre el período 2024-2025 aumentan los llamados por una gobernanza tecnológica a nivel global, Reino Unido, Francia y Corea del Sur lanzan cumbres temáticas en torno a una IA ética y segura.

Cabe señalar que, ante esta constelación de países y organismos regionales e internacionales, se suman iniciativas como la de la Unión Internacional de las Telecomunicaciones, Foro Económico Mundial y las del Consejo de Europa.

Entre los principios comunes que definen la gobernanza ética de la inteligencia artificial podemos mencionar las siguientes áreas:

- Transparencia
- Responsabilidad
- No discriminación
- Privacidad y protección de datos
- Sostenibilidad
- Seguridad

La gobernanza ética de la IA se inició como una reflexión académica y evolucionó a una prioridad global, más aún cuando los gobiernos comenzaron a ver amenazada su seguridad nacional. El mal uso de la IA alteraba la vida política de los países, desestabilizaba gobiernos,

democráticamente elegidos, y tiene la capacidad de influir en la opinión pública a través de campañas de desinformación.

Decimos que los debates sobre una IA ética y su gobernanza se inició siendo un debate académico, porque durante la década los 1940, Norbert Wiener, comenzó a reflexionar y estudiar sobre los efectos sociales de los sistemas automatizados. En 1942, las “Tres Leyes de la Robótica de Asimov” creó un marco de ficción que influyó en la discusión ética de las máquinas autónomas. Entre las décadas de los años 1970-1990, conceptos como bioética y ética informática surgen y contribuyen al debate académico sobre la necesidad de una ética en los avances tecnológicos. Entre los principales dilemas prácticos estaban casos como los algoritmos de vigilancia y sus sesgos algorítmicos; la automatización, uso de drones y armas autónomas, el reconocimiento facial y la privacidad de los datos, entre otros.

Cada uno de los conceptos y casos antes descritos, nos demuestra que la humanidad vive una nueva etapa dentro de su historia. Así mismo, la historia de las relaciones internacionales también entra a un nuevo período, el cual es liderado por las Big Tech Companies, por ser los nuevos actores globales y los actores geopolíticos que marcan el paso de un nuevo orden mundial basado en las tecnologías.

Considerando que, las *Big Tech Companies*, tienen un poder cuasi soberano dado que gestionan datos, usan algoritmos y plataformas a nivel global, donde los gobiernos no tienen ni la capacidad financiera ni los recursos humanos para competir con ellos, y que las leyes que pueden crear (para regular su comportamiento, así como sus bienes y servicios que produzcan) siempre serán anacrónicas, debido al vertiginoso ritmo de aparición de nuevas tecnologías. Por tanto, los gobiernos pueden utilizar la Techplomacy para poder ir al paso de las Big Tech Companies y tratar de influir dentro de la Gobernanza Global Tecnológica⁴⁶.

Por otro lado, la Techplomacy, puede también promover acuerdos multilaterales para establecer normas comunes a nivel globales. Los avances normativos creados por la Unión Europea, como el *Digital Services Act*⁴⁷ o el EU AI Act⁴⁸, pueden servir de referencia para diseñar acuerdos globales.

46 Torres Jarrín, M. (2023). Rethinking EU-CELAC Interreligiousism in the Digital World: Techplomacy as a Foreign Policy Instrument for Global Tech Governance, in Gardini, G.L.. The redefinition of the EU presence in Latin America and the Caribbean, Peter Lang, 199-213.

47 European Commission, The Digital Service Act.

<https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act>

48 European Parliament (2025). EU AI Act.

<https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

Otro tema que está surgiendo y que es de gran interés, al Estado Nación, es el concepto de soberanía digital de los Estados frente a la concentración del poder de las *Big Tech Companies* en las ciudades *tech hub* como Silicon valley, Beijing o Shenzhen, entre otros. Las tecnologías emergentes y las tecnologías disruptivas pueden desestabilizar la seguridad global, si no se gobiernan adecuadamente. Se entiende por soberanía digital a la capacidad de un Estado Nación o institución para controlar y regular sus propios datos, infraestructuras digitales y decisiones tecnológicas, sin depender de actores externos, como puede ser el caso de la dependencia que tienen los Estados Nación con las *Big Tech Companies*.

Los gobiernos comienzan a cuestionarse la necesidad de desarrollar una autonomía estratégica digital, ciberseguridad, y control de la IA y la computación en la nube, dado que la infraestructura digital la poseen y controlan las *Big Tech Companies*, y es cuando surgen los debates sobre la privacidad de datos y vigilancia digital.

Los casos vigilancia masiva en los Estados Unidos, evidencia la dependencia de servidores y servicios digitales extranjeros, países como Alemania, Francia y China comienzan a hablar explícitamente de soberanía digital, de ahí que la Unión Europea, decide promover la creación del Reglamento General de Protección de Datos y la Estrategia Digital Europea. Países como China y Rusia abogan por un control estatal del ciberespacio.

Si al inicio del siglo pasado, en 1904, cuando Halford John Mackinder, presenta su teoría geopolítica del *Heartland*⁴⁹, defendiendo la relevancia del estudio al espacio territorial en términos de dominación y ejercicio de poder, y cuya tesis de Mackinder es quien gobierne Europa del Este gobernará el *Heartland*, quien gobierne el *Heartland* gobernará la Isla Mundo, y quien gobierne la Isla Mundo gobernará el mundo. Entonces tendríamos que decir, que ahora existe una nueva región no geográfica, llamada ciberespacio, la cual es desarrollada, gestionada y controlada por las *Big Tech Companies*, pero quien controle el ciberespacio controlara el mundo⁵⁰.

Por todo lo antes expuesto, la Techplomacy, es un instrumento de política exterior que puede servir de herramienta clave para regular las *Big Tech Companies*, promover el uso ético de la inteligencia artificial⁵¹ y proteger la paz y la seguridad mundial.

49 Mackinder, H. J. (1904). The geographical pivot of history. *The Geographical Journal*, 23(4), 421–437.

50 Argumosa Pila, J. Garofali Acosta, A. and Torres Jarrín, M. (Editors) (2021). *Emerging Markets & Emerging Powers. The future of the Asia-Pacific in the New World Order*. Tirant Lo Blanch.

51 Torres Jarrín, M. (2021). La UE & la gobernanza ética de la inteligencia artificial: Inteligencia artificial & diplomacia. Vol. 48 (2021). Cuadernos Salmantinos de Filosofía, Salamanca: Universidad Pontificia de Salamanca.

Inteligencia artificial y protección de datos personales: el caso del RGPD de la Unión Europea y sus implicaciones para América Latina y el Caribe

Dra. Mónica Velasco Pufleau⁵²

*Directora de la Cátedra sobre Relaciones
Unión Europea-América Latina y el Caribe del
European Institute of International Studies (Suecia)*

En abril de 2016, el Parlamento Europeo (PE) y el Consejo de la Unión Europea (UE) adoptaron el Reglamento General de Protección de Datos (RGPD),⁵³ en tanto colegisladores en esta materia, con fundamento en lo previsto en el Artículo 16 del Tratado de Funcionamiento de la UE (Moraes, 2018, p. 36). Desde el 25 de mayo de 2018, el RGPD es de aplicación directa en todos los Estados Miembros (EM) de la UE. Por ser un texto pertinente a efectos del Espacio Económico Europeo (EEE), este reglamento se aplica también a tres países de la Asociación Europea de Libre Comercio; es decir, a Islandia, Noruega y Liechtenstein.⁵⁴ El apartado 1 del Artículo 4 del RGPD define a los “datos personales” como:

52 Directora de la Cátedra sobre Relaciones Unión Europea-América Latina y el Caribe del *European Institute of International Studies* desde marzo de 2021. Tiene un doctorado en Derecho, especialidad Estudios Internacionales, por la Universidad de Barcelona y un postdoctorado por la Universidad de Luxemburgo. La autora agradece a la Comisión Nacional de Protección de Datos (“*Commission nationale pour la protection des données*”) de Luxemburgo por su valiosa información y clarificaciones proporcionadas como parte de sus actividades de formación sobre el RGPD. El idioma original de esta contribución es el español. La autora agradece a los servicios del SELA por su traducción al inglés.

53 Reglamento (UE) 2016/679 del PE 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 (Reglamento general de protección de datos) (Texto pertinente a efectos del EEE). Diario Oficial de la UE (DOUE) L 119/1, 4 de mayo de 2016. El procedimiento legislativo completo para la adopción del RGPD está disponible en: [https://oeil.secure.europarl.europa.eu/oeil/en/procedure-file?reference=2012/0011\(COD\)](https://oeil.secure.europarl.europa.eu/oeil/en/procedure-file?reference=2012/0011(COD)) (último acceso: 29 de mayo de 2025).

54 Suiza es país miembro de la Asociación Europea de Libre Comercio, pero no del EEE.

toda información sobre una persona física identificada o identificable («el interesado»); se considerará persona física identificable toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un identificador, como por ejemplo un nombre, un número de identificación, datos de localización, un identificador en línea o uno o varios elementos propios de la identidad física, fisiológica, genética, psíquica, económica, cultural o social de dicha persona.⁵⁵

Por su parte, el apartado 2 del citado Artículo 4 del RGPD establece como definición del “tratamiento” de datos personales:

cualquier operación o conjunto de operaciones realizadas sobre datos personales o conjuntos de datos personales, ya sea por procedimientos automatizados o no, como la recogida, registro, organización, estructuración, conservación, adaptación o modificación, extracción, consulta, utilización, comunicación por transmisión, difusión o cualquier otra forma de habilitación de acceso, cotejo o interconexión, limitación, supresión o destrucción;

La adopción del RGPD buscó garantizar que las personas físicas –referidas como “interesados” en el RGPD en su versión en español y “*data subjects*” en su versión en inglés– contaran con una protección más estricta y armonizada en lo que respecta al tratamiento de sus datos personales en toda la UE y el EEE,⁵⁶ al tiempo que las empresas disfrutaran de condiciones de competencia equitativas (Consejo de la UE, 2024). Para ello, el RGPD introdujo una serie de innovaciones respecto al régimen de protección de datos personales preexistente bajo la conocida como “Directiva de Protección de Datos” de 24 de octubre de 1995,⁵⁷ que fue derogada por el RGPD. Estas innovaciones incluyen el fortalecimiento de los derechos de las personas físicas para ejercer un control efectivo sobre sus datos personales, así como de las competencias de las autoridades de supervisión para garantizar el disfrute de dichos derechos. Por ejemplo, el RGPD otorga en su Artículo 17 una base legal específica al llamado “derecho al olvido”, como parte del derecho de supresión (Kohl, 2023, pp. 750-754). Bajo ciertas condiciones, el derecho al olvido permite a la persona interesada solicitar al proveedor de un motor de búsqueda en línea que elimine uno o varios enlaces a páginas web desde la lista de resultados obtenidos como producto de una búsqueda realizada a partir de su nombre (Comité Europeo de Protección de Datos –CEPD–, 2020, p. 4).⁵⁸

55 Sobre un análisis relacionado con la distinción entre “datos personales” y “datos no personales” en la literatura, ver Finck & Pallas (2020).

56 Sobre la flexibilidad otorgada a los EM en la implementación del RGPD, ver Yuliyanova Chakarova (2019).

57 DOUE L 281/31, 23 de noviembre de 1995.

58 Sobre los antecedentes de este derecho en la jurisprudencia europea, ver *Google Spain SL y Google Inc. v Agencia Española de Protección de Datos (AEPD) y Mario Costeja González*, C131/12, ECLI:EU:C:2014:317.

Adicionalmente, resulta ilustrativo que las autoridades nacionales de protección de datos cuentan ahora con la posibilidad de imponer multas administrativas más significativas, en caso de incumplimiento con las disposiciones del RGPD (Ruohonen & Hjerpe, 2022). Estas multas pueden alcanzar hasta 20 millones de euros o, tratándose de una empresa, el equivalente al 4% del volumen de su negocio total anual correspondiente al ejercicio financiero anterior, cuando dicho incumplimiento tenga que ver, por ejemplo, con los principios relativos al tratamiento de datos personales, los derechos de las personas físicas o la transferencia de datos personales a países terceros u organizaciones internacionales (Artículo 83 del RGPD). Hasta la fecha, algunas de las mayores multas administrativas han sido impuestas a raíz de las actividades de las grandes tecnológicas como Meta, Amazon y Google –esta última integrada en el conglomerado Alphabet– (Data Privacy Manager, 2025).

