Jornadas de Investigación y Análisis

Inteligencia Artificial,
¿Moda pasajera o cambio permanente de paradigma?

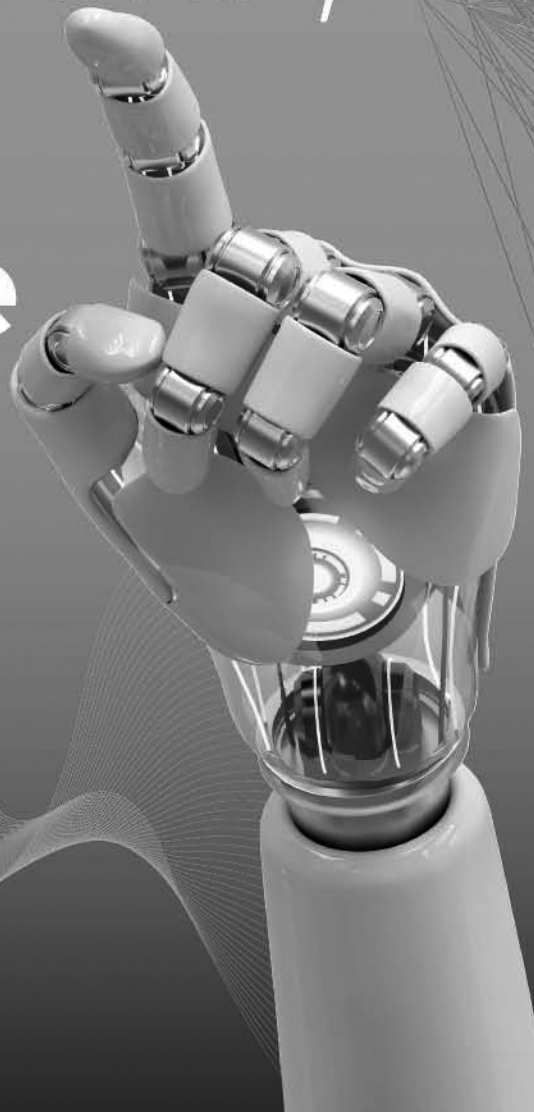
UNIVERSIDAD DE COSTA RICA

PROSIC

Programa Institucional
Sociedad de la Informacion
y el Conocimiento

Jornadas de Investigación y Análisis

Inteligencia Artificial,
¿Moda pasajera o cambio permanente de paradigma?



UNIVERSIDAD DE COSTA RICA

PROSIC

Programa Institucional
Sociedad de la Informacion
y el Conocimiento

Universidad de Costa Rica. Programa Sociedad de la Información y el Conocimiento. Memoria de las Jornadas de Investigación IA, ¿moda pasajera o cambio permanente del paradigma?/Valeria Castro Obando/editora. Programa Institucional Sociedad de la Información y el Conocimiento, Universidad de Costa Rica. - San José, C.R.: Prosic, Universidad de Costa Rica, 2025.

173 pp.

ISBN 978-9968-510-31-8

1. Conociendo lo que es la inteligencia artificial. 2. Potencialidades y casos de uso de la inteligencia artificial. 3. Riesgos y dilemas éticos de la inteligencia artificial. 4. Debates sobre la regulación de la inteligencia artificial. 5. Avances y oportunidades de la inteligencia artificial en Costa Rica. Universidad de Costa Rica. Prosic.

Memoria de las Jornadas de Investigación y Análisis IA,
¿moda pasajera o cambio permanente del paradigma?

Alejandro Amador Zamora
Coordinador

Valeria Castro Obando
Editora

Dariel Amador Pérez
Wilson González Gaitán
Asistentes de Investigación y apoyo editorial

Keilor Angulo Blanco
Diseño y Diagramación

Septiembre 2025

Índice

Presentación

Alejandro Amador Zamora

Los hitos que pueden dar el banderazo de salida a la revolución digital datan de finales de los años setenta, con el lanzamiento del primer microprocesador comercial (1971) y las computadoras personales accesibles al público (a finales de la década de los 70). Mucho ha avanzado la tecnología desde entonces: la proliferación de Internet, los teléfonos móviles, el auge de las redes sociales y más recientemente, la aceleración en la adopción de medios digitales durante la pandemia del Covid-19.

En los últimos años nos encontramos frente a una nueva ruptura del paradigma digital con la aparición de la Inteligencia Artificial (IA), particularmente con el auge de la IA Generativa. Para contextualizar temporalmente lo anterior, vale la pena recordar que ChatGPT, uno de los modelos de IA Generativa más conocidos en la actualidad, se lanzó oficialmente al público el 30 de noviembre de 2022 y en poco tiempo alcanzó millones de suscriptores, superando por mucho a plataformas como Facebook, Instagram y Tiktok.

La IA representa, sin lugar a duda, uno de los desarrollos tecnológicos más transformadores del siglo XXI, con implicaciones en todo nuestro entorno: ciencia, industria, educación, economía, arte y entretenimiento, entre otros. Promete cambios que ya se están dando a mayor o menor medida, como la automatización de procesos y la expansión de modelos generativos. La IA trae consigo promesas y oportunidades, pero también retos, tanto para el mundo como específicamente para nuestro país.

Este documento reúne un conjunto de ponencias presentadas en las *Jornadas de Investigación "IA, ¿moda pasajera o cambio permanente del paradigma?* Un evento llevado a cabo en el 2024 por el Programa Sociedad de la Información y el Conocimiento (Prosic) con el objetivo de fomentar la discusión sobre las repercusiones que la IA está generando en nuestras sociedades.

Las contribuciones aquí compiladas exploran la IA desde una serie de perspectivas multidisciplinarias que abarcan temas variados como el impacto, su origen histórico, la regulación y su uso responsable, por nombrar algunos. Como se han caracterizado las Jornadas de Prosic a lo largo de los años, esperamos que la publicación de estas memorias ayude a fomentar el diálogo interdisciplinario y la reflexión crítica sobre la temática recordando siempre que la tecnología nunca será un fin en sí mismo, sino una herramienta para mejorar la calidad de la condición humana.

Alejandro Amador Zamora

Coordinador

Programa Sociedad de la Información y el Conocimiento

1

Introducción

Valeria Castro Obando

Hoy la inteligencia artificial (IA) ha dejado ser una tecnología futurista para convertirse en un elemento más de nuestra cotidianeidad. Si bien existen múltiples formas de definirla, esta puede ser concebida como el conjunto de sistemas que simulan "la inteligencia humana en máquinas que están programadas para realizar tareas que normalmente requieren a una persona"[1]. Estas, además, de imitar el pensamiento racional, tienen la capacidad de analizar grandes volúmenes de datos a partir de los cuales pueden reconocer patrones, tomar decisiones y adaptar la información a nuevas circunstancias distintas a las iniciales. En consecuencia, el terreno de la IA es sumamente amplio y abarca distintas áreas tales como el aprendizaje automático, el procesamiento del lenguaje natural, la robótica y la visión artificial, entre muchos otros.

Aunque, la definición anterior pueda parecer muy ajena, todos los días interaccionamos con la IA. Por ejemplo, al utilizar asistentes de voz como Siri o Alexa, plataformas de streaming, redes sociales y cuando accedemos a todo tipo de servicios digitales generalmente se utilizan herramientas de IA, a pesar de que no se tenga consciencia de ello. Es más, en este contexto, de creciente digitalización constantemente "cedemos nuestros datos [y] ponemos en marcha un proceso de entrenamiento que influirá en las decisiones futuras sobre recomendaciones"[2], las aplicaciones y servicios alcanzado un grado de personalización y eficiencia nunca vistas.

Sin embargo, no fue hasta el lanzamiento de ChatGPT en noviembre del 2022, cuando resurgió el interés por la IA, particularmente por la *inteligencia artificial generativa* (IA generativa), una tecnología que se distingue de sus antecesoras por ser capaz de aprender temas complejos (como lenguajes de programación, química, arte o biología) y reutilizar los datos que se le suministra en su entrenamiento para resolver problemas y producir contenido nuevo como ideas, textos, imágenes, videos o música[3]. Estas potencialidades han llamado la atención de empresas e industrias, las cuales han visto en la IA una oportunidad para automatizar tareas repetitivas, incrementar la eficiencia, alcanzar más precisión y productividad y crear nuevos productos y servicios sustentados en esta tecnología[4].

A partir de esto, innovaciones como los asistentes virtuales, los softwares para reconocimiento de imagen y voz y los sistemas de recomendación se están utilizando con mayor frecuencia a lo interno de los sectores productivos. De hecho, se cree que la IA podrá revolucionar los servicios de salud, el transporte, la producción, la educación, la agricultura[5], la ban-

1 MIT Sloan Management Review. 9 de junio de 2023. ¿Qué es la Inteligencia Artificial y por qué todo mundo habla de ella? ¿Qué es la Inteligencia Artificial y por qué todo mundo habla de ella? (mitsloanreview.mx)

2 Delgado, M. (20 de diciembre de 2023). ¿Por qué es importante que sepas de inteligencia artificial? National Geogprahic https://www.nationalgeographic.com.es/ciencia/por-que-es-importante-que-sepas-inteligencia-artificial_21257

3 Amazon Web Services. (2023). ¿Qué es la IA generativa? AWS ¿Qué es la IA generativa? - Explicación de la inteligencia artificial generativa - AWS (amazon.com)

4 Ibíd MIT, 2023.

5 Organización de Naciones Unidas. (8 de noviembre de 2023). El poder de la inteligencia artificial y sus desafíos en el marco de las Naciones Unidas. Centro Regional de Información. El debate de la Inteligencia Artificial en la ONU. (unric.org)

ca y los sistemas de energía, tal y como las hemos conocido hasta ahora.

Por tal motivo, diversas organizaciones internacionales se han dado a la tarea de examinar este fenómeno para conocer tendencias, generar proyecciones e identificar fortalezas y debilidades en las empresas y los países. Lo anterior indica la importancia atribuida al tema de IA, no obstante, no todo son proyecciones positivas, pues es difícil predecir las consecuencias que la IA podrá generar en la economía y las sociedades[6].

Uno de los aspectos que más preocupa es su impacto en los mercados laborales pues se cree que podría llevar a una reconfiguración y al reemplazo de ciertas ocupaciones. El Fondo Monetario Internacional (FMI) indica que cerca del 40% de los trabajos en el mundo estarán expuestos a la IA. A pesar de eso, también se espera que la IA complemente el trabajo que realizan las personas. Por eso, se considera fundamental que la IA sea gestionada de forma tal que esta permita una revitalización de la productividad, pero no conduzca a una profundización de la desigualdad entre sectores, países y a lo interno de estos[7].

Para esto, parece esencial garantizar que los sistemas de IA sean desarrollados y usados de una forma que no violente los derechos humanos, respete la privacidad y prevenga daños. Esto ha llevado a reflexiones y dilemas éticos con respecto a la forma como se diseñan y entrenan los algoritmos de IA, su transparencia y explicabilidad, así como los procedimientos de recolección, almacenamiento y datos que los alimentan. A partir de esto, ha surgido un acalorado debate sobre la pertinencia de regular o no estos desarrollos tecnológicos y si los gobiernos tienen la responsabilidad de generar este tipo de regulaciones[8]; al tiempo que también se discuten qué áreas y alcances podría abarcar una regulación de esta naturaleza.

Ante este panorama resulta necesario reflexionar la IA, no sólo como una tecnología, sino también como un fenómeno transformador y disruptivo que requiere de una mirada crítica que resalte los posibles impactos ético-sociales y económicos que tendrá la creciente integración de la IA en nuestras sociedades. Desde esta óptica, el Programa Sociedad de la Información y el Conocimiento (Prosic) decidió dedicar sus Jornadas de Investigación y Análisis 2024 al estudio de la IA. Con este objetivo se desarrollaron las *Jornadas de Investigación IA, ¿moda pasajera o cambio permanente del paradigma?* como un espacio de discusión multisectorial para abordar los impactos que la IA está generando actualmente en nuestras sociedades.

6 Cazzaniga, M., Jaumotte, F., Longji, L., Giovanni, M., Augustus, J., Pizzinelli, C., Rockall, E., & Tavares, M. (2024). Gen-AI: Artificial Intelligence and the Future of Work. IMF Staff Discussion Note SDN2024/001, International Monetary Fund, Washington, DC.

7 Georgieva, K. (16 de enero de 2024). La economía mundial ha transformado por la inteligencia artificial ha de beneficiar a la humanidad. IMF Blog. La economía mundial transformada por la inteligencia artificial ha de beneficiar a la humanidad (imf.org)

8 Jabbour, G. (16 de junio de 2023). Mareas normativas: así es como se está regulando la Inteligencia Artificial. EXPANSION. Así es como se está regulando la Inteligencia Artificial a nivel global (expansion.mx)

Durante los tres días del evento, expertos en inteligencia artificial y representantes gubernamentales, de organizaciones privadas, organismos internacionales y de sociedad civil contribuyeron al análisis colectivo de la evolución, las potencialidades, los riesgos y los desafíos suscitados por la IA para los Estados, las empresas y las organizaciones. Es por eso que ante la multiplicidad y riqueza de los puntos vista externados en la actividad, resulta necesario que se documenten dichos aportes para que se constituyan en insumos útiles para la consulta pública, sea con fines académicos o para nutrir procesos de toma de decisión.

Bajo esta intención, la **Memoria de las Jornadas de Investigación y Análisis IA, ¿moda pasajera o cambio permanente del paradigma?** nació como un producto de conocimiento que pretende plasmar la diversidad de enfoques, discusiones y puntos de vista que surgieron en el evento. A este efecto, el documento está integrado por 16 ponencias, las cuales están articuladas en 5 ejes de reflexión distintos en los que se abordan las bases conceptuales, historia y origen de la IA; los casos de uso de esta tecnología; los riesgos y dilemas éticos; los debates y tensiones regulatorias sobre la IA; y los avances y oportunidades que esta ofrece en Costa Rica.

De ese modo, la Memoria inicia planteando que, para un aprovechamiento efectivo de la inteligencia artificial se requieren condiciones específicas que faciliten los procesos de integración de dicha tecnología. Sin embargo, no basta con incorporar la IA en distintas áreas o digitalizar y automatizar procedimientos, sino que se requiere habilitar recursos (capital humano, datos e infraestructura) que potencien su desarrollo, fortalezcan las capacidades en investigación, desarrollo e innovación (I+D+i) de los Estados y contribuyan a disponer de una institucionalidad y un marco regulatorio propicio a la IA. Bajo esta consideración la ponencia de **Loreto Aravena Mori** ofrece una mirada integral en la que se examinan los avances para el desarrollo, la adopción y la gobernanza de la IA de 19 países latinoamericanos a través del *Índice Latinoamericano de Inteligencia Artificial* (ILIA) del Centro Nacional de Inteligencia Artificial (CENIA) de Chile.

Los hallazgos de esta medición muestran que la región ha realizado esfuerzos para robustecer las infraestructuras digitales, impulsar la formación de talento especializado y crear regulaciones que faculten el desarrollo ético y responsable de la IA. No obstante, de la mano de estos avances, se observan desafíos vinculados con la brecha digital, el desigual acceso a conectividad de alta calidad, la baja capacidad computacional, la limitada oferta formativa en IA avanzada, una marcada brecha de género en la investigación científica y un bajo nivel de innovación (medido por patentes, startups tecnológicas y aplicaciones tecnológicas desarrolladas). Esta radiografía, más que informar sobre el estado de situación de la región en materia de IA, constituye un marco de referencia para orientar la construcción de políticas que atiendan a las brechas señaladas y potencien los beneficios de la IA.

En esta línea, **Johan Montero Granados** resalta las potencialidades de la IA como una tecnología capaz de efectuar tareas que asemejan la inteligencia humana; lo que en su criterio,

optimizará todo tipo de procesos, haciéndolos más eficientes e innovadores. Gracias a esto y a la vinculación con otras tecnologías disruptivas -como las redes 5g y el IoT- se abren las puertas para revolucionar campos como la agricultura y las telecomunicaciones. Empero, más allá de estos beneficios, el autor nos recuerda que la adopción de la IA debe ser un proceso informado, estratégico y orientado a resultados; además de apegarse a principios éticos, estándares de seguridad y de sostenibilidad. Para esto, se requiere de acciones conjuntas entre gobiernos, reguladores y empresas.

De la mano de estas consideraciones, el recorrido histórico ofrecido por **Henry Lizano Mora** nos deja ver que la IA no es un tema nuevo y que, por el contrario, sus cimientos teórico-conceptuales han estado presentes desde la antigüedad. Bajo esta perspectiva se destacan los periodos de auge y de retrocesos en el desarrollo de la IA, así como los elementos que permitieron el resurgimiento del campo (el *big data*, el aprendizaje profundo y la aparición de modelos generativos). Con base a esto, se resaltan los principales desafíos que se enfrenta con el avance de la IA, destacándose retos como el desempleo, los dilemas en la toma de decisiones autónomas y el antropomorfismo; y subrayando, la necesidad de impulsar los desarrollos tecnológicos con una mirada humana e integral que considere los diversos aspectos que intervienen en los procesos de adopción.

Para **Óscar Rojas Badilla**, una de las cuestiones centrales a las que debe prestársele atención es al crecimiento exponencial de los datos digitales ya que eso incrementa las necesidades de almacenamiento y genera una enorme presión en términos de la infraestructura y la sostenibilidad de los Centros de Datos. Este elemento clave para el desarrollo de la IA, consume cerca del 3% de la electricidad mundial y produce el 0,5% de las emisiones de carbono, lo que obliga a buscar soluciones ante la creciente necesidad de los Centros de Datos. En este contexto, el autor se refiere a la experiencia de la empresa BJumper con el fin de ejemplificar la forma como la IA puede optimizar las operaciones de los Centros de Datos al facilitar la automatización de operaciones críticas y aplicar análisis predictivos que pueden mejorar la eficiencia energética, disminuyen los errores humanos y mitigan el impacto ambiental.

Aplicar este tipo de mejoras a las infraestructuras requeridas para el desarrollo de la IA, permite el aprovechamiento de dicha tecnología en áreas muy diversas, como el campo de la salud. En esta rama en particular, la IA puede agilizar el análisis de grandes volúmenes de datos, automatizando procesos y haciendo los diagnósticos médicos más precisos. Eso no solo supone la posibilidad de contar con servicios de salud más eficientes y personalizados, sino que también representa una oportunidad para impulsar nuevos nichos productivos. En ese sentido, la intervención de **Rodrigo Herrera** ejemplifica un caso de innovación tecnológica nacional en el que la startup costarricense AInnovatech desarrolló la solución VisionAI, un modelo de IA capaz de diagnosticar la retinopatía diabética a partir del análisis de imágenes de retina y realizar la detección temprana de otras enfermedades.

Por otro lado, la IA también está transformando la comunicación y la manera como se genera contenido, especialmente por el creciente uso de la inteligencia artificial generativa (IA generativa). En la mirada de **Óscar Alvarado Rodríguez**, este cambio tecnológico trae consigo desafíos en materia de transparencia, veracidad y autoría del contenido que es generado con estas herramientas, lo que nos obliga a implementar mecanismos para mitigar y prevenir riesgos como la desinformación, la pérdida de empleos y la reproducción de sesgos en los modelos de IA. Para el autor esto evidencia que el desarrollo tecnológico no es un proceso neutral y que por el contrario, este se ve afectado por el contexto, las ideas, los prejuicios e ideas preconcebidas que tienen quienes diseñan sistemas de IA; lo que ocasiona consecuencias negativas.

En este contexto cuestiones como la protección de los derechos humanos y la seguridad de la información resultan esenciales para orientar el desarrollo y la creación de sistemas de IA. Aunado a ello, **Luis Arturo Martínez Vásquez** considera necesario analizar los impactos psicopolíticos que produce el uso de herramientas de IA, sobre todo por la masiva recopilación de datos, un fenómeno que mercantiliza la experiencia humana, limita la privacidad de las personas en los entornos digitales y puede reforzar el control social en nuestras sociedades. En su opinión esto se puede ver reforzado por la adopción de políticas de privacidad confusas y la falta de una legislación efectiva, lo que a su parecer restringe el derecho a la autodeterminación informativa.

Por lo anterior, Martínez propone una postura ética centrada en el bien común, que revalorice la privacidad, siente responsabilidades y propicie un modelo de IA en el que la libertad y la dignidad humana sean pilares centrales no sacrificados en función de la rentabilidad. El texto de **Andrea Vera Celeste** complementa esta visión al señalar que el diseño e implementación de los sistemas de IA deben integrar criterios éticos y una gestión de riesgos que permita considerar situaciones como la discriminación algorítmica, la falta de transparencia, la manipulación de datos y las amenazas a la autonomía individual a lo largo de todo el ciclo de vida de la IA. Un enfoque como este considera los efectos no deseados que pueden surgir cuando la ética es omitida o aplicada de una manera superficial. Para prevenir esto, la autora sugiere la adopción de frameworks para la gestión de riesgos, la transparencia en el uso de datos, la capacitación continua del personal técnico y la integración de principios éticos que orienten el desarrollo tecnológico hacia el respeto de los derechos humanos y el bienestar colectivo.

Con esta referencia **Gloria Guerrero** nos recuerda que el desarrollo de la IA debe ser pensado local y regionalmente porque ello permite entender mejor los impactos, las oportunidades y los riesgos que esta tecnología puede generar en contextos específicos. Desde esta perspectiva, analiza el desempeño de América Latina en el *Índice Global sobre IA Responsable* (GIRAI); para evidenciar avances en la regulación y el desarrollo de marcos éticos, y encontrar falencias en la transparencia, la seguridad y la inclusión en el diseño e implementación de la IA. Para solventar estos vacíos, deben vincularse diversos actores a estos procesos, pues esta es la única manera de que

se construya una visión ética y contextualizada de la IA que se alinee con los retos que enfrenta el Sur Global.

Pero para lograr una implementación responsable y justa de la IA, se le debe prestar especial atención a los esfuerzos regulatorios que realizan los países y las regiones. En esta postura se enmarca la ponencia de **Natalia González Alarcón**, quien examina las iniciativas regionales impulsadas en América Latina, destacando los esfuerzos orientados a consolidar una gobernanza regional que refleje las prioridades locales del área (como las declaraciones de Santiago y Montevideo); así como el creciente interés de los Estados por discutir si se requiere impulsar regulaciones en IA.

Aunque esta preocupación responde a los riesgos que puede generar la IA, emitir cualquier regulación no resulta una tarea fácil. Los rápidos avances de esta tecnología dificultan la aplicación de enfoques más tradicionales en materia regulatoria y obligan a reconocer la complejidad y particularidades de esta tecnología. Para **Daniel Rodríguez Maffioli**, la IA es un fenómeno multidimensional que se expresa en los niveles tecnológico, social, filosófico y moral, cuyas repercusiones tan diversas nos invitan a evitar abordajes superficiales y a profundizar en el análisis de los impactos e implicaciones que puede tener la IA. Asumir una posición como esta permite que nos cuestionemos qué fundamenta la necesidad de regular o no la IA, qué debe regularse, de qué forma y cómo. Esta visión resalta la importancia de adoptar enfoques regulatorios más flexibles que favorezcan la experimentación y eviten la creación de leyes que pueden quedar obsoletas rápidamente.

En paralelo, otras de las áreas que suelen mencionarse con frecuencia cuando se plantea el tema de las regulaciones de la IA tienen que ver con propiciar la innovación, proteger los derechos humanos, garantizar la privacidad y la seguridad de la información, entre muchas otras. Ante, la variedad de aristas que pueden considerarse, **Fabián Gamboa Hernández** plantea la necesidad de enfocarse en examinar el impacto que la IA está ocasionando en el ámbito de los derechos de autor pues los avances para producir textos e imágenes generan interrogantes con respecto a la autoría de estas obras y si estas pueden ser protegidas por derechos de autor.

En la óptica de Gamboa esto sugiere desafíos ético-jurídicos relacionados con cuestiones como el uso de obras protegidas como datos de entrenamiento para los modelos de IA, la posibilidad de que un ente no humano pueda ser considerado (o no) como autor y la ultrasuplantación (apropiación del contenido producido con IA).

Otro aspecto poco asociado a la IA, pero muy relacionado con esta tecnología son las plataformas digitales. Estas no sólo han cambiado el mundo laboral al digitalizar gran número de actividades y precarizar el empleo; sino que también han llevado a una adaptación de los marcos regulatorios a las nuevas formas de trabajo surgidas en el marco de estas transformaciones. Bajo este enfoque, **Ronald Saénz Leandro** analiza la regulación de plataformas en América Latina e identifica avances y vacíos que demuestran que la normativa de la región ha tendido a ser parcial, poco integral y ha sido influenciada por coyunturas sociales, políticas y económicas.

Ante la excesiva atención prestada a los aspectos técnicos o económicos, **Henry Lizano Mora**, propone una mirada antropológica hacia el avance acelerado de la IA en la que más allá de los ciclos de sobre expectación tecnológica, se recupere la centralidad de lo humano en las discusiones sobre las tecnologías emergentes. Poner el foco en esto permite entender las transformaciones radicales que están ocurriendo en diferentes ámbitos, mostrando no sólo las ventajas, sino también las afectaciones negativas (como el antropomorfismo digital) que plantean nuevos dilemas éticos y exige garantizar que las transiciones tecnológicas adopten un enfoque humano, inclusivo y ético.

Lo anterior resulta especialmente relevante sobre todo por la creciente digitalización de la gestión pública, una tendencia que ha llevado a la constante utilización de distintas tecnologías de la información y la comunicación (TIC). En este contexto, la integración de la IA ha tomado gran auge por las potencialidades que ofrece; sin embargo, si bien esta tecnología contribuye a incrementar la eficiencia, la transparencia y añade de mayor valor agregado a los servicios, **Jorge Umaña Cubillo** nos recuerda que su inclusión en la Administración Pública debe ser realizada bajo estándares, políticas y estrategias alineadas con principios éticos y los derechos humanos.

Desde este punto de vista, Umaña examina las directrices establecidas en la *Carta Iberoamericana de Inteligencia Artificial* (2023) resaltando la forma como esta herramienta puede orientar a los gobiernos en la implementación de la IA en el sector público mediante principios como la transparencia, la rendición de cuentas, la equidad, la protección de datos y la cooperación internacional. Siendo esencial que los Estados cuenten con herramientas que orienten en el uso, aprovechamiento y desarrollo de la IA, la Memoria cierra con la intervención de **Marlon Ávalos Elizondo**, quien se refiere a la *Estrategia Nacional de Inteligencia Artificial* (ENIA) y la presenta como experiencia en la que a partir del reconocimiento de la IA como un catalizador de desarrollo, se logró adoptar un instrumento para orientar el desarrollo e implementación de la IA de una forma ética, segura y centrada en las personas en Costa Rica, desde la colaboración inter y multisectorial.

Es así como la diversidad de textos incluidos en esta Memoria demuestra la importancia de continuar impulsando espacios de discusión sobre la inteligencia artificial, ya que la rápida evolución de esta tecnología continuará impactando la economía, la educación, el transporte, las comunicaciones, la salud e inclusive las relaciones humanas. En ese sentido, estas Memorias, más que recoger las experiencias compartidas en las Jornadas de Investigación del 2024, pretenden visibilizar la multiplicidad de oportunidades y desafíos que nuestras sociedades enfrentarán durante las próximas décadas a nivel ético, legal y de seguridad.

Es por eso que consideramos que este producto de conocimiento servirá para identificar vacíos, resaltar fortalezas, señalar recomendaciones y orientar en el desarrollo de investigaciones futuras, la creación de soluciones y el diseño de regulaciones basadas en la evidencia. Con esta intención, deseamos que este documento aporte a los debates sobre el de-

sarrollo de los sistemas de IA en Costa Rica desde una mirada crítica que ayude a enfrentar un futuro que muy probablemente cada vez, sea más como digital, interconectado e incierto y en el que deben encontrarse nuevos caminos para promover el debate y la construcción colectiva del conocimiento.

Valeria Castro Obando

Investigadora y Coordinadora de las

Jornadas Anuales de Investigación del Prosic

1 El índice latinoamericano de inteligencia artificial (ILIA): un panorama de la inteligencia artificial en la región

Loreto Aravena Mori

Loreto Aravena Mori. Coordinadora Ejecutiva del índice Latinoamericano de la Inteligencia Artificial (ILIA).

Periodista y máster en Arquitectura y Urbanismo de la U. Politécnica de Cataluña, con más de 20 años de experiencia en creación y edición de contenidos en medios de comunicación escritos y radiales. Como directora de Comunicaciones ha posicionado los objetivos estratégicos de organizaciones encargadas de impulsar el desarrollo económico sostenible y justo de Chile. En su actual cargo como coordinadora ejecutiva del Índice Latinoamericano de Inteligencia Artificial, lidera y coordina este informe, desde su concepción hasta su publicación final.

19

Loreto Aravena Mori

Resumen

La ponencia expone los principales resultados de la segunda edición del Índice Latinoamericano de Inteligencia Artificial (ILIA) una medición desarrollada por el Centro Nacional de Inteligencia Artificial (CENIA) para medir el avance de la inteligencia artificial (IA) en 19 países de América Latina y el Caribe (ALC) a través de tres dimensiones clave: factores habilitantes, investigación-desarrollo-adopción y gobernanza. Con esto el ILIA busca orientar a los tomadores de decisiones de cada país en la generación de políticas públicas de IA, fomentar las oportunidades de cooperación regional y promover un desarrollo ético y equitativo de la IA.

Palabras clave: Inteligencia artificial, América Latina y el Caribe, gobernanza digital, innovación tecnológica, políticas públicas.

Introducción

Durante los últimos años, la inteligencia artificial (IA) se ha convertido en una herramienta para potenciar todo tipo de transformaciones en las sociedades actuales, lo que supone beneficios, pero también plantea desafíos complejos para los países. Este acelerado avance tecnológico no sólo exige la adopción de marcos de acción éticos, responsables y contextualizados, sino también la necesidad de impulsar la creación de herramientas que ayuden a evaluar el estado de desarrollo y la preparación de los Estados ante la IA.

Es desde esta perspectiva que el Centro Nacional de Inteligencia Artificial (CENIA) desarrolló el Índice Latinoamericano de Inteligencia Artificial (ILIA) como una medición para que los países de América Latina y el Caribe (ALC) comprendan su grado de preparación tecnológica, identifiquen brechas y oportunidades y establezcan políticas públicas basadas en evidencia. Esto en un contexto que indica que la región enfrenta serios retos en materia de infraestructura, talento humano, regulación y políticas públicas digitales.

Por tal motivo, la presente ponencia ahonda en el contexto, la metodología y los principales hallazgos del ILIA para su edición 2024. Los resultados del índice indican los logros y las brechas que posee ALC en este ámbito, estableciendo comparaciones entre países. Además, resalta la importancia de fortalecer el ecosistema regional de IA y de posicionar al área como un actor activo dentro de los debates globales sobre la IA.

¿Qué es el CENIA y cuál es su misión?

El Centro Nacional de Inteligencia Artificial (CENIA) nació como una iniciativa de un grupo de investigadores de las

cuatro principales universidades del país (Universidad Católica, Universidad de Chile, Universidad Técnica Federico Santa María y la Universidad Adolfo Ibáñez) que, inspirados por la publicación de la *Primera Política Nacional de Inteligencia Artificial* del 2021, buscaban promover el desarrollo y el uso responsable y ético de la IA en Chile. A partir de esto, postularon a un concurso de financiamiento basal para centros científicos y tecnológicos capaces de desarrollar proyectos de alto impacto a largo plazo y que otorga la Agencia Nacional de Investigación y Desarrollo (ANID) por plazos de 5 a 10 años. Gracias a que se ganó dicho concurso se pudo establecer el CENIA hasta llegar a constituirse en una corporación privada que hoy es capaz de generar sus propios ingresos.

Durante el 2021, el CENIA inició sus funciones oficialmente y desde ese momento, ha trabajado para desarrollarse en tres áreas de acción:

1. Investigación de excelencia: el CENIA está conformado por más de 72 investigadores y 100 estudiantes de 10 universidades chilenas. Además, el centro concentra la producción académica chilena en el campo de la IA, pues ha producido más de 90% de las publicaciones del país en el tema. Las investigaciones del centro se realizan en cinco áreas distintas: 1) aprendizaje profundo para visión y procedimiento de lenguaje natural, 2) IA simbólica, 3) la IA inspirada en el cerebro, 4) aprendizaje de máquinas para la física y 5) IA centrada en el ser humano.
2. Transferencia tecnológica: en este ámbito el centro actúa como un laboratorio de investigación y desarrollo (I+D) que entrega soluciones de IA al Estado, la industria y la sociedad civil. Este tipo de servicios pretenden beneficiar a las personas y fortalecer el sistema local de innovación basado en tecnologías de IA.
3. Vinculación con el medio: para esto, el CENIA implementa proyectos y experiencias para acercar la IA a la sociedad con iniciativas que incluyen la alfabetización, la capacitación de la fuerza laboral hasta la generación de espacios que permitan comprender mejor el alcance y uso responsable de la IA.

Origen del ILIA

La extensa variedad de temáticas que se trabajan desde el CENIA refleja el interés del centro de convertirse en un referente latinoamericano en IA. Es por ello, que bajo esta intención se consideró oportuno el establecimiento de proyectos de alcance regional, planteándose la idea de crear una medición que sirviera para que los países puedan identificar necesidades y registrar avances en materia de IA, con el fin de que sirva como un un bien público regional que permite dialogar sobre el desarrollo de la IA con base en evidencia.

Así fue como nació el Índice Latinoamericano de Inteligencia Artificial (ILIA), como un proyecto que no sólo involucra a los investigadores e investigadoras del CENIA, quienes se encargan de recopilar información cuantitativa y cualitativa del estado de avance de la IA; sino que también integra a distintos actores del mundo académico privado y estatal para que

aporten con sus puntos de vista a la construcción de esta herramienta. En ese sentido el ILIA es un instrumento que permite al CENIA generar puntos de encuentro con cuatro mundos: el público, el académico, la sociedad civil y la empresa. No solo porque los involucra durante el proceso de construcción del informe, sino porque a estas capas les entrega información relevante sobre una IA que busca estar más cerca de las personas, a su servicio. Sin medición de la IA es difícil saber dónde generar puntos de colaboración entre países y entre las distintas capas involucradas en el ecosistema de IA.

Una contraparte esencial del proyecto del ILIA ha sido la Comisión Económica para América Latina y el Caribe (Cepal), quién por segundo año consecutivo trabajó con el CENIA en el documento que provee información relevante para la toma de decisiones y la proyección de oportunidades de crecimiento económico para los países de América Latina y el Caribe. Otras organizaciones importantes han sido la Organización de Estados Americanos, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (Unesco), la Unión Europea (UE) y las instituciones financieras encargadas de promover el desarrollo socioeconómico en ALC, como el Banco Interamericano de Desarrollo (BID) y el Banco de Desarrollo de América Latina y el Caribe (CAF).

¿Por qué medir la IA?

Tener una medición de esta tecnología permite identificar el estado de situación y el progreso de la IA en la región, así como determinar su nivel de preparación, sirviendo para orientar los esfuerzos estratégicos y de inversión de los países. Por ello, el ILIA busca constituirse en un instrumento de robustez metodológica que analice de forma integral los ecosistemas de la IA, ofrezca un monitoreo de indicadores a lo largo del tiempo y detecte tanto las brechas como las oportunidades de mejora.

El ILIA tiene una taxonomía, una manera de estructurar la entrega de los datos cuantitativos y cualitativos de los ecosistemas de IA en la región. Por eso, se ordena en torno a dimensiones, las que se componen de subdimensiones y éstas, a su vez, indicadores. Estos tienen su expresión más granular en los subindicadores, que son los que finalmente recogen los datos brutos de distintas variables.

Por su parte, los datos brutos se obtienen de fuentes secundarias y primarias, transparente y legítimas, todas ellas respaldadas además por la consultoría de la CEPAL este año para actualizar y fortalecer todos los indicadores y subindicadores del índice. Tres dimensiones componen el ILIA: 1) Factores Habilitantes; 2) Investigación, Desarrollo y Adopción; y 3) Gobernanza. Cada dimensión se desglosa en subdimensiones, indicadores y subindicadores, los cuales se normalizan en una escala de 0 a 100 puntos.

En la dimensión de **factores habilitantes** se evalúan los elementos y condiciones tecnológicas que facilitan el desarrollo efectivo de los ecosistemas de IA y para ello, se abordan tres subdimensiones:

- Infraestructura: relacionada con el soporte que permite que la IA progrese y que considera aspectos -o indicadores- como la *Conectividad, el Cómputo y Dispositivos*.

- Datos: referida a aquellos datos públicos y fiables disponibles para ser utilizados en innovaciones en IA y que se miden a través de un indicador llamado Barómetro de Datos, el que a su vez está compuesto de cuatro subindicadores que miden y entregan puntaje: *Disponibilidad* (acceso a datos abiertos), *Capacidad* (para recopilar, descargar, procesar, usar, y compartir los datos), *Gobernanza* (datos fiables, completos y transparentes) y su *Uso e Impacto.*

- Talento Humano: examina las competencias que tienen las personas o profesionales de un país para diseñar, desarrollar e implementar soluciones basadas en IA, y que se mide a través de indicadores como *Alfabetización, Formación Profesional* y *Talento Humano Avanzado.*

En la segunda dimensión sobre **Investigación, desarrollo, innovación y adopción** aborda los avances en investigación, I+D y adopción de la IA en los sectores público, privado y académico. Para poder medirlos, se separan en tres subdimensiones:

- Investigación: referida a la capacidad de cada país para generar nuevo conocimiento, lo que se mide a través de la cantidad de *Publicaciones en IA,* cuán activa es la *Investigación en IA de los investigadores* (si publican seguido en los últimos años) y la *Productividad de investigadores (as)*, además del *Impacto de sus contenidos* sobre otros papers científicos en IA. Y otras variables como *la Presencia de centros de investigación de IA;* la *Proporción de autoras en esta disciplina;* la *Investigación consistente en IA,* y la *Participación en main tracks y side events de conferencias A+.*

- Innovación y Desarrollo: que evalúa el grado de *Desarrollo* de softwares y patentamientos, la *Productividad open source, la Calidad open source, y su Calidad, y* el grado de Innovación que alcanza cada país, lo que se mide a través de subindicadores como *Empresas de IA, Empresas Unicornio, el Número de Inversiones Privadas en I+D*, *Desarrollo de Aplicaciones* y *Gasto en I+D,* entre otras.

- Adopción: que analiza la integración de la IA a nivel de Industria y Gobierno de cada país. A nivel industria, mide la adopción de la IA en las actividades económicas que transforman materias primas en productos y en las que hay generación de bienes y servicios, y lo hace a través subindicadores como *Trabajadores en sectores de alta tecnología, Fabricación de tecnología mediana y alta, y Proporción del valor añadido de fabricación de tecnología mediana y alta*.

En la dimensión de gobernanza, se evalúan las normas, las reglas, las políticas y los mecanismos que regulan el uso ético

y responsable de la IA. Este pilar se compone de tres subdimensiones con indicadores y subindicadores:

- Visión e Institucionalidad: que considera las *Estrategias de IA vigentes,* el *Involucramiento de la sociedad en la construcción de éstas* y la *Institucionalidad* que las ejecuta.

- Vinculación Internacional: referida a la *Participación en organismos internacionales*, por parte de los países, en instancias de discusión sobre regulación de IA y la *Participación en la definición de estándares*, que reflejan su adhesión a mecanismos de gobernanza globales.

- Regulación: referida a las *normativas vigentes*, al compromiso de las naciones con la *Ciberseguridad* y a los pasos que éstas han seguido en Ética y Sustentabilidad.

En la edición 2024 del ILIA se examinaron 19 países de ALC (Argentina, Bolivia, Brasil, Chile, Colombia, Costa Rica, Ecuador, México, Panamá, Paraguay, Perú y Uruguay, se sumaron países como Cuba, El Salvador, Guatemala, Honduras, Jamaica, República Dominicana, Venezuela), 7 más que en el 2023.

Metodología del ILIA

En 2024, el índice incorporó a siete nuevos países, sumando un total de 1 y se expandieron sus variables de medición de 49 a 76. Se emplearon fuentes secundarias validadas, priorizando la comparabilidad regional. Además, se incluyó un Comité Técnico Asesor con representantes de organismos multilaterales como CEPAL, UNESCO, BID, CAF y la OEA.

Hay ciertas consideraciones metodológicas que el índice tiene en cuenta, tanto en esta edición como lo hizo en la del año pasado. En primer lugar, debe mencionarse que se priorizó el uso de fuentes secundarias de carácter público, transparentes y legítimas. Además, se decidió que estas debían brindar información sobre los 19 países estudiados. En consecuencia, se consideraron fuentes como: ITU Datahub, el Speedtest de Ookla, el Global Connectivity Index 2020 para obtener cifras de Nube y los informes del HPC Observatory, entre otras.

Por otro lado, el enfoque del ILIA es asegurar la comparabilidad de los datos entre los diferentes países para darle validez al estudio, por lo que en aquellos casos en los que la información pública solo estaba disponible para un número limitado de ellos, se optó por no utilizar dichos datos.

Se usó una escala de 0 a 100 para estandarizar los valores de diferentes indicadores en una escala común, lo que permite la comparación directa. Los datos originales de los elementos

medidos, en bruto, tienen unidades, rangos, o magnitudes diferentes y **se normalizaron** para transformar los valores a una escala uniforme y permiten la comparación relativa entre países de América Latina y el Caribe[1].

Este año se priorizaron ciertos elementos (dimensiones, subdimensiones, indicadores) en la construcción del índice, porque determinan la **influencia relativa de cada uno en el resultado final.** Por ejemplo, es muy importante el peso de Factores Habilitantes este año, porque sin esos elementos no se puede avanzar en el desarrollo de la IA. Asimismo, debe aclararse que los puntajes no son extrapolables fuera de América Latina, porque cada región tiene su propio contexto cultural, económico y político.

Mientras en 2023 se evaluó el estado de avance de la IA en **12 países** (Argentina, Bolivia, Brasil, Chile, Colombia, Costa Rica, Ecuador, México, Panamá, Paraguay, Perú y Uruguay) en 2024 se sumaron Cuba, El Salvador, Guatemala, Honduras, Jamaica, República Dominicana, Venezuela, llegando a un total de 19. También se actualizaron los subindicadores, que son las variables que se miden y entregan puntaje. Respecto del año 2023, aumentaron de 49 a 76 las variables a medir. Por último, debe mencionarse que se adhirieron más contrapartes, recomendadas por la Cepal y por organizaciones prestigiosas del ecosistema y que poseen conexiones a nivel latinoamericano.

1 La mayor parte de los datos son a escala de 0 a 100, pero hay otras de categorizaciones 0 a 1. (Ejemplo: países que no tienen estrategia de IA=0 y países que tienen= 1).

Resultados principales: Pioneros, adoptantes, exploradores

El ILIA indica que los países que tienen mayor madurez en sus diferentes dimensiones y que por tanto son considerados como **pioneros** fueron Chile (con una puntuación de 73,07), Brasil (69,30) y Uruguay (64,98). Estos destacan no sólo por su avance en la implementación de tecnologías basadas en la IA, su progreso en la generación de talento especializado y en los ámbitos de productividad científica y capacidad de innovación, sino también porque están orientando sus estrategias nacionales hacia la consolidación y expansión de estas tecnologías en todos los sectores de su economía y sociedad.

Por su parte, en el grupo de los países **adoptantes** se encuentran Argentina, Colombia, México y República Dominicana pues están usando la IA en los sectores productivos, los servicios y las administraciones públicas, pero de manera incipiente. En el ámbito de la investigación, han logrado avances significativos en IA, aunque todavía no la escalan al nivel de los países pioneros. Asimismo, en lo que respecta a políticas de fomento de la IA, están desarrollando estrategias de IA.

En el caso de los Estados **exploradores** debe señalarse que estos son los países con los puntajes más bajos, ya que están en las primeras etapas de sondeo de la IA, desarrollando capacidades básicas en esta área.

Cuando se analizan los resultados por dimensión se evidencia que los países con mejores puntajes en la dimensión de *Factores Habilitantes* fueron Chile con una puntuación de 64,60 y Uruguay con 60,70. A estos le siguen Argentina, Brasil, Costa Rica y México, que superan el promedio regional de 40,26 puntos. En contraste, los otros 13 países se ubican alrededor o por debajo de este promedio. De igual modo, el promedio regional no está por sobre los 50 puntos, lo que refleja el desafío que tiene por delante la región en aspectos básicos para empujar la IA como: la Conectividad, que en el área apenas sobrepasa los 50 puntos; en Cómputo, que es bajo con 21 puntos; y en Talento Humano, que tiene 39 puntos.

Pese a que en promedio ha aumentado en un 100% la concentración de talento en IA en la fuerza de trabajo de América Latina y el Caribe en los últimos ocho años, ningún país ha alcanzado los niveles evidenciados en países del Norte Global en el mismo período de tiempo. También persisten brechas estructurales relacionadas con los déficits en infraestructura de cómputo de alto rendimiento (HPC), la velocidad de conectividad, el talento humano avanzado y la equidad de género en investigación.

La participación de mujeres en investigación de IA varía significativamente entre países de la región, aunque en términos generales, Latinoamérica y el Caribe se encuentran significativamente por debajo del 50% esperado de mujeres en la investigación en IA en todos los países considerados en este índice. Desde 2010 hasta 2023, el porcentaje de participación femenina se ha mantenido prácticamente invariable. Hay una insuficiencia de esfuerzos para reducir la brecha de género en la mayoría de los países.

Las mediciones de Talento Humano arrojan que Chile lidera con 74,30 puntos y Uruguay le sigue con 62,11 y que son los únicos dos países que superan la barrera de los 60 puntos; mientras que en alfabetización Chile, Uruguay y Argentina están a la cabeza. Esto indica que la **Educación temprana en ciencia** debe iniciar desde una edad temprana de modo que ello estimule el desarrollo de capacidades que impulsen la IA. Por ello, es urgente incentivar políticas fiscales que enfoquen la atención en el pensamiento crítico, pensamiento computacional y vocaciones de STEM en la escuela.

Sobre el talento humano avanzado, se analizaron las universidades que ofrecen programas de doctorado y magíster en IA para evaluar la capacidad de cada país para formar capital humano avanzado en IA. Solo **9 países** de los 19 considerados en el estudio, cuentan con **programas de magíster en IA de excelencia internacional,** pero su acceso está fuertemente limitado por la baja cantidad y cobertura. El puntaje promedio regional es de 10,48 puntos, destacándose Uruguay con 100 puntos, equivalente a cinco programas de magíster, seguido por Chile con 38,36 puntos.

En doctorado el promedio es aún más bajo con 4,99 puntos. Solo tres países tienen programas de PhD en universidades del Ranking QS[2]: Brasil, México y Chile. Este último país

2 Los indicadores referentes al Ranking QS buscan mostrar la presencia de programas de formación altamente competitivos en el marco global.

muestra un puntaje de 75 puntos, porque en el país ha habido una política de incentivos a esta especialización desde el gobierno central, que destina fondos para aquello.

Por su parte, en materia de Investigación, Desarrollo y Adopción se aprecia que el desempeño regional es de 47,46 puntos en promedio, destacando naciones como Brasil, Chile, Uruguay y México, con puntuaciones de 79,17; 75,21; 66,68 y 66,20. Las puntuaciones son reducidas sea por la insuficiente inversión en este ámbito o por la falta de incentivos.

Sobre el ecosistema en maduración el ILIA indica que más allá de las diferencias de magnitud, todos los países cuentan con al menos dos investigadores consistentes en IA, 11 de los 19 presentan centros de investigación en IA dependientes de alguna universidad o privado. A ellos se suma el aumento del número de publicaciones con relación al año anterior. Además, 5 países cuentan con más de dos centros de investigación en IA: Argentina, Brasil, Chile, México y Perú. Esto representa un contraste con respecto a los resultados de 2023, cuando se identificaron tres países.

Estos centros no sólo impulsan sus líneas de trabajo en la investigación y la innovación, sino que además facilitan la formación de talento especializado, lo que permite comprender el nivel de sofisticación y madurez del ecosistema regional en esta área estratégica.

Con respecto al nivel de innovación que alcanzan los países en IA, se midió la cantidad de patentes, estrechamente relacionadas con una alta capacidad de transformar innovaciones científicas y académicas en soluciones concretas asociadas a problemas públicos y el sector privado. Por lo mismo, refleja un entorno de colaboración robusto entre academia, industria y gobierno.

En el transcurso del último año, el panorama es bastante similar a la medición de 2023. México con 100 puntos (equivalente a 4,22 patentes de IA por millón de habitantes) y Brasil con 90,78 puntos (3,84 patentes de IA por millón de habitantes). El resto de los datos muestran un panorama desalentador, de una región con baja productividad en esta área, lo que podría explicarse por la baja inversión en I+D+i.

En la región, la creación de startups privadas que han alcanzado una valoración de **1.000 millones de dólares o más**, sin haber salido a la bolsa, es baja, con 31 puntos promedio. El año pasado solo se midieron las Inversiones entrantes en IA, es decir, en todo el flujo de capital destinado a empresas o proyectos relacionados con la IA. Este año, se sumaron cuatro subindicadores que buscan analizar la escalabilidad de las iniciativas y el compromiso de cada país con la investigación y desarrollo de productos y servicios comerciales derivados de la IA.

Por otro lado, en relación con la gobernanza, que es la dimensión que comprende a los elementos y variables que generan

certezas frente a una tecnología revolucionaria, como lo es la IA, se examinaron aspectos como las hojas de ruta que trazan los países en IA, los mecanismos y aproximaciones que establecen límites (protección de datos, ciberseguridad), los marcos regulatorios, los estándares comunes frente a sesgos y discusión regulatoria a nivel regional. Desde esta perspectiva destacan Brasil, pero sobre todo Chile por la alta importancia que ese país le ha dado a la creación e implementación de la Política Nacional de IA y las metas que se trazó para cumplir desde el año en que fue lanzada (2021).

Sobre las estrategias de IA, se observa que hay muchos países que ni siquiera cuentan con una herramienta de esta naturaleza. Hasta noviembre del 2024, sólo 7 países contaban con un instrumento de este tipo: Colombia, Brasil, Chile, Perú, Argentina, Uruguay, y República Dominicana. Un promedio bajo en comparación con el contexto global, donde cerca del 30% de países cuenta con dichos instrumentos. El ritmo del avance tecnológico y las fuertes implicancias del rezago en la promoción, implementación y regulación de esta tecnología, urgen a que más países de la región adopten políticas integrales en esta materia para ir más allá de la discusión para hacer un diagnóstico.

En este ILIA 2024 se amplía el ámbito de medición de la regulación en relación con la primera versión del ILIA, incorporando una revisión de los proyectos de ley en discusión y vigentes en cada uno de los 19 países, junto con un análisis crítico de la aproximación regulatoria que se evidencia tras este proceso. Considerando la naturaleza de la tecnología, en esta edición del índice, se profundiza en el análisis de la regulación en materia de ciberseguridad y ética.

Se incluyó un nuevo indicador de Ética y Sustentabilidad, que integra la información del Índice Global de IA Responsable (GIRAI) y del Índice de Preparación de Redes (NRI) para caracterizar la situación de los países tanto en el ámbito de justicia ambiental y social, como de la ética y respeto de los derechos humanos. A partir de esto, se identificó que no existe una posición en la regulación de América Latina y que las propuestas normativas no tienen un hilo común, sino mucha creatividad. Hay varias inspiradas en la perspectiva de riesgos de la UE y otras que no, y que apuntan a ciertos elementos como regular uso de fake news, o estafas telefónicas. Por tanto, las aproximaciones regulatorias son desiguales y carecen de coordinación regional, lo que genera una fragmentación en los lineamientos que se están haciendo.

La diversidad de aproximaciones es un elemento que añade una dificultad adicional a la participación en los espacios señalados. Hay oportunidad de generar una unificación en una posición regional y que nos permita participar unificadamente en espacios de discusión a nivel global. La Vinculación Internacional se construye a partir de la participación de los países en instancias relevantes para el establecimiento de la gobernanza global y de la incidencia que éstos logran en esos espacios.

Conclusiones

El Índice Latinoamericano de Inteligencia Artificial (ILIA) se ha consolidado como una herramienta estratégica para comprender y evaluar el avance de la inteligencia artificial en América Latina y el Caribe. Su enfoque integral permite identificar no solo los progresos, sino también las brechas que enfrenta la región en términos de infraestructura, talento humano, adopción tecnológica y gobernanza.

El ILIA 2024 evidencia que América Latina y el Caribe se encuentran en un proceso de maduración en sus ecosistemas de IA. Aunque algunos países avanzan con solidez, otros requieren inversiones estructurales en infraestructura, formación y gobernanza. Hay grandes desafíos estructurales, como el acceso desigual a conectividad de calidad, la limitada capacidad de cómputo, la escasa formación avanzada en IA y una persistente brecha de género en la producción científica. Asimismo, el estudio revela una gobernanza fragmentada en la región, evidenciada por la ausencia de estrategias nacionales en muchos países y por enfoques regulatorios diversos y poco coordinados.

A pesar de esto, el índice no solo permite mapear estas realidades, sino que también habilita la cooperación, la comparación y el aprendizaje mutuo. Ante el avance acelerado de la IA a nivel global, es urgente fortalecer capacidades locales, cerrar brechas y construir una posición regional unificada para incidir en los marcos globales de gobernanza tecnológica.

Finalmente, el ILIA pone en evidencia que la inteligencia artificial no es solo un desafío técnico, sino una oportunidad histórica para que América Latina defina una agenda propia en el ámbito digital. Esto implica asumir un compromiso político y estratégico que articule inversiones, fortalezca capacidades locales y posicione a la región como un actor activo en los debates globales sobre el futuro de la tecnología y sus impactos en la vida social.

2 Inteligencia artificial para ecosistemas digitales

Johan Montero Granados

Johan Montero Granados. CTO Government, Technical & Sales Support para Latam Norte y Caribe, Ericsson.

Ingeniero Eléctrico con una maestría en Gestión de Proyectos y más de 20 años de experiencia en el sector de Telecomunicaciones. Actualmente, ocupa el cargo de Director de Tecnología (CTO) para Cuentas Gubernamentales en Ericsson, donde gestiona las operaciones técnicas relacionadas con Redes de Acceso Radioeléctrico en América Latina Norte y el Caribe. En su rol, colabora con los principales actores del sector en la región para promover el portafolio y la visión estratégica de Redes de Acceso Radioeléctrico de Ericsson. Sus esfuerzos están centrados en aprovechar el potencial transformador de las redes inalámbricas y la tecnología 5G, fomentando oportunidades de desarrollo para operadores, gobiernos, industrias y la sociedad en general.

Su trayectoria en Ericsson abarca más de 19 años, durante los cuales ha ocupado diversos cargos, principalmente en Operaciones y Entrega de Proyectos en América Latina. Esta amplia experiencia le ha dotado de un profundo conocimiento de la industria de telecomunicaciones y de sus desafíos en constante evolución.

31

Johan Montero Granados

Resumen

La inteligencia artificial (IA) comprende un conjunto de tecnologías capaces de realizar tareas sin la intervención humana, como el reconocimiento de la voz, el procesamiento del lenguaje natural y la toma de decisiones. Bajo esta perspectiva, el presente documento ahonda en la noción de la IA, destacando sus características y los conceptos clave asociados con el fin de analizar los marcos de aplicación prácticos de la IA en sectores como la agricultura y las telecomunicaciones. El artículo finaliza subrayando la importancia de la colaboración entre las empresas, los reguladores y los gobiernos para garantizar el uso ético, seguro y eficiente de la IA.

Palabras clave: inteligencia artificial, automatización, telecomunicaciones, redes móviles, marco de aplicación práctico.

Introducción

Los avances tecnológicos asociados a la IA la han convertido en una de las tecnologías más revolucionarias del siglo XXI, por su capacidad para desempeñar tareas tradicionalmente humanas y las potencialidades para transformar distintas actividades socio-productivas. Si bien la IA no es un concepto reciente, esta ha evolucionado a lo largo de más de siete décadas hasta haberse popularizado en los últimos años gracias a la publicación de herramientas como Chat GPT. Eso ha propiciado un avance acelerado en la adopción de este tipo tecnologías en nuestras vidas cotidianas, lo que muestra el impacto que están teniendo las mejoras en las capacidades cognitivas y funcionales de las máquinas.

Uno de los pilares fundamentales de la IA moderna es el aprendizaje automático (machine learning), un enfoque que permite que los sistemas actúen de forma autónoma mediante el análisis de datos. Junto con la Inteligencia artificial generativa y los procesos de automatización, la IA ha impulsado el desarrollo de aplicaciones para distintos sectores de la economía, lo que aporta beneficios en términos operativos y contribuye a la reducción de costos y a mejorar la calidad de los servicios ofrecidos.

No obstante, la integración de IA en distintos sectores plantea desafíos significativos en relación con el uso de datos, la evaluación costo-beneficio y la ética tecnológica. Por tal motivo, la implementación efectiva de estas soluciones requiere un análisis cuidadoso del contexto, los objetivos y los recursos disponibles. Además, requiere de marcos regulatorios claros que acompañen su desarrollo y garanticen la protección de la privacidad, la seguridad y los derechos de las personas usuarias.

¿Qué es la inteligencia artificial?

La inteligencia artificial (IA) comprende al conjunto de sistemas y programas que son capaces de realizar tareas que nor-

malmente tienen una intervención o requieren de inteligencia humana. A través de la generación de algoritmos, la consolidación de datos y de muchas otras estrategias que están inmersas en este tipo de herramientas (por ejemplo, Gemini o Chat-GPT), se pueden desarrollar una amplia variedad de tareas que incluyen: el reconocimiento de voz, el procesamiento del lenguaje natural, la toma de decisiones, el aprendizaje automático y la percepción visual, entre otras.

Con base a lo anterior, la IA también puede ser considerada como una rama de la informática que se ocupa de los sistemas que pueden razonar, aprender y actuar de forma autónoma. Es importante aclarar que el concepto de IA no es nuevo, sin embargo, es durante los últimos tres años que adquirió mayor auge y visibilidad gracias a la publicación de Chat-GPT. En realidad, la IA se viene estudiando desde hace más de 70 años y aplicaciones como Alexa y Siri (se fundamentan en IA) han sido integradas con total naturalidad en nuestra cotidianeidad.

Dentro del campo de la IA, uno de los conceptos fundamentales es el **Machine Learning** (o aprendizaje automático), un enfoque que permite a las máquinas aprender a partir de datos sin intervención humana directa, lo que marca un punto de inflexión en la forma en que se desarrollan las comunicaciones y el procesamiento de información.

Un desarrollo más reciente es la llamada **inteligencia artificial generativa** (IA generativa), la cual se caracteriza por el uso de algoritmos que se retroalimentan de sus propias respuestas. De este modo, generan contenido o soluciones cada vez más integrales y precisas, en respuesta a las consultas realizadas. Relacionados con esta capacidad se encuentran los **modelos de fundación** (denominados *Foundation Large Language Models*), los cuales manejan grandes cantidades de datos para alimentarse, procesar la información y aprender de ellos mismos y generar respuestas adicionales, que estén mejoradas o sean complementarias a las que ya se están recibiendo por dichos sistemas. Estos modelos constituyen la base sobre la cual se construyen muchas de las herramientas actuales de IA.

Por último, un componente esencial en el desarrollo de la inteligencia artificial es la **automatización**. Este proceso elimina la intervención manual en tareas específicas mediante algoritmos o sistemas para que los procesos se ejecuten de manera autónoma. Esto no solo optimiza los sistemas, sino que también actúa como un elemento transversal dentro del ecosistema de la IA, al estar presente en prácticamente todos los modelos y algoritmos existentes.

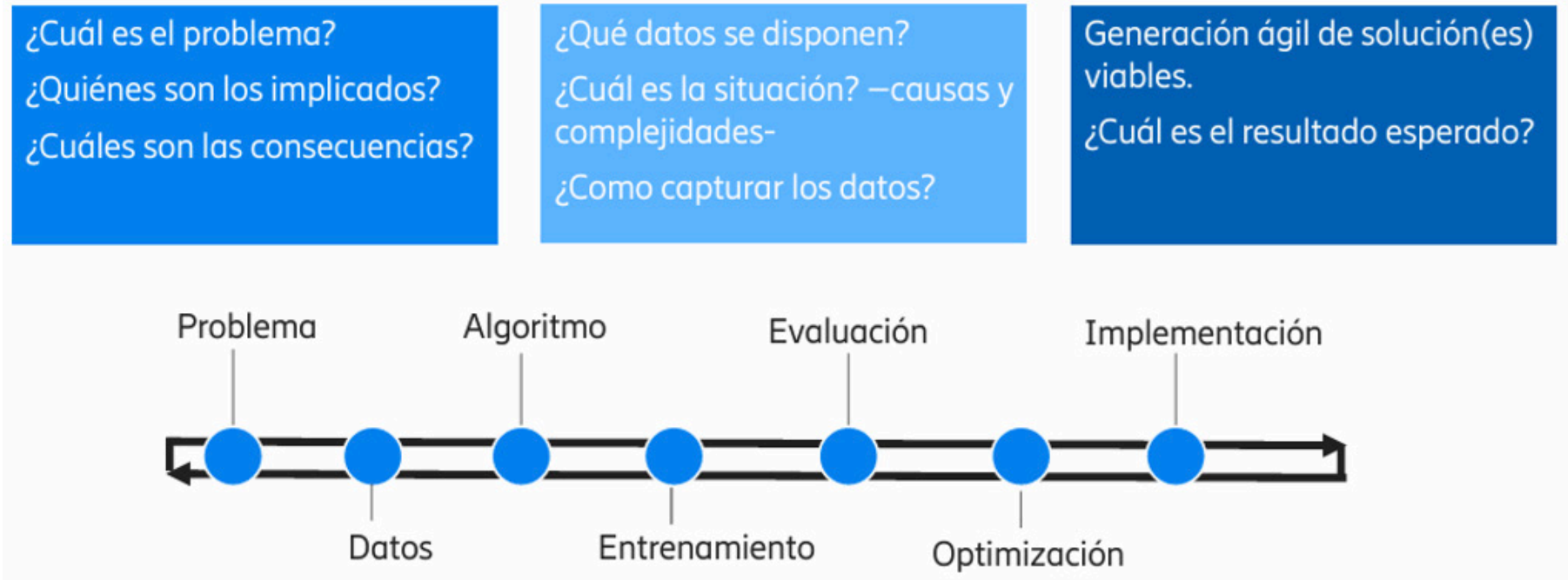
Marco de aplicación práctico de la IA

Las distinciones previas ayudan a comprender cuál es el marco de aplicación práctica de la IA. Y es que conforme avance la IA, es muy posible que esta tecnología pueda ser integrada a todo. Conforme nuestra capacidad de procesamiento de

datos se incremente, es de esperar que los sistemas sean más rápidos y veloces a la hora de almacenar datos, en responder y procesar dichos datos, pues más aplicaciones serán desarrolladas con IA. Sin embargo, impulsar este tipo de transformaciones también implica un costo-beneficio.

Cuando se desea integrar la IA a lo interno de una organización se debe evaluar el problema o situación que se pretende solventar. Además de examinar cuáles son las implicaciones y consecuencias de incluir la IA como un medio para solucionar la situación, pues si el problema es pequeño podría ser que no valga la pena realizar un gran desarrollo, sobre todo si se considera el factor costo-eficiencia. En estos escenarios resulta más útil emplear un modelo/método más tradicional que recurrir a uno en donde se requiere aplicar modelos complejos que demandan mayor inversión.

Figura 2.1. Marco de aplicación práctico de IA

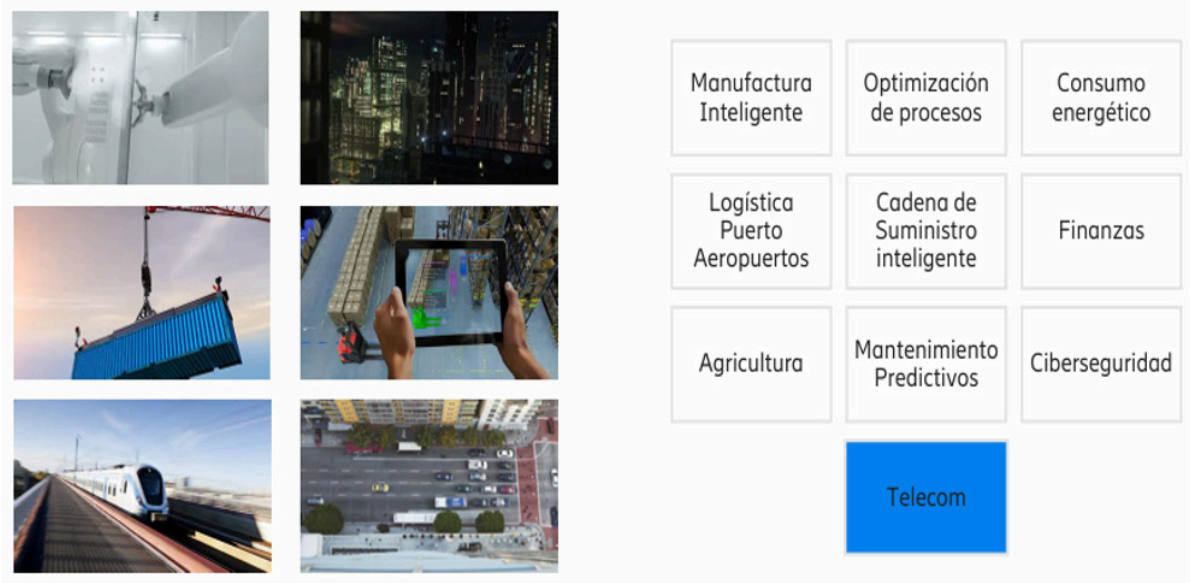


Fuente: Ericsson, 2024.

Otra cuestión importante tiene que ver con los datos que se tienen disponibles, el método de recolección y la forma como estos retroalimentan las herramientas. A pesar de que hasta cierto punto estos modelos son cognitivos (es decir, que aprenden de ellos mismos, como en el caso de la IA generativa) si los datos de entrenamiento son errados, los resultados generados posiblemente serán erróneos.

Tener claridad sobre estos aspectos nos permite entender cuáles aplicaciones de IA podrán resultar de mayor utilidad para la organización. Esto va a ser un aspecto que varíe en función del tipo de actividad que desempeñe, además, que considerará la variedad de herramientas y campos de aplicación que tiene la IA (por ejemplo, para la manufactura inteligente, la optimización de procesos, el consumo energético, la logística, la cadena de suministros y herramientas para la agricultura de precisión, entre otros).

Figura 2.2. Aplicaciones de IA en Industria y Producción



Fuente: Ericsson, 2024.

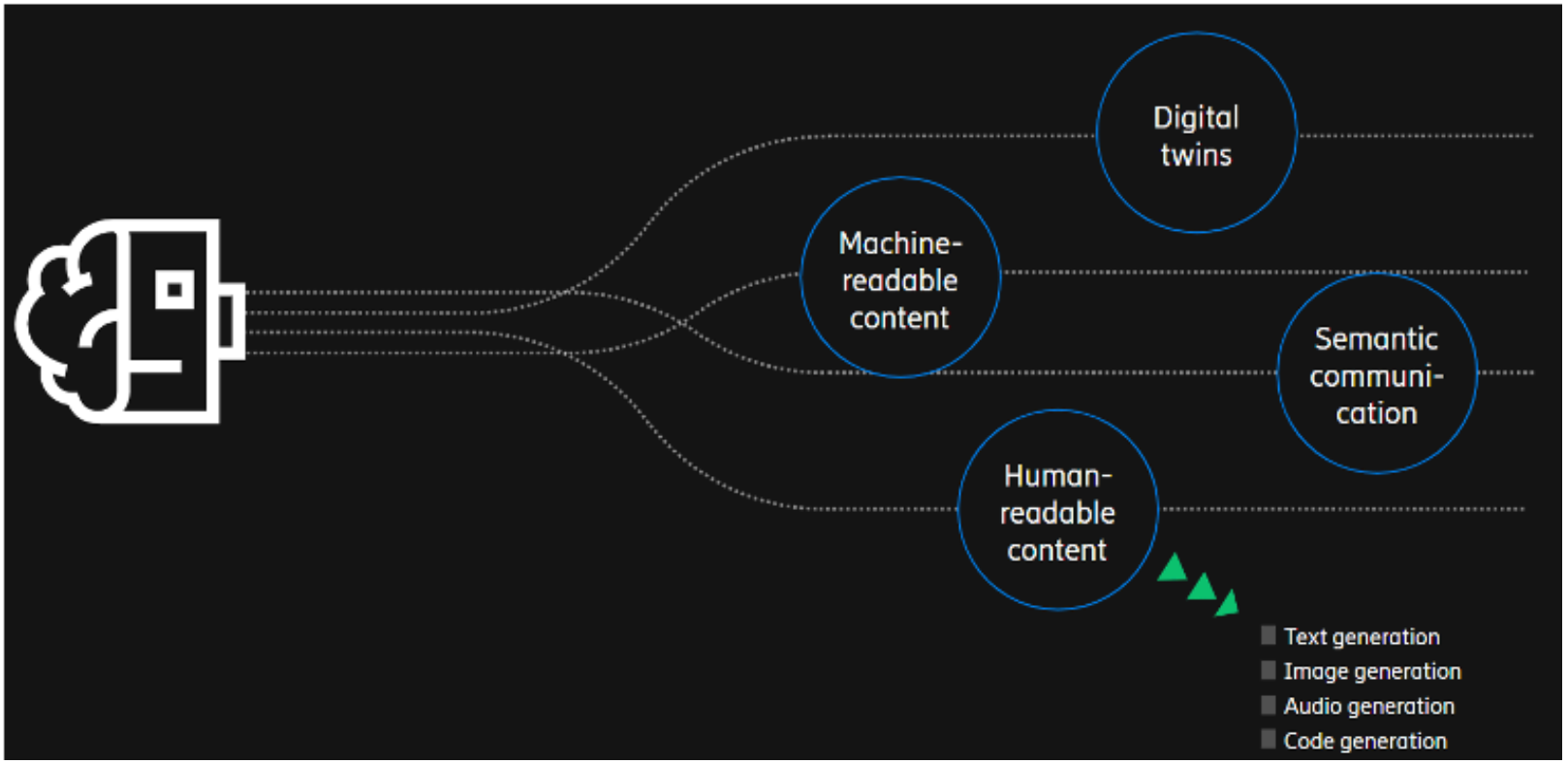
Actualmente, hay más casos de uso que los ilustrados en la figura 2.2. y de estos quizás las aplicaciones para el campo de la agricultura son algunas de las más importantes para el mercado costarricense. En esta área la implementación de sistemas de monitoreo para la aplicación de fertilizantes y la distribución de recursos hídricos mediante sensores, la comunicación máquina a máquina y la automatización de drones puede facilitar la labor agrícola y de regadío, sin intervención humana. Con ello se puede optimizar el uso de recursos y programar estas actividades para cumplir objetivos específicos; además, se pueden establecer indicadores para medir la eficiencia energética y el consumo de recursos.

Otra de las grandes áreas de aplicación de la IA, es la de las telecomunicaciones. En este campo, existe un modelo aplicativo hacia las redes de telecomunicaciones que está integrado por 3 pilares principales: 1) los servicios de comunicación, 2) redes empresariales o redes privadas y 3) las medidas para asegurar la prestación de servicios diferenciados y de mejor calidad. Estos pilares pueden ser fortalecidos mediante la implementación de modelos de automatización y aplicativos de IA.

Con los avances que traerán las próximas generaciones móviles, es muy posible que surjan nuevos casos de uso, se elevan las expectativas de calidad de los servicios por parte de las personas usuarias y que se complejicen los aplicativos. Por eso, las redes de telecomunicaciones tendrán que ser más inteligentes para manejar la creciente complejidad, lo que implica cambiar modelos tradicionales y empezar a introducir conceptos y tecnologías nuevas para la gestión, prevención y control de las redes de telefonía móvil.

Ante este contexto, los operadores de telecomunicaciones podrán utilizar la IA como un soporte esencial para desarrollar aplicaciones que contribuyan en la eficiencia, la redundancia y aseguren el correcto manejo de los casos de uso que tenemos. Asimismo, la IA puede ser usada para regular el consumo energético mediante algoritmos que ayuden a predecir el tráfico de datos generado por las personas usuarias, con el fin que de los sistemas se apaguen o enciendan en función de las demandas que se identifiquen en los sistemas. Además de esta optimización energética, la IA también se puede emplear para monitorear posibles fallas y amenazas, así como para emitir alertar cuando se considere necesario.

Figura 2.3. IA generativa: entregando valor en las telecomunicaciones - Las cuatro vías principales a través de las cuales la IA generativa puede aportar valor en las telecomunicaciones



Fuente: Tomado de https://www.ericsson.com/en/blog/2023/11/how-generative-ai-is-transforming-telecom

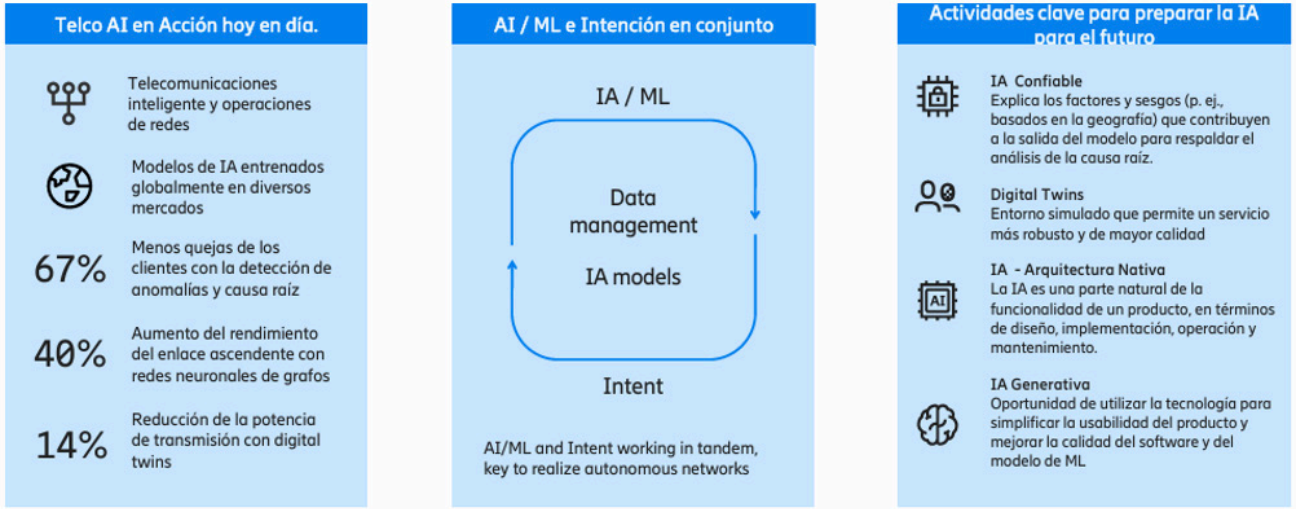
La IA puede optimizar los recursos de acceso como el espectro radioeléctrico al habilitar aplicaciones que ayudan en el control de interferencias, incrementan la potencia y permiten propagar la señal en diferentes áreas, lo que mejora la continuidad y la calidad de los servicios de telecomunicaciones. Gracias a esto se mejora la percepción de servicio por parte del cliente, lo que puede incentivar a los operadores a desarrollar de nuevos casos de uso de la IA. A su vez, las mejoras impulsadas por la IA en las redes de telecomunicaciones, ha-

cen que servicios como la cirugía y la atención médica por vía remota sean una realidad.

Otra de las áreas de aplicación de la IA, son los gemelos digitales. Esta tecnología está siendo muy utilizada por los operadores europeos por cuestiones regulatorias y requisitos preparatorios para la prestación de servicios de salud. Esto permite que el sistema realice simulaciones para que el servicio comercial se vea menos interrumpido, evitando afectaciones para los clientes; lo que muestra que muchos de los aplicativos de IA destinados al sector de las telecomunicaciones también pueden estar enfocados en transformar la experiencia de las y los usuarios finales.

También existe el enfoque de las **Redes Impulsadas por Intención** (*Intent Driven Autonomous Network* en inglés) que consiste en redes que se programan dentro de los sistemas de gestión de las redes, para lo cual se definen variables de operación que el sistema interpreta y empieza a regular sin la intervención manual de personas. A partir de esto las redes pueden cumplir con objetivos específicos, lo que hace que se consideren las redes como programables. Este tipo de programación se puede realizar en entornos de alto tráfico (como los venues deportivos) en los que se establecen umbrales de ocupación y el sistema se regula de tal manera que empieza a tomar decisiones de balanceo de tráfico para asegurar que el sitio tenga el mejor servicio posible.

Tabla 2.4. Experiencia de la IA en Redes impulsadas por "intención"



Fuente: Tomado de https://www.ericsson.com/en/ai/ai-in-networks

Catalizar los beneficios que ofrece la IA en el sector de las telecomunicaciones, requiere del trabajo conjunto con los reguladores y los gobiernos. Esto es clave para desarrollar aplicaciones de IA que sean seguras, éticas y confiables, en las que se protejan los datos (por ejemplo, de señalización, movilización, preferencias y los principales portales consultados de la persona usuaria) que circulan a través de las redes de telefonía móvil. La inteligencia artificial debe estar alineada con las políticas y regulaciones establecidas en el país, de modo que se tenga claridad sobre lo que la IA puede realizar, de dónde procede y cuáles son sus consecuencias. Además, la rápida evolución de esta tecnología nos obliga a tener que mantenernos actualizados con las tendencias de los grandes desarrolladores.

¿Y qué ha realizado Ericsson[1] en el tema?

Actualmente, la compañía cuenta con más de 500 proyectos a nivel mundial desarrollados en IA y más de 34 mil empleados capacitados en IA quienes se encargan de desarrollar aplicativos enfocados en la estrategia y operaciones de los operadores del sector de telecomunicaciones. Aunado a ello, la empresa cuenta con portal público (telecom.ai) en el que se encuentran casos de uso puntuales y abiertos, además, de White Papers sobre la evolución de la IA en el campo de las telecomunicaciones y otros sectores productivos.

La compañía pública dos veces al año el "Erisson Mobility Report", el cual durante más de una década, ha compartido las últimas tendencias y datos de la industria de las telecomunicaciones, proyecciones futuras para 5G, el tráfico de datos móviles, las suscripciones móviles, la cobertura de la población, FWA, IoT y más[2].

Conclusiones

La evolución progresiva de la inteligencia artificial ha logrado un nivel de madurez tecnológica que está potenciado su integración transversal en múltiples sectores productivos, lo que abre oportunidades para impulsar el desarrollo de soluciones innovadoras para solventar problemas complejos. En este contexto, tanto el aprendizaje automático como la IA generativa se han convertido en componentes clave del ecosistema digital actual, al permitir la operación autónoma de los sistemas de IA a partir de datos y ofrecer aplicaciones que pueden ser aprovechados por áreas tan diversas como la agricultura, la manufactura y las telecomunicaciones.

A pesar de los beneficios que la IA puede aportar, resulta indispensable que la implementación de esta tecnología con-

1 Ericsson es una empresa del sector de las telecomunicaciones, pionera en el mercado de despliegue de redes de quinta generación. Se estima que más del 50% del tráfico que hoy circula en el mundo es realizado mediante redes que han sido desplegadas por Ericsson. La empresa tiene presencia en más de 120 países en los que se han instaurado redes.