Es relevante mencionar que los principios relativos al tratamiento previstos por el RGPD se refieren a: (a) la licitud, lealtad y transparencia, (b) la limitación de la finalidad, (c) la minimización de datos, (d) la exactitud, (e) la limitación del plazo de conservación, (f) la integridad y confidencialidad y (g) la responsabilidad proactiva (Artículo 5 del RGPD). Por su parte, además del derecho al olvido ya mencionado como parte del derecho de supresión, el RGPD establece como derechos de las personas físicas: (i) el derecho a ser informado, (ii) el derecho de acceso, (iii) el derecho de rectificación, (iv) el derecho a la limitación del tratamiento, (v) el derecho a la portabilidad de los datos, (vi) el derecho de oposición y (vii) el derecho a no ser objeto de decisiones individuales automatizadas, incluida la elaboración de perfiles (Artículos 12 al 22 del RGPD). Estos derechos corresponden a una serie de obligaciones impuestas a los responsables y encargados del tratamiento de los datos personales, según lo dispuesto en el Capítulo IV del RGPD.

Por ser un reglamento con carácter “ómnibus”, la aplicación del RGPD es multisectorial, imponiendo obligaciones tanto a entidades públicas como privadas, independientemente del sector al que pertenezcan (Schwartz, 2013, p. 1974).⁵⁹ Dentro de la estrategia digital de la UE, esto significa que el RGPD actúa como una base normativa transversal que garantiza la protección de los datos personales en todos los ámbitos de la transformación digital. Entre estos ámbitos está el de la inteligencia artificial (IA), objeto de la conocida como “Ley de IA”,⁶⁰ que constituye el primer marco jurídico exhaustivo sobre IA en el mundo (Comisión Europea –CE–, 2025). En palabras de la CE, la Ley de IA “establece un conjunto claro de normas basadas en el riesgo para los desarrolladores y los implementadores de IA en relación con los usos específicos de la IA” (Ibid.). Estos riesgos, definidos como “la combinación de la probabilidad de que se produzca un perjuicio y la gravedad de dicho

⁵⁹ Sobre las excepciones previstas para la aplicación del RGPD, ver por ejemplo el Considerando 19 del RGPD.

⁶⁰ DOUE L 2024/1689, 12 de julio de 2024.

perjuicio” (Artículo 3.2 de la Ley de IA), se engloban en cuatro niveles distintos: riesgo inaceptable, alto riesgo, riesgo limitado y riesgo mínimo.

En cuando a la relación entre la Ley de IA y el RGPD en materia de protección de datos personales, se puede observar de manera ejemplificativa lo siguiente:

- 1)** Existe alineación entre el espíritu general y los principios fundamentales de la Ley de IA y los del RGPD. Notablemente, esto incluye la prioridad concedida por ambos instrumentos al respeto de los derechos fundamentales. En el caso de la Ley de IA, por ejemplo, su Artículo 1 estipula como parte del objeto de la Ley la promoción de una IA centrada en el ser humano, garantizando los derechos fundamentales previstos en la Carta de Derechos Fundamentales de la UE. Por su parte, el Artículo 1 del RGPD menciona como parte de su objeto la protección de los derechos y libertades fundamentales de las personas físicas y, en particular, su derecho a la protección de los datos personales. En cuanto a los principios, un ejemplo significativo está en el principio de transparencia, especialmente en relación con los sistemas de IA calificados de alto riesgo (ver, entre otros, Artículo 13 de la Ley IA).
- 2)** La Ley de IA entra en vigor sin perjuicio de las obligaciones y derechos establecidos por el RGPD, elevando la protección relativa al tratamiento de los datos personales en algunos casos, pero sin limitarla (Considerandos 8-10 de la Ley de IA). En este sentido es ilustrativo el apartado 5 del Artículo 10 de la Ley de IA sobre “Datos y gobernanza de datos”. Este apartado establece que, además de las disposiciones del RGPD, deben de cumplirse con una serie de condiciones específicas para tratar excepcionalmente categorías especiales de datos personales,⁶¹ cuando dicho tratamiento sea estrictamente necesario para garantizar la detección y corrección de los sesgos asociados a los sistemas de IA de alto riesgo de conformidad con lo dispuesto en el apartado 2, letras f) y g), de dicho Artículo 10.
- 3)** En estrecha relación con lo expuesto previamente, las disposiciones de la Ley de IA, en tanto legislación sectorial, y las del RGPD, en tanto legislación omnibus, se refuerzan mutuamente. Como se puede ver en el caso citado en el inciso anterior, respecto al uso de categorías especiales de datos personales, la Ley de IA fundamenta largamente sus disposiciones en las disposiciones del RGPD. En cuanto al RGPD, este garantiza una protección general a las personas físicas respecto al tratamiento de sus datos personales, que resulta fortalecida gracias a la especificidad regulatoria de la Ley de IA, misma que proporciona un nivel adicional de protección a los interesados.

⁶¹ Sobre estas categorías, ver Artículo 9 del RGPD.

Fuera de la UE y el EEE, el RGPD ha tenido un fuerte impacto en los mercados globales, especialmente el relacionado con la economía digital (Bradford, 2020, p. 131). Asimismo, ha inspirado la adopción o actualización de marcos legales relativos a la protección de la privacidad y los datos personales en numerosos países alrededor del mundo (Greenleaf, 2021, pp. 4-5). Un caso emblemático en América Latina es el de Brasil, cuya Ley General de Protección de Datos vigente (*“Lei Geral de Proteção de Dados Pessoais”*, por su nombre en portugués)⁶² refleja un alto grado de “europeización”, resultado de la influencia del RGPD (Filipec et al., 2022, p. 94). Durante el proceso de adopción de esta ley, se reconocieron no sólo motivaciones económicas y comerciales para armonizar la legislación brasileña con la de la UE, sino también una necesidad social para reforzar la protección del derecho a la privacidad y otros derechos fundamentales en el país (Ibid., pp. 93-94). De esta forma, la *“Lei Geral de Proteção de Dados Pessoais”* refleja la esencia misma del modelo europeo que, en comparación con otros modelos promovidos por China o Estados Unidos, busca poner al ser humano al centro de la transformación digital, dando prioridad a la protección de sus derechos fundamentales (Bradford, 2023, p. 9). Entre estos derechos se encuentra, precisamente, el derecho a la protección de datos personales previsto en el Artículo 8 de la Carta de Derechos Fundamentales de la UE.⁶³

Más allá de su papel en la convergencia regulatoria y el establecimiento de estándares internacionales, el RGPD ejerce efectos extraterritoriales significativos para las entidades responsables o encargadas del tratamiento de datos personales que se encuentran fuera de la UE/EEE, incluidas aquellas establecidas en América Latina y el Caribe (ALC). Para entender estos efectos es indispensable referirse al Artículo 3 del RGPD, relativo al “Ámbito territorial”, mismo que a la letra dice:

- 1.** El presente Reglamento se aplica al tratamiento de datos personales en el contexto de las actividades de un establecimiento del responsable o del encargado en la Unión, independientemente de que el tratamiento tenga lugar en la Unión o no.
- 2.** El presente Reglamento se aplica al tratamiento de datos personales de interesados que residan en la Unión por parte de un responsable o encargado no establecido en la Unión, cuando las actividades de tratamiento estén relacionadas con:
 - a)** la oferta de bienes o servicios a dichos interesados en la Unión, independientemente de si a estos se les requiere su pago, o

⁶² Ley Nº 13.709, de 14 de agosto de 2018.

⁶³ Sobre la evolución de este derecho en la UE, ver González Fuster (2014).

- b)** el control de su comportamiento, en la medida en que este tenga lugar en la Unión.
- 3.** El presente Reglamento se aplica al tratamiento de datos personales por parte de un responsable que no esté establecido en la Unión sino en un lugar en que el Derecho de los Estados miembros se aplique en virtud del Derecho internacional público.

A partir del contenido de dicho Artículo 3, pueden identificarse dos supuestos analizados con más detalle en esta contribución.⁶⁴ El primero, previsto en el apartado 1, sujeta a un establecimiento situado en la UE/EEE al cumplimiento del RGPD, pero también a aquel ubicado fuera de dicho espacio, cuando este realice un tratamiento de datos personales en el contexto de las actividades del establecimiento localizado en la UE/EEE. Es decir, este supuesto incorpora el “criterio del establecimiento”, por citar a la CEPD (2019, p. 5).

En relación con lo anterior, es importante señalar que toda transferencia de datos personales fuera de la UE/EEE deberá observar lo previsto en el Capítulo V del RGPD, relativo a las “Transferencias de datos personales a terceros países u organizaciones internacionales”, para asegurar su legalidad. Por ejemplo, una forma de realizar transferencias válidas bajo el RGPD requiere que exista una “decisión de adecuación” adoptada por la CE en relación con países terceros, siguiendo el procedimiento previsto en el Artículo 45 del RGPD. Si esta decisión existe, las transferencias de datos personales realizadas al país o países terceros en cuestión desde la UE/EEE podrán tener lugar sin autorización o garantía adicional específica.

Hoy por hoy, sólo Argentina y Uruguay cuentan con decisiones de adecuación vigentes en toda ALC. Estas decisiones fueron adoptadas, respectivamente, en 2003 y 2012 (Bradford, 2020, p. 150). Ambas decisiones fueron además confirmadas por la CE como parte de la primera revisión del funcionamiento de las decisiones de adecuación publicada el 15 de enero de 2024. Junto con estas decisiones, la CE confirmó otras decisiones con países como Suiza, Nueva Zelanda, Israel y Canadá (CE, 2024a). En el caso de Brasil, existen negociaciones con vistas a la adopción de una decisión de adecuación con la UE que, en palabras del Servicio Europeo de Acción Exterior (SEAE, 2024), “permita que la información se mueva sin restricciones, fomentando así el comercio, la economía digital, la investigación y en general la cooperación entre la UE y Brasil”.

Es fundamental tener en cuenta que el incumplimiento de las disposiciones previstas para las transferencias internacionales de datos desde la UE/EEE conlleva consecuencias importantes para el responsable o el encargado del tratamiento de los datos. Un ejemplo destacado es la multa impuesta en 2023 a Meta Platforms Ireland Limited, que ascendió a 1,200 millones de euros. Esta multa, la más elevada jamás impuesta bajo el RGPD, fue resultado de una inves-

⁶⁴ Sobre el tercer supuesto, previsto en Artículo 3.3 del RGPD, ver CEPD (2019, pp. 25-26).

tigación llevada a cabo sobre Facebook por la autoridad nacional de datos de Irlanda; país donde Meta tiene su sede central en la UE (European Data Protection Board –EDPB, 2023). Asimismo, resulta relevante mencionar, aunque sea muy brevemente, que las transferencias internacionales de datos han sido objeto de intensas controversias jurídicas. Es emblemática la serie de litigios conocida como “saga Schrems”, que culminó con la invalidación de dos decisiones de adecuación con Estados Unidos por parte de la Corte de Justicia de la UE en 2015 y 2020, respectivamente (Tzanou, 2021).

Por su parte, el supuesto descrito en el apartado 2 del Artículo 3 se relaciona con el “criterio de la selección de destinatarios”; es decir, interesados que residan en la UE/EEE (CEPD, 2019, p. 15). Esto significa que la ausencia de un establecimiento en la UE/EEE no implica necesariamente que el tratamiento de datos personales por una entidad establecida fuera de este territorio quede excluido de la aplicación del RGPD. De hecho, este tratamiento quedará cubierto, de acuerdo con los incisos a) y b) del apartado en comento, cuando las actividades de tratamiento se relacionen con la oferta de bienes o servicios a dichos interesados o con el control de su comportamiento en la UE/EEE. En estos casos, los responsables y encargados del tratamiento deberán de cumplir con las obligaciones previstas para dicho tratamiento dentro de los marcos jurídicos de sus respectivos países, además de con lo previsto por el RGPD (Ibid., pp. 15-16, 25).