2 Para ahondar más en estos reportes se recomienda consultar: https://www.ericsson.com/en/reports-and-papers/mobility-report

sidere el contexto y objetivos que se desean alcanzar ya que no siempre es rentable o eficiente la aplicación de modelos complejos para resolver problemas menores. Por eso es necesario que se evalúen aspectos como el costo-beneficio, la disponibilidad y calidad de datos, así como el acceso a recursos técnicos, entre otras cuestiones.

Por otro lado, lo expuesto en el artículo evidencia que el sector de las telecomunicaciones es uno de los que mayores beneficios ha obtenido con la integración de la IA en sus operaciones. Esto sobre todo por el uso de algoritmos para optimizar las redes, gestionar el tráfico de datos, prevenir fallos y mejorar la calidad de los servicios. El éxito en la adopción de distintos aplicativos de IA, abre nuevas posibilidades para automatizar y personalizar la gestión de las infraestructuras de telecomunicaciones mediante tecnologías como los gemelos digitales y las redes impulsadas por intención.

3 Del concepto a la revolución: un viaje histórico

Henry Lizano Mora

Henry Lizano Mora. Investigador del Centro de Investigación Observatorio del Desarrollo (CIOdD) y Jefe de Tecnologías de Información de la Oficina de Servicios Generales (OSG) de la Universidad de Costa Rica (UCR).

Máster en Sistemas de Información (ITCR) y Candidato a Doctor en Administración del (ITCR). Se ha desempeñado como director del Centro de Informática y profesor de Ingeniería Industrial y Administración Pública desde 2006, todos roles en la Universidad de Costa Rica (UCR). La experiencia incluye la dirección de numerosos cursos de grado, postgrado y colaboraciones con organizaciones privadas y públicas. Ha participado en proyectos de investigación en Gestión de Procesos de Negocio, Transformación Digital, Automatización Robótica de Procesos, Inteligencia Empresarial e Innovación; así como, la participación en diversas conferencias y programas sobre Gestión de Procesos de Negocio tanto a nivel nacional como internacional.

Henry Lizano Mora

Resumen

La siguiente ponencia ofrece una aproximación histórico-conceptual de la inteligencia artificial (IA), entendiéndola como un fenómeno con raíces filosóficas ancladas en la antigüedad y cuya evolución más reciente se encuentra en los modelos generativos de última generación. Desde esta perspectiva, el texto examina los fundamentos teóricos que antecedieron al nacimiento de la IA como una disciplina formal, destacando hitos específicos que dan cuenta de los principales periodos de crecimiento y estancamiento de este campo del conocimiento.

El análisis concluye identificando algunos de los desafíos a los que se enfrentan las sociedades ante la creciente integración de estas tecnologías, por lo que se considera esencial que el desarrollo tecnológico avance con un enfoque centrado en las personas, procurando el bienestar colectivo y el desarrollo sostenible de la IA.

Palabras clave: Inteligencia artificial, historia de la tecnología, filosofía, ética tecnológica, modelos generativos.

Introducción

Hoy la inteligencia artificial (IA) ha dejado de ser vista como un campo técnico y se ha constituido en una preocupación transversal ya que al ser una de las tecnologías más disruptivas de nuestro tiempo, afecta a todos los ámbitos de la vida humana con repercusiones que se expresan a nivel social, económico, político y cultural. No obstante, el desarrollo de esta tecnología no ha ocurrido de manera espontánea ni puede ser considerado como un fenómeno exclusivamente contemporáneo pues tiene su origen en preguntas filosóficas milenarias con respecto a la mente, la razón y la inteligencia (Copeland & Proudfoot, 2007; Kakas, 2025; MacLennan, 2009).

Esto resalta la importancia de analizar el desarrollo de la IA desde una perspectiva histórica ya que ello permite entender la evolución del concepto hasta nuestros días, destacando su base filosófica y reflexionando sobre las implicaciones actuales y futuras que podrá tener esta tecnología en las próximas décadas. Una aproximación como esta ayuda a dimensionar las capacidades actuales de la IA e identificar dilemas sociales, éticos e inclusive epistemológicos(Chang, 2020)(Huang & Smith, 2006)(Kubassova et al., 2021a) y (Luger, 2022).

Por lo tanto, esta ponencia ahonda en los debates filosóficos de la Antigüedad hasta llegar a los avances más recientes de la IA con los modelos generativos como los GPT (Vaswani et al., 2017). Para ello, se analizan los hitos clave en el desarrollo de la IA y se destacan: el papel de la lógica formal y el empirismo moderno, los fundamentos matemáticos del siglo XX, los inviernos de la IA y su resurgimiento con el Big Data y el aprendizaje profundo. Aunado a esto, se examinan los esfuerzos que Latinoamérica ha realizado en materia de IA en el ámbito regional y académico, lo que evidencia que la IA puede ser una

construcción diversa y dialógica en contextos distintos al del norte global.

Asimismo, más que presentar un abordaje técnico, el texto busca abordar preguntas relacionadas con las implicaciones sociales de la inteligencia artificial. De ese modo, se plantean interrogantes como los siguientes: ¿cómo deben regularse estas tecnologías?, ¿qué tipo de sociedades estamos construyendo al integrar la IA?, ¿de qué forma garantizaremos que estas tecnologías sean para el bienestar humano? Volver a reflexionar sobre el origen filosófico-histórico de la IA, debe ser la ruta para pensar de forma crítica y responsable el futuro de esta tecnología.

Fundamentos filosóficos de la inteligencia artificial

Cuando se habla de inteligencia artificial es común que se suela hacer referencia a la prueba de Turing (Turing, 1950) y que este momento sea visto como el evento que acabó por acuñar el término; sin embargo, para entender las raíces de la IA, se debe retroceder aún más en la historia y explorar sus orígenes filosóficos. Desde la antigüedad, diversos filósofos se cuestionaron la naturaleza de la mente humana y la inteligencia y como esta última, puede ser entendida y definida. Pensadores como Platón, Sócrates, Aristóteles e incluso Carlo Magno reflexionaron temas y propusieron teorías y teoremas que hasta hoy siguen siendo pilares del pensamiento científico actual (Kakas, 2025).

Los aportes de Aristóteles (384-322 a.C.) fueron trascendentales porque sus ideas sentaron las bases del empirismo moderno y contribuyeron al desarrollo de la lógica formal al crear la noción de los silogismos, un tipo de razonamiento a partir del cual se producen generalizaciones/conclusiones con base a enunciados particulares. Por ejemplo, si se afirma que los humanos son mortales y Sócrates es mortal, se deduce que Sócrates es humano. Los silogismos han sido clave para el desarrollo del pensamiento científico, no sólo porque se siguen utilizando cuando se aplica el método científico, sino también porque se encuentran en los cimientos del pensamiento lógico y de la concepción moderna de la inteligencia (Copeland & Proudfoot, 2007).

Al llegar la edad moderna, destacan figuras como Francis Bacon, que en 1605 introdujo el empirismo científico, René Descartes quién dijo la famosa frase "Pienso, luego existo" (Descartes, 1637) que fue adoptada como una máxima del racionalismo, Groffi Leibniz, cuyo aporte sobre el sistema binario influyó en la computación actual (con los unos y ceros) y George Boole que en el siglo XIX, estableció las bases de la lógica booleana (Boole, 1854), que es esencial para el funcionamiento de los algoritmos que usamos en la actualidad. Todos estos introdujeron conceptos que anticipación a la lógica computacional.

Fundamentos matemáticos y técnicos del siglo XX

Los fundamentos matemáticos de la IA en el siglo XX, se desarrollaron a partir de los aportes de McCulloch y Pitts quienes en 1943 propusieron el primer modelo matemático de una red neuronal (McCulloch & Pitts, 1943). Gracias a esto, nació el concepto de perceptrones o redes neuronales integradas por múltiples capas que permiten calcular probabilidades y procesar información (Rosenblatt, 1958). Otro hito importante fue la publicación del texto "Cibernética" en 1948 por parte de Norbert Wiener y en el que se definieron principios clave para la creación de sistemas automatizados y autorregulados, lo que influyó en la forma como fue evolucionando la IA hasta la actualidad (Wiener, 1948).

De igual modo, no puede olvidarse la contribución que realizó Alan Turing, el cual en 1950 (Turing, 1950) planteó su famoso test para determinar si una máquina puede imitar el comportamiento humano de tal modo que la persona no pueda identificar si está interactuando con una persona y/o una inteligencia artificial. Esto marcó un punto de inflexión en los debates sobre la inteligencia de las máquinas y de hecho, en 1980 estas ideas fueron debatidas por John Seals con el experimento del famoso cuarto chino en el que se evidenció lo complejo que es el pensamiento humano y se cuestionó la capacidad de las máquinas para replicar la manera como el cerebro humano procesa la información y razona (Searle, 1980). A pesar de esto, el campo de la IA fue creciendo de manera paulatina.

Nacimiento formal de la inteligencia artificial

La noción de inteligencia artificial fue usada por primera vez en 1956 durante la Conferencia de Dartmouth (McCarthy et al., 1955), un espacio muy relevante pues significó el inicio formal del estudio académico de la IA. Posteriormente, en 1958 el término volvió a ser referenciado cuando se desarrolló el perceptrón (Rosenblatt, 1958), el cual fue considerado como la primera red neuronal e incentivó el estudio de las redes neuronales multicapas. Algunos años más tarde, el investigador John Wesenbaum creó el primer chatbot llamado *Eliza* en 1964 (Weizenbaum, 1966), un programa que simulaba conversaciones humanas. Su impacto fue tal que originó el efecto Eliza, un fenómeno psicológico que hace que las personas le atribuyan cualidades humanas a sistemas automatizados y crean que están interactuando con un ente pensante.

Entre otras contribuciones importantes al campo de la IA se encuentra el trabajo de Lotfi Zadeh, quien en 1965 introdujo el concepto de lógica difusa (Zadeh, 1965), lo que sirvió para representar razonamientos más flexibles y similares al pensamiento humano. Por otro lado, en 1969, la Universidad de Stanford logró construir el primer robot controlado por computadora, marcando un avance significativo en la integración entre inteligencia artificial y robótica (Scheinman, 1969).

En Costa Rica, el filósofo Claudio Gutiérrez fue pionero al desarrollar los primeros algoritmos de IA en el país. Para ello, usó la computadora Matilda de IBM, la cual había sido integrada en la Universidad de Costa Rica (UCR) en 1968 con distintos propósitos; lo que constituye un hito relevante en la historia tecnológica del país y evidencia como a lo largo del tiempo se han gestado distintos aportes en diversas partes del mundo (Romero, 2024).

Invierno y resurgimiento de la IA

En la década de 1970 sucedió el primer invierno de la inteligencia artificial, un periodo que alude a una etapa de desencanto y estancamiento en el campo. Esto ocurrió por las críticas surgidas en Reino Unido ante el incumplimiento de las expectativas generadas entre 1950 y 1960 (Lighthill, 1973). Las limitaciones tecnológicas y de hardware impidieron alcanzar los avances prometidos, lo que provocó una reducción significativa en los presupuestos para la investigación en IA. A pesar de esto, la comunidad académica de las universidades públicas mantuvo sus esfuerzos de investigación en IA; ya que, desde una perspectiva comercial, no había suficientes incentivos para invertir en el área, pues las iniciativas de I+D suelen generar retornos, pero muy a largo plazo.

La década de 1980 trajo consigo un nuevo impulso con la aparición de los sistemas expertos como el X-con de Vigio Aikidmen (McDermott, 1982), la quinta generación de las computadoras en Japón (Moto-Oka, 1982) y el algoritmo de propagación de retroceso propuesto por Joule-Bering-Hilton (Rumelhart et al., 1986). Si bien esto supuso un progreso significativo para el área, las infladas expectativas de la época llevaron a un nuevo invierno (McCorduck, 2004). No fue hasta la llegada del Big Data en el 2012, luego que un año antes se acuñó el termino por (Manyika et al., 2011), que resurgió el interés por estudiar la IA.

Durante las últimas décadas, la IA ha logrado avances muy relevantes. Por ejemplo, en 1997 la supercomputadora Deep Blue logró derrotar al mundial de ajedrez Garry Kasparov (Campbell et al., 2002); mientras que, en el 2006, Geoffrey Hinton propuso el Deep Learning (aprendizaje profundo) (Hinton et al., 2006) y (LeCun et al., 2015) con lo que se produjo una revolución en la manera como las máquinas aprenden desde grandes volúmenes de datos. Algunos años después, en el 2011 IBM presentó el sistema Watson, el cual ganó el concurso Jeopardy (Ferrucci et al., 2010) y en ese mismo momento, aparecieron las primeras asistentes virtuales (Siri, Alexa y Cortana) creadas para interactuar con dispositivos móviles a través de la voz (Cheyer et al., 2010; Horvitz, 2014; Pradhan & Sarikaya, 2017).

La era de los modelos generativos y la IA contemporánea

Después de la salida comercial de los asistentes virtuales, se generaron diversos avances tecnológicos que permitieron lle-

gar a los actuales modelos generativos de IA. Para comprender este proceso hay que regresar al 2012 cuando se publicó *AlexNet*, un avance que transformó la visión por computadora (Krizhevsky et al., 2012). Este periodo se considera como el inicio de una nueva era industrial, la Industria 4.0 (Kagermann et al., 2013), que junto con el auge del Big Data (Manyika et al., 2011) y la ciencia de datos, contribuyeron a desarrollar la IA moderna (Cleveland, 2001; Dhar, 2013). Dos años después, Google compró la empresa DeepMind, con lo que se le dio un nuevo impulso al campo de la IA y se abrieron nuevas fronteras en la investigación y en el desarrollo de aplicaciones. Las potencialidades promovidas por estos avances permitieron que en el 2015 se consolidaran herramientas como *TensorFlow*, una plataforma de código abierto que permite el desarrollo y entrenamiento de modelos de aprendizaje profundo (Abadi et al., 2016) y más adelante potenciado por PyTorch (Paszke et al., 2019).

En este periodo también surgieron aplicaciones como *AlphaGo* (promovido por DeepMind) que venció a los campeones humanos de Go (Silver et al., 2016), la arquitectura *Transformer* creada por el equipo de Ashish Vaswani en la publicación *Attention Is All You Need* en 2017 (Vaswani et al., 2017)the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Esta última se considera como la base teórica de los sistemas GPT (Generative Pre-trained Transformer), cuya primera versión fue habilitada en el 2018 y tuvo un importante impacto. Cabe señalar que los transformadores generativos son la tecnología sobre la cual operan herramientas modernas como ChatGPT y otras aplicaciones de IA conversacional, es decir, es la base de los Grande Modelos de Lenguaje LLM por sus siglas en inglés (Brown et al., 2020) y es así como llegamos al lanzamiento de ChatGPT por parte de OpenAI (OpenAI, 2022).

Debe aclarase que las inteligencias artificiales generativas no revientan la lógica computacional básica, sino que siguen empleando el proceso clásico de entrada, procesamiento y salida. No obstante, lo que marca la diferencia es la forma como se utiliza dicha lógica para que la IA sea más práctica y accesible para la sociedad.

En esta nueva era de la IA generativa, predominan los modelos GPT, los cuales se caracterizan por el uso de miles de millones de parámetros. Si bien algunos intentan comparar dichos parámetros con las neuronas humanas y/o con los procesos cerebrales como la sinapsis, esto puede llevarnos a errores de interpretación como con el conocido efecto Eliza (Weizenbaum, 1966) o el antropomorfismo (en el que se atribuyen rasgos humanos a los sistemas automatizados) (Epley et al., 2007). Ahora bien, el conjunto de tecnologías que fundamentan a la IA generativa comprende distintas áreas, entre las que pueden mencionarse:

- **Aprendizaje automático (Machine Learning):** es el entrenamiento de modelos para detectar patrones, realizar correlaciones y generar respuestas (Mitchell, 1997).

- **Aprendizaje profundo (Deep Learning):** permite el procesamiento de datos complejos y la extracción de significados más sofisticados (LeCun et al., 2015).

- **Modelos multimodales:** tienen capacidad para reconocer y generar voz, texto, imágenes y otros tipos de información (Radford et al., 2021).

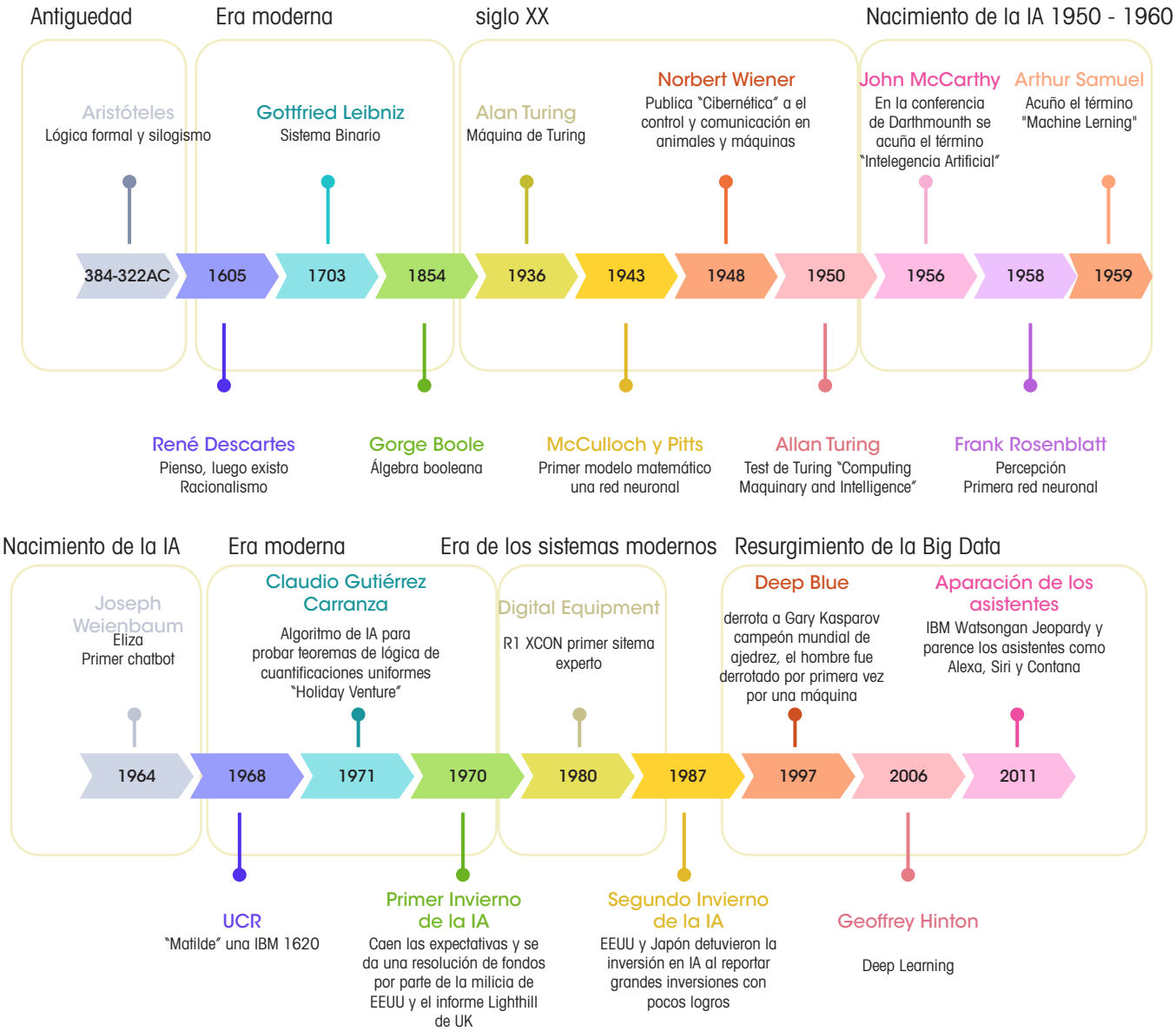
Al analizar cómo funciona internamente una IA generativa, se evidencia que la misma está compuesta por tres módulos principales, según se afirma en distintas publicaciones de los desarrolladores. Entre estos destaca el sistema *ChatGPT*, que incorpora:

- **El Prompt:** Prompt o Instrucción es el texto y/o comando que la persona usuaria introduce como entrada. De este surge la ingeniería de prompts, una técnica que sirve para afinar las instrucciones y obtener mejores resultados (Liu et al., 2023).

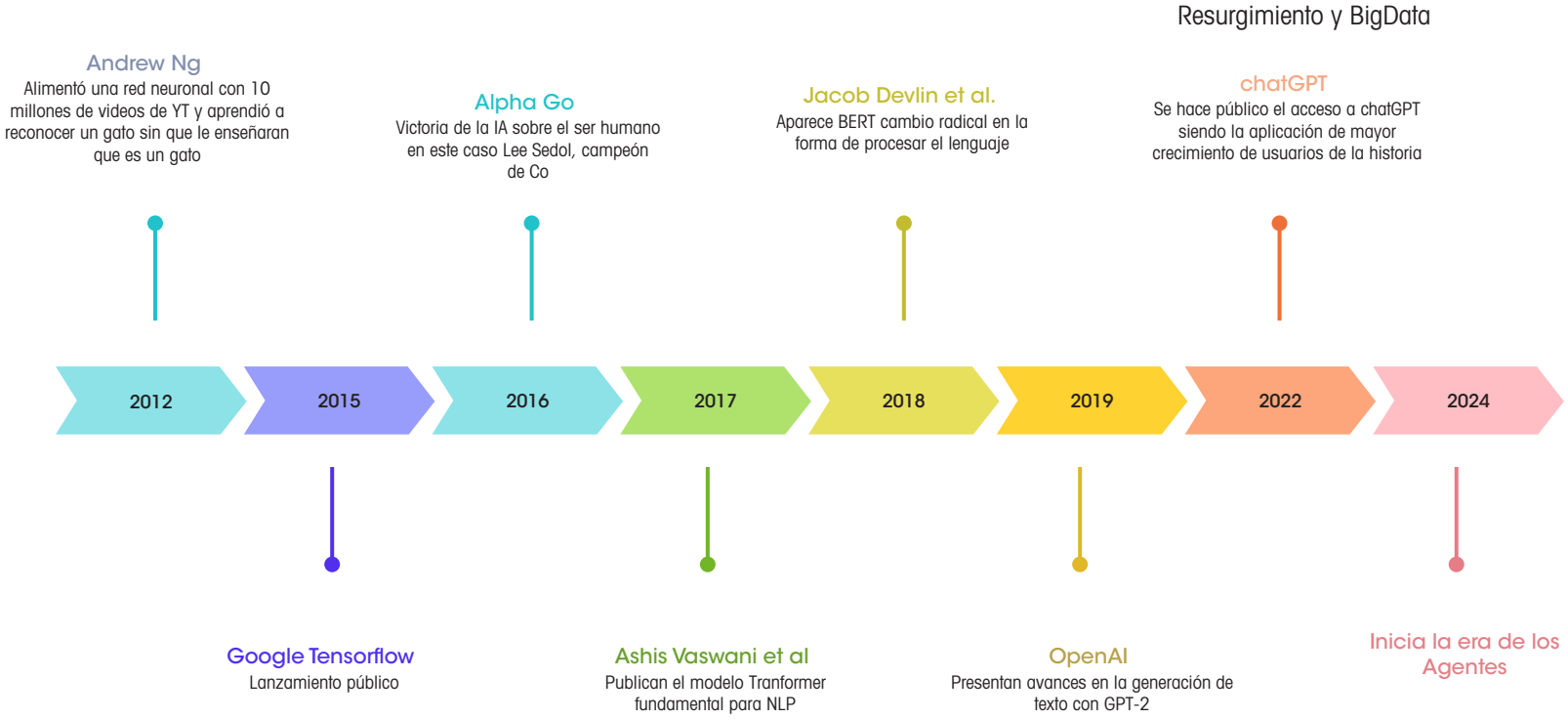
- **La tokenización**: que refiere a la transformación del texto en datos vectoriales para que el modelo pueda interpretarlos (Sennrich et al., 2016).

- **El gran modelo del lenguaje (LLM)**: se encarga de comparar el prompt con el corpus de información. En ciertos casos, se aplica el reforzamiento con retroalimentación humana, una fase adicional en la que se mejora la calidad y coherencia de las respuestas que son producidas (Brown et al., 2020).

Figura 3.1. Historia de la IA



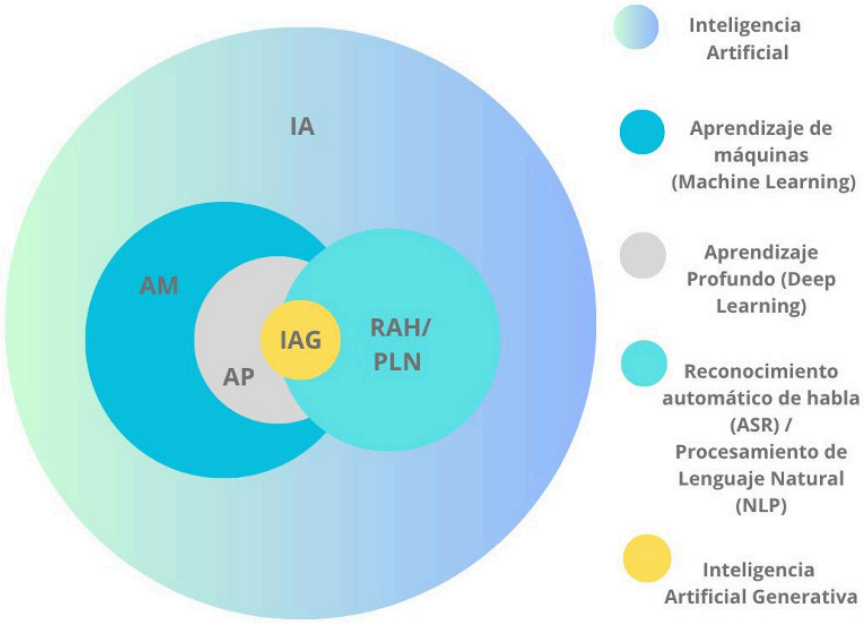
Era de la IA Moderna



Fuente: Elaboración propia, 2024.

Conjuntamente, todos los elementos de este sistema trabajan para producir textos que sean comprensibles, relevantes y contextualmente adecuados con relación a lo que se le solicita. Por tanto, la IA generativa más que replicar la lógica básica, la transforma para convertir a la IA en una herramienta poderosa y socialmente influyente (Kubassova et al., 2021b).

Figura 3.2. IA Generativa



Fuente: Elaboración propia con base a Kubassova et. al, 2021b.

Conclusiones

El recorrido histórico ofrecido a través de esta ponencia evidencia que el desarrollo de la IA es un fenómeno en el que se interconectan la filosofía, la lógica, las matemáticas y la ingeniería. Es por eso que para comprender su evolución se debe asumir una mirada mucho más amplia y reflexiva a través de la que se reconozca que los avances de la IA. Además, trae consigo desafíos que merecen atención. Entre estos se encuentran el desempleo, la automatización de tareas y los dilemas éticos que estas tecnologías plantean.

Por tanto, a futuro deben discutirse conceptos como la *singularidad tecnológica* y la posibilidad de alcanzar una inteligencia artificial general. Aunque estas ideas entusiasman e inducen debates, deben ser abordadas desde una perspectiva crítica que reflexione sobre las implicaciones sociales, filosóficas y humanas de la IA.

Finalmente, esta aproximación conceptual busca resaltar que la IA no es un fenómeno completamente nuevo: sus raíces filosóficas se remontan a pensadores de la antigüedad. Lo que sí es nuevo es su accesibilidad global y su capacidad para transformar vidas. En ese sentido, la tecnología debe enfocarse en generar valor público, en servir a las personas y en promover un desarrollo centrado en lo humano. Imaginemos cómo será nuestra convivencia con máquinas inteligentes en los próximos años y trabajemos para que esa relación sea ética, justa y beneficiosa para todos.

Referencias

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., … Zheng, X. (2016). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. https://www.tensorflow.org/

Boole, G. (1854). *An Investigation of the Laws of Thought, on which are founded the Mathematical Theories of Logic and Probabilities*. Walton and Maberly. https://archive.org/details/investigationofl00booliala

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language Models are Few-Shot Learners. *arXiv preprint arXiv:2005.14165*. https://arxiv.org/abs/2005.14165

Campbell, M., Jr, A. J. H., & Hsu, F. (2002). Deep Blue. *Artificial Intelligence*, *134*(1-2), 57-83. https://doi.org/10.1016/S0004-3702(01)00129-1

Chang, A. (2020). History of Artificial Intelligence. *Intelligence-Based Medicine*. https://doi.org/10.1016/b978-0-12-823337-5.00002-0

Cheyer, A., Kittlaus, D., & Gruber, T. (2010). Siri: A Virtual Personal Assistant. *Proceedings of the AAAI Spring Symposium on Intelligent Agents*. https://www.aaai.org/ocs/index.php/SSS/SSS10/paper/view/1081

Cleveland, W. S. (2001). Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics. *International Statistical Review*, *69*(1), 21-26. https://doi.org/10.1111/j.1751-5823.2001.tb00477.x

Copeland, B., & Proudfoot, D. (2007). *Artificial intelligence: History, foundations, and philosophical issues*. 429-482. https://doi.org/10.1016/B978-044451540-7/50032-3

Descartes, R. (1637). *Discours de la méthode pour bien conduire sa raison, et chercher la vérité dans les sciences*. Jan Maire. https://gallica.bnf.fr/ark:/12148/bpt6k47461r

Dhar, V. (2013). Data Science and Prediction. *Communications of the ACM*, *56*(12), 64-73. https://doi.org/10.1145/2500499

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). *On Seeing Human: A Three-Factor Theory of Anthropomorphism* (Vol. 114). Psychological Review. https://doi.org/10.1037/0033-295X.114.4.864

Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A., Lally, A., Murdock, J. W., Nyberg, E., Prager, J., Schlaefer, N., & Welty, C. (2010). Building Watson: An Overview of the DeepQA Project. *AI Magazine*, *31*(3), 59-79. https://doi.org/10.1609/aimag.v31i3.2303

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*, *18*(7), 1527-1554. https://doi.org/10.1162/neco.2006.18.7.1527

Horvitz, E. (2014). Cortana: A Personal Assistant for Windows Phone. *Proceedings of the AAAI Conference on Artificial Intelligence*. https://www.microsoft.com/en-us/research/publication/cortana-personal-assistant-windows-phone/

Huang, T. W., & Smith, C. (2006). *The History of Artificial Intelligence*.

Kagermann, H., Wahlster, W., & Helbig, J. (2013). *Recommendations for Implementing the Strategic Initiative INDUSTRIE 4.0: Securing the Future of German Manufacturing Industry*. acatech – National Academy of Science and Engineering. https://en.acatech.de/publication/recommendations-for-implementing-the-strategic-initiative-industrie-4-0/

Kakas, A. (2025). Aristotle's Original Idea: For and Against Logic in the era of AI. *ArXiv*, *abs/2503.12161*. https://doi.org/10.48550/arXiv.2503.12161

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS 2012)*, *1*, 1097-1105. https://papers.nips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html

Kubassova, O., Shaikh, F., Melus, C., & Mahler, M. (2021a). History, current status, and future directions of artificial intelligence. En *Precision Medicine and Artificial Intelligence* (pp. 1-38). Elsevier. https://doi.org/10.1016/B978-0-12-820239-5.00002-4

Kubassova, O., Shaikh, F., Melus, C., & Mahler, M. (2021b). History, Current Status, and Future Directions of Artificial Intelligence. En *Precision Medicine and Artificial Intelligence* (pp. 1-38). Elsevier. https://doi.org/10.1016/B978-0-12-820239-5.00002-4

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444. https://doi.org/10.1038/nature14539

Lighthill, J. (1973). *Artificial Intelligence: A General Survey*. Science Research Council. https://doi.org/10.1017/S0269888900007795

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, Prompt, and Predict: A Systematic Sur-

vey of Prompting Methods in Natural Language Processing. *ACM Computing Surveys*, *55*(9), 1-35. https://doi.org/10.1145/3560815

Luger, G. (2022). A Brief History and Foundations for Modern Artificial Intelligence. *Int. J. Semantic Comput.*, *17*, 143-170. https://doi.org/10.1142/s1793351x22500076

MacLennan, B. (2009). *History of Artificial Intelligence Before Computers*. 1763-1768. https://doi.org/10.4018/978-1-60566-026-4.CH277

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity. *McKinsey Global Institute*. https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

McCorduck, P. (2004). *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence* (2.ª ed.). A.K. Peters Ltd.

McCulloch, W. S., & Pitts, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *The Bulletin of Mathematical Biophysics*, *5*(4), 115-133. https://doi.org/10.1007/BF02478259

McDermott, J. P. (1982). R1: A Rule-Based Configurer of Computer Systems. *Artificial Intelligence*, *19*(1), 39-88. https://doi.org/10.1016/0004-3702(82)90021-2

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.

Moto-Oka, T. (1982). Fifth Generation Computer Systems. *Proceedings of the 1982 IFIP Congress*, 1-10. https://archive.org/details/fifthgenerationc0000unse

OpenAI. (2022, noviembre). *Introducing ChatGPT*. https://openai.com/blog/chatgpt

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, 8024-8035. https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html

Pradhan, S., & Sarikaya, R. (2017). Amazon Alexa: A Voice Service for Everyone. *Proceedings of the 1st International Workshop on Conversational Agents*, 1-2. https://doi.org/10.1145/3073565.3073576

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning Transferable Visual Models From Natural Language Supervision. *arXiv preprint arXiv:2103.00020*. https://arxiv.org/abs/2103.00020

Romero, J. Á. R. (2024). Sobre el pensamiento del Dr. Claudio Gutiérrez Carranza. Un algoritmo de Inteligencia Artificial para probar teoremas de lógica de cuantificación uniforme. *Revista de Filosofía de la Universidad de Costa Rica*, *63*(165), 211-227. https://doi.org/10.15517/revfil.2024.58416

Rosenblatt, F. (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review*, *65*(6), 386-408. https://doi.org/10.1037/h0042519

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533-536. https://doi.org/10.1038/323533a0

Scheinman, V. (1969). *Stanford Arm: First Computer-Controlled Robotic Arm*. https://web.stanford.edu/~learnest/ee392o/history.html

Searle, J. R. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences*, *3*(3), 417-424. https://doi.org/10.1017/S0140525X00005756

Sennrich, R., Haddow, B., & Birch, A. (2016). Neural Machine Translation of Rare Words with Subword Units. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1715-1725. https://doi.org/10.18653/v1/P16-1162

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Driessche, G. V. D., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484-489. https://doi.org/10.1038/nature16961

Turing, A. (1950). Computing Machinery and Intelligence (1950). *Ideas That Created the Future*. https://doi.org/10.1093/oso/9780198250791.003.0017

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Atten-*

tion Is All You Need (No. arXiv:1706.03762). arXiv. http://arxiv.org/abs/1706.03762

Weizenbaum, J. (1966). ELIZA—A Computer Program for the Study of Natural Language Communication Between Man and Machine. *Communications of the ACM*, *9*(1), 36-45. https://doi.org/10.1145/365153.365168

Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. MIT Press. https://archive.org/details/cyberneticsorcon00wien

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, *8*(3), 338-353. https://doi.org/10.1016/S0019-9958(65)90241-X

4 Inteligencia artificial para la sostenibilidad de los data centers

Óscar Rojas Badilla

Óscar Rojas Badilla. Director de Negocios, Innovación e IA, Bjumper.

Es un destacado especialista en administración de empresas de tecnología, con una sólida trayectoria en el ámbito de tecnologías emergentes para la gestión de centros de datos y automatización de infraestructuras críticas. Actualmente se desempeña como Director de Estrategia de Inteligencia Artificial y Desarrollo de Negocios en Bjumper. Su experiencia abarca desde la dirección general hasta la expansión de mercados internacionales, respaldada por una pasión constante por la innovación y la implementación de la industria 4.0 dentro del modelo de negocio.

Óscar Rojas Badilla

Resumen

La sostenibilidad en la industria de los data centers enfrenta retos significativos debido al crecimiento exponencial de datos y su impacto ambiental, se estima que para 2030, los Data Centers (centros de datos) podrían representar entre el 3% y el 4% de las emisiones globales de dióxido de carbono (CO_2), debido al aumento en la demanda de procesamiento y almacenamiento de información (Carcar et al., 2024). Este incremento en la generación de datos y su correspondiente impacto ambiental subraya la necesidad de implementar prácticas sostenibles en la gestión de estas infraestructuras críticas.

La inteligencia artificial (IA) se perfila como una solución clave para abordar estas problemáticas, permitiendo optimizar recursos, reducir el consumo energético y mitigar las emisiones de carbono. Este documento explora cómo la IA contribuye a la sostenibilidad en los data centers mediante el análisis de la solución ThinkData de Bjumper, un sistema avanzado de análisis de datos multifuente. Además, se detalla el marco normativo y los estándares internacionales aplicables, como la Directriz de Eficiencia Energética (EED) y la ISO 30134-2, enfatizando su rol en la promoción de prácticas sostenibles en este sector crítico.

Palabras Clave: Inteligencia Artificial, Sostenibilidad, Data Centers, Eficiencia Energética, Normativas Ambientales.

Introducción

En la era digital, el volumen de datos generados globalmente ha crecido de forma exponencial. En 2020, se produjeron 64.2 zettabytes de datos, marcando un récord histórico, impulsado por el aumento del trabajo, la educación y el entretenimiento desde casa durante la pandemia de COVID-19. Se proyecta que esta cifra supere los 181 zettabytes para 2025 y alcance 2,142 zettabytes en 2035 (Statista, 2023). A pesar de este crecimiento, solo un 2% de los datos generados en 2020 fue almacenado para uso futuro, lo que subraya el desafío de mantener la infraestructura de almacenamiento al ritmo del crecimiento y refuerza la necesidad de adoptar soluciones innovadoras y sostenibles en la gestión de datos.

Figura 4.1. Crecimiento exponencial de datos

Fuente: Elaboración propia.

Este crecimiento desmesurado está impulsado por tendencias como la digitalización, el internet de las cosas (IoT) y el uso intensivo de plataformas basadas en la nube. Sin embargo, esta explosión de datos plantea desafíos significativos para los data centers, que son las infraestructuras críticas responsables de almacenar, procesar y gestionar esta información.

Los data centers consumen alrededor del 3% de la electricidad global y generan el 0,5% de las emisiones de carbono (Data Centres & Networks - IEA, s. f.). Gran parte de este consumo energético proviene de fuentes no renovables, lo que agrava la crisis climática. Además, los sistemas de enfriamiento y distribución de energía en los data centers a menudo son ineficientes, lo que incrementa aún más el impacto ambiental. Ante este panorama, la industria se enfrenta a una encrucijada: cómo satisfacer la creciente demanda de datos mientras se minimiza el impacto ambiental.

El desafío de la sostenibilidad en los data centers

El impacto ambiental de los data centers se deriva principalmente de su consumo energético, uso del agua y emisiones de gases de efecto invernadero (GEI). Un rack típico con una potencia instalada de 1.5 kW consume la misma cantidad de energía que una vivienda unifamiliar ocupada por cuatro personas. Este dato ilustra la magnitud del problema cuando se considera que muchos data centers operan con cientos o incluso miles de racks.

Además, los sistemas de climatización y distribución energética son responsables de hasta el 40% del consumo total de un data center (Campos Corporación, 2023). Estas ineficiencias no solo aumentan los costos operativos, sino que también amplifican las emisiones de carbono. Por ejemplo, un data center con 100 racks puede emitir más de 300 toneladas de CO_2 al año, lo que equivale a las emisiones de 110 automóviles de gasolina típicos.

El consumo de agua también es una preocupación significativa. Los sistemas de enfriamiento en los data centers requieren grandes volúmenes de agua, lo que genera un impacto adicional en regiones donde este recurso es limitado. En promedio, un rack de 1.5 kW utiliza alrededor de 158 m^3 de agua al año, una cantidad comparable al consumo anual de una vivienda.

La inteligencia artificial como solución transformadora

La inteligencia artificial (IA) se ha convertido en una herramienta clave para enfrentar los desafíos que afectan la sostenibilidad y la eficiencia en los data centers. Su capacidad para optimizar procesos, automatizar operaciones y realizar análisis predictivos en tiempo real permite a los operadores de data centers no solo reducir el impacto ambiental de estas instalaciones, sino también mejorar su rentabilidad y desempeño operativo.

Optimización de procesos mediante IA

Una de las aplicaciones más relevantes de la IA en los data centers es el monitoreo continuo y en tiempo real de parámetros críticos como la temperatura, la distribución de energía y el uso del agua. Por ejemplo, mediante algoritmos avanzados de aprendizaje automático, la IA puede analizar millones de puntos de datos por segundo, detectar patrones y prever problemas antes de que se conviertan en incidentes críticos. Esto resulta especialmente útil en sistemas de climatización, responsables de hasta el 40% del consumo energético de los data centers (Ministerio Para la Transición Ecológica y el Reto Demográfico, s. f.).

El monitoreo basado en IA permite que los sistemas de enfriamiento se ajusten automáticamente según la carga de trabajo del Data Center y las condiciones ambientales. Esto reduce el consumo de energía y asegura una operación eficiente sin comprometer la integridad de los equipos. Según estudios recientes, implementar IA en la gestión de la climatización puede reducir hasta un 15% el consumo energético general de los data centers (BOE.es - DOUE-L-2024-80715 Reglamento Delegado (UE) 2024/1364 de la Comisión, de 14 de Marzo de 2024, Relativo A la Primera Fase Del Establecimiento de un Régimen de Evaluación Común de la Unión Para Data Centers., s. f.)

Automatización de operaciones

Además de optimizar procesos, la IA automatiza tareas repetitivas y críticas, como la redistribución de cargas de trabajo entre servidores. Esta automatización no solo mejora la eficiencia operativa, sino que también reduce los errores humanos, que son responsables de hasta el 70% de las interrupciones en los data centers según la Uptime Institute. Por ejemplo, Google ha implementado sistemas de IA para optimizar la distribución de energía en sus data centers. Estos sistemas han logrado reducir el PUE (Power Usage Effectiveness) en un 30%, marcando un precedente en el uso de la inteligencia artificial para la sostenibilidad.

Análisis predictivo para la toma de decisiones

Otro aspecto transformador de la IA es su capacidad para realizar análisis predictivos. Al prever demandas futuras de recursos, como energía o capacidad de almacenamiento, los operadores pueden planificar con antelación y evitar cuellos de botella. Esto no solo mejora la experiencia del cliente, sino que también reduce los costos asociados con el sobredimensionamiento o la falta de capacidad. En ese sentido, los modelos predictivos permiten calcular las fluctuaciones en la carga de trabajo y anticipar aumentos en el consumo de energía, ajustando los sistemas para minimizar desperdicios. De acuerdo con Statista, este tipo de análisis puede mejorar

la eficiencia energética en un 20%, alineando las operaciones con metas de sostenibilidad global.

Impacto en la sostenibilidad

La adopción de la IA no solo mejora la eficiencia operativa, sino que también tiene un impacto directo en la sostenibilidad. Al optimizar el uso de recursos como energía y agua, los data centers pueden reducir significativamente sus emisiones de carbono y cumplir con normativas internacionales como la ISO 14064, que establece directrices para la medición y reporte de emisiones de gases de efecto invernadero.

La IA se presenta como una solución transformadora que no solo aborda los desafíos actuales de los data centers, sino que también los posiciona para un futuro más eficiente y sostenible. Las capacidades de optimización, automatización y análisis predictivo permiten a los operadores maximizar el rendimiento, reducir costos y cumplir con sus objetivos de sostenibilidad. Estos avances reafirman el papel de la IA como un catalizador en la transición hacia un entorno digital más respetuoso con el medio ambiente.

Think Data de Bjumper

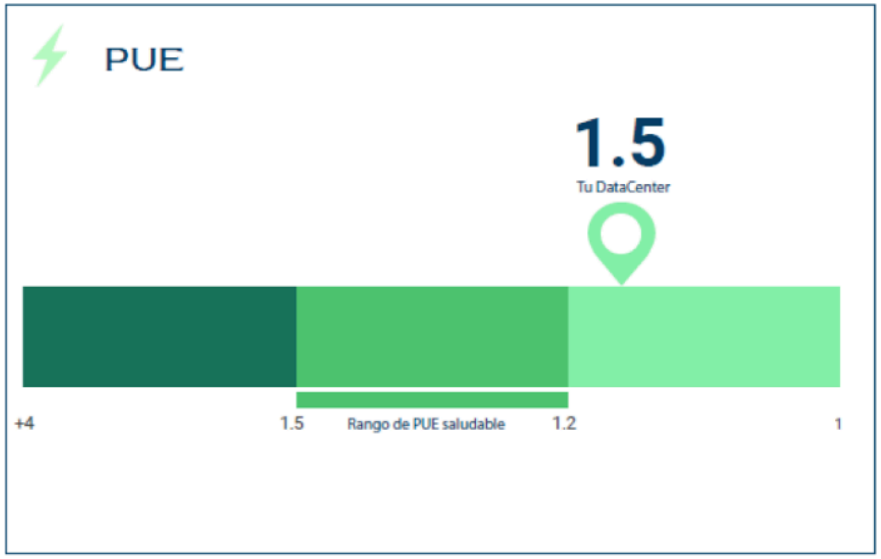
ThinkData, una solución innovadora desarrollada por Bjumper, que utiliza inteligencia artificial (IA) para optimizar operaciones en los data centers y cumplir con estándares internacionales de sostenibilidad.

Optimización mediante indicadores de sostenibilidad

Los indicadores clave utilizados por ThinkData permiten un análisis integral de las operaciones:

- **PUE (Eficiencia del Uso de Energía):** Relación entre la energía total consumida y la destinada a los equipos de TI. Un PUE bajo refleja un uso energético eficiente. Según datos del informe de sostenibilidad, ThinkData ayuda a los data centers a alcanzar valores cercanos a 1.2, considerados óptimos para nuevas instalaciones.

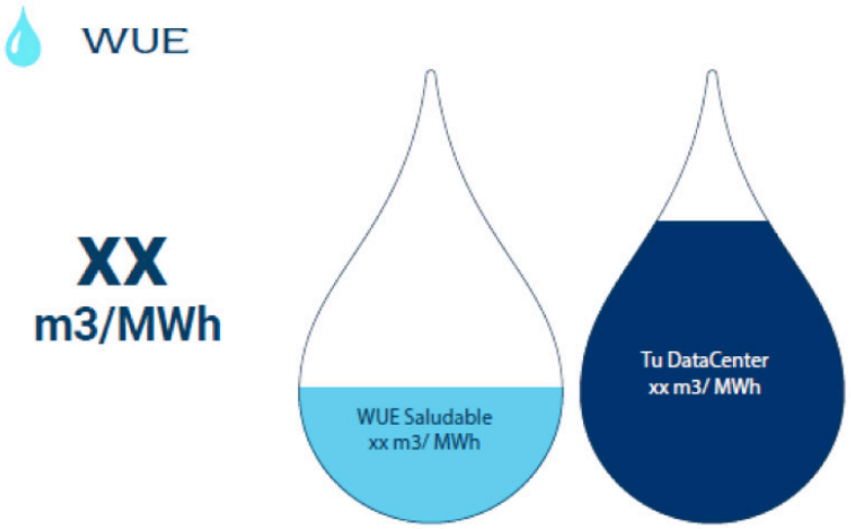
Figura 4.2. Visualización del PUE



Fuente: Elaboración propia.

- WUE (Eficiencia del Uso del Agua): Evalúa el consumo de agua en función de la energía utilizada. ThinkData no solo mide este indicador, sino que también identifica áreas de optimización en los sistemas de enfriamiento, que representan un alto porcentaje del uso de agua en estas infraestructuras.

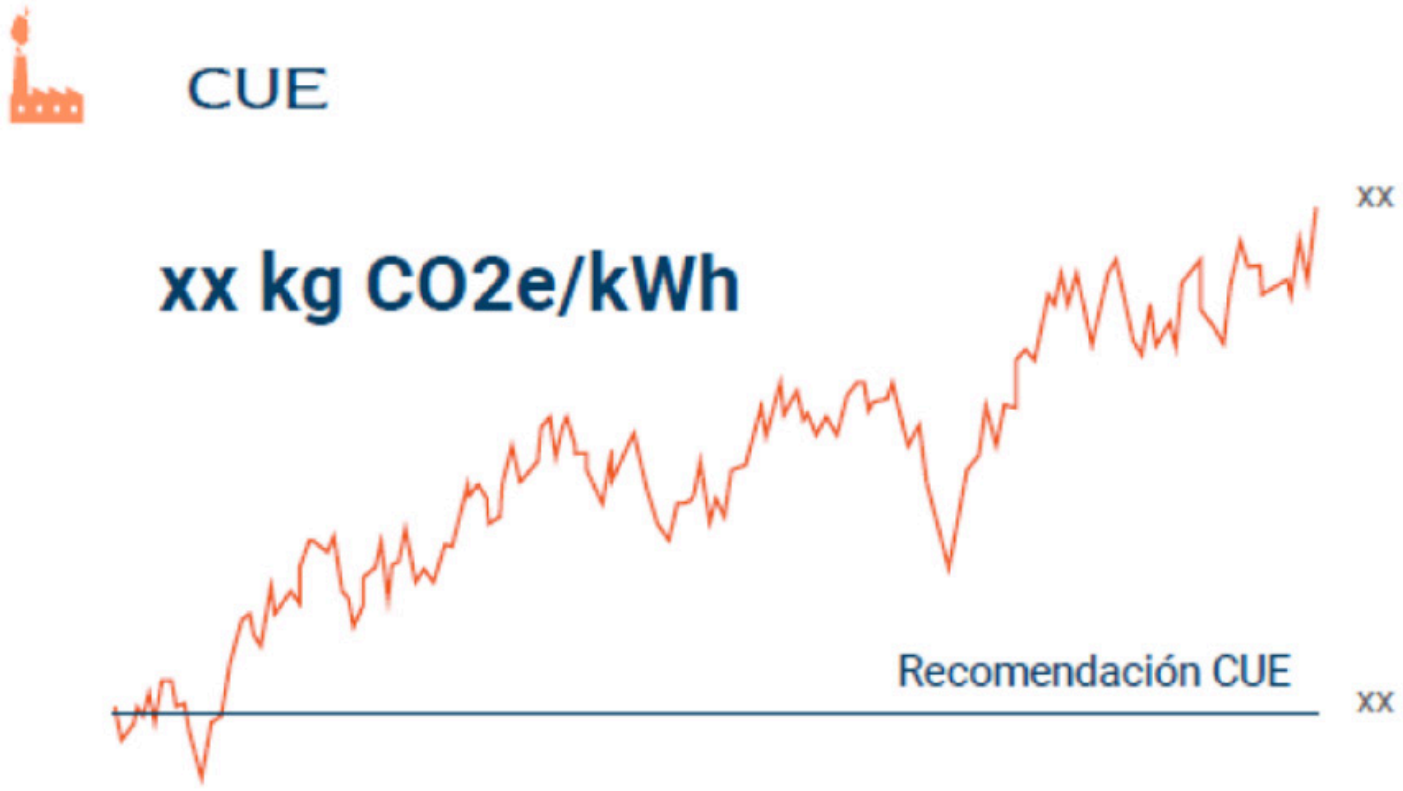
Figura 4.3. Visualización del WUE



Fuente: Elaboración propia.

- CUE (Eficiencia del Uso del Carbono): Mide las emisiones de dióxido de carbono por kilovatio hora. ThinkData permite a los operadores reducir significativamente estas emisiones al identificar fuentes ineficientes y promover el uso de energías renovables.

Figura 4.4. Visualización del CUE



Fuente: Elaboración propia.

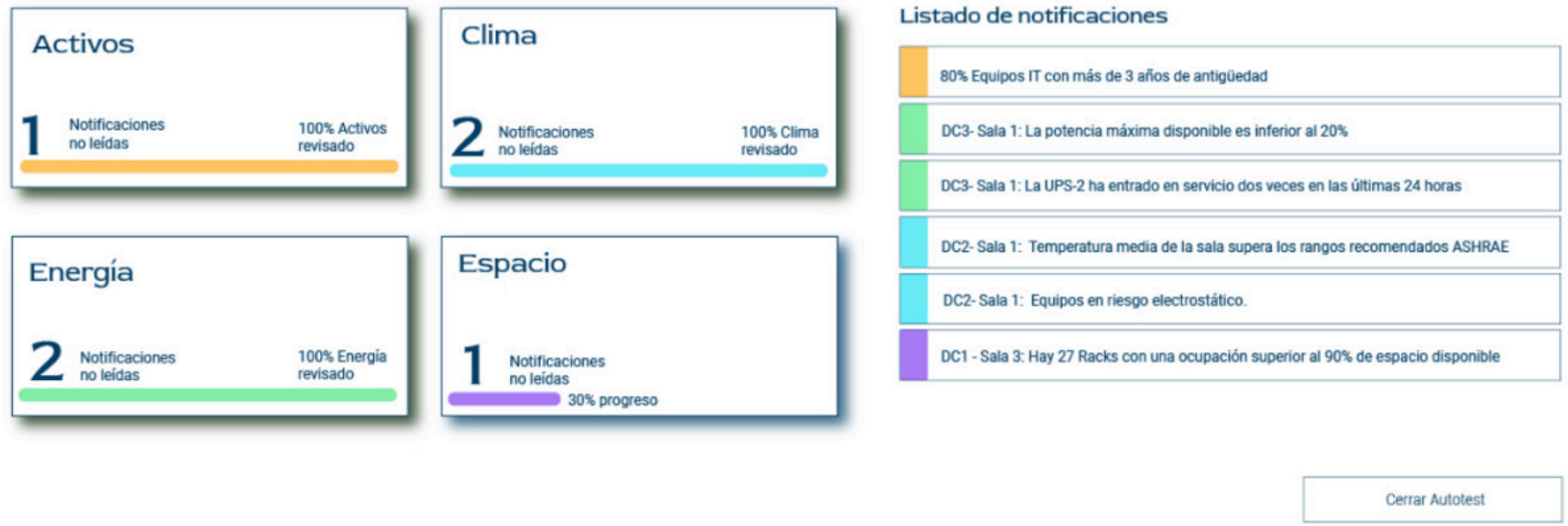
Aplicación de modelos predictivos y prescriptivos

ThinkData es una plataforma avanzada que integra inteligencia artificial para optimizar la operación de los data centers. A través de su enfoque en la recopilación, análisis y simulación de datos críticos, ThinkData ofrece soluciones específicas para abordar los retos operativos y sostenibles del sector.

Chequeo automático del Data Center

ThinkData realiza un análisis detallado y automatizado de las áreas clave de un data center, incluyendo activos, condiciones climáticas internas, consumo de energía y uso del espacio. Este proceso genera un conjunto de notificaciones basadas en los datos recolectados, lo que permite identificar áreas críticas que requieren intervención. El sistema categoriza los resultados para facilitar la evaluación y priorizar acciones con lo mejora la eficiencia operativa; además, el enfoque proactivo asegura que las instalaciones mantengan un rendimiento óptimo mientras cumplen con las normativas ambientales y de sostenibilidad.

Figura 4.5. Autotest con IA de Think Data



Fuente: Elaboración propia.

Comparación del rendimiento del Data Center frente al sector

Una funcionalidad destacada de ThinkData es su capacidad para comparar indicadores clave de desempeño con promedios sectoriales y nacionales ya que:

- Permite evaluar el **PUE (Power Usage Effectiveness)** del data center frente a los estándares promedio de la industria en su país.
- Analiza la **temperatura media** dentro del data center, proporcionando información crítica sobre el desempeño de los sistemas de enfriamiento.
- Monitorea el crecimiento en el número de racks instalados, permitiendo identificar tendencias en relación con otras empresas del sector durante los últimos seis meses.
- Estas comparaciones ayudan a posicionar estratégicamente a los data centers dentro de su sector, ofreciendo perspectivas sobre su eficiencia y competitividad.

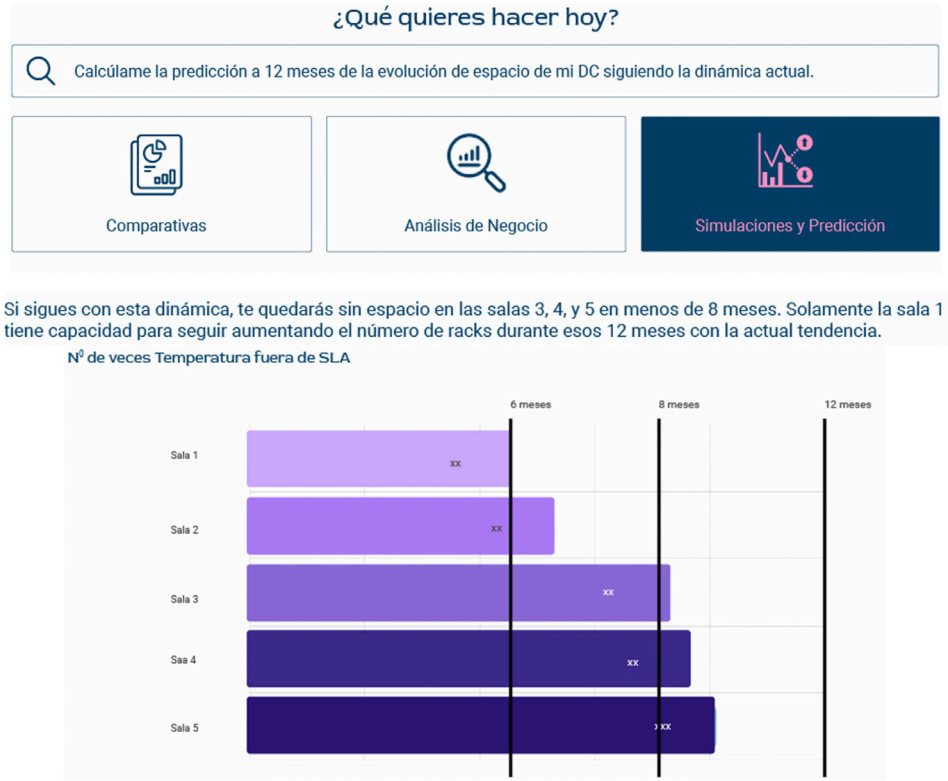
Figura 4.6. Comparativa Data Centers de Think Data

Fuente: Elaboración propia.

Simulaciones y predicciones basadas en IA

ThinkData utiliza modelos predictivos para anticipar la evolución operativa de los data centers. En ese sentido, permite realizar simulaciones sobre el crecimiento y las demandas futuras de espacio, energía y activos, basadas en tendencias históricas y datos en tiempo real. Asimismo, las simulaciones incluyen recomendaciones específicas para la optimización del uso de recursos, anticipando escenarios potenciales y facilitando la planificación estratégica. Estas capacidades no solo ayudan a mitigar riesgos, sino que también permiten a los operadores de data centers tomar decisiones informadas que maximicen el rendimiento a largo plazo.

Figura 4.7. Simulaciones y Predicciones de Think Data



Fuente: Elaboración propia.

El Informe de sostenibilidad de ThinkData: una solución integral para cumplir con normativas internacionales

el informe de sostenibilidad generado por ThinkData representa un avance significativo para los operadores de data centers que deben cumplir con normativas internacionales estrictas, como la **Directriz de Eficiencia Energética (EED)** y la **Corporate Sustainability Reporting Directive (CSRD)**. Diseñado para automatizar procesos y garantizar el cumplimiento regulatorio, este informe integra métricas ambientales críticas, evaluaciones de impacto local y datos de alta calidad.

Automatización y generación de informes de cumplimiento

Gracias a la monitorización constante de parámetros clave como energía, agua y emisiones, ThinkData facilita la generación automática de informes que cumplen con las obligaciones regulatorias establecidas por la Unión Europea (UE). La automatización elimina errores manuales y acelera la recopilación y presentación de datos, permitiendo a los operadores concentrarse en la optimización de las operaciones.

Obligaciones para los data centers

El marco regulatorio europeo ha impuesto requisitos específicos que los Data Centers deben cumplir y ThinkData ofrece soluciones alineadas con estas exigencias:

1. Publicación de Datos: De acuerdo con el artículo 12 de la **Directiva (UE) 2023/1791**, los Data Centers están obligados a publicar anualmente la información descrita en el Anexo VII de la directiva. Este informe debe estar disponible antes del **15 de mayo de cada año**. El informe de ThinkData simplifica este proceso al generar automáticamente los datos requeridos gracias a su sistema de monitorización multi-fuente e incluir métricas precisas de **PUE (eficiencia energética)**, **WUE (eficiencia del uso del agua)** y **CUE (eficiencia del carbono)**. Esto asegura que los Data Centers puedan cumplir con la publicación de datos de forma puntual y sin necesidad de procesos adicionales.
2. Reporte a la Base de Datos Europea: Según el **Reglamento Delegado (UE) 2024/1364**, los Data Centers deben reportar información energética detallada a la base de datos que será creada por la Comisión Europea. La fecha límite inicial es el **15 de septiembre de 2024** y a partir de entonces, el 15 de mayo de cada año. ThinkData facilita el reporte al proporcionar: informes estructurados y listos para ser enviados directamente a la base de datos, así como datos normalizados de acuerdo con los estándares europeos y reduciendo con ello, la

carga administrativa para los operadores.

Contenido del Informe de Sostenibilidad

El informe de ThinkData aborda áreas clave para garantizar el cumplimiento normativo y mejorar la sostenibilidad operativa, entre las cuales pueden mencionarse las siguientes:

- **Métricas Ambientales:** Se incluyen datos sobre el consumo de energía, uso de agua y emisiones de gases de efecto invernadero (GEI). Estas métricas son fundamentales para medir el desempeño sostenible del data center y alinear sus operaciones con normativas internacionales como la ISO 14064.

- **Evaluaciones de Impacto Local:** ThinkData analiza factores adicionales como el uso de recursos, el impacto del ruido en áreas cercanas y la biodiversidad, proporcionando una visión integral del efecto del data center en su entorno.

Datos de Alta Calidad: Los datos recopilados se clasifican en tres niveles de precisión:

- **Medidos:** Datos recolectados directamente mediante sensores y equipos especializados.

- **Calculados:** Basados en fórmulas y parámetros técnicos definidos.

- **Estimados:** Aproximaciones realizadas cuando los datos medidos o calculados no están disponibles.

La clasificación asegura que cada informe sea confiable y adecuado para tomar decisiones estratégicas fundamentadas.

Ventajas de ThinkData en el cumplimiento normativo

1. **Automatización Completa:** La integración de datos permite que los informes se generen automáticamente sin intervención manual.
2. **Ahorro de Tiempo:** Al eliminar la necesidad de recopilación manual, los operadores pueden cumplir con las fechas límite de las regulaciones de manera más eficiente.
3. **Reducción de Errores:** El uso de datos medidos y calculados asegura un nivel alto de precisión, minimizando los riesgos asociados con informes inexactos.

4. **Cumplimiento Normativo Simplificado:** La estructura del informe está diseñada para alinearse directamente con los requisitos de la Comisión Europea, asegurando que se cumplan las normativas sin esfuerzo adicional.

Marco normativo y estándares internacionales

el avance hacia la sostenibilidad en los data centers está respaldado por un marco normativo robusto y estándares internacionales. Entre las normativas más relevantes se encuentran la Directriz de Eficiencia Energética (EED) y la Directriz de Informes de Sostenibilidad Corporativa (CSRD) de la Unión Europea. Estas normativas exigen a las empresas monitorear y reportar su desempeño en términos de sostenibilidad, estableciendo metas claras para la reducción del consumo energético y las emisiones de carbono.

En términos de estándares técnicos, la norma ISO 30134-2 proporciona un marco para medir y evaluar la eficiencia energética mediante el PUE. Este estándar permite a los operadores comparar el desempeño de sus data centers y monitorear mejoras a lo largo del tiempo. Por otro lado, la norma ISO 14064 ofrece directrices para la medición y reporte de emisiones de GEI, lo que garantiza la transparencia y la responsabilidad en las prácticas sostenibles.

Conclusión

La sostenibilidad en los data centers no es solo un imperativo ambiental, sino también una oportunidad estratégica para mejorar la eficiencia operativa y reducir costos. La inteligencia artificial, como lo demuestra ThinkData, ofrece soluciones prácticas y efectivas para enfrentar los desafíos de sostenibilidad. Sin embargo, el éxito de estas iniciativas depende en gran medida del cumplimiento de normativas y estándares internacionales, que proporcionan un marco estructurado para la implementación de prácticas sostenibles.

Referencias

Admin. (2023, 24 mayo). Tendencias en la climatización de Centros de Datos - Campos. Campos. https://campos-corporacion.com/tendencias-climatizacion-centros-de-datos/

BOE.es - DOUE-L-2024-80715 Reglamento Delegado (UE) 2024/1364 de la Comisión, de 14 de marzo de 2024, relativo a la primera fase del establecimiento de un régimen de evaluación común de la Unión para centros de datos. (s. f.). https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-80715

Carcar, S., Carcar, S., & Carcar, S. (2024, 9 noviembre). Las dos transiciones económicas clave, rumbo a la colisión: por qué la digitalización puede torpedear la lucha climática. El País. https://elpais.com/economia/negocios/2024-11-09/las-dos-transiciones-economicas-clave-rumbo-a-la-colision-por-que-la-digitalizacion-puede-torpedear-la-lucha-contra-el-cambio-climatico.html

Centros de datos. (s. f.-a). Ministerio Para la Transición Ecológica y el Reto Demográfico. https://www.miteco.gob.es/es/energia/eficiencia/centros-de-datos.html

Centros de datos. (s. f.-b). Ministerio Para la Transición Ecológica y el Reto Demográfico. https://www.miteco.gob.es/es/energia/eficiencia/centros-de-datos.html

Data centres & networks - IEA. (s. f.). IEA. https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks

Data growth worldwide 2010-2025 | Statista. (2023, 16 noviembre). Statista. https://www.statista.com/statistics/871513/worldwide-data-created/

Ministerio para la Transición Ecológica y el Reto Demográfico. (s. f.). Ministerio Para la Transición Ecológica y el Reto Demográfico. https://www.miteco.gob.es/

Tecnologías efectivas para reducir la huella de carbono en los centros de datos de América Latina – https://yulder.co/. (s. f.). https://yulder.co/tecnologias-efectivas-para-reducir-la-huella-de-carbono-en-los-centros-de-datos-de-america-latina/

thinkdataeficiencia | Bjumper. (s. f.). Bjumper. https://www.bjumper.com/thinkdataeficiencia

5 Cambiando vidas con la inteligencia artificial

Rodrigo Herrera

Rodrigo Herrera. Co-fundador y Chief Technology Officer (CTO) de Ainnovactech.

Ha sido galardonado como Dr. Honoris Causa por el Claustro Doctoral en Barcelona. Con una formación en ingeniería en sistemas, ostenta un Máster en Big Data & Business Intelligence. Rodrigo lidera en el país la organización sin fines de lucro Saturdays.AI, encargada de democratizar el conocimiento de la Inteligencia Artificial. Su inclinación hacia tecnologías emergentes, como el machine learning y el deep learning, lo ha posicionado en la vanguardia del desarrollo de software para detección temprana de enfermedades en AinnovaTech. Ha tenido un papel relevante en la academia, desempeñando roles como director, profesor y coordinador de programas en áreas de Ciencia de Datos, Inteligencia Artificial y Transformación Digital en universidades nacionales e internacionales.

Rodrigo Herrera

Resumen

Esta ponencia se enfoca en mostrar la manera como la inteligencia artificial (IA) está transformando a los distintos sector productivos y especialmente al campo de la salud, gracias a la integración de soluciones innovadoras de toda índole. Desde esta perspectiva, el documento examina el caso de la startup costarricense AInnovatech, la cual ha venido realizado desarrollando distintas soluciones tecnológicas que buscan contribuir a la detección temprana de enfermedades como la retinopatía diabética, el Alzheimer y el hígado graso, entre otras. A partir de estas aplicaciones, se reflexiona sobre el enorme potencial que ofrece la IA para mejorar los diagnósticos médicos y con ello, la calidad de vida de las personas.

Palabras clave: Inteligencia artificial, diagnóstico médico, detección temprana, retinopatía diabética, innovación tecnológica.

Introducción

Hoy la inteligencia artificial (IA) nos deslumbra con las capacidades que ofrece esta tecnología, pues esta ha dejado de ser una promesa futurista para convertirse en una herramienta que está impulsando la transformación de los distintos sectores productivos a lo interno de nuestras sociedades. Su capacidad para analizar grandes volúmenes de datos, y automatizar procesos y tareas rutinarias, ha impulsado la aparición de novedosos productos y servicios cada vez son más innovadores, eficientes y personalizados.

Todo esto representa una revolución tecnológica que está abriendo las puertas para desarrollar nuevos modelos de negocio basados en aplicaciones poco exploradas, particularmente en áreas en las que la precisión y la rapidez son vitales. Y es que justamente, uno de los ámbitos en los que ambas potencialidades son requeridas es el campo de la salud, donde la IA podrá convertirse en una herramienta para generar diagnósticos más certeros. Esto resulta sumamente valioso en un contexto en el que cada día millones de personas son diagnosticadas con enfermedades crónicas u otros padecimientos graves que requieren de la detección temprana y una atención pronta.

Es desde este punto de vista, que la ponencia busca ahondar en la experiencia de la startup costarricense AInnovatech[1] al considerarla como un caso de innovación aplicada al desarrollar soluciones de IA enfocadas en el diagnóstico temprano de enfermedades a través del análisis de imágenes médicas.

1 Dado el tipo de soluciones que desarrolla la empresa, esta facilita servicios a optometristas, oftalmólogos, aseguradoras, laboratorios y hospitales. La empresa está constituida en Costa Rica y los Estados Unidos; además, tiene presencia en México, Chile e India.

¿Cómo aplicar la IA al campo de la salud?

Actualmente, más de 500 millones de personas padecen diabetes y de estos, 30 de 100 llegan a desarrollar retinopatía diabética. Si esta enfermedad no es detectada a tiempo, las personas van perdiendo la vista hasta llegar al estado de ceguera y en muchos casos, la pérdida de visión ocurre cuando muchas personas aún están en edad laboral. Se estima que de 100 casos de ceguera, 95 pueden ser prevenidos si se detectan de forma temprano. Sin embargo, el 49% de los diagnósticos de la enfermedad suelen ser errados, lo que evidencia la necesidad de mejora en los procedimientos de detección y diagnóstico.

Lo anterior motivó la creación de un modelo de IA (*VIsionAI*) con el que se puede tomar una imagen de la retina y posteriormente, dicha imagen es subida a una plataforma que en pocos segundos, determina si una persona posee retinopatía diabética u otro tipo de retinopatías. Este modelo tiene la capacidad de clasificar las imágenes y de segmentarlas a partir de un banco de miles de imágenes que han sido previamente validadas por especialistas en el campo médico. De ese modo, el diagnóstico efectuado por el modelo permite tener una certeza del 96%.

El mismo tipo de tecnología puede ser utilizada para identificar el riesgo cardiovascular u otras enfermedades por medio de las imágenes médicas que se toman de la retina o de un procedimiento como el fondo de ojo. Debido a que la toma de imágenes médicas con cámara resulta costoso y complejo por la pericia que se debe tener para capturar adecuadamente las diferentes concavidades del ojo humano, las soluciones de IA pueden subsanar estas dificultades. Es así como, las soluciones de IA pueden contribuir a la detección temprana de enfermedades de una forma rápida y accesible, al tiempo que evitan errores en el diagnóstico y permiten contar con tamizajes más certeros.

Actualmente, la empresa se encuentra trabajando en otras innovaciones que le permitan la detección de otras enfermedades (por ejemplo, del Alzheimer) de modo que ello faculte que a través de imágenes médicas, las y los pacientes puedan recibir un diagnóstico temprano y con ello, puedan adoptar las medidas necesarias para mejorar su calidad de vida. En esta línea, AInnovatech también está llevando a cabo un prototipado de una cámara especial para realizar la toma de imágenes de fondo de ojo de forma automática y haga que el proceso sea más rápido y simple. Además, ha implementado un kit para la detección de pre-diabetes, riesgo cardiaco, disfunciónón renal y hépatica dentro de la misma solución VisionAI.

Conclusiones

Los casos expuestos en la ponencia evidencian que la IA podrá llegar a constituirse en un instrumento fundamental para impulsar la innovación en sectores como el de la salud, ya que su capacidad para procesar datos con alta precisión y ofrecer diagnósticos rápidos, se posiciona como un medio para disminuir errores, optimizar el uso de recursos y mejorar la atención médica. Además, en contextos en los que la disponibilidad de especialistas sea limitada, este tipo de herramientas puede contribuir a facilitar el acceso a los servicios de salud, lo que muestra el enorme potencial de la IA para transformar la forma como han operado los sistemas sanitarios hasta ahora.

Por otro lado, el caso de la startup AInnovatech indican que la aplicación práctica de la IA puede llevar a la creación de soluciones específicas a problemas de salud pública urgentes. En ese sentido, debe recordarse que el modelo de Vision AI no solo mejora la detección temprana, sino que también representa una alternativa eficiente, accesible y alternativa a los métodos tradicionales de tamizaje. Gracias a este tipo de innovaciones no sólo se introducen mejoras operativas al sistema de salud, sino que también se estimula la aparición de nuevas oportunidades de negocio, al tiempo que se fortalece la economía del conocimiento.

Asimismo, la versatilidad que pueden adquirir algunas de estas soluciones tecnológicas para que sean aplicadas con otros fines o en otros campos, debe ser visto como una oportunidad para fomentar esfuerzos de investigación y el desarrollo (I+D) que no olviden la importancia de potenciar soluciones tecnológicas que generen mayor valor público y se alineen con el bienestar social. Aunado a ello, la experiencia de AInnovatech también pone sobre la mesa la necesidad de promover ecosistemas que impulsen el desarrollo científico-tecnológico desde contextos locales con proyección global.

6 Reflexiones éticas sobre la IA generativa desde la comunicación

Óscar Alvarado Rodríguez

Óscar Alvarado Rodríguez. Docente e investigador de la Escuela de Ciencias de la Comunicación Colectiva.

Realizó su maestría en Interacción Humano-Computador en la Universidad de Uppsala en Suecia y su doctorado en Ciencias de la Ingeniería con especialidad en Ciencias de la Computación en la Universidad Católica de Lovaina en Bélgica. Sus intereses actuales de investigación se centran en cómo interactuamos con tecnologías altamente influenciadas por sistemas de inteligencia artificial, el diseño de sus interfaces para las personas usuarias y sus implicaciones en su modo de vida

Óscar Alvarado Rodríguez

Resumen

Los avances facilitados por la inteligencia artificial (IA) generativa, así como su creciente integración en la vida cotidiana, plantean desafíos éticos de diversa naturaleza que deben ser abordados desde una mirada multi e interdisciplinaria. Bajo esta óptica, esta ponencia propone una reflexión desde el campo de la comunicación en torno al uso, los riesgos y las implicaciones sociales de estas tecnologías. A partir de preguntas clave, se identifican elementos críticos que deben ser considerados en un marco ético que promueva un uso justo, transparente y humanamente responsable de la IA.

Palabras clave: inteligencia artificial, inteligencia artificial generativa, ética, comunicación, sesgos algorítmicos.

Introducción

La creciente utilización de la inteligencia artificial (IA) generativa ha provocado notables transformaciones en el terreno social, económico y cultural. Herramientas como ChatGPT, DALL-E y otras semejantes, han hecho que capacidades que antes estaban restringidas a especialistas, ahora estén al alcance de cualquier persona. Sin embargo, el que este tipo de servicios sean tan accesibles ha generado preguntas fundamentales sobre la responsabilidad, los derechos, la veracidad y la representación en la creación de mensajes y difusión de la información.

En este campo de la comunicación surgen inquietudes específicas con respecto al uso de la IA para producir contenidos y con ello, transformar el trabajo informativo, así como los marcos éticos que orientan estas prácticas. Es así como la presente ponencia pretende ofrecer una reflexión crítica sobre algunos de los principales desafíos éticos que plantea la IA generativa desde la perspectiva comunicacional.

La inteligencia artificial generativa y sus aplicaciones en la Comunicación

En términos generales, los sistemas de inteligencia artificial (IA) requieren de un corpus o conjunto de datos que es usado para entrenar a los sistemas de IA, así como de un modelo algorítmico a través del cual se procesa ese conjunto de datos para simular respuestas inteligentes. Por su parte, la IA generativa está sustentada en modelos algorítmicos específicos que, partiendo de un corpus de contenido previo, permiten la generación de textos, imágenes, audios y video novedosos, actualmente muy utilizados en la generación de contenido.

Tal versatilidad ha llevado a que la IA cada día sea más integrada en diversas actividades y campos del conocimiento. El área de la comunicación no es la excepción, ya que desde este ámbito la IA generativa está siendo usada con distintos propósitos. Por ejemplo, varios medios de comunicación han

experimentado con IA generativa para sustituir a las personas redactoras. Así mismo, la IA se ha empleado en campañas publicitarias, la generación de audiovisuales y la creación y recomendación de contenido personalizado en diversas plataformas.

Retos actuales con la IA y sus dilemas éticos para la Comunicación

Tales beneficios obligan a reflexionar sobre cómo, cuándo y de qué manera debería utilizarse la IA cuando se realizan tareas asociadas a la comunicación. Para ello, vale la pena contextualizar algunos de los retos actuales que siguen suscitándose con la IA.

En primer lugar, se debe tener en cuenta que los sistemas de IA se describen como cajas negras de gran complejidad, llenas de modelos estadísticos y de estrategias algorítmicas que producen resultados sobre los cuales no siempre hay suficiente claridad de donde provienen o cómo se generaron. Aunque existen soluciones que permiten entender si un nodo específico de alguna red neuronal se activa más ante un dato u otro, dicho tipo de soluciones magnifica el problema en lugar de solventarlo, pues recurre a otras soluciones de IA igual de opacas para determinar esa respuesta, manteniendo con ello el problema de la transparencia.

En segundo lugar, los *sesgos* presentes en los sets de entrenamiento, las *pareidolias* y el *asertivismo* con que estas tecnologías tratan de convencernos de un resultado (que podría estar sesgado o no ser ciertos), resultan retos adicionales. El que respuestas incorrectas sean presentadas con gran seguridad por parte de estos sistemas compromete su uso con mayor responsabilidad en situaciones en las que la precisión es muy importante, como en el caso de las actividades periodísticas o informativas. Como los sistemas de IA son creados por personas y entrenados por corpus de datos también influenciados por personas, los sesgos propios del ser humano se trasladan a estos sistemas por todos sus frentes, lo que recalca la importancia de desarrollar marcos éticos que regulen tanto el entrenamiento de los modelos como su aplicación.

En este contexto, no hay que olvidar que **la IA afecta directamente diversos grupos de personas y poblaciones**. En ese ámbito, es muy evidente el sesgo que tienen muchos de los modelos de IA generativa, por ejemplo, cuando se le solicita producir imágenes y estos tienden a generar o reconocer más fácilmente a personas de tez blanca. Por tanto, al usar estas herramientas deberíamos verificar si esos modelos sobre o sub representan más a unas poblaciones sobre otras en los resultados de estos modelos.

En tercer lugar, una de las grandes preocupaciones que genera la IA es la posible *pérdida de empleos* por el reemplazo. Si bien es posible que muchas ocupaciones desaparezcan, también ocurrirá una transformación del empleo ante la integración de las distintas tecnologías de IA en nuestro diario vivir.