Con ciertas excepciones, las obligaciones de los establecimientos ubicados fuera de la UE/EEE, cuyas actividades de tratamiento de datos personales se enmarcan en el apartado 2 del Artículo 3 del RGPD, incluyen la designación por escrito de un representante en la Unión en uno de los EM en que estén los interesados cuyos datos personales se traten en el contexto de una oferta de bienes o servicios, o cuyo comportamiento esté siendo controlado (Artículo 27 del RGPD).⁶⁵ A falta de este representante, las autoridades de protección de datos de distintos EM, según sea aplicable, podrán ser competentes para conocer de incumplimientos a las disposiciones del RGPD por tales establecimientos. En su estudio “*Challenges to the extraterritorial enforcement of data privacy law – EU case study*”, Michal Czerniawski y Dan Svantesson (2023) aportan un ejemplo interesante relacionado, precisamente, con el uso de la IA. En particular, se refieren al caso de Clearview AI, una compañía basada en Estados Unidos que presta servicios de reconocimiento facial para apoyar tareas de aplicación de la ley. En el presente, su base de datos tiene más de 60 mil millones de imágenes, según la página de Internet de la propia compañía (<https://www.clearview.ai/>). Czerniawski y Svantesson señalan que, entre 2022 y 2023, las autoridades de protección de datos de cuatro EM –Italia, Grecia, Francia y Austria– adoptaron decisiones en contra de Clearview AI. En sus decisiones, “concluyeron que la empresa infringió el RGPD al realizar reconocimiento

⁶⁵ Sobre este tema, ver Considerando 80 y Artículo 4.17 del RGPD, así como CEPD (2019, pp. 26-31).

facial en fuentes públicas de la web y prohibieron su procesamiento posterior”.⁶⁶ Como consecuencia, la compañía recibió multas por la cantidad total de 60 millones de euros (p. 140). Adicionalmente, una multa de 30.5 millones de euros fue impuesta a Clearview AI por la autoridad de protección de datos de Países Bajos en 2024 (Data Privacy Manager, 2025).

De igual manera, otro ejemplo importante en el ámbito de la IA se refiere al de la compañía OpenAI (<https://openai.com/>), también basada en Estados Unidos y desarrolladora de ChatGPT. En este caso, a falta de un representante nombrado de conformidad con el citado Artículo 27 del RGPD, varias autoridades de protección de datos de los EM iniciaron investigaciones sobre el (in)cumplimiento de la compañía con este reglamento, tras el lanzamiento de ChatGPT en 2022.⁶⁷ Para coordinar estas investigaciones nacionales, el CEPD creó un grupo de trabajo sobre ChatGPT (conocido como “*ChatGPT Taskforce*”, por su nombre en inglés), que publicó un primer informe en mayo de 2024 (EDPB, 2024). Poco después, por ejemplo, la autoridad de protección de datos italiana impuso una multa a OpenAI por 15 millones de euros, ordenándole además llevar a cabo una campaña de información de 6 meses para promover la comprensión y la concienciación pública sobre el funcionamiento de ChatGPT (Garante, 2024).

Como se ha dicho previamente (Velasco Pufleau, 2024), la Alianza Digital entre la UE y ALC ofrece una oportunidad estratégica para abordar de manera conjunta los desafíos y beneficios de la transformación digital, incluyendo en materia de protección de datos e IA. Esta voluntad de cooperación se ha materializado, por ejemplo, en la creación de espacios como el Diálogo Político de Alto Nivel sobre Gobernanza de Datos, celebrado en Montevideo en 2024. En dicho encuentro se abordaron temas clave como la regulación de la protección de datos personales, el rol de las autoridades independientes de supervisión, las transferencias internacionales de datos y la intersección entre la IA y los datos (CE, 2024b; SEAE, 2024). Avanzar hacia una gobernanza de datos basada en principios y valores compartidos, como el respeto a los derechos fundamentales y el fortalecimiento de la democracia, resulta esencial para garantizar un bienestar justo, sostenible e inclusivo en ambas regiones.

⁶⁶ La traducción al español en la correspondiente publicación del SELA ha sido realizada por la autora.

⁶⁷ De acuerdo con el CEPD, OpenAI tiene un establecimiento en la Unión sólo desde el 15 de febrero de 2024. Las múltiples investigaciones antes referidas son válidas hasta esa fecha (EDPB, 2024, p. 5).

Referencias

Bradford, A. (2020). *The Brussels Effect: How the European Union Rules the World*. Oxford: Oxford University Press.

Bradford, A. (2023). *Digital Empires: The global battle to regulate technology*. Oxford: Oxford University Press.

CE (2024a). *Informe de la Comisión al Parlamento Europeo y al Consejo sobre la primera revisión del funcionamiento de las decisiones de adecuación adoptadas con arreglo al artículo 25, apartado 6, de la Directiva 95/46/CE*. Doc. COM/2024/7 final, Bruselas.

CE (2024b). *Diálogo político de alto nivel UE-ALC sobre la gobernanza de los datos*. https://international-partnerships.ec.europa.eu/eu-latin-america-and-caribbean-digital-alliance/eu-lac-digital-alliance-high-level-policy-dialogue-data-governance-2024-09-30_en?pre-flang=es&etrans=es (último acceso: 29 de mayo de 2025).

CE (2025). *Ley de IA*. <https://digital-strategy.ec.europa.eu/es/policies/regulatory-framework-ai> (último acceso: 29 de mayo de 2025).

CEPD (2019). *Directrices 3/2018 relativas al ámbito territorial del RGPD (artículo 3)*. https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_3_2018_territorial_scope_es.pdf (último acceso: 29 de mayo de 2025).

CEPD (2020). *Directrices 5/2019 sobre los criterios del derecho al olvido en los casos de motores de búsqueda en virtud del RGPD (1.ª parte)*. https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201905_rtbsearchengines_afterpublicconsultation_es.pdf (último acceso: 29 de mayo de 2025).

Consejo de la UE (2024). *Protección de datos en la UE*. <https://www.consilium.europa.eu/es/policies/data-protection> (último acceso: 29 de mayo de 2025).

Czerniawski, M. & Svantesson, D. (2023). Challenges to the extraterritorial enforcement of data privacy law – EU case study. En M. Brinnen et al. (eds.), *Dataskyddet 50 år – historia, aktuella problem och framtid* (pp. 127-153). Visby: eddy.se ab.

Data Privacy Manager (2025). 20 biggest GDPR fines so far [2025]. <https://dataprivacymanager.net/5-biggest-gdpr-fines-so-far-2020/> (último acceso: 29 de mayo de 2025).

EDPB (2023). *1.2 billion euro fine for Facebook as a result of EDPB binding decision*. https://www.edpb.europa.eu/news/news/2023/12-billion-euro-fine-facebook-result-edpb-binding-decision_es (último acceso: 29 de mayo de 2025).

EDPB (2024). *Report of the work undertaken by the ChatGPT Taskforce*. https://www.edpb.europa.eu/system/files/2024-05/edpb_20240523_report_chatgpt_taskforce_en.pdf (último acceso: 29 de mayo de 2025).

Filipec, O. et al. (2022). Personal Data Protection in Brazil: How Much Europeanization? *International and Comparative Law Review*, 22(2), 81-104.

Finck, M. & Pallas, F. (2020). They who must not be identified—distinguishing personal from non-personal data under the GDPR. *International Data Privacy Law*, 10(1), 11-36.

Garante (2024). *ChatGPT, The Italian Data Protection Authority closes the preliminary investigation*. <https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/10085432#english> (último acceso: 29 de mayo de 2025).

González Fuster, G. (2014). *The Emergence of Personal Data Protection as a Fundamental Right of the EU*. Cham: Springer.

Greenleaf, G. (2021). Global Data Privacy Laws 2021: Despite COVID Delays, 145 Laws Show GDPR Dominance. *Privacy Laws & Business International Report*, 169, 1, 3-19.

Kohl, U. (2023). The Right to be Forgotten in Data Protection Law and Two Western Cultures of Privacy. *International & Comparative Law Quarterly*, 72(3), 737-769.

Moraes, C. (2018). The European Parliament and Transatlantic Relations: Personal Reflections. En E. Fahey (ed.), *Institutionalisation Beyond the Nation State: Transatlantic Relations: Data, Privacy and Trade Law* (pp. 31-38). Cham: Springer.

Ruohonen, J. & Hjerpe, K. (2022). The GDPR enforcement fines at glance. *Information Systems*, 106, 101876.

Schwartz, P. M. (2013). The EU-U.S. Privacy Collision: A Turn To Institutions And Procedures. *Harvard Law Review*, 126(7), 1966-2009.

SEAE (2024). *La UE y sus socios de América Latina y el Caribe aprovechan el uso responsable de los datos*. https://www.eeas.europa.eu/node/445574_es (último acceso: 29 de mayo de 2025).

Tzanou, M. (2021). Schrems I and Schrems II: Assessing the Case for the Extraterritoriality of EU Fundamental Rights. En F. Fabbrini, E. Celeste & J. Quinn (eds.), *Data Protection Beyond Borders: Transatlantic Perspectives on Extraterritoriality and Sovereignty* (pp. 99-116). Oxford: Hart Publishing.

Velasco Pufleau, M. (2024). La Alianza Digital como Oportunidad para Profundizar las Relaciones entre la Unión Europea y América Latina y el Caribe. En *Inteligencia Artificial y Diplomacia: Las Relaciones Internacionales en la Era de las Tecnologías Disruptivas* (pp. 89-100). Caracas: Sistema Económico Latinoamericano y del Caribe (SELA). <https://sela.org/wp-content/uploads/2024/10/ia-diplomacia-ia-las-relaciones-internacionales-en-la-era-de-las-tecnologias-dirruptivas.pdf> (último acceso: 25 de junio de 2025).

Yuliyanova Chakarova, K. (2019). *General Data Protection Regulation: Challenges Posed by the Opening Clauses and Conflict of Laws Issues*. European Union Law Working Papers, 41. Stanford/Vienna: Transatlantic Technology Law Forum.

Inteligencia Artificial (IA) y Sistemas Judiciarios: el desafío de la Justicia Numérica

Naiara Posenato,
*Profesora en la Universidad de Milán e
Investigadora del European Institute of International Studies (Suecia)*

1. Introducción

Entre los múltiples impactos que la Inteligencia Artificial (IA) ha generado en el ámbito jurídico, uno de los ejes de discusión más relevantes se refiere a su utilización en los sistemas judiciales. Para algunos sectores, la IA se presenta como una panacea capaz de resolver los problemas estructurales que, en mayor o menor medida, afectan a los sistemas judiciales en todo el mundo: el exceso de litigiosidad y la lentitud en la toma de decisiones. En el extremo opuesto del espectro, se proyecta una visión distópica en la que un *juez robot* sustituiría por completo al magistrado humano.

El fenómeno del uso de la IA en los sistemas judiciales ha sido impulsado por factores coyunturales, como la aceleración del proceso de digitalización de la justicia durante la pandemia de COVID-19, y sobre todo por la evolución reciente y la difusión global —a partir de noviembre de 2022— de los llamados *Grandes Modelos de Lenguaje* (*Large Language Models*, LLMs). Desde entonces, herramientas inteligentes han comenzado a ser empleadas en sistemas judiciales de todo el mundo, aunque de forma muchas veces desordenada y heterogénea.

Con el objetivo de orientar a los operadores jurídicos en el uso adecuado de tecnologías basadas en IA, diversos órganos, estados e instituciones, tanto nacionales como internacionales, han adoptado marcos éticos y normativos. No obstante, dicho marco regulatorio

se presenta como fragmentado, mutable y complejo, dividido entre normas generales y regulaciones específicas, con diferentes grados de vinculatoriedad y en constante evolución.

Un primer esfuerzo de sistematización del tema requiere distinguir entre las aplicaciones de la IA en el sector judicial en general y su empleo específico en la función decisoria, dado que los riesgos asociados a cada uso difieren sustancialmente. Del mismo modo, el análisis del marco ético y normativo puede beneficiarse de la identificación de ciertas tendencias emergentes, ya sea en función de su evolución temporal, como también de su especialización respecto al uso de IA por parte de los sistemas judiciales.

2. Principales funciones de la inteligencia artificial para el sistema judicial y la función jurisdiccional

La IA puede potencialmente intervenir en diversas etapas del proceso judicial, abarcando un amplio espectro de aplicaciones.

En un extremo, se emplea para automatizar actividades operativas, como la tramitación procesal y la emisión de decisiones simples o interlocutorias (*replacement technology*). En el otro extremo, puede asumir funciones más complejas, involucrando el razonamiento jurídico propiamente dicho, evaluando hechos y pruebas para, potencialmente, elaborar sentencias (*disruptive technology*). En este sentido, se distingue entre la aplicación de la IA al sistema judicial en general y su utilización específica en la función decisoria.

La comparación del uso actual de la IA por parte de los sistemas judiciales de diversos ordenamientos jurídicos⁶⁸ demuestra que las *replacement technologies* han sido aplicadas en distintos niveles del proceso⁶⁹, desempeñando funciones clasificables en diferentes categorías.