En ese sentido, debemos dejar de enfocarnos en un futuro distópico y en su lugar debemos preguntarnos: ¿quiénes están usando a la IA?, ¿cuáles son sus objetivos con estas herramientas?, ¿cómo son entrenadas?, ¿qué modelos se escogen para entrenar esos datos? Este tipo de interrogantes pueden ayudarnos a encauzar los dilemas a los que nos enfrentamos a la hora de utilizar estas tecnologías.

En cuarto lugar, otro tema que debe considerarse son *los debates sobre los derechos de autor y conexos*, pues no queda claro si éstos son infringidos cuando se utiliza IA generativa, ni tampoco se ha esclarecido quién será el responsable de la infracción. Ante esto, sería responsable: ¿la empresa que desarrolló el modelo de IA o las personas que utilizan el modelo y/o quienes crean las imágenes, videos y texto a partir de modelos de IA generativa? Además, ¿es la persona que consume el sistema de IA la autora de lo que genera, es la empresa que implementó el modelo, o las personas que generaron el contenido previo para entrenar el modelo?, o ¿acaso es la persona autora del prompt que se utilizó para crear ese resultado? En esta línea, se presupondría que estos servicios de IA deberían haber solicitado permiso para usar esos datos de contenido previo de entrenamiento. Vale la pena preguntarse si es ético o lícito utilizar modelos de IA generativa que no hayan solicitado ese permiso.

En quinto lugar, otras dudas éticas que podemos hacernos en el campo de la comunicación son: ¿debe una persona indicar a sus públicos cuando alguno de sus productos fue creado con IA generativa y cuál fue el modelo utilizado?, ¿en qué situaciones esta indicación debería ser un requerimiento obligatorio?

En este contexto, es necesario establecer a *la persona humana, ya sea las que implementaron el modelo o la persona que lo consume, como la responsable de los resultados de cualquier modelo de IA*, para evitar atribuirle la responsabilidad al modelo o a la tecnología. Además, se debería explicitar siempre cuando algún contenido o resultado que sea generado por la IA para que el público tenga conocimiento y valore el consume de ese contenido en su debida dimensión.

Finalmente, es importante reflexionar sobre la importancia de crear responsabilidades penales si se usa la IA para la difusión de información falsa o en contra de derechos humanos y el desarrollo de actividades de ciberterrorismo o de cibercrimen. Para eso, se deben establecer mecanismos para permitir *auditorías previas o revisiones posteriores del set de entrenamiento y de los modelos de IA*.

Es necesario recordar que el desarrollo de la tecnología debe estar siempre centrado en la persona, procurando el potenciamiento de sus capacidades en esta era de cambio tecnológico. En ello resulta indispensable que se *priorice la autonomía de las personas y la protección de sus datos durante el desarrollo de los modelos de IA*.

Conclusión

La inteligencia artificial generativa es una de las transformaciones tecnológicas más importantes de los últimos años, no sólo por la creciente capacidad para automatizar diverso tipo de tareas, sino también por su integración en todo tipo de espacios. A pesar de que la IA brinda beneficios y múltiples oportunidades, esta tecnología no está exenta de retos a nivel ético, social, económico y político.

Para la comunicación, la IA suscita retos que obligan a repensar de una forma crítica la producción, la circulación y el consumo de contenidos. Al incorporar sistemas que no son neutrales, éstos pueden transmitir los intereses, los sesgos y las limitaciones humanas de las personas que los crean o que los utilizan a los procesos comunicativos e informativos. Esto evidencia desafíos particulares desarrollados en los párrafos anteriores, pero también muestra que su abordaje no debe ser realizado desde una única mirada, privilegiando los análisis técnicos de la herramienta.

Estas discusiones deben integrar diversas perspectivas multi e interdisciplinarias de modo que ello realmente nutra y contribuya en la construcción de marcos éticos que promuevan la equidad, la transparencia y, sobre todo, el uso responsable de estas herramientas. A partir de eso, debemos centrar las discusiones en las personas que diseñan, utilizan, regulan y son afectadas por estas tecnologías, pues éstas no son neutrales y podrían generar consecuencias graves para nuestras sociedades y democracias.

7

Inteligencia artificial, privacidad y ética del bien común en el capitalismo de la vigilancia

Luis Arturo Martínez Vásquez

Luis Arturo Martínez Vásquez. Doctor en Filosofía por la Universidad de Granada, España.

Obtuvo mención Cum Laude en su tesis de doctorado sobre Antropología filosófica latinoamericana. Además, es, bachiller en Filosofía y Humanidades, licenciado en Docencia con énfasis en Filosofía y Máster en Educación. Cursó estudios de maestría y doctorado en Filosofía en la Universidad de Costa Rica (UCR). Cuenta con una especialización en Entornos Virtuales de Aprendizaje.

Actualmente se desempeña como asesor académico en la Vicerrectoría de Docencia de la Universidad de Costa Rica, dedicado a temas de ética de la Inteligencia Artificial. Además, forma parte del cuerpo editorial del proyecto Ellacuría Obras Completas de la Universidad de Granada, pertenece al grupo de Investigación interdisciplinaria FiDIAS: Ficción, Desinformación, Inteligencia Artificial y Subjetividades y es miembro de LIBERESP: Laboratorio Iberoamericano de Ética de la Salud Pública, nodo Costa Rica.

Entre sus obras escritas se encuentran: Interacción humano-máquina y ontologías relacionales: Principios para una ética de la Inteligencia Artificial. En AAVV. Filosofía: Crisis y Perspectivas. Guayacán, 2024, "Soñar os hará libres. Los retos del transhumanismo a los estudios humanísticos". Kénosis, 2024, "La "tierra y gente" del paraíso lascasiano: Apuntes sobre la trans-formación de la otredad en la conquista de América Latina" (UNA, 2019), "La realidad humana como problema. Hacia una antropología filosófica latinoamericana de la liberación" (Estudios, 2023), The liberating philosophy of Ignacio Ellacuría. Historical reality, humanism and praxis. (L. A. Martínez Vásquez, R. Carrera Umaña, & L. R. Díaz Cepeda, Eds.). (Lexington Books, 2024).

Luis Arturo Martínez Vásquez

Resumen

Desde la perspectiva del capitalismo de vigilancia, esta ponencia aborda las implicaciones psicopolíticas que derivan de la integración masiva de los modelos de inteligencia artificial (IA) en la actualidad. A partir de esto, se problematiza la transformación de la persona humana en una mercancía de datos, se cuestiona la erosión de la privacidad como un valor fundamental de nuestras sociedades y se propone una reflexión sobre la autodeterminación informativa y la ética del bien común como vías para exigir mayor responsabilidad por parte de los desarrolladores de la IA.

Palabras clave: Inteligencia artificial, privacidad, capitalismo de la vigilancia, ética del bien común, autodeterminación informativa.

Introducción

Con la integración masiva de la Inteligencia Artificial (IA) en muchos ámbitos de la vida cotidiana los sujetos han experimentado beneficios en términos de la eficiencia, la conectividad y el acceso a la información; sin embargo, con el incremento en la cantidad de datos personales e información que son recopilados y procesados por los sistemas de IA surgen cuestionamientos con respecto a la forma como se concibe y gestiona la privacidad.

El auge del Big Data y de las técnicas de minería de datos, han contribuido a establecer dinámicas de vigilancia constante en la que la información personal pasa a ser un recurso explotable. Con ello, se trasciende lo meramente técnico y se da lugar a nuevas configuraciones del poder y control. Ante este contexto estamos frente al surgimiento de formas novedosas de control social, en el que los datos mejoran los servicios digitales, pero también son utilizados para modificar y predecir las conductas humanas humanas. Aquí emergen una serie de dilemas ético-jurídicos y antropológicos, que si bien han estado presentes en otros contextos, poseen características específicas que merecen la pena analizarse.

En ese sentido, se aborda el impacto psicopolítico que producen los procesos de datificación de la realidad y la mercantilización de la experiencia humana en el capitalismo de vigilancia a partir de una mirada filosófica. Bajo esta perspectiva se analizan las tensiones entre el desarrollo tecnológico, la privacidad, la autonomía y el bien común; al tiempo que se cuestiona la viabilidad de la privacidad ante entornos digitales que están regulados por algoritmos predictivos. Desde este posicionamiento, se plantean limitaciones para la economía de datos y se ofrecen criterios éticos para contar con una regulación más equitativa en la que los desarrolladores de IA asuman las consecuencias indirectas de sus decisiones.

Esta reflexión se enmarca en los debates actuales sobre la regulación y ética de la IA, por lo que pretende contribuir a la formación de marcos jurídicos más sólidos, sobre todo en sociedades en las que las propuestas legislativas siguen sin abordar con la debida profundidad la protección de la privacidad. Ante una economía en la que la persona se convierte en un insumo más para generar predicciones y lucro, se debe repensar la relación entre la tecnología y la libertad desde una ética del bien común que priorice la dignidad y la autonomía individual.

Datificación de la realidad y mercantilización del sujeto

El creciente auge por el desarrollo de sistemas que integran la inteligencia artificial en los diferentes ámbitos de la vida social ha traído consigo un concomitante crecimiento por la captación y almacenamiento de los datos de usuarios que utilizan estos sistemas. La razón que motiva esta relación es en primer lugar técnica, toda vez que los datos son el insumo que utilizan los modelos y algoritmos de IA para ser creados, entrenados y que estos puedan estar en continuo funcionamiento. El uso de estos datos permite que los sistemas de IA propicien tareas como la desambiguación lingüística, la clasificación y el etiquetado de los elementos clasificados por intereses; pero también ha motivado una búsqueda constante de utilidad mediante la expansión masiva de la recolección de datos.

Así, desde inicios del siglo XXI quienes han desarrollado estos modelos, han encontrado también en la recolección, el almacenamiento y la distribución de los datos masivos, un campo de mercantilización que en los últimos años ha implicado varios miles de millones de dólares con proyecciones de tasas de crecimiento significativas. De esta manera, quienes desarrollan la IA, al recolectar estos conjuntos de datos han propiciado una dinámica económica bastante solvente a partir de las interacciones que las personas generan cuando ingresan a una página web o a una aplicación desde cualquier dispositivo digital, en la mayoría de las veces, sin claridad por parte del usuario de los extremos de estos mecanismos.

Este proceso de comercialización ha sido denominado **datificación de la realidad** (Mayer- Schönberger y Cukier, 2013; Ábrego y Flores, 2021), ya que cada aspecto de la experiencia humana es traducido a datos, generando una dinámica económica a partir de la creación de perfiles algorítmicos. Con eso se produce una mercantilización en la que se le atribuye un valor a las personas a partir de los datos que ofrecen, quienes se transforman en "usuarias" y "sujetos de datos", se convierten en insumos para los sistemas predictivos y se utilizan para influir en los comportamientos y conductas de otras, lo que erosiona la privacidad y termina configurando una psicopolítica de control: se trata de lo que Zuboff (2013) ha denominado "capitalismo de la vigilancia".

Por lo anterior no es de extrañar que cuando un usuario escribe algunas palabras específicas sobre algún tema en cualquier motor de búsqueda, prontamente comienza a encontrar

anuncios en redes sociales y/o al visitar otros sitios electrónicos que tienen relación con el tema de ese interés. Así entendida, más que una comodidad en la experiencia del usuario, la predicción de anuncios actúa como mecanismo que mina la privacidad y mercantiliza la realidad para que los datos sean comercializables de manera cada vez más masiva.

Con esto, la experiencia humana, termina transformándose en materia prima frente a esta dinámica mercantilizadora que permite la mejora de servicios y predictibilidad de ciertas decisiones. En ese sentido, la colección de nuestros datos refleja un proceso de degradación de la conducta humana porque su uso se convierte en un instrumento psicopolítico que se utiliza para pronosticar nuestras maneras de pensar y actuar a cambio de lo que se denomina una mejora en la experiencia del usuario.

Ante este panorama, resulta relevante profundizar en la idea de la privacidad y analizar si efectivamente esta puede ser cumplida en los modelos algorítmicos. También vale la pena cuestionarse si la privacidad desde este contexto resulta deseable, o si, por el contrario estamos de acuerdo en ceder nuestra privacidad a cambio de mejores condiciones en la innovación de estos modelos. Esta pregunta no es retórica en tanto permite reconsiderar el valor de la privacidad frente a una cultura de conveniencia, predicción y automatización.

La privacidad como autodeterminación informativa

Determinar lo que entendemos por privacidad es un reto. Tabani (2007) después de estudiar distintas teorías de la privacidad ha identificado 4 grandes categorías; a saber, aquellas definiciones que plantean la privacidad como no intromisión, las que abogan por el aislamiento físico, las que sugieren limitar el acceso a nuestra información en ciertos contextos y las definiciones que optan por comprenderla como control sobre la información.

Para efectos de propiciar una idea más contextualizada, se plantea pensar la privacidad como el **derecho a la autodeterminación informativa**: lo que implica, saber de manera directa qué datos estamos cediendo cuando utilizamos o se implementa un modelo de IA; así como tener control sobre quién accede a nuestra información personal. En este marco, surge una preocupación ética fundamental: ¿es posible mantener la privacidad en la era de la inteligencia artificial? Y aún más ¿es deseable hacerlo?

Crítica a las políticas de autorregulación y a la ética epidérmica

La respuesta a las preguntas recientemente planteadas no es sencilla de resolver, en tanto las políticas actuales de privacidad impulsadas por las empresas suelen ser confusas, ambiguas y en algunos casos, innegociables. Unido a lo anterior es identificable una tendencia a que las políticas empresariales estén centradas en el desarrollo de los procesos de autorregulación, lo que significa que, ante la ausencia de normativa específica, las grandes compañías han creado sus propios códigos éticos y normas de privacidad. En ese sentido, se observa un predominio de la autorregulación empresarial sin normativa externa robusta, aunque cada vez hay mayor interés por generar regulaciones de la tecnología, como en el caso de la reciente promulgación de la Ley de Inteligencia Artificial de la Unión Europea (UE), que se está constituyendo como un modelo a seguir en otras latitudes.

Por otra parte, en lo que compete a los procesos de información a las personas usuarias de estos modelos, existen mecanismos de divulgación de políticas de privacidad que, ya sea por el tipo de lenguaje o el acceso a la documentación correspondiente, no permiten una información directa a las personas usuarias de las consecuencias que tiene aceptar las cookies de una página web o ingresar la información personal en un formulario de una aplicación. Este tipo de mecanismos impide a los usuarios valorar los riesgos y beneficios que tiene este tipo de interacción, así como qué tanto de la privacidad se cede a las compañías.

Simultáneamente, se ha venido desarrollando desde estos mismos ámbitos una "ética epidérmica" de la Inteligencia Artificial, superficial e instrumentalizada y que no profundiza en lo que entendemos por ética de la IA, a pesar de que suele referirse a sesgos o conflictos específicos. Este tipo de ética ha sido promulgada por personas que no han tenido una preparación específica en el ámbito de la ética lo que hace que aparezcan conceptos, explicaciones y teorías "éticas" que han sido desarrolladas desde el campo de la IA, pero que están vaciadas de contenido y realmente no examinan a fondo las implicaciones derivadas del desarrollo de estos modelos ni pretenden la protección de las garantías de quienes utilizan estos mecanismos. Esta instrumentalización mina la posibilidad de tener una ética verdaderamente transformadora y que resguarde a las personas de posibles conflictos que surjan en este ámbito.

Propuesta desde una ética del bien común

Frente a este escenario y teniendo en consideración que en la historia del pensamiento occidental la privacidad ha tenido altos y bajos, siendo algo muy preciado en ciertos momentos y en otros no tanto, es menester propiciar espacios de reflexión sobre las implicaciones que la datificación de la realidad y privacidad

tiene en la vida de los seres humanos. Es por eso, que se propone la necesidad de un equilibrio en el entendido de que la ética de la privacidad no puede ser comprendido como un valor absoluto, sino que, por el contrario, debe ser parangonado con el beneficio colectivo que pueda tener la sesión de ciertos datos.

Y en este sentido, también será necesaria una diferenciación entre la privacidad de los individuos y la privacidad de los colectivos. Esto implica pensar que, en lugar de que se rija por las reglas del mercado, la toma decisiones sobre la privacidad debería analizarse desde el beneficio individual y colectivo que tenga, así como sus implicaciones.

El segundo aspecto a tomar en cuenta es la relación de la privacidad con la autonomía, es decir, con la capacidad de ser seres independientes y por tanto, la libertad. Es por eso que debe comprenderse que la idea de privacidad merece de un posicionamiento tal que permita la toma de decisiones con autonomía de frente a aquellas estructuras sociales que truncan las opciones libres y voluntarias de los individuos.

En tercer lugar, si bien esto podría verse como una propuesta utópica, se debe valorar la posibilidad de establecer una ruptura parcial con esta economía de los datos, en el entendido del evidente desequilibrio en la relación que existe ente la persona usuaria y el modelo de IA. La persona usuaria tiene poco acceso a los datos y no sabe muy bien cómo utilizarlos, mientras que en la contraparte se desarrolla una comercialización de datos personales que pareciera ser irrestricta.

Por tanto, se debe pensar en las políticas sobre las cookies y los rastreadores que se despliegan cuando los usuarios ingresan a ciertas páginas o espacios y los mecanismos para que estos sean accesibles para todo tipo de población, tanto desde la manera en la que se presenta el consentimiento -en muchos casos es casi implícito- como desde el mecanismo con que se explican las implicaciones que esto tiene para los individuos.

Conclusiones

La Inteligencia Artificial ha catalizado un proceso de datificación de la realidad, lo que afecta negativamente no sólo a la privacidad de las personas, sino que también pone en entredicho su autonomía. Esto se evidencia en el tipo de políticas actuales de autorregulación que predominan y que resultan insuficientes para garantizar el ejercicio pleno de la libertad y la autonomía de los individuos. Esto es relevante en tanto, al ceder los datos, estamos facilitando información que tiene que ver con nuestras esperanzas y alegrías, pero también con el resguardo de información tan relevante como nuestros historiales médicos y bancarios, entre otros. Todos estos son explotados con fines de lucro en contra de nuestro interés.

Por tal motivo, se requiere avanzar hacia la adopción de marcos normativos más sólidos, tanto desde el ámbito de la autorregulación de las empresas como del de las normativas nacionales e internacionales, de tal manera que se pueda de-

fender la privacidad desde un enfoque que la incorpore como parte del derecho de la autodeterminación informativa. Junto con esto es necesario que se aplique una ética de la Inteligencia Artificial profunda, que se nutra de la interdisciplinariedad y esté basada en la perspectiva del bien común.

Referencias

Ábrego, S., & Flores, M. (2021). Datificación, subjetividad y poder: Una crítica al capitalismo digital. Universidad Nacional Autónoma de México.

Mayer-Schönberger, V., & Cukier, K. (2014). Big data: La revolución de los datos masivos (D. Gortázar, Trad.). Turner Publicaciones.

Tabani, M. (2007). La privacidad en contexto: Tecnología, política y la integridad de la vida social (E. Sanabria, Trad.). Gedisa.

Unión Europea. (2024, 12 de julio). *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). Diario Oficial de la Unión Europea, L 1689, pp. 1–144.

Zuboff, S. (2019). La era del capitalismo de la vigilancia: La lucha por un futuro humano frente a las nuevas fronteras del poder (J. J. Delgado, Trad.). Paidós.

8 Ética y Riesgos en Ciberseguridad

Andrea Vera Celeste

Andrea Celeste Vera. Entusiasta con más de 20 años de experiencia en Tecnología y Seguridad de la Información. Panelista en varios congresos y conferencias nacionales e internacionales.

Magister en Ciberseguridad (Universidad de Valencia). Ingeniera en Sistemas de Información. Diplomatura en Seguridad de la Información. Certificaciones internacionales: CISA (Certified Information Systems Auditor) y CEH (Certified Ethical Hacking), ISO/IEC 27001:2022 Internal Auditor.

Andrea Celeste Vera

Resumen

El avance de la inteligencia artificial (IA) plantea múltiples desafíos éticos y riesgos de ciberseguridad que deben ser abordados desde una perspectiva integral. En ese sentido, esta ponencia explora el papel de la ética en el ciclo de vida de los sistemas de IA como un mecanismo que previene riesgos no deseados como la exposición y la pérdida de privacidad entre otros. Desde esta perspectiva, se destaca la importancia de implementar una evaluación de riesgos basada en frameworks internacionales que ayuden a mitigar las amenazas sobre los modelos de IA. El texto finaliza con una serie de recomendaciones para integrar principios éticos universales y promover el desarrollo seguro, transparente y responsable de la IA.

Palabras clave: Inteligencia artificial, ética en inteligencia artificial, ciclo de vida de la IA, riesgos algorítmicos, evaluación de riesgos.

Introducción

La inteligencia artificial (IA) está transformando la forma en que se toman decisiones, se gestionan recursos y se ofrecen servicios en diversos sectores. Estos desafíos son aún más críticos en áreas como la ciberseguridad, donde la IA puede tanto mitigar como amplificar los riesgos existentes. Es así que a la hora de abordar el tema de las regulaciones y la ética en la inteligencia artificial (IA) también debe considerarse otra dimensión crítica, la cual está representada por los riesgos asociados a la ciberseguridad en los sistemas de IA. Y es que durante los últimos años, el desarrollo e implementación de estas tecnologías ha ocurrido de forma exponencial, con lo que se han generado oportunidades significativas en múltiples sectores. Sin embargo, dicho crecimiento también trae consigo desafíos de carácter ético y por supuesto, a nivel de la ciberseguridad.

Tales retos no deben ser ignorados, sobre todo si se considera que la integración de la IA en los procesos sociales, económicos y administrativos implica tomar decisiones que afectan directamente a las personas y comunidades. Por tanto, esto exige la adopción de un enfoque responsable y ético desde la concepción y diseño de un sistema de IA hasta su implementación. En este contexto, es imprescindible que nos preguntemos qué entendemos por ética en la IA, cómo se aplica esto en la práctica y qué riesgos derivan de su ausencia.

La ética en la IA no sólo se trata de una cuestión teórica o filosófica, sino una necesidad práctica que atraviesa todo el ciclo de vida de los modelos y algoritmos. Desde el diseño hasta la puesta en producción, la ética debe estar presente en cada decisión técnica, en el uso de los datos, en la transparencia de los procesos y en la rendición de cuentas ante los impactos sociales que puede generar la IA.

Junto a los desafíos éticos, se suman los riesgos técnicos y organizacionales derivados de la utilización de modelos de IA. La ciberseguridad, los sesgos algorítmicos, el perfilamiento indebido y la falta de control sobre las bases de datos uti-

lizadas, son apenas algunas de las amenazas que debemos identificar y mitigar. En ese sentido, esta ponencia pretende ofrecer una visión integral que articule principios éticos con la gestión de riesgos, utilizando marcos de referencia reconocidos y proponiendo recomendaciones prácticas para un desarrollo responsable de la inteligencia artificial.

¿Qué entendemos por ética en la inteligencia artificial?

La ética en la inteligencia artificial no es un concepto nuevo, pero ha cobrado especial relevancia debido al impacto creciente de esta tecnología en la sociedad. En términos generales, hablamos de una rama de la ética que analiza y evalúa los dilemas morales surgidos del uso de la IA. Este interés ha generado un volumen importante de literatura, estudios e investigaciones, y es fundamental que tengamos claro a qué nos referimos cuando hablamos de ética aplicada a esta tecnología. Pero, ¿qué abarca la ética en la IA? Los criterios éticos en la IA abarcan múltiples dimensiones, entre ellas:

- El **análisis y tratamiento de datos**. No hay que olvidar que los datos son los que alimentan y se usan para entrenar a los diferentes modelos.

- El **diseño y funcionamiento de los algoritmos**. Estos pueden haber sido creados por una organización, persona y/o adoptados de algún proveedor.

- Las **prácticas tecnológicas**. Estas son implementadas por las organizaciones a nivel interno y/o por terceros.

La ética debe estar presente en todo el ciclo de vida de desarrollo del software, es decir, desde el momento en que se concibe un modelo o algoritmo hasta su implementación. Aunque estos softwares suelen ser desarrollados con otros tipos de tecnología, la ética tiene que estar presente en cada una de estas etapas. Es así como podemos visualizar el desarrollo de sistemas de IA como un ciclo de vida en espiral, similar al desarrollo de software tradicional, pero con el componente ético atravesando cada fase (ver figura 8.1.).

Figura 8.1. Ciclo de vida de la inteligencia artificial

Fuente: Elaboración propia.

Este ciclo es continuo y evoluciona conforme el sistema alcanza mayor madurez. En el principio de la espiral tenemos una madurez más baja y a medida que vamos ciclando en la espiral, debemos alcanzar mayor madurez. Para lograr esto se requiere la integración de principios éticos, buenas prácticas y políticas, de modo que la ética esté presente como un eje transversal en todas las etapas. Con eso nos aseguramos de que la inclusión de criterios éticos no quede exclusivamente en manos del área técnica, ya que este sector a veces puede perder de vista el foco de la ética en las etapas del ciclo de vida del modelo, lo que puede hacer que salgan a producción productos que no estaban testeados correctamente.

Retos éticos y efectos no deseados

Ignorar la ética en el desarrollo de IA puede acarrear serias consecuencias cuando no se incluye de forma integral en los sistemas de IA. Los principales efectos no deseados y riesgos que pueden generarse son:

- Amenazas a la libertad y la responsabilidad individual.
- Discriminación, sesgos y exclusión social.
- Perfilamiento algorítmico indebido.
- Conflictos entre personalización y el bien colectivo.
- Uso de bases de datos masivas sin control o validación.
- Opacidad en los datos usados para entrenar modelos[1].
- Desinformación o falta de transparencia hacia el usuario.

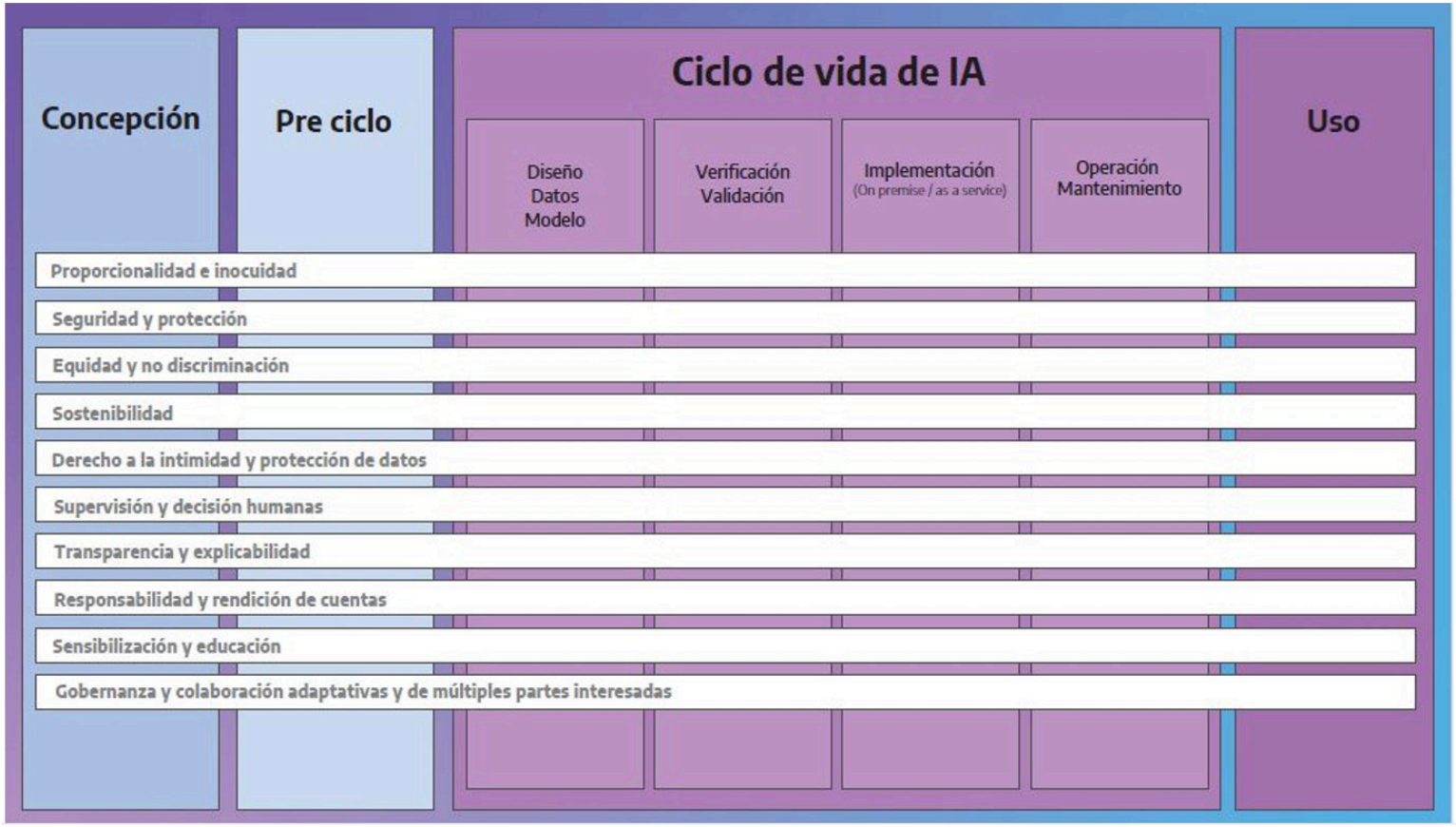
Principios éticos universales según la UNESCO

Para abordar estos retos, contamos una infinidad de principios que pueden ser analizados y discutidos en función de sus alcances y planteamientos. Sin embargo, lo importante es que se apliquen principios éticos que sean universales y sólidos, como los que propone la Organización de Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). Desde la propuesta de esta organización, cuando se crea un modelo/algoritmo, debería de existir un pre-ciclo en el que se detecten los riesgos potenciales y a partir de ahí, se deben aplicar los siguientes principios universales:

1 Advertir a los usuarios qué datos fueron utilizados para entrenar ciertos modelos basados en inteligencia artificial es muy importante y muy pocas organizaciones tienen en cuenta a la hora de comunicar dentro de los servicios que ofrece qué datos se utilizaron para entrenar esos modelos o abrir la fuente de datos para que el usuario final pueda testear o validar esos datos con los que entrenaron al modelo o al algoritmo.

- Transparencia.
- Justicia y equidad.
- Rendición de cuentas.
- Privacidad.
- Sostenibilidad.

Figura 8.2. Aplicación de los principios éticos de la UNESCO en el ciclo de vida de la IA



Fuente: UNESCO. (2022). Recomendación sobre la ética de la inteligencia artificial. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura.

Cada organización, entidad u organismo puede desarrollar sus propios principios, según su contexto, sin embargo, estos siempre deben basarse en normativas y estándares internacionales. Además, no se debe perder de vista que el objetivo de estos principios es que los mismos nos acompañen en cada una de las etapas del desarrollo de una IA.

Evaluación de riesgos en inteligencia artificial

Además de la ética, en el desarrollo de la IA también se deben considerar los riesgos técnicos y de seguridad asociados al uso de IA. Para ello, se puede emplear un *mapa de riesgo* que, basado en cualquier framework de riesgo, permite estimar los riesgos a partir del análisis de amenazas externas que podrían impactar el contexto interno de una organización. En ese caso, debemos preguntarnos si una amenaza externa que no logre ser controlada por mi organización podría afectar los activos clave de mi organización. Es por eso que se deben evaluar los controles y medidas destinadas a mitigar vulnerabilidad y proteger activos como los sistemas de IA.

Tabla 8.1. Clasificación del riesgo

Riesgos para la continuidad y la calidad de las operaciones	Seguridad física: Algunos sistemas con componentes de IA manejan aspectos asociados a la seguridad física (safety), no sólo de bienes tangibles como las mercancías y las instalaciones, sino incluso de las personas y otros seres vivos. Un error en este ámbito puede llegar a poner en peligro la vida.
	Para las operaciones: La mala calidad de la información puede llevar a malas decisiones operativas, así los peligros de una información sesgada o incompleta o inventada (riesgo de alucinación) proporcionada por una IA, puede presentar un grave riesgo a las operaciones e incluso a la continuidad del Negocio.
Riesgos de reputación y cumplimiento legal y ético	Riesgo de discriminación: Daños representativos y de asignación que pueden influir en que se perpetúen los estereotipos y prejuicios sociales.
	Automatización y daños ambientales: La capacitación y operación de LLM requiere mucha capacidad de cómputo, lo que comporta altos costos ambientales derivados del consumo de energía.
	Riesgos legales: Como los derivados de la utilización de datos de carácter personal fuera de las finalidades, o en número desproporcionado o sin respetar la duración acordada, entre otros.
	Riesgo de toma de decisiones no éticas: Si a la IA generativa se da la posibilidad de tomar decisiones aparece el riesgo de que las mismas sean inadecuadas o no éticas.
Riesgo de confidencialidad	Pueden comprometer la confidencialidad al filtrar información clasificada e inferir información confidencial, esto aplica tanto a datos personales como a cualquier otro tipo de información clasificada.

Fuente: Elaboración propia.

Con base a los controles y/o contramedidas que se establecen para determinado tipo de activo, se puede calcular el riesgo, comparando el impacto que podría producirse una vez que se materialice una vulnerabilidad con la probabilidad de que esa amenaza ocurra a lo interno de la organización. En síntesis, en un mapa de riesgo para un sistema de IA puede valorar aspectos como:

- Amenazas externas (no controlables).
- Impacto potencial sobre nuestros modelos o algoritmos.
- Controles y medidas de mitigación.
- Probabilidad de ocurrencia.

Esta metodología permite evaluar el nivel de riesgo que implica cada modelo de IA dentro del contexto organizacional. Para establecer nuestra evaluación del riesgo se pueden utilizar distintos frameworks que nos ofrecen grandes compañías tecnológicas como Mitre. Esta última cuenta con un marco muy robusto en el que se consideran los riesgos que trae aparejado el advenimiento de la IA. Otro de los frameworks que también puede ser utilizado es el de OWASP, el cual suele ser muy utilizado en el ciclo de desarrollo de software y desde el 2021, empezó a orientarse a los modelos de machine learning. Ambos frameworks permiten establecer estrategias de evaluación, mitigación y monitoreo continuo.

Para una gestión responsable de los riesgos en IA, se recomienda:

- Usar cualquier marco de gestión de riesgos y aplicar medidas de control de riesgo para: (i) evaluar y gestionar los riesgos de desplegar IA, incluyendo cualquier potencial impacto adverso para los individuos; (ii) decidir sobre el nivel apropiado de involucramiento en la toma de decisiones ayudada por IA, y (iii) manejar el modelo de entrenamiento de IA y el proceso de selección.
- Hacer mantenimiento, monitoreo, documentación y revisión de los modelos de IA que han sido desplegados, con miras a tomar remediaciones en caso de ser necesarios.
- Revisar canales de comunicación e interacciones con los stakeholders para brindar divulgación y canales de retroalimentación efectivos.
- Asegurar que el personal relevante que lidia con sistemas de IA esté adecuadamente entrenado. Otros miembros del personal cuyo trabajo requiera la interacción con el sistema de IA debe estar entrenado para al menos estar alerta de y sensible sobre los beneficios, riesgos y limitaciones al utilizar IA, para que sepan cuándo alertar a los expertos en la materia dentro de sus organizaciones.

- Evaluar el impacto de los modelos en individuos y comunidades.

- Definir el nivel de autonomía permitido a la IA.

Conclusiones

En conclusión, la ética en inteligencia artificial debe entenderse como un eje transversal que atraviesa todo el ciclo de vida de los modelos y algoritmos, desde su diseño inicial hasta su implementación y mantenimiento. No se trata de un aspecto accesorio, sino de una responsabilidad central que influye directamente en la calidad y la legitimidad de los sistemas desarrollados. La omisión de principios éticos puede derivar en consecuencias negativas que no solo generan daños sociales, sino que también pueden comprometer la credibilidad y sostenibilidad de las organizaciones que implementan inteligencia artificial.

Para mitigar estos riesgos, es imprescindible integrar marcos robustos de evaluación como MITRE ATLAS y OWASP AI/ML Top 10, que permiten identificar amenazas, vulnerabilidades y definir controles efectivos. Asimismo, la transparencia en el uso de datos y en los procesos de toma de decisiones automatizadas debe constituir una prioridad, permitiendo a los usuarios conocer qué datos se han utilizado y cómo se toman las decisiones que los afectan.

A esto se suma la necesidad de formar y capacitar de manera continua a los equipos técnicos en principios éticos y en prácticas de gestión de riesgos. Finalmente, para garantizar que la IA beneficie a la sociedad, es indispensable combinar principios éticos con estrategias efectivas de ciberseguridad y control, los principios propuestos por organismos internacionales como la UNESCO ofrecen un marco sólido que puede ser adaptado por cada organización en consonancia con los derechos humanos y los estándares globales. El desarrollo ético y seguro de la inteligencia artificial es, en última instancia, una construcción colectiva que requiere compromiso, conciencia y acción desde múltiples niveles. América Latina tiene una oportunidad única de adoptar marcos responsables que prioricen la equidad, la sostenibilidad y el respeto por los derechos humanos en esta nueva era tecnológica.

Referencias

Comisión Europea. (2020). White Paper on Artificial Intelligence - A European approach to excellence and trust. Retrieved from European Commission: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST Special Publication No. AI.100-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1

International Organization for Standardization. (en desarrollo). *ISO/IEC DIS 23894 – Information technology – Artificial intelligence – Risk management*. https://www.iso.org/standard/77304.html

International Organization for Standardization. (en desarrollo). *ISO/IEC CD 42001.2 – Information technology – Artificial intelligence – Management system*. https://www.iso.org/standard/81229.html

International Organization for Standardization. (en desarrollo). *ISO/IEC DTR 24368 – Information technology – Artificial intelligence – Overview of ethical and societal concerns*. https://www.iso.org/standard/77412.html

International Organization for Standardization. (en desarrollo). *ISO/IEC DTR 27563 – Impact of security and privacy in Artificial Intelligence*. https://www.iso.org/standard/82738.html

International Organization for Standardization. (en desarrollo). *ISO/IEC AWI 12792 – Information technology – Artificial intelligence – Transparency taxonomy of AI systems*. https://www.iso.org/standard/85932.html

Organización para la Cooperación y el Desarrollo Económicos. (2019). *Principios de la OCDE sobre inteligencia artificial*. https://oecd.ai/es/ai-principles

9 Uso responsable de la IA, una visión latinoamericana

Gloria Guerrero

Gloria Guerrero. Directora ejecutiva de la Iniciativa Latinoamericana de Datos Abiertos (ILDA).

Es Licenciada en Relaciones Internacionales por el Tecnológico de Monterrey y Maestra en Políticas Públicas por la Hertie School of Governance, en Berlín, Alemania. Su trabajo se centra en la intersección de datos, tecnología y derechos humanos. En los últimos 10 años ha trabajado en México, Reino Unido, Alemania, y distintos países de Latinoamérica liderando proyectos de digitalización, gobierno abierto e innovación, así como iniciativas cívicas para promover la participación y la defensa del espacio cívico.

Desde abril de 2023, es Directora Ejecutiva de la Iniciativa Latinoamericana por Datos Abiertos (ILDA), en este rol promueve el desarrollo de vínculos, conocimiento y comunidades para contribuir al desarrollo inclusivo de América Latina. Es también coordinadora de la red Data for Development, D4D, una alianza de organizaciones de investigación del Sur Global, para movilizar conocimiento sobre el uso de los datos y la inteligencia artificial.

Gloria Guerrero

Resumen

Este artículo examina los desafíos y oportunidades que ofrece la inteligencia artificial (IA) para América Latina, bajo una perspectiva que enfatiza los derechos humanos y la ética. Desde esta mirada, se considera que la IA no es una tecnología neutral, ya que sus procesos, decisiones y desarrollos están profundamente marcados por los sesgos y valores de quienes la diseñan; lo que evidencia la necesidad de contar con mecanismos que aseguren rendición de cuentas y transparencia.