Una primera función es de naturaleza *informativa*. En el ámbito de la promoción del acceso a la justicia, especialmente en el apoyo al usuario del sistema judicial, la IA puede proporcionar información sobre los servicios disponibles, en relación con la posición específica de un solicitante, a través de asistentes virtuales (*chatbots*). Por ejemplo, en 2022, Austria implementó una herramienta de esta naturaleza denominada JUSTITIA⁷⁰, destinada a propor-

68 Para obtener información actualizada sobre los sistemas de inteligencia artificial y otras herramientas avanzadas de ciberjusticia utilizadas por los poderes judiciales en Europa, puede consultarse el **Resource Centre on Cyberjustice and AI**, coordinado por la CEPEJ: <https://public.tableau.com/app/profile/cepej/viz/ResourceCentreCyberjusticeandAI/AITOOLSINITIATIVESREPORT?publish=yes>. Todos los enlaces han sido consultados por última vez el 2 de junio de 2025.

69 Según la UNESCO, los principales usos de la inteligencia artificial por parte de operadores judiciales se refieren a la búsqueda de información jurídica y técnica (43 %) y al apoyo en la redacción de documentos mediante chatbots (28 %). La encuesta, realizada entre septiembre y diciembre de 2023, incluyó a cerca de 36.000 participantes de más de 160 países. Cfr. Gutiérrez, J. D. (2024), **UNESCO Global's Judges Initiative: Survey on the use of AI Systems by Judicial Operators**, París, 8, disponible en <https://unesdoc.unesco.org/ark:/48223/pf0000389786>

70 Cfr. <https://justizonline.gv.at/jop/web/home>

cionar soporte al público en general en el sector judicial, equipada con un asistente virtual que informa a los usuarios sobre los servicios disponibles en el sistema judicial austríaco. En la misma línea, se destaca el uso de la IA para generar información jurídica en un lenguaje sencillo y más accesible al justiciable, como a *Guia Prático da Justiça*, desarrollada en 2023, en Portugal buscando agilizar la interacción de los usuarios con la justicia en el tema del matrimonio y el divorcio.⁷¹

La función desempeñada por la IA también puede ser *inclusiva*, cuando su aplicación en los órganos de justicia se refiere, por ejemplo, al procesamiento de lenguaje natural (video, audio e imágenes). Potencialmente, es capaz de ofrecer soporte a personas con discapacidades, así como efectuar reconocimiento de voz y a la transcripción de procesos judiciales o su traducción⁷². Para esta última aplicación, se desarrolló en India el sistema SUVAS (*Vidhik Anuvaad Software*)⁷³, que traduce decisiones y órdenes judiciales a nueve idiomas locales diferentes, presentado por la Corte Suprema en noviembre de 2019.

Una función de naturaleza *conectiva* también ha sido desempeñada por instrumentos de presentación electrónica de documentos (*e-filing*) que, con la intervención de la IA, pueden asignar los casos a los jueces de forma automatizada, siguiendo criterios predefinidos, con el fin de garantizar mayor imparcialidad y equilibrio.⁷⁴

Avanzando hacia un nivel superior de complejidad, la IA también puede ser utilizada en la resolución de disputas en línea: es el caso de los denominados sistemas de resolución de disputas en línea (*Online Dispute Resolution - ODR*), que emplean tecnologías para la resolución de controversias entre las partes⁷⁵. Algunos ordenamientos nacionales ya operan en estos entornos totalmente digitalizados en plataformas especiales de ODR y con la utilización de IA para determinadas tareas, como Canadá, con el *Civil Resolution Tribunal*⁷⁶; el Reino Unido, a través del *Money Claim Online*⁷⁷; o los Países Bajos, con el *Rechtwijzer* en el ámbito de las disputas en materia de consumo.⁷⁸

71 Cfr. <https://justica.gov.pt/Servicos/Guia-pratico-da-Justica>

72 En junio de 2023, el juez argentino José Osvaldo Ledesma utilizó ChatGPT para simplificar el lenguaje de una sentencia de derecho de familia, haciéndola más comprensible. La sentencia incluía, en su parte final, un resumen de fácil lectura elaborado por la IA, como refiere Gutiérrez, J. D. (2024), AI technologies in the judiciary: Critical appraisal of Large Language Models in judicial decision making. **Handbook on Public Policy and Artificial Intelligence**, Edward Elgar Publishing, 323 ss.

73 Cfr. <https://www.drishtijudiciary.com/current-affairs/supreme-court-vidhik-anuvaad-software-suvas>

74 Práctica aplicada en diversos ordenamientos de Europa del Este, cfr. Škop M., Merényi, M., Turashvili, K. Izdebski, Ker-
eš, D., Ilić, V. (2021), **ALGOVrithms 2.0 – State of Play: Usage on Algorithms by the Governments in Czechia, Georgia, Hungary, Poland, Serbia and Slovakia**, ePaństwo Foundation.

75 A propósito, veja-se Uncitral (2023), **Technology and the Future of Online Dispute Resolution (ODR) Platforms for Consumer Protection Agencies**, <https://unctad.org/system/files/official-document/tcsditcinf2023d5en.pdf>

76 Cfr. <https://civilresolutionbc.ca/>

77 Cfr. <https://www.gov.uk/make-court-claim-for-money/make-claim>

78 Cfr. <https://rechtwijzer.nl/>

La protección de los datos personales, mediante actividades de anonimización, también se ha llevado a cabo a través del empleo de IA en diversos sistemas judiciales europeos, como Alemania, Italia y Suiza, entre otros⁷⁹. En este sentido, la IA desarrolla una *función* que es simultáneamente *protectora* en lo que respecta a la difusión de datos personales e información sensible, y *promotora de la transparencia del poder judicial* y de la comprensión del derecho, aspectos que, a su vez, son prodrómicos a la seguridad jurídica.

Gran parte del uso de la IA en los sistemas judiciales se refiere a la asistencia en el desempeño de las actividades procesales en sentido amplio, mediante el ejercicio de una *función administrativa*. Los sistemas de IA pueden automatizar la clasificación de casos, categorizar documentos e incluso analizar preliminarmente las peticiones, acelerando el flujo de trabajo judicial. En el ámbito de la investigación, facilitan el procesamiento, la clasificación y el análisis de datos para fines como búsquedas en documentos jurídicos, en fuentes legislativas, doctrinales y jurisprudenciales, por ejemplo, a través de búsquedas de naturaleza semántica (no solo literales), extrayendo del contexto nuevas palabras clave. La IA también puede sintetizar documentos y generar resúmenes automáticos, extraer referencias e identificar metadatos, información esencial y mejorar la capacidad de investigación jurídica.⁸⁰

En definitiva, distinta de los supuestos anteriormente mencionados es la hipótesis de las denominadas *disruptive technologies*, en las que la inteligencia artificial asume —aunque de forma parcial— funciones que tradicionalmente corresponden a la actividad decisoria del juez. Este caso se distingue de los anteriores por referirse al uso directo de la inteligencia artificial por parte del juez, siendo, naturalmente, la función que suscita mayores interrogantes éticos y jurídicos en torno a su propia legitimidad y aplicación.

En este ámbito se inserta la llamada *jurimetría o justicia predictiva*, mediante la cual la IA es capaz de identificar patrones que posteriormente se utilizan para visualizar, simular o incluso prever los posibles resultados de litigios futuros. Su aplicación en el ejercicio de la función jurisdiccional permite desarrollar modelos de decisión capaces de reproducir el abanico de decisiones judiciales adoptadas en relación con una determinada cuestión.

Este tipo de aplicación de la IA es muy polémico y involucra posturas diferentes de los sistemas judiciales: Desde hace tiempo, el sistema judicial estadounidense emplea programas

79 Véase, para un análisis de los instrumentos utilizados para la anonimización, TERZIDOU, K. (2023) Automated Anonymization of Court Decisions: Facilitating the Publication of Court Decisions through Algorithmic Systems. **Proceeding of the Nineteenth International Conference on Artificial Intelligence and Law**, New York, ICAIL, 297 ss.

80 Tareas que han sido realizadas con el apoyo de la inteligencia artificial, por ejemplo, en el Tribunal de Justicia de la Unión Europea, cfr. CJEU (2024), **Artificial Intelligence Strategy**, 10 ss., disponible em <https://curia.europa.eu/jcms/upload/docs/application/pdf/2023-11/cjeuaistrategy.pdf>

conocidos como *risk assessment tools* para evaluar la peligrosidad o el riesgo de reincidencia de personas condenadas y recientemente, ha desarrollado varios algoritmos específicamente diseñados para estimar el riesgo de reincidencia, entre ellos el sistema denominado PATTERN (*Prisoner Assessment Tool Targeting Estimated Risk and Needs*).⁸¹

En sentido opuesto, con la reforma del *Code de l'organisation judiciaire* de 2019, el legislador francés impuso severas restricciones a tales prácticas, estableciendo la prohibición de reutilizar los datos relativos a magistrados contenidos en decisiones judiciales con la finalidad de evaluar, analizar, comparar o predecir sus actuaciones profesionales, sean efectivas o presuntas, bajo pena de sanción pecuniaria⁸². En la misma dirección se sitúa la decisión del Tribunal Constitucional Federal de Alemania, de 2023, que declaró la inconstitucionalidad de disposiciones contenidas en las leyes de seguridad pública de los Länder de Hesse y Hamburgo, las cuales autorizaban el análisis automatizado de datos personales con fines de prevención del delito⁸³.

Un estudio publicado en 2023 señala que en América Latina se han registrado varios casos en los que jueces afirmaron haber utilizado inteligencia artificial como apoyo —y, en ciertos aspectos, en sustitución parcial— para la elaboración de decisiones judiciales. Entre ellos destaca el caso del juez colombiano Juan Manuel Padilla García⁸⁴, quien en enero de 2023 declaró haber consultado Open ChatGPT para fundamentar una sentencia sobre el derecho a la salud de un niño con autismo. La sentencia fue impugnada ante la Corte Constitucional, que en agosto de 2024 dictó una decisión extensa⁸⁵, concluyendo que no se vulneró el debido proceso, ya que el uso de IA tuvo un carácter meramente auxiliar (*ad adiuvandum*) y no sustituyó la función decisoria del juez.

3. Riesgos de la inteligencia artificial para el sistema judicial y la función decisoria

Si bien es cierto que la aplicación de la IA al sistema judicial puede representar un avance significativo para el funcionamiento global de la justicia, también conlleva riesgos considerables, especialmente cuando se examinan algunas de sus funciones más sensibles y controvertidas.

⁸¹ Cfr. <https://www.bop.gov/inmates/fsa/pattern.jsp>.

⁸² Cfr. **Loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice**, art. 33, disponible en <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000038261631>.

⁸³ BVerfG, **Urteil des Ersten Senats vom 16. Februar 2023**, online: <https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/2023/02/rs202302161bvr154719.html>. Traducción al inglés disponible en <https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/2023/02/rs202302161bvr154719en.html>.

⁸⁴ Rama Judicial de Colombia Juzgado 1º laboral del Circuito Cartagena, Radicado No. 13001410500420220045901, **Sentencia No. 032, de 30-01-2023** (disponible en <https://forogpp.com/wp-content/uploads/2023/01/sentencia-tutela-segunda-instancia-rad.-13001410500420220045901.pdf>)

⁸⁵ Corte Constitucional – República de Colombia, **Sentencia T-323/24**, de 2 de agosto de 2024, disponible en <https://www.corteconstitucional.gov.co/relatoria/2024/T-323-24.htm>.

En lo que respecta a la *función informativa*, el uso de inteligencia artificial incorpora por su propia naturaleza el riesgo de difusión de datos personales sensibles o de secretos comerciales. Este riesgo, presente en toda aplicación de sistemas inteligentes, adquiere una especial relevancia en el ámbito judicial, debido a la naturaleza particularmente delicada de los datos tratados. La potencialidad de daño se acentúa en los casos que involucran a personas en situación de vulnerabilidad, como menores de edad o víctimas de delitos sexuales.⁸⁶

Más allá de la evidente simplificación para el justiciable que representa la centralización de la información en bases de datos, persiste la dificultad de que los datos disponibles en línea no siempre son “completos”, debido a que la digitalización del acervo histórico judicial constituye un proceso largo, costoso y aún inconcluso en muchos sistemas judiciales nacionales. Incluso en ordenamientos como el brasileño —que dispone de uno de los sistemas de *data lake* procesal más avanzados del mundo (la Plataforma Codex)⁸⁷—, apenas una cuarta parte de los tribunales nacionales ha integrado efectivamente sus propios datos procesales a dicha plataforma⁸⁸. A ello se suma, en términos de *inclusividad*, el obstáculo que representa el desconocimiento tecnológico de ciertos sectores de la población, que limita el acceso efectivo a la información jurídica en formato judicial.⁸⁹

La implementación de sistemas de IA en los tribunales con función de *administración* permite, según algunos, reducir de forma considerable los costes operativos de la justicia. En efecto, la optimización del tiempo de jueces y funcionarios constituye una ventaja evidente del uso de la inteligencia artificial al permitirles concentrarse en tareas de mayor complejidad. Sin embargo, ello no implica necesariamente que el ahorro económico; más bien, podría sostenerse lo contrario: para que dicha gestión interna sea efectiva, resulta imprescindible invertir en una estructura organizativa adecuada, personal especializado y programas sistemáticos de alfabetización digital y capacitación técnica para todos los actores involucrados. El diseño y la gestión interna de las herramientas de IA directamente por parte del propio Poder Judicial, así como el uso de sistemas alojados internamente (*hosted systems*) se perfilan como una solución adecuada, tanto para mitigar los riesgos asociados al uso de tecnologías no calibradas para el contexto jurídico y cultural específico, como para prevenir la opacidad funcional derivada de los denominados *black boxes*, protegidos por derechos comerciales o industriales.⁹⁰

86 Cabe señalar, al respecto, Rast, C. (2023), Cybersecurity Threats to the Judiciary. *Judges Journal*, 62:3.

87 Cfr. <https://www.cnj.jus.br/sistemas/plataforma-codex/>.

88 Para más detalles consúltese Posenato, N. (2025), *Tecnologia e Justiça: sobre a aplicação da Inteligência Artificial (IA) no Sistema Judiciário*. Ludovico, G., Nahas, T., **Diritti fondamentali, Lavoro e Nuove Tecnologie**, Milano, Milan University Press, 72-76.

89 Para profundizar el tema se vea v. Koenig, S. S. (2024), *Dispositifs audiovisuels, information juridique et judiciaire, règlement en ligne des différends. Impact du Numérique sur la Justice*, 1. Institut des études et de la recherche sur le droit et la justice, Paris.

90 En este sentido vease CEPEJ (2025), **Reflections of the AIAB on the use of artificial intelligence in judicial systems**,

En segundo lugar, en el ámbito del proceso judicial, muchas de las actividades llamadas “administrativas” pueden impactar concretamente en la definición de los derechos. La clasificación de casos, por ejemplo, implica interpretación jurídica, y cuando es realizada por sistemas de IA, se basa en patrones predefinidos y consolidados, lo que limita su capacidad para incorporar nuevos criterios. Este riesgo se agrava cuando la clasificación condiciona, por ejemplo, el acceso a la justicia, como en hipótesis de admisibilidad de recursos.⁹¹

En lo que respecta específicamente a la *función decisoria*, argumentáse que el uso de la inteligencia artificial puede contribuir a la previsibilidad de las decisiones judiciales. Al identificar y aplicar a los casos futuros patrones coherentes extraídos de grandes volúmenes jurisprudenciales, la IA favorece una mayor estabilidad y uniformidad en la aplicación del Derecho. Esto permite que las partes tengan una comprensión más clara de los posibles desenlaces de sus litigios, y que se apliquen criterios homogéneos a casos semejantes —lo que resulta particularmente relevante en tribunales con alta carga procesal y múltiples salas que pueden interpretar la norma de forma divergente.

No obstante, pueden formularse diversas observaciones al respecto. En primer lugar, la certeza no está siempre garantizada, ya que es bien sabido que la mayoría de los sistemas de IA generativa incorporan un cierto grado de aleatoriedad, lo que les permite generar respuestas distintas ante una misma pregunta (*prompt*), dependiendo del momento en que se formule o de mínimas variaciones en su redacción. Eso sin mencionar el riesgo de alucinaciones, un problema específico de los modelos de lenguaje de gran escala (LLMs): se generan textos aparentemente coherentes y convincentes que, sin embargo, contienen información falsa o inexistente, incluso cuando el sistema ha sido alimentado exclusivamente con datos correctos.⁹²

Este riesgo está aún agravado por la opacidad respecto a cómo está programado el algoritmo y cómo se vinculan los datos subyacentes, que puede traducirse en la imposibilidad de

Doc. CEPEJ-AIAB (2025)1Rev5, disponible en <https://rm.coe.int/cepej-aiab-2025-1rev5-en-reflections-of-the-aiab-on-the-use-of-artific/1680b42a56n>.

91 Un ejemplo ilustrativo es el sistema brasileño, donde la evaluación de la “repercusión general”, criterio para admitir recursos extraordinarios (en materia constitucional) podría quedar restringida a precedentes ya establecidos, impidiendo la entrada de nuevas tesis constitucionales aún no consolidadas jurisprudencialmente, con perjuicios tanto para el justiciable como para el desarrollo del derecho. Ese riesgo se evita porque, según declaración del Tribunal Supremo Federal (STF), el indicativo siempre es validado o confirmado durante la efectiva apreciación del caso concreto por parte de los ministros. Véase <https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=471331&ori#:~:text=O%20Victor%20C3%A9%20uma%20intelig%C3%Aancia,aplica%20ao%20caso%20dos%20autos>.

92 Un caso muy conocido tuvo lugar ante el Tribunal de Distrito de los Estados Unidos para el Distrito Sur de Nueva York, *Mata v. Avianca, Inc.*, y culminó con la sentencia del 22 de junio de 2023, **Mata v. Avianca, 22-cv-1461(PKC)**, accesible en <https://casetext.com/case/mata-v-avianca-inc-2>. Menos conocido, pero análogo fue el caso **Felicity Harber v. The Commissioners for His Majesty's Revenue and Customs** [2023] UKFTT 1007 (TC), disponible en <https://www.bailii.org/uk/cases/UKFTT/TC/2023/TC09010.html>.

verificar, si una interpretación normativa proviene efectivamente de una fuente doctrinal o jurisprudencial autorizada, comprometiendo así la legitimidad de la argumentación jurídica. En segundo lugar, la lógica de la calculabilidad puede conducir a que la IA, en el actual grado de evolución, elabore decisiones judiciales siguiendo un razonamiento meramente reproductivo, con el efecto negativo de cristalizar el Derecho e impedir su evolución. Ya no se concibe el razonamiento jurídico como un mero silogismo, un cálculo objetivo y previsible: las decisiones judiciales se ven cada vez más impulsadas a adoptar enfoques interdisciplinarios, consecuencialistas y contextuales, que permitan incorporar las transformaciones sociales, societales, económicas, medioambientales e incluso tecnológicas que afectan el entorno normativo contemporáneo.⁹³

4. El uso de la inteligencia artificial por el sistema judicial y por los jueces: principales marcos regulatorios

Los primeros documentos de alcance verdaderamente universal y con vocación generalista para la regulación de la inteligencia artificial fueron la Recomendación de la OCDE sobre la Inteligencia Artificial⁹⁴, de 2019, que estableció principios fundamentales como la transparencia, la rendición de cuentas, la solidez técnica y la inclusión; y la Recomendación sobre la Ética de la Inteligencia Artificial de UNESCO (2021)⁹⁵, que introdujo una perspectiva humanista y basada en derechos, subrayando la necesidad de evitar usos discriminatorios o desproporcionados de estas tecnologías.

Estas iniciativas revelaban un enfoque ético-principialista y una aproximación general a la inteligencia artificial, sin una diferenciación por sectores. Una excepción significativa fue el Consejo de Europa que, en 2018, a través de la Comisión Europea para la Eficiencia de la Justicia (CEPEJ), publicó la *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, el primer instrumento sectorial específicamente orientado a promover el uso de la IA en el ámbito judicial.⁹⁶ Desde una perspectiva ética, esta carta establece su adhesión a cinco principios fundamentales: respeto de los derechos fundamentales, no discriminación, calidad y seguridad, transparencia e imparcialidad y supervisión humana.

93 En el mismo sentido, Cour de Cassation (2025), **RAPPORT. Préparer la Cour de cassation de demain. Cour de cassation et intelligence artificielle**, 24, accesible en https://www.courdecassation.fr/files/files/Publications/IA%20-%20Rapport%202025/Rapport_IA_2025_Web.pdf.

94 OECD, *Recommendation of the Council on Artificial Intelligence*, adoptada en 22/05/2019, disponible en <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>.

95 UNESCO, Recomendación sobre la ética de la Inteligencia Artificial, adoptada en 23 de noviembre de 2022, disponible en https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa.

96 Consultable en <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>

Un cambio de paradigma en la reglamentación de la IA comenzó a configurarse con la actuación normativa de la Unión Europea, especialmente a partir de la adopción del Reglamento (UE) 2016/679, conocido como Reglamento General de Protección de Datos (RGPD)⁹⁷, y, de forma más reciente, con el Reglamento (UE) 1689/2024, denominado *AI Act*, en vigor desde el 1º de agosto 2024.⁹⁸

El RGPD, si bien no está específicamente enfocado en la IA, incluye normas relevantes para su utilización en el ámbito judicial, como el artículo 22, que regula las decisiones automatizadas y establece que “el interesado tendrá derecho a no ser objeto de una decisión basada *únicamente* en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos sobre él o le afecte significativamente de modo similar”.

Una innovación sustancial en esta materia ha sido introducida por el *AI Act*⁹⁹. Este nuevo marco normativo, complejo y de alcance general, regula el uso de tecnologías de inteligencia artificial con fines no-militares en el ámbito europeo, siendo aplicable tanto a operadores establecidos en la UE como a proveedores, públicos o privados, situados fuera del territorio europeo, siempre que el usuario del servicio se encuentre en la UE o en el Espacio Económico Europeo. Por ello, se ha sostenido que este instrumento se inscribe en un “nuevo constitucionalismo europeo” en materia digital, orientado hacia la adopción de normas vinculantes que conforman una nueva noción de orden público, basada en la soberanía digital.¹⁰⁰

En lo que respecta al sector judicial, según el artículo 6(2) del *AI Act*, los sistemas de IA enumerados en el Anexo III se consideran de alto riesgo. En particular, el punto 8 de dicho anexo se refiere a los sistemas de IA utilizados en la administración de justicia y en procesos democráticos, incluyendo aquellos destinados a ser empleados por autoridades judiciales o en su nombre para asistir en la investigación e interpretación de hechos y normas jurídicas, así como en su aplicación a casos concretos, o en métodos alternativos de resolución de conflictos. Sin embargo, el propio artículo 6(3) prevé una excepción significativa: un sistema de IA mencionado en el Anexo III no se considerará de alto riesgo si no representa un riesgo

97 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 (Reglamento general de protección de datos), consultable en <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32016R0679>

98 Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo de 13 de junio de 2024 por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican los Reglamentos (CE) n.º 300/2008, (UE) n.º 167/2013, (UE) n.º 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial), consultable en <https://eur-lex.europa.eu/legal-content/PT/ALL/?uri=CELEX:32024R1689>

99 En la misma línea, la UE adoptó los llamados **Digital Markets Act** e o **Digital Services Act** (2022), así como el **Data Governance Act** (2023).

100 Estos conceptos se desarrollan en profundidad en el volumen Santosuosso, A., Sartor, G. (2024), **Decidere con l'IA. Intelligenze artificiali e naturali nel diritto**, Bologna, Il Mulino.

significativo para la salud, la seguridad o los derechos fundamentales, en particular si no influye *de forma sustancial* en el resultado de la decisión.

Una evolución normativa similar puede observarse en el propio Consejo de Europa, que en 2024 adoptó la Convención Marco sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de Derecho.¹⁰¹ Se trata de un instrumento jurídicamente vinculante —y no meramente de un conjunto de orientaciones éticas— destinado a regular el impacto de la inteligencia artificial en el conjunto de las instituciones democráticas, incluidos los sistemas judiciales; sin embargo, aún no ha entrado en vigor.

La evolución de las regulaciones nacionales revela una tendencia creciente hacia la especialización en la disciplina del uso de la inteligencia artificial en el ámbito judicial, aunque mediante instrumentos de naturaleza diversa —normativos, administrativos o jurisprudenciales— que, si bien no siempre adoptan la forma de leyes específicas, buscan establecer controles cuando estas tecnologías inciden en el acceso a la justicia o en la toma de decisiones judiciales.