Con este propósito y a modo de evidencia, se analiza el Índice Global de IA Responsable (GIRAI). Los resultados de esta medición muestran importantes falencias en las regulaciones y herramientas de políticas públicas para una IA responsable, sobre todo en lo que respecta a la transparencia. Es por eso que, ante tal contexto, se destaca la necesidad de incluir las perspectivas locales y de fomentar la participación de múltiples actores en la regulación y el diseño de estas tecnologías. Además, se reitera la importancia de contar con una gobernanza robusta que fomente el desarrollo de la IA centrado en los derechos humanos, la seguridad, la equidad y la sostenibilidad.

Palabras clave: inteligencia artificial, ética, gobernanza de datos, derechos humanos, América Latina.

Introducción

Hoy la inteligencia artificial (IA) se ha convertido en una tecnología clave para fomentar el desarrollo digital a nivel global; sin embargo, su adopción suscita desafíos de carácter ético, social y político en regiones como América Latina. Solo porque la IA no es una tecnología neutral, sino también porque la apropiación de esta requiere de un proceso crítico y responsable en el que se asegure que: 1) su integración responde a las necesidades locales, 2) se comprende su funcionamiento técnico, y que 3) respete los derechos humanos.

Estos retos se incrementan ante la marcada desigualdad socioestructural que predomina en la región y obliga a valorar cuestionamientos relacionados con el origen, el manejo y la gestión de los datos que alimentan a los sistemas de IA. En este contexto, cuestiones como la transparencia y la cultura de gobernanza de datos deben ser priorizadas para asegurar la protección de derechos y la equidad en el acceso y el uso de la IA.

A partir de esto, el texto analiza tanto las promesas como las implicaciones negativas del uso de la IA en la región, así como las iniciativas orientadas a una gobernanza responsable y ética de estas tecnologías, destacando la orientación plasmada en el Índice Global sobre Inteligencia Artificial Responsable (GIRAI). Los hallazgos de esta herramienta permiten medir y promover la adopción de marcos regulatorios éticos y apuntan a la necesidad de construir un enfoque latinoamericano de la IA que considere las particularidades de la región, coloque en el centro a las personas y sus derechos.

¿Cómo funciona la inteligencia artificial?

Las tecnologías no son neutrales y en el caso particular de la inteligencia artificial (IA) hay que recordar que esta es diseñada por personas que tienen sesgos, ideas, estereotipos y visiones del mundo que son insertadas al momento de crear estas tecnologías. A pesar de esto, no siempre se pueden identificar tan fácilmente dichos sesgos ya que la IA suele operar como una caja negra que toma decisiones basadas en modelos que analizan cantidades enormes de información. Estas cajas negras funcionan con base en algoritmos, datos y la potencia computacional (una infraestructura que permite que estas tecnologías funcionen).

Los algoritmos son como las instrucciones que sigue el sistema para dar una respuesta, mientras que los datos conforman al conjunto de información que se usa para entrenar a esos algoritmos. Los datos son el insumo base con que trabajan estas tecnologías, y debemos entender y recordar que los datos también tienen un valor ético, político y humano. No solo se trata de números o de bases de datos masivas ya que, dichos números representan personas y realidades distintas.

Los chips de procesamiento y de los centros de almacenamiento son la parte física de la tecnología, los cuales demandan recursos como agua, espacio, energía. Este tipo de consideraciones deben ser tomadas en cuenta cuando se aboga por el uso masivo de la IA, ya que eso va a tener un impacto en el planeta y en la cantidad de recursos naturales que se usaran para entrenar a dichas tecnologías.

Las promesas de la IA

Las grandes promesas de la IA se asocian a la capacidad para aumentar la eficiencia, es decir, optimizar el desarrollo de tareas repetitivas, muchas veces de carácter administrativo y que puede liberar tiempo en los equipos de las organizaciones y empresas. Esto permite que se reduzcan costos y faculta el análisis de datos para tener proyecciones más precisas al incrementar el volumen de los datos que son analizados. Dichos modelos pueden mejorar los procesos de toma de decisiones y el manejo de recursos, además representan una oportunidad de negocios para las pequeñas y medianas empresas que siempre se encuentran en desventaja frente a la cantidad de recursos y tecnologías con las que suelen trabajar.

La IA también nos permite un alto nivel de personalización y una mejora en la experiencia de la persona usuaria frente a los servicios públicos. El campo del marketing, está utilizando la IA como nunca, lo que se evidencia en varios de los algoritmos presentes en las redes sociales cuyo nivel de sofisticación ha sobrepasado las predicciones sobre nuestros gustos y preferencias[1]. De igual modo, la IA puede contribuir al desarrollo

1 A pesar de la utilidad, este tipo de funcionalidades también representa un riesgo, pues herramientas con capacidades como estas pueden incidir en los procesos electorales.

de mejores modelos de gobernanza de datos e incrementar la transparencia y la participación ciudadana.

¿Qué implicaciones negativas surgen al utilizar la IA?

Los datos son el elemento esencial con el cual funciona la IA, por lo que toca preguntarnos quién controla dichos datos, a quienes pertenecen, quienes reciben esta información, cómo la analizan, quién decide qué funciona y qué no. Además, ¿qué sucede si un algoritmo se equivoca?, ¿Cómo se asigna responsabilidad?

En estos casos no queda claro si **la responsabilidad** es del programador, la empresa, la plataforma y/o la persona usuaria. Estas interrogantes son las que debemos poner sobre la mesa, sobre todo si se trata de un servicio público que brinda el Estado. En este caso, ¿el responsable es el Estado? Además, esto es un aspecto importante para tomar en cuenta porque las instituciones públicas ya están usando la IA y eso puede tener un impacto directo en los derechos de ciertas comunidades o personas. De hecho, los sesgos de diversa naturaleza (estadísticos, de información y representación) pueden hacer que ciertos beneficios sociales les sean negados a estas poblaciones.

Por otro lado, en cuanto a la **seguridad y la privacidad de los datos**, cabe cuestionarse ¿cómo se puede proteger la identidad de las personas?, ¿cómo gobernamos estos datos que hoy están tomando decisiones sobre nuestras vidas?, y ¿cómo generamos mecanismos de confianza y de uso ético de estas tecnologías para que no reflejen sesgos? Esto es de especial relevancia ya que dependiendo de cómo recolectemos y clasifiquemos los datos, estos pueden o no, contribuir a perpetuar discriminación para ciertas poblaciones.

De igual modo, cuando los sesgos se presentan en aspectos más técnicos, los algoritmos pueden tener errores y generar fallas de forma predeterminada. Sin embargo, si dichas fallas no son corregidas, se corre el riesgo de que continúen estas dinámicas erróneas. Ante esto, surge el desafío de ¿cómo fomentar el uso responsable de la IA en América Latina para que este se alinee a las particularidades de la región como sur global[2]?

Claves para impulsar el uso responsable de la IA en América Latina

La Iniciativa Latinoamericana de Datos Abiertos (ILDA) contribuyó con la iniciativa del Índice de IA responsable o GIRAI

2 El Sur global es una expresión que se refiere a los países que se encuentran regiones (como América Latina, África y Asia) que no pertenecen a los países más desarrollados en materia tecnológica y que sin duda han tenido un crecimiento desigual frente a la innovación.

(Global Index on Responsable AI[3]) en Latinoamérica. Con este propósito, se realizó la recolección de datos para identificar lo qué está pasando en la región, mapear vacíos y avances. Esta información ha llevado al diseño de un modelo de implementación sustentado en 3 dimensiones distintas:

1. Los derechos humanos en la IA: busca que los países adopten medidas que protejan los derechos humanos a la hora de adoptar la IA.
2. La gobernanza responsable: mide la forma como los países establecen herramientas que les ayudan a evaluar los sistemas de gobernanza de la IA mediante estándares éticos, leyes y/o marcos de política pública.
3. Capacidades para adoptar una IA responsable: determina cómo se está avanzado en los compromisos y en las agendas de desarrollo vinculadas a la IA y los objetivos de desarrollo sostenible (ODS). Por tal motivo, se le presta especial atención a cuestiones relacionadas con la igualdad de género y la reducción de pobreza, entre muchas otras cuestiones transversales.

Estas dimensiones no pretenden **medir** el nivel de implementación de las estrategias de inteligencia artificial, sino **la perspectiva de derechos humanos** incluida en estas herramientas. ¿Y qué tanto se ha incorporado esta mirada? Los resultados del índice muestran que aunque muchos países cuentan con una estrategia de IA o están creando nuevas leyes, el tema de la ética y los derechos humanos, no siempre está implícito en las mismas.

3 Corresponde a una medición liderada por un Research ICT Africa con el apoyo del International Research Development Centre and AI4DAfrica.

Asimismo, al evaluar la seguridad y robustez de los sistemas de IA a nivel global, se evidencia un bajo avance, ya que un porcentaje muy bajo de países (29 de 100) cuenta con evidencias reales de que se están desarrollando sistemas de seguridad robustos y confiables. Esto evidencia una deuda en el tema y contrasta enormemente con la tendencia a proteger mayormente a los derechos de propiedad intelectual y a los datos. La protección debe ser entendida como un sistema en el que se tutelen diferentes aspectos y no se ponga en riesgo a las personas cuando se utilizan algoritmos.

También se observa una propensión a establecer regulaciones en áreas específicas, sin embargo, hay una falta de regulaciones que garanticen la transparencia, especialmente cuando los sistemas de IA son adoptados en el sector público. En este caso, son muy pocos los países latinoamericanos que poseen este tipo de regulaciones.

Por otro lado, existen distintas conversaciones a nivel regional sobre cómo entender y apropiarnos con una visión latinoamericana de la inteligencia artificial. Por ejemplo, la Organización de Estados Americanos (OEA) está generando un proceso de alto nivel que, de ser aprobado por los países miembros, concluirá con la publicación de los lineamientos sobre gobernanza de datos y algoritmos para el próximo año. También existen dos declaraciones (la de Santiago y la de Montevideo) que han sido promovidas por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (Unesco) y el Banco de Desarrollo de Fomento de América Latina y el Caribe (CAF), en las que se buscan establecer mecanismos regionales para fomentar el uso ético de la IA.

Conclusiones

Desde ILDA se considera necesario que se utilice el Índice de Inteligencia Artificial Responsable como un insumo para identificar lo que está sucediendo a nivel global, pero también para fomentar el desarrollo de herramientas de explicabilidad que brinden mayor transparencia sobre estas tecnologías y su funcionamiento. Además, sin duda, se necesitan mejorar las regulaciones que afectan el uso de datos por privados, así como aspectos relacionados con la privacidad y la seguridad.

Por otro lado, se debe garantizar la diversidad en los equipos que desarrollan estas tecnologías y que las están implementando, así como pensar en la forma de como integrar las perspectivas multisectoriales con el fin de entender los diferentes riesgos y posibilidades de estas tecnologías.

De igual modo, resulta indispensable que la academia, la sociedad civil y el sector privado se involucren en las discusiones sobre qué es lo que implica esta visión ética de la IA. Asimismo, hace falta ahondar en los detalles específicos que se relacionan con cuestiones como la ciberseguridad, la gobernanza de datos, la interoperabilidad, la inclusión, la transformación digital verde y el uso de la IA en el sector público. Estos temas deben ser el punto de partida para reflexionar el futuro de las tecnologías, desde un enfoque que reduzca los riesgos y garantice el involucramiento activo de América Latina a estos procesos de cambio tecnológico global.

10

Inteligencia artificial, una conversación social no técnica: Perspectivas desde la UNESCO

Natalia González Alarcón

Natalia González Alarcón. Coordinadora de la agenda ética e inteligencia artificial en América Latina de la Organización de Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO).

Economista y especialista en política tecnológica. Más de 10 años de experiencia trabajando en organismos gubernamentales y multilaterales en políticas de innovación, transformación digital, uso ético de datos y tecnologías emergentes. Actualmente se desempeña como coordinadora de la agenda de ética e inteligencia artificial (IA) en América Latina en la UNESCO. Anteriormente, se desempeñó como investigadora en el GovLab de la Universidad de Nueva York (NYU), enfocándose en temas de gobernanza de datos e IA, y trabajó en el Banco Interamericano de Desarrollo (BID) como la coordinadora de la iniciativa fAIr LAC, la primera alianza regional para el uso ético y responsable de tecnología en América Latina y el Caribe. Natalia es economista y máster en Economía de la Universidad de Los Andes, Colombia, y máster en Políticas Públicas de la Universidad de Georgetown.

Natalia González Alarcón

Resumen

Con la rápida expansión de la inteligencia artificial (IA), especialmente de herramientas de IA generativa como ChatGPT, han surgido cuestionamientos con respecto a los potenciales riesgos y retos que trae esta tecnología a nivel social, ético y político. Por tal motivo, los países, la academia, las organizaciones internacionales y multilaterales se han dado a la tarea de responder a estos interrogantes, creando herramientas para abordar los impactos en la creciente integración en diferentes sectores socio-productivos y disminuir riesgos derivados.

Es así como la presente ponencia propone analizar la IA no solo desde una perspectiva técnica, sino también abordarla con un enfoque social, centrado en la gobernanza, la inclusión, los derechos humanos y la sostenibilidad ambiental. Desde esta perspectiva se resaltan los riesgos que la IA genera en contextos desiguales, así como el papel que tiene la regulación y la necesidad de que se promuevan marcos éticos aplicables en la práctica. Finalmente, se presentan iniciativas y herramientas que buscan impulsar una respuesta coordinada desde América Latina.

Palabras clave: Inteligencia artificial, inteligencia artificial generativa, ética, gobernanza digital, desigualdad tecnológica.

Introducción

La rápida evolución de la inteligencia artificial (IA) ha desencadenado una transformación acelerada en distintos ámbitos de la vida pública y privada. Hoy sorprende la velocidad con la que fue adoptada la IA generativa ChatGPT, que alcanzó un millón de usuarios en menos de una semana de haber sido lanzada en noviembre del 2022. Para ponerlo en perspectiva, Instagram demoró 2,5 meses en alcanzar esa cifra, Facebook 10 meses, Twitter 24 y Netflix 41. Además, entre 2017 y 2024 se tuvo un crecimiento en gran escala de las herramientas y sistemas de IA.

Esta ponencia plantea que la conversación sobre IA debe ser comprendida como un fenómeno profundamente social, que requiere respuestas integrales desde los derechos humanos, la equidad de género, la sostenibilidad y la inclusión. Desde la perspectiva de la UNESCO, se busca una aproximación global, colaborativa y responsable frente al desarrollo y uso de estas tecnologías disruptivas.

La exposición de la IA generativa ha sido acompañada por un incremento en las inversiones. Según el *AI Index Report* 2025, la inversión privada en IA generativa alcanzó los 33,9 mil millones de dólares en 2024, lo que representa un incremento del 18,7 % respecto al año anterior y es más de ocho veces superior al registrado en 2022. Actualmente, la IA generativa representa más del 20% del total de la inversión privada en IA, lo que refleja su papel central en la agenda tecnológica glo-

bal. Sin embargo, este dinamismo también supone desafíos éticos, sociales y ambientales.

La expansión de su uso en la toma de decisiones, en muchos casos sin marcos claros de gobernanza ni regulaciones adecuadas de privacidad, pone de relieve la urgencia de establecer mecanismos para mitigar riesgos como los sesgos algorítmicos o los impactos negativos no intencionados. A pesar de esto, el debate sobre la IA a veces se torna predominantemente técnico, ignorando reflexiones sobre las implicaciones sociales, éticas y políticas de integrar la IA en diversas áreas. Por tal motivo, esta ponencia busca reflexionar sobre los retos que se enfrentan en materia de IA desde un enfoque crítico centrado en América Latina y considerando los avances y las desigualdades históricas que afectan el acceso, uso y gobernanza de la tecnología.

Auge e impacto de la IA: beneficios, riesgos y desigualdades

Con la masiva adopción de ChatGPT se generó un punto de inflexión que generó un crecimiento exponencial de nuevos sistemas, especialmente de grandes modelos de procesamiento de lenguaje natural. Empresas como OpenAI, Google, Microsoft y Meta incursionaron en una carrera por impulsar sistemas y aplicaciones de IA que superen a los de sus competidores. Este dinamismo se evidencia en que, entre 2019 y finales de 2024, Microsoft ha invertido más 14.000 millones de dólares en OpenAI, mientras que Meta ha seguido mejorando y lanzando nuevas versiones de su modelo LLaMA.

La competencia por generar modelos de IA cada vez más veloces y precisos está cambiando radicalmente el panorama empresarial e inclusive, la manera como se toman las decisiones estratégicas. Esta tendencia se refleja en una encuesta realizada por Salesforce en 2023, en la que el 67% de los directivos del sector de tecnologías de la información afirmó que planea priorizar la adopción de IA generativa en los próximos 18 meses, y un 33% la considera una prioridad máxima. Sin embargo, el mismo estudio muestra el lado opuesto de la balanza: un 79% de los encuestados manifestó preocupación por los riesgos que esta tecnología podría representar para la seguridad, y un 73% señaló su inquietud por los posibles sesgos en los resultados generados.

Esta dualidad, entre los beneficios y los riesgos de la IA, está atravesada por una dimensión adicional: la desigualdad. La manera en que esta tecnología puede contribuir a cerrar brechas o, por el contrario, profundizarlas depende en gran medida del acceso. Aunque en América Latina la conectividad ha crecido a un ritmo cercano al 8% anual durante los últimos años, aún hay cerca de 230 millones de personas en la región no tienen acceso a Internet. Además, la mayoría de estas conexiones son móviles, lo que impide que se pueda acceder de una forma sostenida a servicios de mayor valor agregado como la telemedicina, la educación a distancia o el trabajo remoto.

A estas desigualdades en el acceso se suman también brechas significativas en el uso de las tecnologías digitales. Un ejemplo claro es la brecha de género: las mujeres y las niñas tenemos un 25% menos de probabilidad que los hombres de saber cómo aprovechar las tecnologías de la información y comunicación (TIC) para actividades que generen valor. Mientras que los hombres suelen usar el Internet para acceder a mercados laborales o servicios financieros, las mujeres tienden a emplearlo en tareas más básicas. Esta brecha digital de género no solo refuerza desigualdades existentes, sino que también puede limitar el aprovechamiento de nuevas tecnologías como la IA, ampliando así las brechas sociales y económicas en lugar de reducirlas.

Finalmente, el crecimiento acelerado de la IA también está teniendo un impacto ambiental significativo. Se estima que para 2026 el consumo energético relacionado con la IA será igual al consumo anual de un país como Bélgica, con un aumento proyectado de hasta 10 veces más que el nivel actual. Sin embargo, estos impactos son desiguales: mientras los centros de datos de Google en países como Finlandia funcionan con un 97% de energías limpias, en Asia los números caen a un rango de 4–18%, utilizando principalmente combustibles fósiles. Este genera impactos diferenciados que suelan afectar de forma más negativa a las poblaciones más vulnerables.

Por lo anterior, puede considerarse que esta aceleración tecnológica está superando la capacidad de los marcos éticos y regulatorios vigentes, poniendo en riesgo que su desarrollo sea seguro, inclusivo y sostenible.

La necesidad de impulsar marcos éticos para la IA

Ante los desafíos mencionados, desde el 2015 se multiplicaron las iniciativas de gobiernos, organizaciones multilaterales y empresas que buscan promover el uso ético de la IA y faltaba un consenso sobre los principios que deberían guiar el desarrollo y uso de esta tecnología. En respuesta de esto, la Organización de Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) impulsó desde el 2021 la *Recomendación sobre la Ética de la Inteligencia Artificial*, la cual fue adoptada por 193 Estados miembros con el fin de crear un instrumento que estableciera principios orientadores -como la transparencia, la responsabilidad, la equidad y el respeto a los derechos humanos- para el desarrollo e implementación de la IA.

Implementar lo establecido en la recomendación constituye el principal desafío, cómo pasar de los principios a la práctica. Pensando en esto, la UNESCO diseñó herramientas como *el Readiness Assessment Methodology* (RAM o *Metodología de Evaluación del Estadio de Preparación*[1]) para analizar el nivel de preparación de los países para un desarrollo y adopción responsable de la IA, y el *Ethical Impact Assessment* (Evaluación de impacto ético) con el que se pretende garantizar la integración de principios éticos en el diseño e implementación de soluciones basadas en IA.

1 Está siendo implementado en más de 60 países a nivel global, y en más de 15 países en América Latina y el Caribe, incluyendo Chile, Colombia, Costa Rica, Perú, México, República Dominicana y Uruguay.

Lo cierto del caso es que hoy ya no se cuestiona si es necesario regular la IA, sino cómo gobernarla de manera efectiva. Este es un debate mucho más amplio que trasciende lo estrictamente normativo y que involucra marcos institucionales, principios éticos, mecanismos de supervisión y coordinación internacional. Diferentes regiones están adoptando enfoques diversos. La Unión Europea (UE), por ejemplo, aprobó la primera Ley de IA basada en los niveles de riesgo, que incluye obligaciones estrictas y sanciones específicas. China ha optado por un modelo estatal centralizado, con un fuerte control gubernamental sobre el desarrollo y uso de la IA. Y ahora Estados Unidos, con su nuevo "AI Action Plan" limita al máximo las regulaciones y promueve la innovación rápida para ganar la carrera de la IA.

A la par de estos enfoques, en todas las regiones del mundo se multiplican los debates y la cantidad de proyectos de ley que buscan regular la IA crecen exponencialmente en los congresos, reflejando una preocupación generalizada por garantizar el uso responsable de esta tecnología.

América Latina como referente en el trabajo regional

A pesar de los múltiples desafíos que enfrenta la región, América Latina ha realizado importantes esfuerzos por consolidar una visión compartida frente a la IA. En alianza con la UNESCO y el Banco de Desarrollo de América Latina y el Caribe (CAF), diversos gobiernos nacionales han celebrado cumbres regionales sobre ética e inteligencia artificial, que han servido como plataforma de diálogo político y técnico. En la segunda de estas cumbres, celebrada en Montevideo en octubre de 2024, se firmó la *Declaración de Montevideo*, un instrumento que no solo consolidó este espacio de trabajo como un mecanismo de coordinación regional, sino que también definió una hoja de ruta con cinco áreas estratégicas: i) gobernanza y regulación, ii) talento y futuro del trabajo, iii) género y diversidad, iv) medio ambiente y v) infraestructura.

Cada uno de estos ejes estableció un conjunto de productos y actividades destinados a crear bienes públicos regionales, tales como herramientas prácticas, marcos de referencia comunes y estrategias para compartir infraestructura y generar capacidades. El objetivo es que estos insumos apoyen a los países de la región en la toma de decisiones y les permitan responder a los retos que plantea la IA.

Conclusiones

La rápida expansión de la IA, en particular de la IA generativa, ha traído múltiples oportunidades, pero también amenaza con exacerbar las desigualdades estructurales, al tiempo que plantea riesgos éticos, sociales y ambientales. Por ello, no debemos limitar el debate a su dimensión tecnológica, porque la IA es, sobre todo, una cuestión social urgente que debe ser abordada para maximizar sus beneficios y mitigar sus riesgos.

Para lograr esto, se debe incluir en las discusiones no sólo a los desarrolladores y empresas tecnológicas, sino también a los gobiernos, la academia, la sociedad civil y todos los sectores de la sociedad.

Solo mediante una participación amplia y diversa será posible construir marcos éticos robustos y mecanismos concretos de gobernanza, adaptados a las realidades locales. La experiencia de la UNESCO propone un enfoque basado en principios éticos, en la participación multisectorial y en herramientas prácticas, buscando abordar los riesgos de forma anticipada y equitativa.

América Latina, a través de sus esfuerzos de cooperación regional, demuestra que es posible avanzar hacia un uso de la IA que respete los derechos humanos, disminuya las desigualdades y promueva el bienestar colectivo. Consolidar una voz latinoamericana fuerte es indispensable para contribuir a la configuración de los modelos globales y garantizar que el desarrollo de la IA se oriente hacia la justicia social, la sostenibilidad y el respeto de los derechos humanos.

11

Regulación de la IA: de expectativas infladas a propuestas factibles

Daniel Rodríguez Maffioli

Daniel Rodríguez Maffioli. Fundador y CEO de Datalex Cyberlaw & Policy.

Daniel es fundador de la firma legal Datalex y consultor internacional en derecho tecnológico, política pública y gestión de riesgos digitales. En los últimos años, se ha posicionado como un referente en inteligencia artificial (IA), ciberseguridad y privacidad en la región, en sus roles de consultor para UNESCO, la OCDE y el Banco Interamericano de Desarrollo (BID). Daniel cuenta con una Maestría en Derecho de la Tecnología y Propiedad Intelectual de la Universidad de Duke (EE.UU.), una Maestría en Regulación Económica de la Universidad Carlos III de Madrid; y es Licenciado en Derecho por la Universidad de Costa Rica.

Daniel Rodríguez Maffioli

Resumen

En los últimos años, el debate sobre la regulación de la inteligencia artificial (IA) ha crecido, aunque sin cuestionar a fondo las razones, las formas y la eficacia real de una potencial regulación. La regulación debe ir más allá de aspiraciones normativas y enfocarse en intervenciones prácticas que orienten el desarrollo y el uso de la IA hacia el interés general. La forma tradicional de legislar, mediante leyes rígidas y de lenta modificación, que presuponen el efecto de su implementación sin datos que lo respalden, resulta inadecuada frente a una tecnología tan dinámica, compleja y de impacto transversal como la IA.

Regular la IA implica múltiples retos: entre ellos, su evolución acelerada, la opacidad de sus algoritmos, la incertidumbre de sus efectos, el uso dual de muchas aplicaciones y la falta de capacidad institucional. Estos desafíos dificultan diseñar normas efectivas y sostenibles en el tiempo. Un ejemplo es la Ley de IA de la Unión Europea (AI Act) que, aunque pionera, ha sido criticada por sus posibles efectos colaterales en la innovación y el desarrollo tecnológico. La experiencia europea evidencia la tensión que se genera entre proteger los derechos fundamentales y no obstaculizar las fuentes de bienestar y progreso prometidas por la IA.

En América Latina, y especialmente en Costa Rica, los esfuerzos regulatorios se encuentran en etapas iniciales, influenciados por modelos europeos, pero sin considerar a fondo las condiciones locales. Los proyectos de ley costarricenses carecen de armonía con otras regulaciones y no incluyen medidas de fomento ni un análisis realista de su aplicabilidad. Por eso, se recomienda establecer marcos adaptativos, basados en la gestión de riesgos y en estándares internacionales, así como fomentar la experimentación e innovación regulatoria, eliminar requisitos innecesarios y fortalecer las capacidades locales a través de órganos técnicos especializados y políticas de impulso a la innovación en IA.

Palabras clave: inteligencia artificial, sistemas de inteligencia artificial, riesgo, regulación, regulación inteligente basada en evidencia.

Introducción

A pesar de que en los últimos años se han incrementado las discusiones sobre los alcances e impacto que puede ocasionar la regulación de la inteligencia artificial (IA), en estas no siempre se analiza detenidamente la verdadera eficacia que podría tener una regulación de esta tecnología. Esto ocurre porque, en general, las discusiones han tendido a centrarse en cuestiones superficiales que se limitan a establecer obligaciones éticas del tipo "deber-ser" o a responder a imperativos categóricos sobre determinados efectos deseables que pretendemos inducir en nuestras sociedades. Dicha orientación contrasta con las características únicas y esenciales de la IA, especialmente su capacidad de aprendizaje autónomo y las crecientes y emergentes habilidades de razonamiento que han comenzado a evidenciar.

Poco importa el "deber-ser" cuando el "ser" -es decir, la realidad social- ya no está configurada por el actuar humano, sino por máquinas inteligentes e impredecibles con agencia propia. Por lo tanto, crear leyes mediante procesos legislativos clásicos, que presumen un resultado poco realista, y sin un análisis profundo de los efectos inmediatos y futuros de la regulación, constituye un enfoque insuficiente y, en muchos casos, contraproducente ante un fenómeno dinámico y radicalmente transformador como la IA.

¿Por qué ocurre esto? Este y otros interrogantes se abordan en los siguientes apartados, con el propósito de ofrecer una reflexión crítica sobre la forma y el alcance que deberían guiar el diseño regulatorio en materia de inteligencia artificial. Finalmente, el artículo plantea una serie de recomendaciones para mejorar los procesos regulatorios en este campo.

Claves para repensar la regulación en IA

Un primer paso fundamental al construir una regulación para la inteligencia artificial consiste en cuestionar las premisas que a menudo se asumen como ciertas. Así, resulta necesario plantearse preguntas como: ¿*debe* regularse la IA?, ¿*por qué* debería regularse?, ¿*puede* efectivamente ser regulada?, ¿*cómo* debería hacerse? Estas preguntas, a las que se le han dado respuestas prematuras e inconsistentes con la realidad, requieren un análisis más profundo.

Por ejemplo, aunque existe consenso sobre algunos riesgos de la IA, es pertinente preguntarse si se han examinado adecuadamente sus efectos colaterales, como la profundización de desigualdades o los sesgos algorítmicos que pueden derivar en nuevas formas de discriminación. Asimismo, resulta necesario valorar si el desarrollo, el uso e implementación de los sistemas de IA debe ser objeto de normas específicas, y en qué medida estas normas deben proteger valores fundamentales como los derechos humanos y la democracia.

La creciente delegación de decisiones en sistemas automatizados es fuente de preocupación, pues se argumenta que ciertas decisiones en la vida de las personas podrían requerir un nivel de supervisión humana. También, debido a los riesgos asociados al uso indebido de la tecnología, como la manipulación de información mediante *deepfakes*, la apropiación no autorizada de identidades digitales o los impactos sobre los mercados laborales a través de la sustitución de puestos de trabajo.

Estos fenómenos reflejan fallos de mercado y efectos colaterales que una eventual regulación podría intentar corregir. Regular la IA, por tanto, no debe ser un fin en sí mismo, sino un medio para neutralizar riesgos, corregir desequilibrios y orientar el desarrollo tecnológico hacia el interés general y la defensa de los valores humanos.

Figura 11.1. Impactos ético -legales de la IA

Fuente: Elaboración propia.

Sin embargo, regular la IA no resulta una tarea sencilla ya que las propias características de esta tecnología condicionan la forma como podemos regularla. Entre los desafíos se encuentran aspectos como:

- La **acelerada evolución de la tecnología**, lo que puede generar obsolescencia regulatoria y poca certeza con respecto a qué se debe regular.

- El **impacto transversal** ocasionado por la diversidad de aplicaciones existentes y los distintos efectos que provocan. La IA no es un campo reservado a las ciencias informáticas y computacionales. Es un fenómeno sociotécnico, filosófico y moral que impregnará y afectará todos los aspectos de la vida, al punto de llegar incluso a entremezclarse con el cuerpo humano a través de la biotecnología y la neurotecnología.

- La **aplicabilidad real en el contexto operativo**. Por ejemplo, si una norma exige a un desarrollador o implementador de IA que su tecnología explique las razones concretas por las cuales una IA actuó de cierta forma o recomendó un determinado curso de acción, dependiendo del tipo de modelo de IA, podría ser imposible dar esas razones, debido a la arquitectura algorítmica profunda y opaca bajo la cual se construyen estos modelos.

- La **carencia de entes reguladores capacitados y con recursos**. Sin una institucionalidad alfabetizada y con suficientes recursos, capaz de aplicar y verificar el cumplimiento de las normas, la regulación se convierte en una aspiración utópica sin razón de ser.

- La **naturaleza distribuida y global de las aplicaciones de IA**. Lo que Costa Rica haga para regular la IA podría no depender completamente del país. Por un lado, la infraestructura de centros de datos que soportan esta tecnología, no están domiciliados en el país. Además, el internet es un espacio abierto y distribuido globalmente, con agentes maliciosos anónimos también distribuidos, por lo que perseguir delitos cometidos a través del uso de IA desde internet, es altamente complejo.

Debe tenerse claro que la IA es una tecnología distinta a lo que conocemos hasta ahora, pues sus potencialidades representan un cambio en el modelo de producción e incluso, plantea un nuevo modo de organización social y laboral. Su desarrollo acelerado, dinámico y cambiante, junto con sus capacidades de autonomía, dificultan que a la misma se le pueda atribuir responsabilidad y/o que pueda ser sujeta a rendición de cuentas, por la compleja cadena de valor involucrada.

También, la IA tiene capacidad de uso dual (por ejemplo, el reconocimiento facial puede usarse para identificar a un delincuente responsable de una violación, pero también para vigilar masivamente a personas inocentes).

Al lado de estos desafíos, la pregunta clave es ¿qué se debería regular? Recientemente, la Unión Europea (UE) aprobó su Ley de Inteligencia Artificial ("AI Act"), una propuesta que ha ganado gran popularidad, por plantear mayor protección de los derechos humanos, aunque a costa de limitar la posibilidad real de generar innovación tecnológica desde Europa. Una norma como la Ley de Inteligencia Artificial de la UE evidencia el tipo de encrucijadas a las que nos enfrentamos a la hora de regular la IA. Toda decisión implica la renuncia a un tema u otro y justo este tipo de discusiones son las que se deben tener a nivel regional y por supuesto, a lo interno de los países.

A su vez, la evolución acelerada de esta tecnología genera el problema de la **obsolescencia regulatoria**, es decir, la pérdida sobrevenida de la eficacia de una ley dentro de un corto periodo de tiempo. Esto ocurrió con la *Ley de Protección de la Persona frente al Tratamiento de sus Datos Personales* (Ley N°8968) de Costa Rica, que fue aprobada en 2011 y sigue vigente hasta la actualidad, pero con limitada eficacia práctica.

Por otro lado, el impacto transversal y multisectorial de la IA provoca una dificultad adicional, ya que es retador diseñar un único conjunto de reglas, que sea igualmente efectivas y proporcionales para cada sector y contexto específico en donde se implemente. Asimismo, debe tomarse en cuenta la aplicabilidad real de ciertas medidas/exigencias legales cuando un sistema o aplicación de IA se encuentra en pleno funcionamiento, ya que constantemente está aprendiendo de los datos que se le alimentan y del entorno donde se despliega, con lo cual, sus parámetros internos están cambiando constantemente.

Todo lo anterior puede llevar a expectativas irrazonables sobre el alcance de la regulación e incluso brindar una falsa sensación de seguridad luego de aprobar una ley para la IA.

Las capas de la IA frente a la regulación

Al pensar en la inteligencia artificial, así como en la "necesidad de regularla" se deben tomar en cuenta las capas que

componen a esta tecnología, pues en cada una de estas es donde podrían gestarse algunos de los riesgos éticos, morales y sociales que se desean mitigar. Desde esta perspectiva, existen al menos 3 capas distintas:

- Infraestructura: que comprende a todos los elementos necesarios para generar el poder de cómputo necesario para que pueda operar un sistema/aplicación de IA. Incluye a los semiconductores (microchips que funcionan como unidades de procesamiento gráfico con capacidad de procesar datos en microsegundos y enviar una respuesta en un tiempo similar), los data centers (que albergan a todas las computadoras que almacenan y procesan la información) y los servicios en la nube.

Algunos cuestionamientos en este ámbito, por ejemplo, tienen que ver con la cantidad de energía y agua que consumen los data center de IA. Por lo tanto, algunos países y sectores proponen generar regulaciones de IA que exijan el cumplimiento de ciertos criterios de eficiencia energética.

- Datos: refiere a los datos que se utilizan para entrenar a los modelos de IA. Son la materia prima de la IA y permiten realizar ajustes a modelos que han sido preentrenados, así como para fines de I+D y para la implementación de sistemas de IA.

En esta capa se pueden experimentar problemas asociados con los derechos de autor o la protección de datos personales, ya que, a la hora de entrenar modelos de IA, las empresas pueden hacer uso de materiales protegidos por los derechos de autor o por las normas de privacidad. Actualmente, empresas como OpenAI se enfrentan a demandas por haber supuestamente incurrido en este tipo de prácticas.

- Aplicaciones: representan el último nivel de la IA y es la capa más cercana a las personas y usuarios, porque es la que les permite entrar en contacto con los sistemas de IA para realizar tareas o funciones específicas. Ejemplos de aplicaciones las encontramos en softwares como ChatGPT y Copilot o en potencialidades que permiten el Internet de las Cosas (IoT), los vehículos autónomos, entre otros.

En esta capa pueden surgir otro tipo de inconvenientes, por ejemplo, chatbots que generan contenido peligroso, discursos de odio o manipulación. También se pueden presentar casos de accidentes y daños materiales cuando la IA es incrustada en medios físicos o hardware, como podría ser un accidente generado por un vehículo autónomo apoyado en IA o una máquina industrial.

Gobernanza mundial de la IA

Sobre la base de estas reflexiones, cada región y país ha optado por una forma de regular la IA. Por ejemplo, Estados Unidos ha abogado por una regulación que favorezca más la

innovación con la promulgación de regulaciones sectoriales y basadas en estándares técnicos, así como en parámetros que garanticen la seguridad por parte de las empresas que desarrollan, gestionan e implementan sistemas de IA. Adicionalmente, se incluye el tema de la gestión de riesgos (de seguridad y discriminación, entre muchos otros). Podemos ejemplificar este enfoque en las regulaciones destinadas al sector financiero y a las normas que rigen la protección de las personas consumidoras[1].

Figura 11.2. Mecanismos para la gobernanza mundial de la IA

Fuente: Elaboración propia.

En el caso europeo, la Ley de Inteligencia Artificial de la UE se puede considerar como la primera regulación integral de la IA en el mundo. En ella se prioriza la protección de los derechos fundamentales y se adopta un enfoque basado en riesgos en el que se establece una clasificación que va desde riesgos inaceptables hasta los aceptables. En línea con esta orientación, la norma crea numerosos requisitos para la IA de alto riesgo (entre las cuales se exigen documentación y sistemas de gestión de la calidad y gestión de riesgos), procedimientos de transparencia y supervisión, junto con un modelo de gobernanza de la IA que se apoya en los reguladores nacionales y europeo. Además, ante potenciales incumplimientos a las disposiciones de la ley se incluyen cuantiosas sanciones.

A pesar de que la Ley de IA de la UE prioriza la protección de los derechos fundamentales, se considera que su cumplimiento tendrá un alto costo, sobre todo porque puede crear expectativas potencialmente irrealizables, al tiempo que podría desincentivar la innovación, la investigación y desarrollo (I+D) y afectar especialmente a Pymes y Startups del segmento de la IA (*Ver Informe Draghi*[2]).

Y en América Latina, ¿qué está pasando? Por el momento, en la región los países han propuesto múltiples proyectos de ley que pretenden regular la IA[3], siendo Perú y El Salvador los únicos que tienen leyes vigentes que regulan la IA de forma integral. Además, se observa una marcada influencia de la Ley de Inteligencia Artificial de la UE y aunque se reconoce el tema como relevante en la región, en los esfuerzos regulatorios rea-

1 En este último caso, este enfoque permitirá que el diseño de los chatbots de IA sea más seguro para las personas menores de edad.

2 Para profundizar en el tema se recomienda ingresar a: https://commission.europa.eu/topics/eu-competitiveness/draghi-report_en

3 Para noviembre del 2024, se estima que en la región había de entre 35 y más de 40 proyectos de ley en América Latina.

lizados se identifica poca consideración hacia la realidad, las necesidades y las capacidades locales.

¿Y qué ha pasado en Costa Rica?

Hasta el 2024 el país contabilizaba tres proyectos de ley en corriente legislativa (ver tabla 11.1.). En estos se evidencia la influencia de la Ley de Inteligencia Artificial de la UE. A pesar de eso, los proyectos de ley no han sido adaptados a la realidad costarricense, y en algunos casos tampoco responden a los estándares internacionales.