En el contexto europeo, además de las experiencias de Alemania y de Francia, España ha avanzado hacia un modelo integral de justicia digital mediante el Real Decreto-Ley 6/2023¹⁰², que contempla el uso de herramientas tecnológicas, incluida la inteligencia artificial, dentro de un marco de garantías procesales y control institucional. El RDL 6/2023 incorpora expresamente el uso de la inteligencia artificial en el ámbito judicial, regulando en sus artículos 56 a 58 las actuaciones automatizadas y asistidas, entre las cuales destaca el artículo 57, que prevé la generación de borradores por sistemas algorítmicos como apoyo a las resoluciones judiciales, siempre bajo control humano y sin menoscabo de la independencia judicial.

En América Latina, Colombia presenta un modelo mixto que combina jurisprudencia y regulación administrativa. En seguida a la sentencia T-323 de la Corte Constitucional, el Consejo Superior de la Judicatura ha emitido el Acuerdo PCSJA24-12243, que establece criterios orientadores para el uso responsable de la IA en el sistema judicial colombiano¹⁰³. Esta normativa introduce una innovación relevante al clasificar los usos de la inteligencia artificial en la función judicial (art. 4), distinguiendo entre: 1) usos en la gestión administrativa

¹⁰¹ Disponible en <https://rm.coe.int/1680afae3c>

¹⁰² Cfr. Real Decreto-ley 6/2023, de 19 de diciembre, por el que se aprueban medidas urgentes para la ejecución del Plan de Recuperación, Transformación y Resiliencia en materia de servicio público de justicia, función pública, régimen local y mecenazgo, consultable en <https://www.boe.es/buscar/act.php?id=BOE-A-2023-25758>

¹⁰³ ACUERDO PCSJA24-12243 16 de diciembre de 2024 “Por el cual se adoptan lineamientos para el uso y aprovechamiento respetuoso, responsable, seguro y ético de la inteligencia artificial en la Rama Judicial”, disponible en https://actosadministrativos.ramajudicial.gov.co/GetFile.ashx?url=%7E%2FApp_Data%2FUpload%2FPCSJA24-12243.pdf

o de soporte a la gestión judicial; 2) usos que requieren revisión detallada de los productos y resultados obtenidos; y 3) usos que exigen especial observancia de principios de transparencia, responsabilidad y privacidad. Aún más recientemente, la Corte Constitucional del país reconoció que el acceso a la información sobre los sistemas algorítmicos, como el código fuente de la aplicación CoronApp, forma parte del derecho fundamental de acceso a la información pública, reafirmando así el principio de transparencia algorítmica en el ámbito estatal.¹⁰⁴

En el caso brasileño, con el objetivo de orientar el desarrollo tecnológico en el ámbito judicial, el Consejo Nacional de Justicia (CNJ) adoptó en 2020 la Resolución n.º 332¹⁰⁵, un marco normativo pionero en América Latina que establece principios éticos y de gobernanza para el uso de la inteligencia artificial en el Poder Judicial. Posteriormente, la Resolución n.º 615/2025 del CNJ ha marcado un nuevo hito al actualizar los lineamientos anteriores y adaptarlos a los desafíos que plantean tecnologías emergentes como los modelos de lenguaje de gran escala (LLMs) y la inteligencia artificial generativa.¹⁰⁶ Así como en el caso de la reglamentación colombiana, entre sus principales innovaciones destaca la clasificación de los sistemas de IA en función del nivel de riesgo que representan para la actividad jurisdiccional.

Se diferencia así entre aquellos sistemas que desempeñan funciones auxiliares, como la organización de información y aquellos que pueden incidir directamente en el proceso decisional, como los utilizados para evaluar pruebas o predecir comportamientos procesales. Estos últimos están sujetos a exigencias más rigurosas en términos de auditoría, trazabilidad y control. La resolución también introduce prohibiciones expresas que reflejan una orientación garantista. Se excluye, por ejemplo, el uso de IA para la predicción de conductas criminales, el reconocimiento emocional, la evaluación de la personalidad de los individuos y la adopción de decisiones automatizadas sin revisión humana efectiva. Este nuevo marco normativo reafirma el compromiso del sistema judicial brasileño con una justicia digital ética, transparente y centrada en los derechos fundamentales.

5. Observaciones finales

Las múltiples formas de utilización de la inteligencia artificial en el ámbito judicial, así como la diversidad de funciones que se le atribuyen, ponen de manifiesto la notable utilidad que

104 CORTE CONSTITUCIONAL, Sala Novena de Revisión, **Sentencia t-067 de 2025**, disponible en <https://www.corteconstitucional.gov.co/relatoria/2025/t-067-25.htm>

105 CNJ, **Resolução nº 332, de 21 de agosto de 2020**. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências, disponible en <https://atos.cnj.jus.br/files/original191707202008255f4563b35f8e8.pdf>

106 CNJ, **Resolução nº 615, de 11 de março de 2025**. Estabelece diretrizes para o desenvolvimento, utilização e governança de soluções desenvolvidas com recursos de inteligência artificial no Poder Judiciário, disponible en <https://atos.cnj.jus.br/files/original1555302025031467d4517244566.pdf>

su aplicación puede ofrecer en este sector. No obstante su enorme potencial, ya se han identificado diversos riesgos y es previsible que surjan otros nuevos dada la rapidez y la magnitud de su desarrollo.

Es importante subrayar que no toda tecnología utilizada en el entorno judicial conlleva el mismo grado de riesgo. En general, los sistemas de IA diseñados para desempeñar funciones meramente auxiliares —que no inciden directamente en la administración efectiva de justicia en casos concretos— presentan un nivel de criticidad reducido. Sin embargo, existen hipótesis en las que la distinción entre funciones decisorias y de simple apoyo no resulta tan clara ni evidente.

El proceso de regulación del uso de la inteligencia artificial revela ciertas tendencias en su evolución, perceptibles a partir de la observación de los sistemas jurídicos que han optado por normar la materia. A nivel internacional, se constata una transición desde marcos basados en principios éticos generales hacia estructuras normativas con un grado creciente de vinculatividad, en algunos casos incluso con características cuasi constitucionales.

En lo que respecta a los ordenamientos jurídicos nacionales, las respuestas normativas al fenómeno de la inteligencia artificial —y particularmente a su aplicación en el ámbito judicial— son marcadamente heterogéneas. No obstante, puede identificarse un principio común: la exigencia de intervención humana en aquellas decisiones que inciden sobre derechos subjetivos, así como la necesidad de que las tareas ejecutadas por sistemas de IA permanezcan bajo supervisión y control humanos.

La Soberanía Digital como pilar estratégico para el desarrollo de América Latina y el Caribe en la era de la Gobernanza Tecnológica Global

Mtro. Gustavo Herrera

Coordinador de Desarrollo Social - SELA

Mtro. Luis Emilio Martínez

Analista de Desarrollo Social - SELA

La humanidad se encuentra en un punto de no retorno histórico, impulsado por un nuevo paradigma vertiginoso como lo es el avance de la Inteligencia Artificial (IA). Esta revolución tecnológica, lejos de ser simplemente un asunto técnico, constituye el nuevo epicentro de la competencia geoeconómica y el campo donde se definirá el futuro del desarrollo de los países. Para la región de América Latina y el Caribe (ALC), este escenario presenta un desafío de enormes proporciones, pero al mismo tiempo representa una oportunidad sin precedentes. La contienda global por establecer las normas que regirán la IA y el poder acumulado por las grandes empresas tecnológicas (Big Tech Companies) no puede ser un debate que la región observe desde la distancia.¹⁰⁷

La evolución de la IA conlleva a avances y retos significativos en la sociedad, en las investigaciones, así como en los procesos históricos, incluyendo las transformaciones en las estructuras sociales, económicas y tecnológicas, en contraste a los avances científicos que van emergiendo, lo que paradójicamente presenta nuevos desafíos a los marcos regulatorios existentes en todas las esferas.

107 McKinsey Global Institute. *The economic potential of generative AI: The next productivity frontier* (2023).<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier>. Informe global que analiza el impacto económico y geopolítico de la IA generativa.

Su progreso representa grandes oportunidades en la gobernanza tecnológica global. Nos encontramos en un umbral tecnológico donde la IA despliega su potencial para transformar industrias, impulsar la innovación y redefinir la propia naturaleza de la inteligencia misma.

Ignorar esta realidad equivale a aceptar un rol inequívoco de meros consumidores de tecnología, de la mano con la adaptación a regulaciones extranjeras, con el riesgo inminente como región de profundizar en las asimetrías y dependencias existentes, ampliando las brechas de desarrollo.¹⁰⁸ Por el contrario, es imperativo articular una respuesta estratégica, proactiva y unificada.

La Soberanía Digital como eje para el desarrollo de ALC es fundamental. Se requiere que su aplicación se realice de manera responsable y beneficiosa para todos. Resulta imprescindible, establecer principios sólidos como la transparencia, equidad, privacidad, responsabilidad, seguridad y respeto por la autonomía humana.

En ese sentido, por medio del presente artículo queremos sostener que para ALC debe ser una respuesta fundamental la construcción de una Soberanía Digital como pilar estratégico para esta nueva revolución tecnológica e industrial.¹⁰⁹ La gobernanza ética de la inteligencia artificial no significa frenar el progreso, sino más bien dirigirlo hacia un futuro más justo y humano.

No se trata de un fin en sí mismo ni de una aspiración al aislacionismo tecnológico, sino de un medio indispensable para ejercer la autodeterminación, gobernar nuestros propios recursos digitales y asegurar que los beneficios de la IA se traduzcan en un desarrollo sostenible, inclusivo y equitativo. Nuestra perspectiva es clara: la soberanía digital es la nueva frontera del desarrollo, y la integración regional es el único vehículo para ampliar nuestra posición al respecto.

El desafío geoeconómico y de desarrollo para América Latina y el Caribe (ALC)

En la región latinoamericana y caribeña se vive en lo que podemos denominar una paradoja digital, por un lado, hemos demostrado una notable capacidad de adopción tecnológica, según los datos de la Unión Internacional de Telecomunicaciones (UIT, 2023) la penetración de internet móvil en varios países de la región supera el 70 %, reflejando una amplia conectividad. Además, el banco de desarrollo de América Latina y el Caribe (CAF, 2023)

108 UNCTAD. *Digital Economy Report 2024: The value of data*. (2024). <https://unctad.org/publication/digital-economy-report-2024> . Documento clave que describe el creciente extractivismo de datos y las desigualdades digitales entre países desarrollados y en desarrollo.

109 El *United Nations Advisory Body on Artificial Intelligence* (2024) señala que la IA constituye una transformación estructural comparable a revoluciones industriales previas, y destaca la urgencia de que regiones como ALC fortalezcan su soberanía digital. United Nations Advisory Body on Artificial Intelligence. (2024). *Governing AI for humanity: Final report*. Naciones Unidas. Disponible en: <https://digitallibrary.un.org/record/4062495>

reporta que el comercio electrónico ha crecido en la región a tasas anuales de dos dígitos durante los últimos dos (2) años, impulsado por la digitalización de empresas y consumidores.

Sin embargo, por otro lado esta modernidad oculta una vulnerabilidad estructural crítica: aunque América Latina y el Caribe es una prolífica generadora de datos y consumidora de servicios digitales, su participación en la propiedad, desarrollo y monetización del ecosistema tecnológico global es marginal (UNTACD, 2024)¹¹⁰, limita sus beneficios y aumenta la dependencia tecnológica.

Esta asimetría se manifiesta en una severa dependencia digital. La infraestructura sobre la que opera la economía digital de la región está abrumadoramente concentrada en un número reducido de actores y geografías. A modo de ejemplo, es preocupante que gran parte de los datos generados en América Latina y el Caribe (ALC) son procesados y monetizados por plataformas digitales extrarregionales, principalmente en Estados Unidos, dando como resultado limitante la capacidad de desarrollar infraestructura digital propia y controlar los flujos de datos.¹¹¹

Esta situación no solo profundiza la dependencia tecnológica de ALC, sino que también plantea desafíos estratégicos para construir un desarrollo digital autónomo. Incluso los centros de datos que operan en la región -como es el caso de Brasil, Chile, Colombia y México- son mayoritariamente propiedad de multinacionales tecnológicas extrarregionales, lo que refuerza la vulnerabilidad estructural en un sector clave para la soberanía digital y la competitividad futura.