Los tres proyectos de ley en corriente legislativa muestran **poca armonía con los demás proyectos de ley existentes** en las áreas de la protección de datos, la ciberseguridad y los servicios digitales. En esta línea debe considerarse que tampoco hay reformas sectoriales que incidan en la protección a la persona consumidora, la propiedad intelectual, el código penal y/o la ley de acceso a la información pública. De igual modo, estas iniciativas de ley reflejan la **falta de un análisis del impacto regulatorio**, lo que hace evidente que no hay claridad sobre qué es lo que se desea regular, si la regulación propuesta será aplicable en la práctica y si habrá costos implicados (y de qué índole) o beneficios (Rodríguez-Maffioli, 2025).

Otro detalle relevante que debe señalarse es que las tres propuestas de norma **incluyen autorizaciones, estudios o registros previos para usos de IA de "alto riesgo" o de "interés público",** cuestión que ni siquiera es contemplada en la Ley de Inteligencia Artificial de la UE. Asimismo, las tres propuestas tampoco integran medidas de fomento para incrementar los factores clave que inciden en la IA como la capacidad de cómputo, la gobernanza, el acceso a datos locales, la educación, la re-capacitación de la fuerza laboral y la promoción de emprendimientos y el uso de IA en las Pymes (que constituyen la mayor parte del parque empresarial del país).

Aunado a esto, la mayoría de los proyectos de ley contienen definiciones que no se basan en estándares internacionales e incluyen conceptos que resultan innecesarios. A esto se le debe prestar especial atención porque una mala definición puede ser atroz (Rodríguez-Maffioli, 2025).

Tabla 11.1. Proyectos de ley en IA presentados en la Asamblea Legislativa de Costa Rica hasta enero del 2025

Número de expediente legislativo	Nombre del proyecto	Observaciones
23 711	Ley de Regulación de la Inteligencia Artificial	-Brinda un marco general adecuado basado en principios generales, derechos fundamentales y en uso sectorial. -Fija objetivos generales, pero da libertad a las empresas para que puedan cumplirlos. -Delega ciertos criterios técnicos a la Sutel, pero esto implicará la necesidad de darle presupuesto y dotarlo de equipo humano especializado. -Se solicita una evaluación de impacto previa para sistemas de IA con alto impacto en los derechos fundamentales, equidad y seguridad, aunque no se concreta, ni se aclara que el estudio fuera solo para sistemas de IA de alto impacto.
23 919	Ley para la Promoción Responsable de la Inteligencia Artificial en Costa Rica	-Texto comprensivo, ordenado y basado en los derechos humanos y la gestión de riesgos. -Se proponen mecanismos de experimentación regulatoria para regular con base a la experiencia y no a partir de suposiciones. -Es excesivamente detallado y reglamentista, lo cual fomenta la obsolescencia regulatoria. -Corre el riesgo de que la regulación se vuelva obsoleta en poco tiempo y que no sea un marco adaptable.
24 484	Ley para la implementación de Sistemas de Inteligencia Artificial	-Establece obligaciones de transparencia en favor de las personas usuarias de los sistemas de IA, la protección de datos personas y la gestión de riesgos que deben tener las empresas al usar IA (aunque lo menciona de manera limitada). -Limita indirectamente el acceso a datos para start-ups y otras empresas pequeñas que quieran innovar con IA. Además, se inclina a favorecer a los titulares de los derechos de autor. -Se usa el concepto de autorización previa para usos en "zonas de impacto primario" (ZIP), lo que no se alinea con los estándares internacionales en la materia. La lista de ZIP incluida es larga. -Prohíbe el uso de obras para entrenar sistemas de IA, sin excepciones por investigación aplicada, usos justos como la parodia, la crítica social y la comedia. Genera un roce con la libertad de expresión.

Fuente: Elaboración propia con base a Rodríguez-Maffioli, 2025.

Para solventar las falencias presentes en estas propuestas de ley se recomienda (Rodríguez-Maffioli, 2025):

1. Eliminar cualquier obligación de registro o autorización previa, salvo usos de la IA en el sector público (debido a principio de legalidad, transparencia u otros).
2. Establecer un marco general que se base principios éticos y en exigir a las empresas que usan IA en sectores o actividades de alto riesgo, que posean un sistema de gestión de riesgos adaptable a la realidad de cada empresa y cada tipo de uso y basado en estándares internacionales (así lo exige la diversidad de aplicaciones y usos, así como la dualidad en el uso de los sistemas de IA).
3. Crear mecanismos de experimentación e innovación regulatoria, como los "sandbox regulatorios" y prototipos de regulación, que permiten construir regulaciones y directrices a partir de experimentos y datos empíricos. Los reguladores sectoriales (SUGEF, SUTEL, ARESEP, Ministerios a cargo de sus sectores, entre otros) pueden prototipar y poner a prueba posibles regulaciones en el mercado, antes de consolidarlas, lo que permite realizar ajustes conforme la tecnología avanza.
4. Revisar las leyes existentes y realizar reformas puntuales para adaptarlas a la IA, bajo la premisa de que las leyes existentes son igual de aplicables para el mundo físico y digital, independientemente del uso de IA.
5. La regulación debe ser adaptativa y en ello puede ayudar el establecimiento de un Consejo Técnico en el Ministerio de Innovación, Ciencia, Tecnología y Telecomunicaciones (Micitt), integrado por diferentes sectores y expertos, encargado de promover lineamientos, hacer investigaciones y dictar políticas y reglas conforme cambia la tecnología. Los "Institutos de IA" promovidos por Estados Unidos y el Reino Unido son ejemplos a seguir.
6. Es necesario que una ley de IA contemple medidas de fomento y habilitación de la IA y que preferiblemente asigne recursos para la ejecución de la Estrategia de IA del país, con el fin de introducir la IA en todos los niveles de la educación costarricense, así como en programas de recapacitación laboral. Costa Rica debe lograr que la mayoría de su población tenga un grado básico de alfabetización y apropiación en inteligencia artificial.

Conclusiones

Ante este panorama, ¿cuál debería ser la ruta que tomemos para regular la IA? Primeramente, se requiere que cualquier regulación de IA sea una **regulación inteligente, adaptativa y basada en evidencia**, que combine medidas de fomento y de protección.

Las primeras deben centrarse en acciones que promuevan la formación de habilidades y la apropiación de la IA, al tiempo que potencien la recapacitación laboral, la adopción de la IA en Pymes y el sector empresarial. En esta línea, también se requiere mejorar el acceso a datos de calidad que permitan el entrenamiento de los sistemas de IA y la investigación y desarrollo (I+D) en esta materia.

Por su parte, las medidas de protección deben incluir intervenciones como la auto-regulación regulada, la estandarización y la gestión de riesgos, el fortalecimiento de las normas sectoriales y de los reguladores sectoriales. Junto con esto, se debe pensar en la delegación de potestades en los reguladores sectoriales y en la posibilidad de establecer un Comité Técnico Asesor multisector a lo interno del MICITT.

La inteligencia artificial trae consigo uno de los mayores cambios de paradigma en la historia de la humanidad. No es posible seguir usando recetas viejas, rígidas y burocráticas para regular un mundo nuevo, dinámico e incierto.

Referencias

Rodríguez-Maffioli, D. (29 de enero de 2025). Comparecencia ante la Asamblea Legislativa. Proyectos de Ley Regulación de la Inteligencia Artificial en Costa Rica (No. 23.771), Ley para la Promoción Responsable de la Inteligencia Artificial en Costa Rica (23.919) y Ley para la Implementación de Sistemas de Inteligencia Artificial (IA) (24.484). Asamblea Legislativa de la República de Costa Rica.

12 Balance entre la inteligencia artificial y los derechos de autor

Fabián A. Gamboa Hernández

Fabián Antonio Gamboa Hernández. Profesional legal de la Unidad de Gestión y Transferencia del Conocimiento para la Innovación.

Licenciado en Derecho por la Universidad de Costa Rica. Ha trabajado en la Defensa Pública del Poder Judicial y, más recientemente, en la Asesoría Legal de PROINNOVA desde 2023, espacio desde el cual ha cursado distintos seminarios impartidos por la Universidad de Panamá y el Instituto Mexicano de la Propiedad Industrial (IMPI), además de haber completado cursos de la Organización Mundial de la Propiedad Intelectual en materia de derechos de autor, patentes, marcas y obtenciones vegetales.

Fabián A. Gamboa Hernández

Resumen

La siguiente ponencia aborda el vínculo y tensiones que han surgido entre la inteligencia artificial (IA) y los derechos de propiedad intelectual, para lo cual se plantea un análisis que se centra en los derechos de autor. Desde esta perspectiva se cuestiona si las obras producidas por los sistemas de IA pueden ser considerados como creaciones susceptibles de ser protegidas por la regulación en propiedad intelectual y si la IA puede ser considerada como autora. En esta línea, se examinan las posibles implicaciones legales que supondría otorgar tal reconocimiento a partir de casos y doctrina que ilustran diferentes posturas sobre el tema.

Seguidamente, se explora la noción anglosajona del *fair use* comparada con la regulación latinoamericana de *usos honrados*, con el fin de discutir la legalidad que puede tener la utilización de contenidos protegidos por derechos de autor sin contar con la autorización previa de los autores originales cuestionando, con ello, qué se entiende por la persona autora desde el ámbito jurídico.

Adicionalmente, se presenta el concepto de la *ultra suplantación* y se diferencia del concepto tradicional del *plagio*, con el fin de visibilizar los desafíos éticos y normativos que el uso de la IA puede generar para el ámbito educativo.

Palabras clave: Inteligencia artificial, propiedad intelectual, derechos de autor, fair use, ultra suplantación.

Introducción

La Inteligencia Artificial ha sido uno de los avances tecnológicos más disruptivos que se ha presenciado en los últimos años, por lo que resulta necesario estudiar -desde una óptica jurídica- los alcances e implicaciones que tiene el uso de esta tecnología frente a los derechos de autor, en virtud de que este ha sido y seguirá siendo un tema de debate en el ámbito de la propiedad intelectual.

Este análisis aborda la relación que hay entre la Inteligencia Artificial (IA) y la protección que concede el ordenamiento jurídico a las obras artísticas y literarias, considerando si las creaciones generadas por estos algoritmos constituyen una obra protegida y quién sería el titular de los derechos generados por la IA. Además, se aborda el concepto del plagio, debido a que hay discusiones sobre la conducta del plagio por el simple uso de la IA, especialmente en instituciones educativas.

A partir del estudio de algunos casos recientes y del ordenamiento jurídico sobre los derechos de autor, se hace un abordaje sobre la intersección que hay entre la tecnología y el derecho, considerando los desafíos que enfrenta la regulación actual para adaptarse a los vertiginosos avances tecnológicos actuales.

Hacia un concepto de la Inteligencia Artificial

En primer lugar, es necesario referirse a un posible concepto de Inteligencia Artificial y determinar la relación que tiene con los derechos de propiedad intelectual. Históricamente, ha sido el ser humano quien ha creado aquellas obras susceptibles de propiedad intelectual, por lo que surge la duda de si realmente lo que produce la Inteligencia Artificial es protegible por derechos de autor.

De acuerdo con una definición de IA proporcionada por ChatGPT -obtenida tras consultar qué se entiende por IA-, se indicó lo siguiente:

La inteligencia artificial (IA) es una rama de la informática dedicada a crear sistemas y algoritmos que pueden realizar tareas **que normalmente requieren inteligencia humana**, como el reconocimiento de patrones, el aprendizaje, la toma de decisiones y el procesamiento del lenguaje natural. Mediante técnicas como el aprendizaje automático, la IA permite a las máquinas mejorar su rendimiento en tareas específicas a partir de grandes volúmenes de datos, sin necesidad de ser programadas explícitamente para cada caso. Existen diferentes tipos de IA, desde la débil, que se especializa en una función limitada, hasta la IA general, **que aspiraría a tener una capacidad cognitiva similar a la humana, aunque aún no se ha alcanzado** (OpenAI, 2024) (destacado intencional).

Esta aspiración hacia capacidades cognitivas similares a las humanas constituye un elemento central para el análisis de la autoría y la titularidad de derechos sobre obras generadas mediante IA, lo cual está intrínsecamente vinculada con aspectos normativos, consideraciones éticas y desafíos en materia de propiedad intelectual.

Cabe señalar que la denominación "artificial" resulta paradójica considerando la metodología de entrenamiento de estos sistemas, debido a que los modelos de IA no poseen capacidad innata de razonamiento, sino que procesan extensos conjuntos de datos preexistentes para generar nuevos contenidos mediante instrucciones específicas ("prompts"). Este proceso de entrenamiento ha generado críticas respecto a lo que algunos autores denominan "plagio de plagios", argumentando que la IA fundamentalmente reconfigura contenido previamente creado por terceros.

En este sentido vale la pena mencionar la posición de Rafael Ángel Herra, en un comentario a las afirmaciones de Noam Chomsky sobre este planteamiento:

> Frente a la universal rapiña que constituye la cultura -rapiña espontánea, necesaria, inevitable en el desarrollo del saber- hay un punto de distinción: si un colega me toma prestada, sin citar mi nombre, una idea o teoría nueva que no se ha fijado por escrito antes de mi invención, o se apropia de un invento técnico o de un descubrimiento como el de la electricidad, estamos en presencia monda y lironda de plagio en sentido duro de rapto,

bajo la sombra de la ley penal.

¿Plagió Edison a Tesla? No sé responder, pero el asunto flota en el ambiente. Para prevenir el robo de inventos precisos, en nuestra época existen las condiciones legales que garantizan la paternidad y los derechos de reproducción. (Herra, 2024).

A partir de lo anterior surge una interrogante principal: ¿constituye la generación de contenido mediante IA, sin la correspondiente citación de las fuentes originales, una infracción a los derechos de autor? Su respuesta debe considerar la estructura tradicional de la propiedad intelectual, que comprende: (1) la propiedad industrial (patentes, marcas, dibujos y modelos industriales); (2) los derechos de autor y derechos conexos que protegen las creaciones literarias y artísticas; y (3) las protecciones sui generis (obtenciones vegetales, secretos industriales) que son aquellas que no se encuadran en las categorías tradicionalmente conocidas.

En este caso, los derechos de propiedad intelectual a los que debe hacerse referencia es a los derechos de autor, pues son aquellos que reconocen a quienes hayan creado obras literarias o artísticas y que reciben protección por la ley, pues -por el momento- la IA ha tenido una fama especial por la generación de textos e imágenes, que tradicionalmente solo era realizada por seres humanos.

Prerrogativas concedidas por los Derechos de Autor

A estas alturas vale la pena cuestionarse ¿cuáles son las prerrogativas que conceden los derechos de autor?, para lo cual se han reconocido: (1) los derechos morales, que son el vínculo estrecho que hay entre el autor y su obra creada y que se materializa en una serie de derechos que incluyen "el derecho a: a) mantener inédita la obra, b) la reivindicación de la obra o derecho de paternidad, c) oponerse a la deformación, mutilación u otra modificación de la obra y d) retirar la obra de circulación" (Miranda, 2014, pp. 32-33) y, por otro lado, (2) los derechos patrimoniales, que son aquellos derechos de explotación económica que tiene el autor en relación con su obra y que puede disponer de ellos como el derecho de reproducción, el derecho de distribución, los derechos de transformación o adaptación de la obra y el derecho de comunicación pública, que se reconocen en el artículo 16 de la *Ley sobre Derechos de Autor y Derechos Conexos*, Ley No. 6683 del 04 de noviembre de 1982.

En este artículo, hay dos derechos específicos que plantean desafíos frente a la inteligencia artificial (IA): el **derecho de paternidad** plantea la cuestión de la titularidad de una obra y quién la generó. Esta problemática es especialmente relevante en cuanto a la determinación del titular al que le pertenece una obra generada por la IA, pues hay discusiones que consideran a la IA como autora de obras y, por tanto, susceptibles de ser plagiadas. En Costa Rica, se han reportado casos en los

que estudiantes han sido acusados de plagio por utilizar inteligencia artificial en trabajos académicos; sin embargo, vale la pena preguntarse si la obra generada por la IA está protegida por derechos de autor y si la IA puede considerarse autor en los términos de la normativa aplicable al respecto, por lo que surge la siguiente pregunta que se responderá más adelante: ¿puede jurídicamente considerarse una obra generado por una IA?, y ¿quién sería el autor en ese caso?

En cuanto a los **derechos patrimoniales**, el derecho de reproducción otorga al autor la capacidad de decidir cómo se utiliza su obra, por lo que esto incluye la facultad de oponerse a reproducciones indebidas de la obra. Esto es particularmente relevante en el entrenamiento de la IA, pues estos modelos se construyen a partir de datos y obras creadas previamente debido a que carecen de una capacidad innata para razonar. Por tal motivo, algunas empresas han utilizado obras protegidas por derechos de autor para entrenar modelos de IA, algo que se ha justificado bajo el concepto del *"fair use"*, un criterio jurisprudencial estadounidense que permite el uso de ciertas obras, como una excepción a la protección, cuando se justifica por el fin o el propósito para el cual se usa la obra, la naturaleza de la obra, si existe o no una copia sustancial de la obra original, y por último, si hay un impacto en el mercado.

Datos de entrenamiento y el *fair use*

A nivel internacional, el *Convenio de Berna para la Protección de las Obras Literarias y Artísticas* de 1886, adoptado por la legislación costarricense mediante la Ley No. 6083, establece el concepto de "usos honrados", que otorga a los países la facultad de permitir la reproducción de obras en casos especiales, siempre que no se afecte la explotación normal de las mismas ni se cause un perjuicio injustificado a los intereses legítimos del autor. De esta manera, se contrasta la visión estadounidense del fair use con la perspectiva latinoamericana de los "usos honrados".

Un ejemplo pone de manifiesto la relevancia de estas discusiones, pues una reciente noticia titula: *M*ás de 15 mil creadores han firmado una declaración en contra del uso de sus obras para entrenar sistemas de *Inteligencia Artificial*:

> En medio de la creciente adopción de inteligencia artificial (IA) generativa en el ámbito creativo, más de 15,000 creadores, entre ellos más de 1,000 músicos, han firmado una declaración exigiendo que las empresas de IA obtengan permiso antes de utilizar sus obras como datos de entrenamiento. Esta declaración, apoyada por artistas destacados como Thom Yorke, Nitin Sawhney y Aurora, señala que el uso de contenido sin autorización amenaza el sustento de los creadores, y plantea un de-

> bate sobre los límites del uso justo y la propiedad intelectual en la era de la IA. (Hernández, 2024).

Un caso relevante en el marco de esta discusión es *Anderson vs. Stability AI* en Estados Unidos. Entre las diversas herramientas de inteligencia artificial disponibles, *Stability AI* se distingue por haber anunciado la capacidad de generar obras al estilo de artistas específicos, proporcionando incluso un catálogo de creadores cuyos estilos podían emularse. A diferencia de ChatGPT, que se especializa en texto, esta plataforma se enfoca principalmente en la generación de imágenes.

En cuanto a la demanda interpuesta se puede indicar lo siguiente:

> La demanda enmendada incluye reclamos de infracción directa de derechos de autor por parte de los artistas, quienes alegan que Stability AI, Midjourney, DeviantArt y Runway se apropiaron ilegalmente de obras protegidas. La demanda original presentada en enero de 2023 alegaba que las empresas violaron la Ley de Derechos de Autor (17 USC § 501) al introducir ilegalmente material protegido por derechos de autor para entrenar sus modelos de IA y generar imágenes derivadas del material protegido . . . Cuando el autor introdujo el texto de solicitud "latas de sopa de tomate al estilo de Andy Warhol", DreamUp generó una serie de imágenes que evocaban las icónicas latas y el logotipo de Campbell's, a pesar de que el autor no introdujo el nombre de Campbell en la solicitud.
>
> Esta imagen resultante evoca fuertemente la pintura de Warhol, Campbell's Soup Cans (1962), una pieza protegida de propiedad intelectual. Sin embargo, los generadores de imágenes de IA no cuentan con un proceso establecido para reconocer los derechos exclusivos de los propietarios de los derechos de autor, en este caso, la Fundación Andy Warhol, para reproducir, compartir y crear obras derivadas de la pintura original. (Williams, 2024)

Aunque este caso aún no ha llegado a su fase final, se espera que esta pueda aclarar en qué medida podría existir una infracción a los derechos de autor. Por un lado, los argumentos de los demandantes -que son artistas- sostienen que hay una apropiación y copia ilícita; que hay una ausencia de consentimiento por parte de los autores y que el entrenamiento se hizo con obras protegidas sin otorgar crédito o compensación. Mientras tanto, las empresas demandadas argumentan que las imágenes generadas por la IA no son copias directas ni sustancialmente similares a las obras protegidas por derechos de autor. Esto porque *Stability AI* y otras empresas utilizan conjuntos de datos como LAION, que simplemente recopila imágenes disponibles en la web a través de enlaces públicos y fundamentan la práctica del entrenamiento en que los modelos de IA funcionan de manera probabilística y no copian elementos específicos de una obra en particular, por lo que no están generando obras derivadas de las creaciones artísticas protegidas por derechos de autor ni se materializa una infracción.

El concepto de obra literaria y artística

El artículo 2 del *Convenio de Berna para la Protección de las Obras Literarias y Artísticas* define las obras literarias o artísticas de la siguiente manera:

> Los términos « obras literarias y artísticas » comprenden todas las producciones en el campo literario, científico y artístico, cualquiera que sea el modo o forma de expresión, tales como los libros, folletos y otros escritos; las conferencias, alocuciones, sermones y otras obras de la misma naturaleza; las obras dramáticas o dramático-musicales; las obras coreográficas y las pantomimas; las composiciones musicales con o sin letra; las obras cinematográficas, a las cuales se asimilan las obras expresadas por procedimiento análogo a la cinematografía; las obras de dibujo, pintura, arquitectura, escultura, grabado, litografía; las obras fotográficas a las cuales se asimilan las expresadas por procedimiento análogo a la fotografía; las obras de artes aplicadas; las ilustraciones, mapas, planos, croquis y obras plásticas relativos a la geografía, a la topografía, a la arquitectura o a las ciencias. (Organización Mundial de la Propiedad Intelectual, 1886).

Todo aquello que encuadre dentro del concepto definido en el artículo transcrito se tiene como obra literaria o artística y constituye uno de los presupuestos básicos para que exista protección por derechos de autor y derechos conexos. Sin embargo, no es el único, pues la obra debe ser original y creada por un autor.

El autor de una obra literaria o artística

El *Convenio de Berna para la Protección de las Obras Literarias y Artísticas* no define lo que se entiende por autor de una obra, por lo que cada país -en el marco de su propia legislación- define lo que se entiende por autor de una obra. Para algunos referentes en materia de Propiedad Intelectual la posición es que

> el principio por el cual únicamente puede ser autor la persona física que crea la obra, responde a la propia naturaleza de las cosas, pues solamente la persona natural tiene la cualidad de crear, por lo que autor sería una persona física (Antequera, 2000, p.109).

Sin embargo, algunos casos han desafiado esta concepción de autor como persona física, entre ellos el caso *Naruto vs. Slater*, en el cual se debatió si un mono podía poseer derechos de autor sobre una fotografía que él mismo tomó. En este litigio, un mono llamado Naruto accionó el obturador de la cámara del fotógrafo David Slater, creando una imagen que se volvió viral. Un grupo activista por los derechos de los animales presentó una demanda en nombre del mono, argu-

mentando que debía reconocérsele al mono como titular de los derechos de autor de la fotografía:

> Posteriormente, en septiembre de 2015, el grupo activista Personas por el Trato Ético de los Animales (PETA) presentó una demanda contra el Sr. Slater ante un tribunal de California en nombre del mono (llamado Naruto en la demanda) para hacer valer el derecho de autor sobre la imagen, alegando que la foto "fue resultado de una serie de acciones voluntarias y decididas por Naruto, sin la ayuda del Sr. Slater, que dieron lugar a obras originales de autor no atribuibles al Sr. Slater, sino a Naruto".
>
> En enero de 2016, el juez de primera instancia desestimó la demanda basándose en que, aunque Naruto hubiera tomado las fotografías "de forma autónoma e independiente", el procedimiento no podía seguir adelante, ya que los animales no tienen personalidad jurídica ante un tribunal y, por consiguiente, no pueden entablar una demanda por infracción de derechos de autor. (Guadamuz, 2018, p. 40-42).

Ahora bien, la legislación del Reino Unido también ofrece una respuesta distinta en cuanto al uso de la Inteligencia Artificial o el uso de máquinas para crear una obra:

> Así el artículo 9.1 de la Copyright, Design and Patents Acts 1988 ("CDPA") establece el principio general de que debe considerarse "autor" de una obra la "persona" que la haya creado. No obstante, el siguiente apartado 3 de ese mismo artículo regula de forma específica las obras creadas por computadores (computer-generated works), e indica que se considerará "autor" de ellas a la persona que haya realizado los arreglos necesarios (necessary arrangements) para la creación de dichas obras. El artículo 178 de la CDPA añade, con respecto a este tipo de obras, que se entenderá que habrán sido generadas por un programa de ordenador en aquellos supuestos en los que no haya un "autor humano".
>
> El régimen jurídico aplicable excluye en todo caso el otorgamiento a estos autores de derechos morales (ex artículo 81 de la CDPA) . . . Por tanto, se asume que, con respecto a las obras creadas por computador, el autor será el programa de ordenador, pero a los efectos de la titularidad de derechos serán considerados "autores" aquellas personas que hayan realizado los arreglos necesarios para que ese programa genere la obra. Con este contexto, es necesario determinar qué ha de entenderse por "arreglos necesarios" para definir quiénes pueden realizarlos. A priori, podrían considerarse incluidos en ese concepto los programadores que hubiera participado en el desarrollo del programa informático a partir del que se genera la obra, incluida la creación del algoritmo (que en sí mismo no es susceptible de protección) y su "alimentación" con parámetros específicos y bases de datos, como ocurre con la IA. (Sanjuán, 2018, p. 86).

En el caso de Costa Rica, la Ley No. 6683 no define lo que se entiende por autor pero el *Reglamento a la Ley de Derechos de Autor y Derechos Conexos*, Decreto No. 24611-J, establece en su artículo 3 una definición jurídica de autor:

1. Autor: Es, salvo disposición expresa en contrario de la Ley, la persona física que realiza la creación intelectual. (Reglamento a la Ley de Derechos de Autor y Derechos Conexos, 1995).

Por lo que en nuestro sistema jurídico el reconocimiento de un autor distinto a una persona física requeriría de una reforma de dicho reglamento o de una nueva ley, pues la definición transcrita parte de una presunción *iuris tantum*.[1]

Plagio frente a la ultra suplantación

Es esencial diferenciar entre el plagio académico tradicional y lo que podría denominarse como "suplantación de autoría mediante inteligencia artificial". En el caso del plagio académico convencional, existe un reconocimiento explícito de que el autor es quien crea una obra intelectual, sin embargo, surge una paradoja epistemológica cuando entra en juego la inteligencia artificial: si esta no puede ser considerada como autor bajo los paradigmas jurídicos actuales, entonces resulta problemático determinar si puede, en sentido estricto, crear una obra.

Al profundizar en esta disyuntiva conceptual, es importante señalar que la denominada "ultra suplantación" ha sido objeto de regulación en el nuevo reglamento de la Unión Europea (UE) sobre inteligencia artificial. Este fenómeno se caracteriza por la apropiación de contenido generado mediante sistemas de inteligencia artificial, presentándolo posteriormente como original y de autoría propia. La distinción entre ambos conceptos -plagio académico tradicional y suplantación mediante IA- no es meramente semántica, sino que constituye una diferenciación sustancial que presenta desafíos significativos para el ámbito académico y científico.

Este fenómeno plantea desafíos en el ámbito normativo de las instituciones educativas, particularmente en lo referente a la implementación de mecanismos regulatorios y en la determinación de sanciones para la y los estudiantes que incurran en estas prácticas. Es por eso que este asunto merece un análisis detallado que considere las implicaciones éticas, jurídicas y pedagógicas de estas nuevas formas de comportamiento académico en la era digital.

Conclusiones

A la luz de los hallazgos, es posible establecer que la Inteligencia Artificial representa un desafío sin precedentes para la normativa actual, dejando en evidencia que los avances tecnológicos pueden representar desafíos éticos y normativos.

En lo que respecta a los derechos de autor frente al entrenamiento de modelos de IA, resulta evidente la urgencia de im-

1 La expresión *"iuris tantum"* se entiende como aquella presunción de carácter legal que admite un determinado hecho como cierto, salvo que se demuestre lo contrario. En la definición de autor, el legislador parte de la premisa de que autor sólo puede ser una persona física, salvo que otra ley indique lo contrario.

plementar reformas legislativas que establezcan criterios claros y precisos, las cuales deben proporcionar seguridad jurídica tanto a autores como a artistas. Esto debe permitirles identificar con certeza cuándo sus derechos están siendo vulnerados y ayudar a establecer las condiciones bajo las cuales las empresas tecnológicas podrían utilizar obras protegidas por derechos de autor para el entrenamiento de sistemas de IA.

Finalmente, las instituciones académicas enfrentan una obligación de actualizar sus instrumentos normativos para abordar las nuevas modalidades de plagio y fraude académico vinculadas a la IA, la cual debe procurar conceptos y definiciones que sean entendidos por toda la comunidad académica, con el fin de implementar buenas prácticas y conductas en la producción intelectual en la era digital.

Referencias

Antequera, R. (2000). *Manual para la enseñanza virtual del Derecho de Autor y los Derechos Conexos*. https://www.collegesidekick.com/study-docs/2806132

OpenAI. (2024). ChatGPT (versión del mayo de 2024) [Modelo de lenguaje GPT-4o]. https://chat.openai.com/chat

Guadamuz, A. (febrero, 2018). ¿Qué nos puede enseñar sobre la legislación de derecho de autor el caso del mono que se hizo un selfi?. *OMPI Revista*. Recuperado 6 de noviembre de 2024, de https://www.wipo.int/edocs/pubdocs/es/wipo_pub_121_2018_01.pdf

Hernández, J. (2024). Más de 15.000 creadores firman declaración rechazando el entrenamiento de IA con sus obras. *Industria Musical*. https://industriamusical.com/mas-de-15-000-creadores-firman-declaracion-rechazando-el-entrenamiento-de-ia-con-sus-obras/

Herra, R. (27 marzo de 2024). Plagio de plagios. *La Nación*.

Ley sobre Derechos de Autor y Derechos Conexos, Ley No. 6683, Asamblea Legislativa de Costa Rica (1982). https://www.pgrweb.go.cr/scij/

Miranda Moya, N. (2014). *La propiedad intelectual derivada de los Trabajos Finales de Graduación en la Universidad de Costa Rica* [Tesis de Licenciatura en Derecho]. Universidad de Costa Rica.

Organización Mundial de la Propiedad Intelectual. (1886). *Convenio de Berna para la Protección de las Obras Literarias y Artísticas*. https://www.wipo.int/wipolex/es/text/283700#P94_12427

Reglamento a la Ley de Derechos de Autor y Derechos Conexos, Decreto Ejecutivo No. 24611-J, Poder Ejecutivo de Costa Rica (1995). https://www.pgrweb.go.cr/scij/

Sanjuán, N. (2018). Inteligencia artificial y Propiedad Intelectual. *Actualidad Jurídica Uría Menéndez*, 52. https://dialnet.unirioja.es/servlet/articulo?codigo=8704504

Williams, S. (febrero, 2024). *AI and Artists' IP: Exploring copyright infringement allegations in Andersen v. Stability AI Ltd*. Center For Art Law. Recuperado 15 de noviembre de 2024, de https://itsartlaw.org/2024/02/26/artificial-intelligence-and-artists-intellectual-property-unpacking-copyright-infringement-allegations-in-andersen-v-stability-ai-ltd/

13

La política de la regulación de plataformas en la era de la IA: el caso de la Gig Economy en América Latina

Ronald Sáenz-Leandro

Ronald Saénz Leandro. Politólogo e investigador doctoral del Internet Interdisciplinary Institute (IN3) de la Universitat Oberta de Catalunya (UOC).

Ha sido profesor e investigador en la Universidad de Costa Rica y actualmente es fellow de la Cátedra en Inteligencia Artificial y Democracia, de la School of Transnational Governance, European University Institute (STG-EUI), Florencia. Actualmente realiza su tesis sobre la política de la regulación de plataformas de la gig economy en América Latina.

Resumen

Esta ponencia explora los cambios que la integración de la inteligencia artificial (IA) está generado en los mercados laborales ante la irrupción cada vez más constante de las plataformas de trabajo en la economía de América Latina. Bajo esta perspectiva, el artículo examina las experiencias de 9 países de la región a partir de un análisis normativo-comparado y consultas con personas expertas. Los resultados de esta investigación revelaron un desarrollo desigual en el ámbito regulatorio, lo que sugiere que la regulación de plataformas es un proceso dinámico y condicionado por factores sociales, políticos y tecnológicos.

Palabras clave: América Latina, gig economy, plataformas digitales, plataformización, regulación.

Introducción

Hoy las plataformas digitales están presentes en todos los sectores de la economía y funcionan como infraestructuras digitales que facilitan, permiten y dan forma a interacciones personalizadas entre personas usuarias y las partes oferentes de un servicio. Para hacerlo, operan a través de programas que realizan la recopilación sistemática, el procesamiento algorítmico, la monetización y la circulación de datos con distintos fines. Es así como estas funciones habilitan la existencia de una amplia gama de plataformas y han contribuido a configurar la llamada "sociedad de las plataformas. Dentro de esta han adquirido especial atención las plataformas de trabajo durante la última década.

Lo anterior muestra el impacto que estas estructuras han generado en el mundo laboral, llevando a una digitalización de actividades, así como a la precarización laboral, particularmente en el caso de las plataformas digitales de reparto y transporte. La trascendencia que han adquirido las plataformas constituye un fenómeno que pone el foco sobre los procesos que se generan en las sociedades a partir de la entrada de las plataformas digitales y los impactos que estas producen en los mercados; además de los conflictos políticos.

Las interacciones que surgen a partir de esto han dado lugar a la **plataformización**. Este es un término que permite estudiar cómo la penetración de estas plataformas en diferentes sectores afecta a los marcos gubernamentales y a los diferentes procesos políticos y económicos en que estas plataformas se desenvuelven. En consecuencia, este concepto sirve para visibilizar las transformaciones que están ocurriendo en el campo de la regulación y la gobernanza.

En este contexto las plataformas digitales son un objeto de estudio central en la actualidad, no sólo por las discusiones regulatorias actuales en las que se evidencia como la innovación tecnológica muchas veces contrasta con los marcos regulatorios que rigen las relaciones laborales y/o el transporte

remunerado; sino también porque conforman lo que se ha venido a denominar como la **gig economy** (concepto que ha sido traducido como la economía de los pequeños encargos, o la economía informal mediada digitalmente). Esta estructura el trabajo a través de nuevas plataformas, las cuales podemos clasificar entre modelos locales y remotos[1].

Durante la década del 2010, en toda América Latina se experimentó una liberalización del mercado de las telecomunicaciones, aunado a la expansión de los teléfonos móviles y de la conectividad móvil de banda ancha, lo que contribuyó a sentar las bases para desarrollar el capitalismo de plataformas en la región. Gracias a esto se incrementó el uso de plataformas digitales, no solo por parte de las personas usuarias, sino también de quienes buscaron utilizarlas como un medio para generar dinero.

Hasta el momento, los estudios sobre plataformas digitales que están disponibles en América Latina ahondan sobre todo en los perfiles socio-laborales de las personas trabajadoras de plataformas y en sus condiciones de trabajo; mientras que otros se concentran en abordar debates jurídicos o a documentar esfuerzos de organización colectiva. A pesar de esto, aún persisten importantes vacíos para sistematizar datos que permitan tener una idea más clara del estado de situación en torno a la regulación de plataformas en la región.

Además, diversos autores han señalado que las agrupaciones de personas trabajadoras de plataformas no han logrado llevar a cabo cambios sustanciales en la legislación del trabajo a nivel regional y el poco avance regulatorio que se tiene se ha dado en el campo de las plataformas de transporte. Ante este panorama, resulta necesario que se realice un mapeo que de manera sistemática responda a las siguientes preguntas: 1) ¿qué se está regulando?, 2) ¿cuáles países están regulando?, 3) ¿sobre qué se está regulando? y 4) ¿qué desafíos está generando la presencia o la ausencia de una regulación de plataformas en algunos países?

Con el fin de responder a estas interrogantes se examinaron 21 documentos de política regulatoria de 9 países latinoamericanos, los cuales fueron emitidos entre 2015 y 2023. Estos textos fueron clasificados a partir de un índice en el que se analizan 5 dimensiones regulatorias que se relacionan con cuestiones como: 1) el acceso al mercado, 2) regulación del empleo y el trabajo, 3) la seguridad y protección de la persona consumidora, 4) los impuestos y 5) gobernanza de datos y la rendición de cuentas algorítmica. Complementariamente se realizaron 10 entrevistas semi-estructuradas con personas expertas de la región tanto de contextos regulados como aquellos que no cuentan con regulaciones de plataforma.

1 Dentro de los modelos locales están los que ya conocemos como las plataformas de entrega, las plataformas de transporte; pero también es posible identificar otro tipo de trabajos remotos que se realizan a través de las plataformas que favorecen el trabajo en la modalidad *freelance*, en donde por ejemplo destaca la plataforma de Amazon Mechanical Turk.

Más allá del algoritmo: plataformas digitales, transformación laboral y vacíos regulatorios

Los hallazgos de este ejercicio revelaron que en los países latinoamericanos donde hay una regulación de plataformas de trabajo para el sector de transporte, se muestran importantes avances regulatorios en la dimensión de acceso al mercado y en la de protección y seguridad de la persona consumidora. Por el contrario, las dimensiones de gobernanza algorítmica y fiscal son en las que se observa menor desarrollo regulatorio en la región.[2]

Con respecto a la dimensión de acceso al mercado, se identificó que hay Estados que imponen requisitos de inscripción y registro; lo que muchas veces tiene que ver con requerimientos que buscan asegurar la seguridad y la protección de la persona consumidora. Sobre la dimensión de gobernanza de datos y rendición de cuentas algorítmica, se determinó que cuando se encuentra regulación en el tema, esta usualmente tiende a buscar el establecimiento y la transparencia de los procedimientos automatizados con medidas como la definición de tarifas y explicaciones sobre el tratamiento de datos personales de las y los usuarios y de los datos agregados de la movilidad a lo interno de las ciudades.

2 Para una versión extendida de los resultados de esta investigación, ver: Sáenz-Leandro, Ronald. "The Battle for TNCs in Latin America: Navigating Policy Trends and Sociotechnical Controversies in the Regulation of Ride-Hailing Platforms." *Tapuya: Latin American Science, Technology and Society*, vol. 8, no. 1, 2025, https://doi.org/10.1080/25729861.2025.2495521.

Por su parte, en la dimensión de trabajo y empleo se identificó un vacío importante que refleja que esta ha sido el área que más ha costado regular; a pesar de ser uno de los ámbitos en los que se tienen mayores debates y es el de mayor atención dentro de las ciencias sociales. La excepción a esto, sería el caso de Chile, donde se aprobó la llamada "*Ley Uber*" (Ley No. 21533), aún cuando la misma no pueda considerarse como lo suficientemente robusta.

Si bien esta norma cubre a todas las plataformas del sector de transporte, aún no ha entrado en vigencia, por lo que se sigue discutiendo la implementación de esta norma. Esta propuesta normativa ha generado gran oposición ya que se estima que cerca de unos 36 000 trabajadores y trabajadoras podrían dejar la actividad cuando se implemente la ley en el país. Este tipo de iniciativas que tiene un carácter más integral suelen estancarse por las dinámicas gubernamentales y el desinterés de los gobiernos para regular realmente al sector. Sin embargo, hay coyunturas políticas que propician la entrada de distintos proyectos de ley dentro de las corrientes legislativas.