Esta concentración se refleja incluso en nuestros mercados regionales. A nivel global, la tres grandes plataformas- Amazon Web Services, Microsoft Azure y Google Cloud- dominaban cerca del 68 % del mercado mundial de servicios de infraestructura en la nube hacia finales de 2024.¹¹²

Esta dinámica se observa también en América Latina y el Caribe, donde el mercado regional muestra una elevada concentración en estos mismos actores, reproduciendo así patrones

110 UNCTAD, *Digital Economy Report 2024*, op. cit., nota 2

111 CAF. (2023). *Digitalización para un futuro mejor: Informe de Economía y Desarrollo (RED 2023)*, p. 69-70. CAF señala que “el grueso de los datos generados por los usuarios de la región es capturado y procesado por plataformas digitales que operan principalmente desde fuera de la región” y que “América Latina y el Caribe presenta una escasa capacidad regional para el desarrollo de infraestructura de datos propia”. Disponible en: <https://www.caf.com/es/especiales/red/red-2023/>.

112 Synergy Research Group. (2025, febrero 6). *Cloud Market Jumped to \$330 billion in 2024 – GenAI is Now Driving Half of the Growth*. Según datos de Synergy, AWS (30 %), Microsoft Azure (21 %) y Google Cloud (12 %) representaban conjuntamente el 68 % del mercado mundial de servicios de infraestructura en la nube pública. Recuperado de <https://www.srgresearch.com/articles/cloud-market-jumped-to-330-billion-in-2024-genai-is-now-driving-half-of-the-growth>

de dependencia tecnológica. Esta situación robustece una vulnerabilidad estratégica tanto en términos de infraestructura, precios, control de flujo de datos y soberanía tecnológica, exponiendo a ALC a marcos regulatorios foráneos, con riesgos asociados como el lock-in tecnológico.¹¹³

La UNTACD (2024)¹¹⁴ advierte que el creciente control de datos y capacidades digitales por parte de un reducido número de actores globales, combinado con la falta de alternativas competitivas, puede reforzar desigualdades y comprometer el desarrollo soberano de economías emergentes. Para ALC, esta situación no solo implica riesgos económicos y de seguridad, sino también una pérdida de margen de maniobra estratégico en la gobernanza digital.

Este fenómeno ha sido conceptualizado como *extractivismo de datos* o *colonialismo digital*.¹¹⁵ Al igual que en siglos pasados se extraían materias primas físicas, hoy el recurso más valioso -los datos- son capturados, procesados y transformados en valor (en forma de algoritmos de IA y servicios digitales) fuera de la región, para luego ser reimportado como producto terminado a un alto costo. Esta dinámica perpetúa asimetrías estructurales norte-sur en la economía digital, siendo no solo un desafío económico sino un riesgo para el desarrollo, ya que coacciona la capacidad de los Estados para diseñar políticas públicas basadas en evidencia propia, desde sistemas de salud predictivos, hasta una gestión urbana eficiente debido a la importancia del refinamiento de macrodatos que cada vez se hace más evidente en la administración pública.

En el mismo orden de ideas, profundiza la desigualdad: la brecha digital ya no se limita al acceso a internet, sino a la capacidad de participar activamente en la economía del conocimiento.¹¹⁶ Sin una estrategia deliberada, la IA podría automatizar empleos de ingresos medios y acentuar

113 El *lock-in tecnológico* refiere a la dependencia que generan las grandes plataformas digitales, restringiendo la capacidad de cambiar de proveedor sin costos elevados o pérdida de funcionalidades (Zuboff, 2019). Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

114 UNCTAD. (2024). *Digital Economy Report 2024: Data-led Development: Realizing the Transformative Potential of Data for the Global South*. Naciones Unidas. La UNCTAD señala que el dominio de plataformas globales en infraestructura digital esencial genera riesgos de *vendor lock-in*, limitando la autonomía tecnológica y la capacidad de los países en desarrollo para definir su propio camino digital. Disponible en: https://unctad.org/system/files/official-document/der2024_en.pdf.

115 Couldry, N., & Mejias, U. A. (2019). *The Costs of Connection: How Data is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press. Los autores desarrollan el concepto de *colonialismo de datos*, enfatizando cómo los flujos globales de datos reproducen patrones históricos de extracción y concentración de poder.

116 UNCTAD. (2024), op. cit. El reporte señala que las nuevas formas de desigualdad digital incluyen la capacidad de procesamiento y análisis de datos, más allá del simple acceso a conectividad.

la polarización social, un riesgo que ya para la región fue advertido por organizaciones como la Organización Internacional del Trabajo (OIT) y el Banco Mundial (2024).¹¹⁷

Un pilar estratégico para el desarrollo sostenible

Frente a este panorama de crecientes asimetrías en el ecosistema digital global, el concepto de soberanía digital emerge como una doctrina estratégica indispensable para ALC. Según el Foro Económico Mundial (2025) se trata de la capacidad de ejercer control sobre el propio destino digital: datos, hardware y software que se genera y utilizan.¹¹⁸

Este término va más allá de crear un *splinternet*¹¹⁹, promover marketing tecnológico o adoptar posturas proteccionistas. Implica la capacidad autónoma y colectiva de los países de la región para formar, controlar y gobernar su ecosistema digital, asegurando que este responda a sus metas de desarrollo sostenible, inclusión y equidad. La soberanía digital es en efecto, una condición habilitante clave, no es un lujo estratégico, para avanzar en la realización de los objetivos de Desarrollo Sostenible (ODS) y la Agenda 2030 en la actual economía del conocimiento.

Sus componentes estratégicos se articulan directamente con los pilares del Desarrollo Sostenible:

- **Infraestructura digital y soberana:** Mas allá de expandir la conectividad, es *sine qua non* avanzar hacia una infraestructura regional robusta. Esto implica aumentar la proporción de centros de datos localizados en la región, fortalecer los puntos de intercambio de tráfico de internet (IXP)¹²⁰ en orden de que la data regional no tenga que viajar a hubs

117 International Labour Organization & World Bank. (2024). *Buffer or Bottleneck? Employment Exposure to Generative AI and the Digital Divide in Latin America and the Caribbean*. Disponible en: <https://www.worldbank.org/en/topic/poverty/publication/generative-ai-and-jobs-in-lac>. El informe advierte que la adopción de IA generativa podría aumentar la polarización en los mercados laborales de ALC si no se adoptan estrategias de adaptación inclusivas.

118 World Economic Forum. (2025, 22 de enero). *What is digital sovereignty and how are countries approaching it?* Recuperado de <https://www.weforum.org/stories/2025/01/europe-digital-sovereignty/>

119 El “splinternet” se refiere a la fragmentación del ciberespacio global en redes nacionales o regionales cada vez más autónomas, impulsadas por regulaciones, censura, infraestructura y estándares tecnológicos diferenciados, lo que deteriora la interoperabilidad global y genera riesgos estratégicos, económicos y técnicos. Recuperado de <https://www.economist.com/the-economist-explains/2016/11/22/what-is-the-splinternet?utm>

120 El desarrollo de puntos de intercambio de tráfico de internet (IXP) es clave para mejorar la eficiencia, reducir costos y fortalecer la soberanía digital en América Latina. CAF (2014) destaca que una infraestructura regional robusta de IXP permite un intercambio más eficiente del tráfico local, disminuyendo la dependencia de rutas internacionales. Agudelo, M., Katz, R., Flores-Roux, E., Duarte Botero, M. C., Callorda, F., & Berry, T. (2014). *Expansión de infraestructura regional para la interconexión de tráfico de internet en América Latina*. CAF. Disponible en: <https://scioteca.caf.com/handle/123456789/522>

extrarregionales. Iniciativas como la RedClara¹²¹ y el proyecto de cable submarino BELLA¹²² son ejemplos claros y constituyen una base para cerrar las brechas territoriales y ofrecer servicios de calidad a las poblaciones rurales y remotas, siendo coherentes con el ODS 9 sobre industria, innovación e infraestructura.

- **Gobernanza de datos para el desarrollo:** Los datos producidos en ALC constituyen un activo estratégico esencial. Ejercer una gobernanza soberana sobre ellos no implica solo fomentar la creación de espacios comunes de datos (data spaces, data trusts) sino proteger los derechos digitales de los ciudadanos potenciando sectores de investigación e innovación claves para los Estados. Ejemplos incluyen el uso de datos de salud anonimizados para prever epidemias o de datos agrícolas para optimizar el uso del agua, entre otras ventajas alienadas con los ODS 2, 3, 6 y 16.
- **Talento y capacidades humanas:** Sin capital humano calificado no habrá soberanía digital posible. Se requiere una transformación educativa profunda que potencie las habilidades STEM (Ciencia, Tecnología, Ingeniería y Matemáticas), así como esfuerzos masivos de reconversión y actualización de la fuerza laboral. Impulsar el talento local en estas materias y otorgar buenos incentivos es clave para revertir la fuga de cerebros, crear empleos de alta demanda, fomentar la inserción laboral y la calidad de servicios con especial atención a la participación de las mujeres en el sector digital.

Regulación inteligente y fortalecimiento institucional: Ejercer soberanía digital requiere de instituciones fuertes, capaces de diseñar e implementar marcos regulatorios eficaces. Esto incluye agencias de protección de datos y ciberseguridad con capacidades reales, así como regulaciones que promuevan la innovación local a través de mecanismos como pueden ser los *sandboxes*¹²³. Una regulación inteligente debe permitir atraer inversiones bajo los términos de la región, proteger la competencia y garantizar la confianza digital.

121 RedCLARA es la red regional de investigación y educación de América Latina, que interconecta a las redes académicas nacionales y facilita la colaboración científica y tecnológica. Ha sido clave en la articulación de proyectos de infraestructura digital como BELLA. RedCLARA. (s.f.). *Quiénes somos*. RedCLARA. Disponible en: <https://www.redclara.net/es/>

122 El proyecto BELLA (Building the Europe Link to Latin America) estableció el primer cable submarino directo entre Europa y América Latina, reduciendo la latencia y fortaleciendo la autonomía digital de la región. Es una infraestructura estratégica para la ciencia, la educación y la economía digital. Comisión Europea. (2021). *BELLA - Building the Europe Link to Latin America*. Disponible en: https://international-partnerships.ec.europa.eu/policies/programming/projects/bella-building-europe-link-latin-america_en

123 Los *regulatory sandboxes* permiten probar marcos normativos para tecnologías emergentes como la IA en entornos controlados. La ONU propone una hoja de ruta para su implementación en países en desarrollo. Pachouri, A. (2023). *How regulatory sandboxes in AI can help improve the technology-policy-society interface*. Naciones Unidas. Disponible en: <https://sdgs.un.org/sites/default/files/2023-05/A12%20-%20Pachouri%20-%20How%20regulatory%20sandboxes%20in%20AI%20can%20help%20improve%20the%20technology-policy-society%20interface.pdf>

En suma, la soberanía digital no es un fin en sí mismo, sino un pilar estratégico del desarrollo sostenible para América Latina y el Caribe (ALC) en este nuevo paradigma de tecnologías disruptivas. La construcción de este pilar requiere acción colectiva, visión de largo plazo y una integración regional más profunda.

Tabla 1. Componentes estratégicos de la soberanía digital y su vinculación con los ODS

Componente estratégico	Ejemplos clave en ALC	ODS
Infraestructura digital soberana	Centros de datos regionales, IXPs, RedCLARA, cable BELLA	9 y 17
Gobernanza de datos para el desarrollo	Data trusts, marcos regionales de gobernanza de datos, protección de datos	3,2,6 y 17
Talento y capacidades humanas	Educación STEM, reconversión laboral, inclusión de mujeres en tecnología	ODS 8, 5
Regulación inteligente e instituciones	Agencias de protección de datos y ciberseguridad, sandboxes regulatorios	16 y 17

Fuente: Elaboración propia con base en CEPAL (2022), RedCLARA (2021), OECD (2023), Brookings Institution (2023), European Parliament (2021).

La Soberanía Digital como instrumento para fortalecer la Gobernanza regional

La Soberanía Digital además de incluir la regulación tecnológica y de datos, debe contener financiación en la innovación, ya que de dicha manera los datos pueden fluir de forma más transparente y libremente, a fin de garantizar la protección en los datos y ofrecer el potencial para generar confianza, transformar y permitir evolucionar las tecnologías digitales, a efecto de fortalecer la gobernanza tecnológica global.

De acuerdo con el Foro Económico Mundial (2025)¹²⁴, los países generalmente coinciden en la necesidad de fomentar las industrias tecnológicas locales, particularmente cuando existen posibles consecuencias significativas para la seguridad nacional. Sin embargo, existen diferentes enfoques en torno a la gobernanza de estas tecnologías y datos. Estos diversos enfoques de la soberanía digital han profundizado la competencia geopolítica entre EE. UU., China y la Unión Europea (UE).

En contraste, América Latina y el Caribe enfrenta distintos desafíos que limitan el avance hacia una soberanía digital plena, entre ellos la debilidad institucional, la falta de uniformidad en las economías de nuestra región, aunado a las diferencias sociales existentes. En ese sentido, se requiere una gobernanza ética robusta que permita disminuir las desigualdades

124 World Economic Forum. (2025, enero 10) *¿Qué es la soberanía digital y cómo la abordan los países?* <https://www.weforum.org/stories/2025/01/europe-digital-sovereignty/>

existentes, con el fin de fortalecer la privacidad y la confianza, permitiendo así que de manera responsable la tecnología beneficie el desarrollo de los países.

Si bien, en los últimos años, ALC ha logrado avances significativos en la digitalización, impulsada tanto por la creciente incorporación de tecnologías en el sector público como por la necesidad de responder a las expectativas de una ciudadanía cada vez más conectada, la región aún enfrenta importantes desafíos para consolidar una transformación digital que sea inclusiva, sostenible y centrada en las necesidades de las personas, lo que demanda esfuerzos continuos y estrategias más integradas para maximizar el potencial del gobierno digital (OCDE/CAF, 2024).¹²⁵

En el mismo orden de ideas, podríamos considerar una aspiración frágil la soberanía digital a escalas individualistas en los Estados de América Latina y el Caribe (ALC) en el mundo multipolar en el que vivimos. Esta solo podrá ser efectiva si se articula en torno a una estrategia regional, robusta, pragmática y orientada al desarrollo económico, social y ambiental.

Ante la creciente concentración del poder digital en manos de grandes plataformas digitales y big tech companies, la acción aislada resulta limitada para garantizar la autonomía digital. Por ello, la única vía realista para materializar esta soberanía es a través de una profunda integración regional. En este escenario, resulta imprescindible la coordinación entre organismos, instituciones y mecanismos internacionales para posicionar dicha agenda, conjuntando esfuerzos y articulando políticas comunes que impulsen la transformación digital con justicia social y cohesión territorial.

Al respecto, proponemos cinco pilares claves para estructurar esta estrategia regional:

- I. Consolidar un mercado único digital regional:** La principal fortaleza en nuestra región es su base demográfica de más de 650 millones de personas. Este potencial demográfico habilita condiciones favorables para un mercado digital integrado el cual permitiría escalar soluciones tecnológicas, aumentar la competitividad y generar economías de escalas. Para ello, se requiere armonizar normativas en ámbitos como protección de datos, ciberseguridad, firmas electrónicas, identidad digital y comercio electrónico, promoviendo marcos interoperables compatibles con estándares internacionales (por ejemplo, GDPR) pero adaptados a las realidades locales. Esta armonización permitiría que una empresa emergente de tecnología en Colombia pueda operar sin fricciones regulatorias en Argentina o México, impulsando la emergencia de campeones digitales regionales y fortaleciendo las cadenas de valor digitales intrarregionales.

125 OECD/CAF (2024), Revisión del Gobierno Digital en América Latina y el Caribe: Construyendo Servicios Públicos Inclusivos y Responsivos, OECD Publishing, Paris, <https://doi.org/10.1787/7a127615-es> OECD/CAF (2024), Revisión del Gobierno Digital en América Latina y el Caribe: Construyendo Servicios Públicos Inclusivos y Responsivos, OECD Publishing, Paris, <https://doi.org/10.1787/7a127615-es>

- II. Inversión en infraestructura digital con propósito social:** más allá de la conectividad básica, es necesario inversiones estratégicas en infraestructura digital soberana. Resultaría viable la creación de un fondo regional de infraestructura digital, con aportes de distintos bancos de desarrollo regionales, bancos públicos, fondos de pensiones e inversionistas privados de la región. Este fondo pudiera servir como base para financiar proyectos prioritarios de carácter regional como cables submarinos regionales, centros de datos de propiedad compartida, puntos de intercambio de tráfico (IXP) para reducir la dependencia de nodos extrarregionales, y especialmente, la expansión de conectividad de última milla a escuelas, hospitales y comunidades rurales. Estos proyectos deben responder a criterios de inclusión social y equidad territorial, priorizando el cierre de brechas digitales.
- III. Ecosistema regional de innovación tecnológica:** En aras de que la transformación digital sea un motor de desarrollo, la innovación debe orientarse a resolver los desafíos estructurales de ALC. Se deben establecer programas de financiamiento e incubación regional, articulados con universidades, centros de investigación y gobiernos locales que promuevan sectores como AgriTech (agricultura resiliente), healthtech (salud digital), EdTech (educación inclusiva), GovTech (gobierno abierto) y FinTech (Inclusión Financiera). Asimismo, las compras públicas deben ser el instrumento de política industrial para priorizar soluciones locales, regionales y extrarregionales fomentando el crecimiento de empresas tecnológicas latinoamericanas y caribeñas.
- IV. Agenda Regional para el Desarrollo de Talento Digital:** La transición digital requiere de un nuevo contrato educativo. Es imprescindible articular una agenda regional de capacidades digitales que armonice currículos en competencias STEM, promueva programas regionales de formación técnica, certifique habilidades digitales y facilite la reconversión laboral frente al avance de la automatización y la inteligencia artificial. Esta agenda debe tener un enfoque total de inclusión social.
- V. Posicionamiento Regional en la Gobernanza Global:** Una región cohesionada en su estrategia digital podrá ejercer mayor influencia en foros multilaterales clave como la ONU, la OCDE o los procesos de la cumbre del futuro. Desde una perspectiva de gobernanza digital, una voz unificada permitirá a ALC negociar en mejores condiciones las regulaciones sobre la IA, gobernanza de datos, flujos transfronterizos, ciberseguridad y derechos digitales, garantizando que dichas normas respondan a los Objetivos de Desarrollo Sostenible y no reproduzcan relaciones y dependencia tecnológica.

En este sentido, la construcción de una soberanía digital articulada regionalmente representa tanto una necesidad como una oportunidad histórica. Requiere avanzar en convergencias normativas, inversiones estratégicas y una agenda compartida que potencie nuestras

capacidades locales. No se trata únicamente de reducir vulnerabilidades externas, sino de construir plataformas propias para la innovación, la inclusión social y la resiliencia institucional. Esto implica pasar del diagnóstico a la implementación, del enfoque fragmentado a la acción coordinada, situando la transformación digital como una política de Estado de alcance regional.

La cuarta revolución industrial¹²⁶ y el auge de la inteligencia artificial han redefinido el tablero geopolítico y geoeconómico mundial. La gobernanza del ecosistema digital se ha convertido en un factor inamovible del desarrollo económico y social. Para América Latina y el Caribe (ALC), no se trata solo de adaptarse al nuevo entorno tecnológico, sino transformarlo en una palanca para superar brechas históricas de desigualdad, dependencia y fragmentación.

En definitiva, se debe asumir el compromiso de contribuir activamente a esta construcción colectiva, en su rol de ente articulador. La integración digital, la autonomía tecnológica y la gobernanza ética de la inteligencia artificial deben consolidarse como pilares fundamentales del nuevo contrato regional de desarrollo sostenible.

Frente al avance acelerado de las tecnologías disruptivas, el tiempo no es un recurso renovable; de hecho, es el recurso económico más escaso del mundo. Por ello, la ventana de oportunidad está abierta, pero no lo estará indefinidamente. El momento de actuar con decisión es ahora.

No se trata del futuro: se trata del presente. La integración, la cooperación, la articulación y la visión compartida no son solo deseables, sino urgentes para nuestra región latinoamericana y caribeña.

126 La Cuarta Revolución Industrial, impulsada por tecnologías como la inteligencia artificial, está reconfigurando el orden geopolítico y geoeconómico global. La OCDE advierte que esta transformación exige nuevos marcos de gobernanza para evitar desigualdades estructurales, mientras que Brookings destaca su impacto disruptivo en la economía mundial y su potencial para profundizar asimetrías entre regiones. OECD. (2023). *Artificial Intelligence and the Future of Skills, Volume 1: Capabilities and Assessments*. OECD Publishing. Disponible en: <https://www.oecd.org/publications/artificial-intelligence-and-the-future-of-skills-volume-1-5f282b90-en.htm> West, D. M. (2018). *The Future of Work: Robots, AI, and Automation*. Brookings Institution. Disponible en: <https://www.brookings.edu/books/the-future-of-work-robots-ai-and-automation/>

Bibliografía

- Agudelo, M., Katz, R., Flores-Roux, E., Duarte Botero, M. C., Callorda, F., & Berry, T. (2014). *Expansión de infraestructura regional para la interconexión de tráfico de internet en América Latina*. CAF. <https://scioteca.caf.com/handle/123456789/522>
- Brookings Institution. (West, D. M.). (2018). *The Future of Work: Robots, AI, and Automation*. <https://www.brookings.edu/books/the-future-of-work-robots-ai-and-automation/>
- CAF. (2018). *Gobernanza de la transformación digital: Propuesta para América Latina*. <https://scioteca.caf.com/handle/123456789/1483>
- CAF. (2023). *Digitalización para un futuro mejor: Informe de Economía y Desarrollo (RED 2023)*. <https://www.caf.com/es/especiales/red/red-2023/>
- Chatham House. (2019, junio). *Tackling the ‘splinternet’: How to preserve the global nature of the internet*. <https://www.chathamhouse.org/2019/06/tackle-splinternet>
- Couldry, N., & Mejias, U. A. (2019). *The Costs of Connection: How Data is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press.
- International Labour Organization & World Bank. (2024). *Buffer or Bottleneck? Employment Exposure to Generative AI and the Digital Divide in Latin America and the Caribbean*. <https://www.worldbank.org/en/topic/poverty/publication/generative-ai-and-jobs-in-lac>
- McKinsey & Company. (2023). *The economic potential of generative AI: The next productivity frontier*. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier>
- OECD. (2023). *Artificial Intelligence and the Future of Skills, Volume 1: Capabilities and Assessments*. OECD Publishing. <https://www.oecd.org/publications/artificial-intelligence-and-the-future-of-skills-volume-1-5f282b90-en.htm>
- OECD. (s.f.). *Development*. <https://www.oecd.org/en/topics/policy-areas/development.html>
- OECD. (s.f.). *Science, Technology and Innovation*. <https://www.oecd.org/en/topics/science-technology-and-innovation.html>
- OECD/CAF. (2024). *Revisión del Gobierno Digital en América Latina y el Caribe: Construyendo Servicios Públicos Inclusivos y Responsivos*. OECD Publishing. <https://doi.org/10.1787/7a127615-es>
- RedCLARA. (s.f.). *Quiénes somos*. <https://www.redclara.net/es/>

Synergy Research Group. (2025, febrero 6). *Cloud Market Jumped to \$330 billion in 2024 – GenAI is Now Driving Half of the Growth*. <https://www.srgresearch.com/articles/cloud-market-jumped-to-330-billion-in-2024-genai-is-now-driving-half-of-the-growth>

The Economist. (2016, noviembre 22). *What is the splinternet?* <https://www.economist.com/the-economist-explains/2016/11/22/what-is-the-splinternet>

UNCTAD. (2024). *Digital Economy Report 2024: Data-led Development – Realizing the Transformative Potential of Data for the Global South*. https://unctad.org/system/files/official-document/der2024_en.pdf

UNCTAD. (2024). *Digital Economy Report 2024: The Value of Data*. Naciones Unidas. <https://unctad.org/publication/digital-economy-report-2024>

UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. <https://unesdoc.unesco.org/ark:/48223/pf0000377071>

United Nations Advisory Body on Artificial Intelligence. (2024). *Governing AI for Humanity: Final Report*. Naciones Unidas. <https://digitallibrary.un.org/record/4062495>

World Economic Forum. (2025, enero 10). *¿Qué es la soberanía digital y cómo la abordan los países?* <https://www.weforum.org/stories/2025/01/europe-digital-sovereignty/>

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.

Pachouri, A. (2023). *How regulatory sandboxes in AI can help improve the technology-policy-society interface* (Policy Brief No. A12). United Nations Department of Economic and Social Affairs. <https://sdgs.un.org/sites/default/files/2023-05/A12%20-%20Pachouri%20-%20How%20regulatory%20sandboxes%20in%20AI%20can%20help%20improve%20the%20technology-policy-society%20interface.pdf>



SISTEMA ECONÓMICO
LATINOAMERICANO
Y DEL CARIBE

Más y mejor Integración

ISBN: 978-980-6458-40-6



www.sela.org



@selainforma