Esas coyunturas pueden ser generadas por escenarios de conflictividad social, por ejemplo, cuando en Costa Rica entraron las plataformas y se produjo un conflicto entre taxistas y conductores de Uber y otras plataformas o durante la pandemia, ya que en ese momento se discutió nuevamente la cuestión de la regulación de las plataformas de transporte y reparto. Otro caso al que se puede hacer mención tiene relación con los movimientos migratorios, particularmente en los países de América del Sur, en donde el trabajo en plataformas se ha vis-

to cooptado por las migrantes venezolanos provocando que las dinámicas migratorias hayan sido utilizadas para politizar proyectos de ley.

Por otro lado, al consultar con las personas expertas sobre los desafíos e impactos que genera la consolidación de las plataformas de transporte, la mayoría coincidió en señalar la necesidad de que se regulen las plataformas en aquellas áreas en las que existen vacíos legales y zonas grises (como ocurre en Costa Rica).

Conclusiones

En síntesis, la investigación evidenció que el enfoque regulatorio está muy centrado en abordar cuestiones como el acceso al mercado y la seguridad de la persona consumidora, lo que muestra una tendencia regional a intervenir mínimamente en el sector. Asimismo, muchas de estas regulaciones tienden a ser bastante superficiales y se identifican importantes áreas de mejora; así como desafíos derivados de la falta de intervención estatal.

Esto ha llevado a la aplicación de un enfoque de normalización y regulación, que es la ruta que han seguido la mayoría de países latinoamericanos que no poseen ningún tipo de normativa y que es típica de los Estados en los que persisten áreas grises. En estos se ha permitido que el mercado se consolide y se aprovechen los vacíos legales. A pesar de eso, también se observa un efecto de difusión institucional del modelo chileno, ya que en algunos países se está tomando como referencia su normativa para emitir regulación de plataformas, aunque haya sido polémica.

Algunas de las personas expertas consultadas, consideran que, aunque la regulación chilena pretende ser un poco más integral, esta carece o no prevé el establecimiento de políticas que permitan abordar cuestiones fundamentales como la continuación de la informalidad laboral más allá de las plataformas de transporte. Por ejemplo, han surgido algunas alternativas informales que ya ni siquiera operan a través de las mismas plataformas de transporte, sino que lo hacen en plataformas sociales como Facebook o WhatsApp. Desde estas aplicaciones se ofrecen servicios de movilidad como la modalidad de "taxi pirata", lo que dificulta la regulación de estas dinámicas de mercados informales fuera de los espacios oficiales de las plataformas de transporte.

En suma, los resultados de la investigación permiten considerar que la regulación de plataformas es un proceso dinámico y no lineal, que no acaba necesariamente con la emisión de una ley y que, por el contrario, cambia con el tiempo, viéndose afectada por la aparición de nuevas partes interesadas y nuevas tecnologías futuras.

14

IA más allá de los algoritmos: el humano en la ecuación en un mundo impulsado por tecnologías emergentes

Henry Lizano Mora

Henry Lizano Mora. Investigador del Centro de Investigación Observatorio del Desarrollo (CIOdD) y Jefe de Tecnologías de Información de la Oficina de Servicios Generales (OSG) de la Universidad de Costa Rica (UCR).

Máster en Sistemas de Información (ITCR) y Candidato a Doctor en Administración del (ITCR). Se ha desempeñado como director del Centro de Informática y profesor de Ingeniería Industrial y Administración Pública desde 2006, todos roles en la Universidad de Costa Rica (UCR). La experiencia incluye la dirección de numerosos cursos de grado, postgrado y colaboraciones con organizaciones privadas y públicas. Ha participado en proyectos de investigación en Gestión de Procesos de Negocio, Transformación Digital, Automatización Robótica de Procesos, Inteligencia Empresarial e Innovación; así como, la participación en diversas conferencias y programas sobre Gestión de Procesos de Negocio tanto a nivel nacional como internacional.

Henry Lizano Mora

Resumen

La inteligencia artificial (IA), especialmente la IA generativa, ha logrado un nivel de evolución que ha propiciado su acelerada adopción en la producción, la educación y la vida social. Sin embargo, la mirada predominantemente técnica-económica ha limitado reflexiones cruciales sobre esos procesos, evitando profundizar en el análisis de cuestiones éticas y relacionadas con el bienestar humano. Es así como desde esta ponencia se propone resaltar el rol de la persona humana en el ecosistema tecnológico emergente, destacando la necesidad de modelos de IA explicables (XAI), políticas públicas inclusivas y marcos normativos éticos que garanticen un desarrollo tecnológico con propósito social. A partir de un análisis de tendencias, investigaciones recientes y experiencias aplicadas en Costa Rica, se propone una IA centrada en las personas como condición fundamental para una transformación digital verdaderamente inclusiva y responsable.

Palabras clave: Inteligencia artificial, inteligencia artificial generativa, tecnologías emergentes, Inteligencia Artificial Explicable, XAI, dimensión humana.

Introducción

Ante el acelerado avance de tecnologías emergentes como la IA, se ha desencadenado una revolución en la forma como se opera e interactúa en las sociedades modernas. Estas transformaciones han sido acompañadas por discursos en los que predominan consideraciones técnico-económicas, pero que muchas veces descuidan las implicaciones sociales y éticas que se pueden producir con esta transición tecnológica. Es por eso que la presente ponencia busca reflexionar el lugar de la persona humana desde una perspectiva crítica que se cuestione las consecuencias de vivir en un mundo cada vez más automatizado y regido por algoritmos.

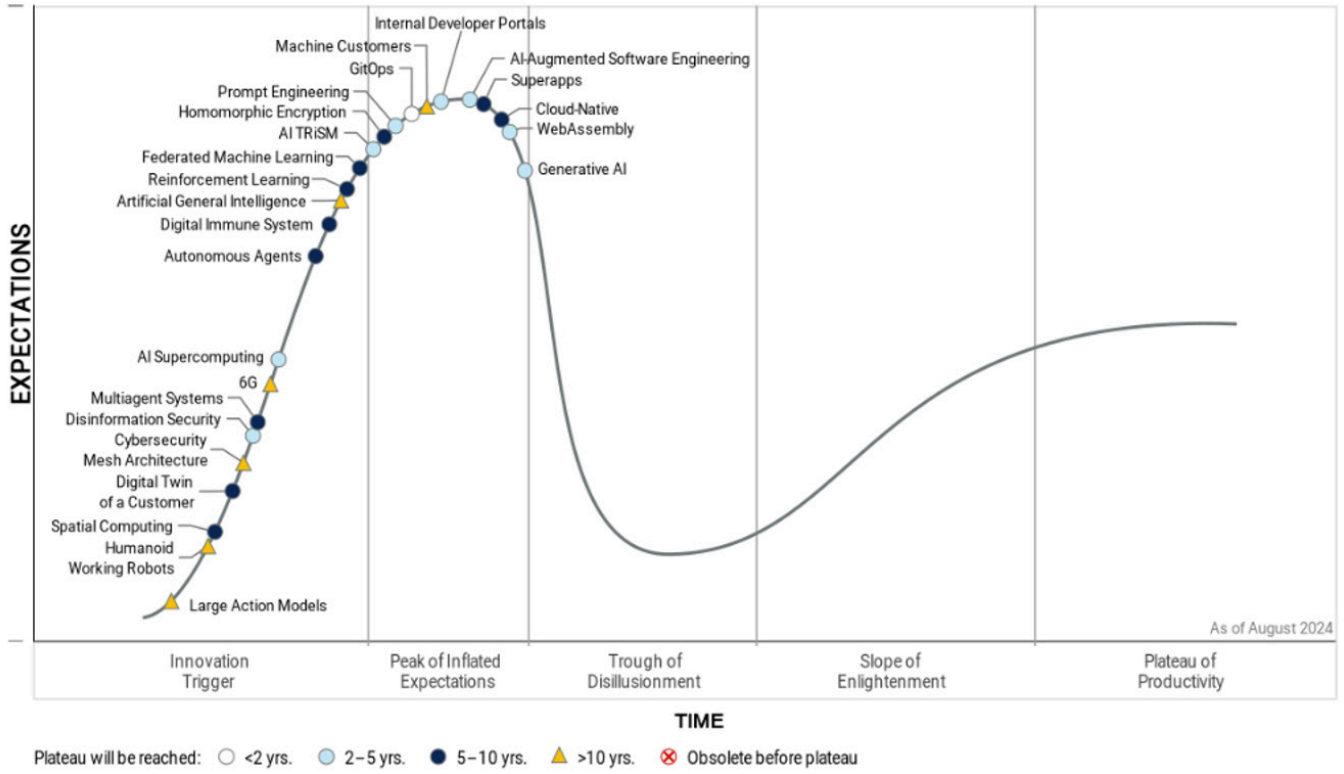
Con base al análisis de ciclos tecnológicos, se examinan tendencias del ecosistema de investigación global en IA, así como las experiencias locales en automatización, lo que permite visibilizar las oportunidades y los riesgos de la IA; especialmente cuando su adopción no considera las emociones, la cultura, el trabajo o la dignidad humana.

Desde esta perspectiva se proponen elementos clave para una IA centrada en las personas, que no solo optimice procesos o genere productividad, sino que también promueva inclusión, explicabilidad, ética y sentido social. Desde el enfoque latinoamericano, se destacan las iniciativas regionales como el LATAM-GPT y el marco de la Organización de Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) como pasos estratégicos hacia ese horizonte.

La curva de las expectativas y el lugar del ser humano

La consultora Gartner ha propuesto, a través del "*hype cycle*" o ciclo de sobre expectación una representación útil para entender y visibilizar las fases que atraviesan las tecnologías emergentes (sobreexpectación, desilusión y eventual estabilización). Cuando aparecen suelen generarse grandes expectativas sobre las mismas y conforme pasa el tiempo se suele llegar a un punto de máximas expectativas que luego tiende a decrecer porque se espera más de lo que terminan ejecutando en la realidad (Gartner, 2024).

Figura 14.1. Hype Cycle (Ciclo de Sobre expectación)



Fuente: Tomado de Gartner, 2024.

En el caso de la IA generativa, la empresa de Gartner indica que el ciclo de sobre expectación subió en el 2023 llegando a su cúspide y durante el 2024 ha comenzado a caer, no porque no entregue beneficios económicos o nuevas expectativas de desarrollo; sino en función del beneficio que está entregando en este momento. Esto significa que la IA generativa se encuentra en una transición desde el punto máximo de expectativa hacia una fase de aplicación porque la IA ya está en uso y se están generando impactos en diversos ámbitos.

Sin embargo, dicho progreso no ocurre de manera aislada, sino que depende de otras tecnologías que complementan a la IA y que actúan como facilitadoras o habilitantes. Entre estas se encuentran:

- Las redes 5G: que aunque han experimentado atrasos en su implementación, son esenciales porque ofrecen ultra baja latencia, lo que es necesario para aquellas aplicaciones que requieran respuesta en tiempo real de misión crítica.

- Cadena de bloques o "Blockchain" y chips especializados con inteligencia artificial: los cuales ayudan a mejorar la seguridad de los datos e incrementan la eficiencia en el procesamiento de información.

- Gemelos digitales: que consisten en recreaciones virtuales exactas de espacios físicos, por ejemplo, de la ciudad de San José y en los cuales se pueden hacer simulaciones del tránsito y pruebas antes de realizar cambios en el entorno real.

- Internet de las cosas (IoT) y realidad virtual inmersiva: son tecnologías que extienden las posibilidades de uso de la IA generativa para aplicarla en entornos cada vez más complejos y realistas.

En consecuencia, la IA no puede ser vista de manera aislada, sino que debe considerarse como una tecnología que forma parte de un ecosistema tecnológico mucho más amplio.

Por otro lado, aunque la IA no es un concepto nuevo, esta marcó un antes y un después porque permitió el entrenamiento de modelos de lenguaje a gran escala y ponerlos al alcance del público en general. Dicho avance reavivó el antropomorfismo, un fenómeno que hace que las personas le atribuyan cualidades humanas a sistemas de IA. Esto ya ocurrió en 1965 cuando se lanzó el programa *Eliza* y hoy está sucediendo con herramientas modernas como *Replika*, una aplicación móvil que simula el comportamiento humano y está disponible para conversar todo el tiempo.

Esta aplicación ha sido objeto de estudios longitudinales en los que se ha evidenciado que con el tiempo, algunas personas usuarias desarrollan vínculos emocionales con esta, llegando incluso a considerarla una pareja sentimental. Asimismo, se ha constatado que algunas personas pagan para acceder a funciones especiales para que la aplicación actúe como pareja y produzca interacciones constantes. Además, algunas redujeron su interacción social al pasar más tiempo con la aplicación Replika (Skjuve et al., 2023).

Esto plantea interrogantes sobre los efectos psicológicos que genera la interacción con algoritmos avanzados que, aunque son sistemas probabilísticos pre entrenados, pueden influir profundamente en la percepción humana. En el 2021, la Comisión Económica para América Latina y el Caribe (CEPAL) abordó el impacto de las tecnologías digitales y emergentes -no solo la IA- en la generación y pérdida de empleos y determinó que si bien la IA parece que tendrá un saldo positivo al crear roles nuevos (como el de ingeniero de *prompts*), también llevará a la desaparición de ciertos puestos tradicionales si no se toman medidas en el desarrollo de competencias y habilidades de la fuerza laboral. (Comisión Económica para América Latina y el Caribe [CEPAL], 2022)

En ese sentido, el principal desafío es que los nuevos empleos no serán ocupados por quienes los perderán ante la automatización, por lo que es fundamental impulsar políticas públicas que capaciten y permitan el desarrollo de habilidades relevantes (de *upskilling* y *reskilling*) que ayuden a la fuerza laboral para adaptarse a los cambios y eviten que las nuevas tecnologías generen exclusión.(CEPAL, 2022)

Impacto emocional de la inteligencia artificial

La IA puede provocar diverso tipo de emociones en las personas, como la aceptación y la esperanza o la ansiedad y el rechazo, entre muchas otras. Japón ha desarrollado un mapa que refleja las emociones que la IA despierta en la población, lo que revela cómo la IA no solo transforma procesos técnicos, sino que también impacta la dimensión humana y social. (Miraikan, 2023)

Asimismo, algunas tecnologías simulan conversaciones de personas fallecidas, usando sus voces e imágenes lo que para algunos podría representar una manera valiosa de conservar recuerdos; mientras que otras pueden percibirlo como algo perturbador o aterrador. Este tipo de sentimientos muestran que la utilización de la IA puede tener profundas implicaciones emocionales y por tanto, requiere de una sensibilidad especial.

Para ejemplificar este impacto, vale la pena referirse a un experimento de automatización realizado en la Universidad de Costa Rica (UCR). Este ejercicio fue realizado en el marco de una investigación sobre transformación digital en la que se implementó la automatización robótica de procesos (RPA) usando herramientas como *Automation Anywhere* [1]en la Oficina de Administración Financiera (OAF). Un proceso que usualmente le lleva a una persona un mes en realizar manualmente fue automatizado y completado en solo tres minutos, identificando, además, errores humanos. La reacción de la persona encargada fue emocional, puesto que pensó que esta tecnología la reemplazaría.

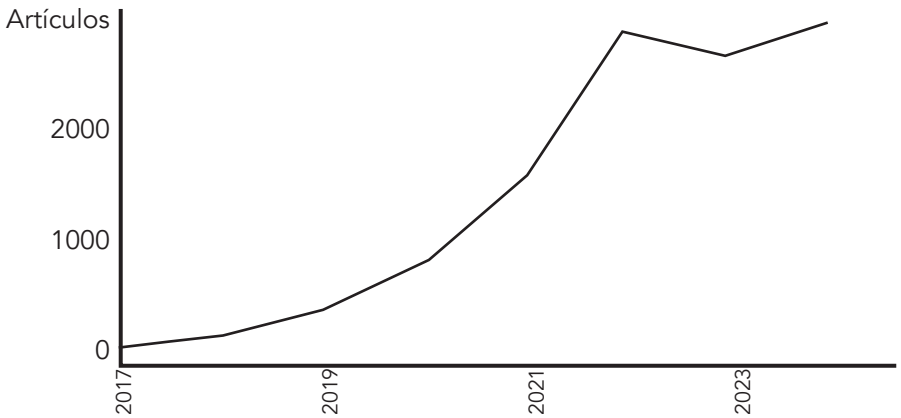
1 Automation Anywhere es una aplicación comercial de terceros para automatización de tareas que usa aprendeizaje automático, disponible en https://www.automationanywhere.com

El resultado de este experimento muestra que introducir tecnologías disruptivas sin considerar el impacto humano puede generar afectaciones emocionales en las personas. Aunque el objetivo de la automatización era liberar tiempo, con la introducción de la herramienta no se tomó en cuenta el efecto psicológico del reemplazo. En consecuencia, la implementación de nuevas tecnologías requiere de una evaluación previa centrada en las personas.

Investigación científica, colaboraciones y geopolítica de la IA

Según nuestra investigación bibliométrica sobre el estado actual de la IA, en la que se analizaron más 10 mil estudios publicados entre el 2017 y 2024, se identificó un crecimiento exponencial durante los primeros años, seguido de una leve caía y una posterior estabilización con una tendencia ascendente. La mayoría de las investigaciones producidas en el periodo se enfocaron en redes neuronales y técnicas de aprendizaje automático. Todo esto sugiere que el avance de la IA continuará durante los próximos años.

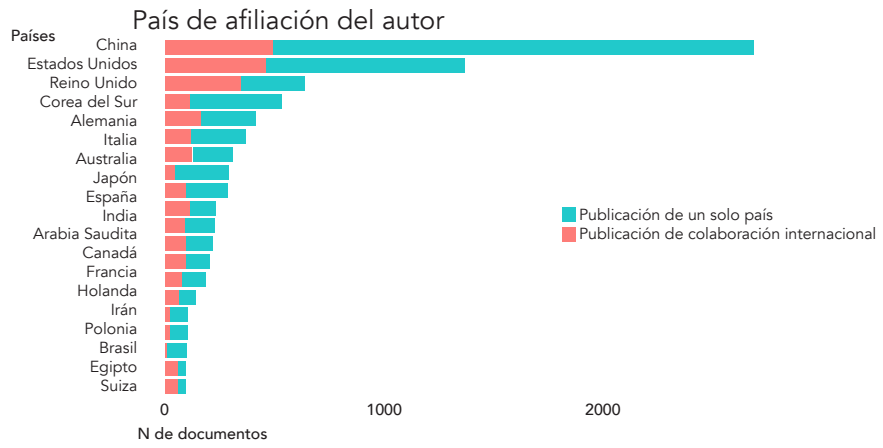
Figura 14.2. Producción científica anual de IA



Fuente: Elaboración propia, 2024.

Por otro lado, al examinar el panorama global de la investigación en IA se constató que las principales universidades que están liderando en el tema se encuentra en China, Estados Unidos y el Reino Unido; aunque China destaca por el volumen de datos (superando a Estados Unidos), a pesar de limitar los estudios internos y limitar sus colaboraciones internacionales. Asimismo, cuando se examinan los resultados por regiones se evidencia que el top de 20 países con mayor producción académica en IA está dominado por países del hemisferio Norte.

Figura 14.3. Investigación de IA por país

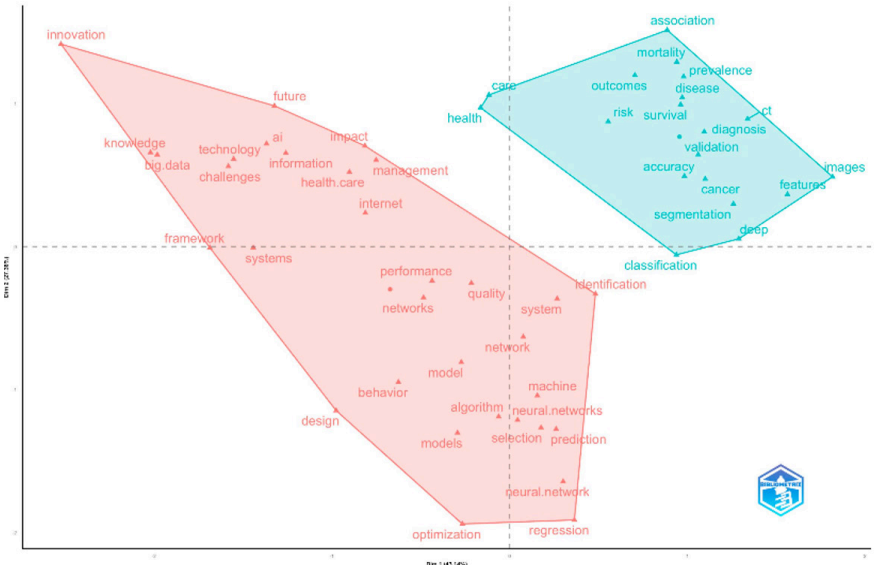


Fuente: Elaboración propia, 2024.

En términos generales, el conocimiento en IA ha experimentado una evolución en la que se visibiliza un crecimiento importante durante las últimas tres décadas. El conocimiento generado puede organizarse en tres clústeres temáticos: 1) las tecnologías, 2) los riesgos y 3) las aplicaciones. Cabe señalar que un análisis factorial confirmatorio destaca especialmente, el tema de los riesgos, particularmente en el área médica, donde se tienen grandes expectativas de precisión.

Otro tema relevantes que los tres artículos científicos más citados durante los últimos ocho años son sobre la *Inteligencia Artificial Explicable* (XAI), una tecnología que busca hacer más comprensibles los resultados de los grandes modelos de lenguaje para los humanos y responder a preguntas como: ¿por qué se tomó una decisión?, ¿qué tan confiable es? Y ¿qué factores influyeron? Se considera que la XAI será vital en áreas críticas como el derecho, las finanzas y la medicina, ya que en estas la confianza en el modelo es esencial(Al & Sağıroğlu, 2024; Alonso, 2020; Darwish, 2022; Minh et al., 2021; Williams, 2021; Zanni-Merk, 2019).

Figura 14.4. Análisis factorial confirmatorio



Fuente: Elaboración propia, 2024.

Con relación a la colaboración internacional en IA debe señalarse que esta se divide en dos bloques: los países alineados con los Estados Unidos y aquellos vinculados con China. En paralelo a estas dinámicas, también se han desarrollado diversos modelos de lenguaje propios, sea por país o por re-

gión. Por ejemplo, en Francia se implementó Mistral, en los países árabes Falcon y en Latinoamérica se ha iniciado con LatamGPT. Sobre este último debe mencionarse que la UCR ha introducido infraestructura de computación de alto rendimiento (HPC) para contribuir con el desarrollo de LATAM GPT, así como para aplacar sesgos geográficos y promover una IA con identidad latinoamericana (LatamGPT, 2025).

De la mano de esto, también se está trabajando en la implementación del Marco de Competencias de la IA para Profesores y Estudiantes de la UNESCO con un enfoque humano para la ética en la aplicación de la IA, el desarrollo pedagógico y la profesionalización docente. Aunado a ello, se están realizando proyectos piloto en pequeñas y medianas empresas costarricenses orientados a integrar la IA en la gestión financiera, la optimización de procesos, la mejora en la atención del cliente y fomentar el aprendizaje organizacional. (Organización de Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO], 2025)

Conclusiones

Hoy la IA ha dejado de ser una promesa futurista para convertirse en una realidad con la que convivimos diariamente. El impacto positivo de las transformaciones que está impulsando dependerá de la capacidad que tengan las sociedades para integrar la tecnología sin perder de vista al ser humano. Por ello, resulta fundamental que se eviten los enfoques meramente técnicos y que en su lugar se adopten principios éticos, de inclusión y transparencia a estos procesos. La IA no debe imponerse a las personas, sino colaborar con ellas.

Por tal motivo es indispensable que las políticas públicas, los marcos regulatorios, la investigación científica y las aplicaciones prácticas avancen hacia una IA explicable, ética y situada en el contexto. Esto implica un compromiso activo con el desarrollo de modelos inclusivos, como LATAM-GPT, que reflejen los valores, lenguajes y necesidades de nuestras comunidades. En última instancia, se trata de que la inteligencia artificial no sustituya nuestra humanidad, sino que la potencie. Solo así podremos construir un futuro tecnológico verdaderamente justo y sostenible.

Referencias

Al, S., & Sağıroğlu, Ş. (2024). A Review of Explainable Artificial Intelligence. *2024 9th International Conference on Computer Science and Engineering (UBMK)*, 310-315. https://doi.org/10.1109/UBMK63289.2024.10773588

Alonso, J. M. (2020). Teaching Explainable Artificial Intelligence to High School Students. *Int. J. Comput. Intell. Syst.*, *13*, 974-987. https://doi.org/10.2991/ijcis.d.200715.003

CEPAL, C. (2022). *Hacia la transformación del modelo de desarrollo en América Latina y el Caribe: Producción, inclusión y sostenibilidad*.

Darwish, A. (2022). Explainable Artificial Intelligence: A New Era of Artificial Intelligence. *Digital Technologies Research and Applications*. https://doi.org/10.54963/dtra.v1i1.29

Gartner. (2024). *Gartner 2024 Hype Cycle for Emerging Technologies Highlights Developer Productivity, Total Experience, AI and Security*. Gartner. https://www.gartner.com/en/newsroom/press-releases/2024-08-21-gartner-2024-hype-cycle-for-emerging-technologies-highlights-developer-productivity-total-experience-ai-and-security

LatamGPT. (2025). *Latam-GPT - Latam-GPT*. https://www.latamgpt.org/en

Minh, D., Wang, H. X., Li, Y., & Nguyen, T. N. (2021). Explainable artificial intelligence: A comprehensive review. *Artificial Intelligence Review*, *55*, 3503-3568. https://doi.org/10.1007/s10462-021-10088-y

miraikan, miraikan. (2023). *AI Map*. https://www.miraikan.jst.go.jp/en/resources/provision/aimap/

Skjuve, M., Følstad, A., & Brandtzæg, P. B. (2023). A Longitudinal Study of Self-Disclosure in Human–Chatbot Relationships. *Interacting with Computers*, *35*(1), 24-39. https://doi.org/10.1093/iwc/iwad022

UNESCO. (2025). *Marco de competencias para docentes en materia de IA*. UNESCO. https://doi.org/10.54675/AQKZ9414

Williams, M.-A. (2021). Explainable artificial intelligence. *Research Handbook on Big Data Law*. https://doi.org/10.4337/9781788972826.00022

Zanni-Merk, C. (2019). *On the Need of an Explainable Artificial Intelligence*. 3. https://doi.org/10.1007/978-3-030-30440-9_1

15

Carta Iberoamericana de Inteligencia Artificial en la Administración Pública: Un llamado a la acción para Costa Rica

Jorge Umaña Cubillo

Jorge Umaña Cubillo. Docente e investigador de la Escuela de Administración Pública.

Jorge es Administrador Público y Actuario. Realizó sus estudios de Maestría en Economía con énfasis en Finanzas y Riesgos. Todo por la Universidad de Costa Rica. Fue miembro del equipo de Gobierno Abierto en la Presidencia de la República de Costa Rica desde 2014 hasta 2018, como Coordinador Nacional de Datos Abiertos, en el despacho del Viceministerio de Asuntos Políticos y Diálogo Ciudadano. Fue coordinador regional de transformación digital, innovación y Gobernanza Abierta en el Trust for the Americas de la Organización de Estados Americanos.

Entre 2018 y 2022 coordinó el área de Datos para el desarrollo del Laboratorio Colaborativo de Innovación Pública (InnovaAP). Actualmente, es asesor de la Alcaldía de la Ciudad de San José en Costa Rica y formador en temas de gobernanza de datos en instituciones públicas.

Jorge Umaña Cubillo

Resumen

Actualmente, la inteligencia artificial (IA) se ha consolidado como un motor de transformación para las administraciones públicas, dinamizando la manera en que operan y generan valor público. Su incorporación ha permitido avances significativos en términos de eficiencia, transparencia y calidad de los servicios que los gobiernos brindan a la ciudadanía. Desde esta perspectiva, el presente análisis examina la importancia de adoptar políticas y estrategias alineadas con las mejores prácticas internacionales en materia de IA.

En este contexto, se toma como referencia la Carta Iberoamericana de Inteligencia Artificial, con el objetivo de comprender su relevancia y el impacto de su contenido para América Latina. Para ello, se analizan los principales documentos internacionales que influyeron en su formulación y se identifican elementos clave para una implementación ética de la IA en el sector público.

Con estos antecedentes, se aborda la situación actual de Costa Rica, destacando la reciente publicación de su Estrategia Nacional de Inteligencia Artificial (ENIA) y los desafíos que enfrenta el país para lograr una implementación efectiva de esta herramienta, así como una integración adecuada de la IA en la gestión pública.

Palabras clave: Administración Pública, Servicios Públicos, Inteligencia Artificial, Carta Iberoamericana de Inteligencia Artificial, Estrategia de Inteligencia Artificial.

Introducción

En la era digital actual, la inteligencia artificial (IA) emerge como una herramienta esencial para las organizaciones públicas ya que permite mejoras significativas en la eficiencia y la calidad de los servicios, así como para hacerlos más transparentes. Esto provoca una revolución de la gestión pública al transformar la forma como se desarrollan y entregan los servicios a la ciudadanía. En este contexto, resulta esencial que se adopten estrategias y políticas de IA que aseguren que esta se alinee con las estrategias, políticas y regulación vinculadas a la transformación digital.

Es por eso que, desde la Administración Pública, la IA está siendo examinada en relación a los marcos de gobernanza que deben implementar los países para regular esta tecnología, así como para implementar estrategias que garanticen un aprovechamiento efectivo y generen valor público en la sociedad. Reflexionar sobre estas cuestiones implica analizar cuáles son las condiciones que deben adoptar los países para impulsar políticas o estrategias de IA.

Para esto, se pueden tomar referencias internacionales y en ello, es usual que se estudien las experiencias de otros Estados con el fin de identificar lecciones aprendidas y buenas prácticas que orienten sobre la ruta de acción que se puede implementar a lo interno de los países. Asimismo, se pueden considerar instrumentos internacionales como las declaraciones, acuerdos o cartas suscritas entre distintos Estados.

En América Latina existen diversas herramientas orientadoras en materia de gobernanza digital, entre las que destacan la Carta Iberoamericana de Gobierno Digital, la Carta Iberoamericana de Participación Ciudadana, la Carta Iberoamericana de Gobierno Abierto y, más recientemente, la Carta Iberoamericana de Inteligencia Artificial. La aprobación de esta última constituye un avance significativo hacia la adopción ética y efectiva de la IA en la región, al tiempo que representa una oportunidad estratégica para Costa Rica.

Bajo esta perspectiva, el presente análisis se basa en la revisión de documentos clave que han influido en la elaboración de la Carta, incluyendo la *Declaración de Buenos Aires* del Centro Latinoamericano de Administración para el Desarrollo (CLAD) del 2019 y la Recomendación del Consejo sobre la IA de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) del 2022. Además, se consideran iniciativas globales, como la *IA para el Bien* de la Organización de Naciones Unidas (ONU) y el Libro Blanco de la Unión Europea (UE). A partir de esto, se ofrece una contextualización de los principios establecidos y se identifican recomendaciones para la regulación de la IA en Costa Rica.

La Carta Iberoamericana de IA

Este documento fue aprobado el 20 de noviembre del 2023 en Varadero (Cuba) por los 23 Estados miembro del CLAD con el respaldo de la Asamblea General de la ONU. Este documento se inspira en la Declaración de Buenos Aires, que destacó por primera vez las oportunidades de la IA en el sector público regional y en otras referencias globales que promueven el uso responsable de la IA. Esto quiere decir que la Carta cuenta con el respaldo de varios países y actores, además, destaca por tener un enfoque basado en los derechos humanos.

La Carta proporciona un marco ético y colaborativo que aborda los desafíos y las oportunidades de la IA, ofreciendo una guía integral para la adopción responsable en las instituciones públicas del país. Además, establece principios fundamentales para la implementación de la IA (como la transparencia, la rendición de cuentas, la ética y los derechos humanos, el acceso y la participación, la seguridad y la privacidad, la innovación, el desarrollo sostenible y la cooperación internacional). Estos principios buscan asegurar que la IA se utilice para el beneficio de la sociedad en general y promover un desarrollo equitativo y sostenible.

Cabe señalar que para la elaboración de este documento se utilizaron varios insumos, entre los que destacan:

- La *Declaración de Buenos Aires* (2019) que menciona por primera vez, algunas de las oportunidades de la IA para el sector público de América Latina.

- La publicación *IA para el Bien de la ONU* (2018), en la que se promovió el uso de la IA con el fin de que este genere valor público.

- Las *Recomendaciones del Consejo sobre la IA de la OCDE* (2022), las cuales establecen buenas prácticas en la materia y orientan la generación y recolección de evidencias.

- El *Libro Blanco sobre la IA* (2020), dado a conocer por la Unión Europea (UE) y en el que se considera como una referencia para situar a las personas en el centro de los procesos de inteligencia artificial.

A partir de estas referencias, la Carta establece directrices comunes que pueden ser utilizadas por los países Iberoamericanos para la aplicación ética de la IA en las Administraciones Públicas y orientar la construcción de estrategias o políticas de IA. Sin embargo, cuando los países cuentan con estrategias de IA, estas herramientas pueden ser fortalecidas con elementos que se señalan en la Carta.

La carta también brinda conocimientos sobre la IA y funge como un mecanismo para compartir experiencias entre los países y facilitar la cooperación y la colaboración. Además, mapea los desafíos (por ejemplo, la opacidad algorítmica[1]) que implica la implementación de la IA a lo interno del sector público, considerando que su abordaje resulta fundamental para no dejar a nadie más y gestionar la IA de una manera positiva y equitativa, que contribuya a expandir el acceso a dicha tecnología. Otra cuestión importante que se aborda tiene que ver con la forma como puede evitarse el uso de la IA para debilitar la democracia.

La identificación de estos desafíos busca posicionar en la agenda pública iberoamericana aquellos factores que limitan una visión favorable de la inteligencia artificial. En este sentido, es fundamental que las políticas públicas se diseñen con el objetivo de evitar o, al menos, abordar dichos obstáculos a través de proyectos e iniciativas que promuevan una integración más efectiva y responsable de la IA en la gestión pública. De la mano de estos desafíos, también se plantean recomendaciones estratégicas en temas de gran relevancia (como el reemplazo del trabajo humano y la generación de mayores brechas digitales) y que contribuyan a adoptar una IA justa, transparente y centrada en las personas.

La Carta Iberoamericana de Inteligencia Artificial enfatiza la necesidad de que el desarrollo y la implementación de sistemas de IA se realicen con transparencia y bajo principios

1 Quiere decir que no hay transparencia en la creación de algoritmos y no se informan a las personas usuarias como fueron elaborados estos.

de rendición de cuentas. Cuando un sistema o algoritmo es responsable de tomar decisiones, resulta esencial esclarecer cómo se adoptan dichas decisiones, qué reglas se aplican y cuál fue el proceso de construcción del sistema. Asimismo, la Carta promueve la adopción de mecanismos concretos que garanticen la rendición de cuentas en el uso de la IA dentro de los procesos públicos.

De igual modo, se considera necesario que la IA sea diseñada y utilizada con **principios éticos** y bajo **estándares de derechos humanos**, que eviten cualquier forma de discriminación. En este ámbito se presta especial atención a cuestiones como el acceso universal a la IA, la seguridad, la privacidad y la protección de datos personales. Este aspecto pretende evitar el uso indebido de los datos, garantizar su calidad y confiabilidad para que no generen sesgos, de modo que los procesos de aplicación sean confiables y produzcan valor público, innovación y desarrollo sostenible.

Asimismo, la Carta establece que la inteligencia artificial debe contribuir al bienestar socioeconómico y a la mejora de la calidad de vida de las personas, mediante aplicaciones e iniciativas sostenibles en ámbitos como la salud, la educación, las finanzas, entre otros. Esta visión se vincula directamente con el ámbito de la Administración Pública, ya que el documento plantea que la IA puede integrarse en las organizaciones públicas para automatizar procesos, generar predicciones, optimizar la planificación y evaluación, fortalecer las estructuras de gobernanza, reorganizar servicios y crear valor público en políticas, servicios y modelos de atención. Además, la IA representa una oportunidad clave para transformar de manera integral el ciclo de las políticas públicas.

¿Y qué ha pasado en Costa Rica?

A partir de lo establecido en la Carta, el país debe trabajar en modelos de gobernanza de datos, así como reflexionar sobre el impacto de la IA en los derechos humanos y el medioambiente. Con base a esto, se deben priorizar cuestiones como las tecnologías limpias, la promoción de sectores económicos clave y en como la IA puede dinamizar la economía, así como en la investigación y desarrollo (I+D) y la innovación. Esta última es de especial importancia porque su fomento puede impulsar la adopción de marcos de trabajo colaborativo para que la IA sea gestionada para superar brechas sociales, optimizar la infraestructura e inclusive, fortalecer la ciberseguridad.

Para lo anterior, se requiere que el país cuente con un marco normativo de IA, así como con leyes que conexas al trabajo de la IA que garanticen la transparencia, la ética y la protección de los derechos humanos a lo interno de la normativa para que ello impulse la adopción de programas de desarrollo de competencias.

La evolución de la IA debe ser acompañada de una mejora en las capacidades y habilidades en la formación académica del estudiantado, así como de las personas que ya forman

parte de la Administración Pública y la ciudadanía en general. En ello resulta esencial que se fortalezcan las capacidades de gobernanza, seguridad y protección de datos, al tiempo que se generan lineamientos tecnológicos y marcos político-estratégicos para establecer un modelo de gobernanza de IA efectivo, duradero y sostenible.

Otro elemento requerido, es la promoción de una innovación responsable mediante incentivos, programas y espacios que conduzcan a un aprovechamiento de la IA para incidir en áreas como la salud y el medio ambiente, al tiempo que fomentan la transparencia y la participación ciudadana. Aunado a esto, se debe aprovechar el conocimiento y referencias internacionales en la materia, bajo un concepto de colaboración.

Recientemente, el Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (Micitt) publicó la Estrategia Nacional de Inteligencia Artificial (ENIA). Este documento establece principios y un marco estratégico que se alinea con las mejores prácticas internacionales, sin embargo, su modelo de gobernanza no resulta lo suficientemente claro ya que se indica que el rector es el Micitt, a través de la Agencia Nacional de Gobierno Digital y la Dirección de Ciberseguridad y ambas instancias aún no habían sido materializadas hasta noviembre del 2024.

La ENIA plantea un enfoque colaborativo, objetivos, áreas estratégicas y acciones. Esta Estrategia evidencia la aplicación de ciertas recomendaciones que establece la Carta como lo son los incentivos a la investigación y el desarrollo (I+D), la generación de capacidades, la promoción del uso de la inteligencia artificial, la integración de la IA para la mejora de los servicios público, el desarrollo de infraestructura, entre otros. Lo anterior demuestra una estrecha alineación entre los principios establecidos en la Carta y los lineamientos propuestos en la ENIA.

El documento también destaca por ser la primera Estrategia de IA que se adopta en Centroamérica, así como por resaltar la necesidad de que haya articulación territorial, es decir, que los beneficios de la IA lleguen a todos los sectores, actores y en general, a la mayoría de la ciudadanía. Sin embargo, para que ese aprovechamiento sea efectivo se requiere asegurar recursos financieros, definir responsables y garantizar la medición y evaluación de cumplimiento de las metas planteadas en la ENIA.

Conclusiones

Actualmente, la inteligencia artificial se presenta como una oportunidad transformadora que las administraciones públicas de la región pueden aprovechar para mejorar la calidad de los servicios, aumentar la eficiencia en la gestión, fortalecer la confianza ciudadana y promover un desarrollo sostenible y equitativo. Sin embargo, para que este potencial se materialice, es imprescindible que la implementación de la IA se base en principios éticos, inclusivos y orientados al bien común. En

este contexto, la Carta Iberoamericana de Inteligencia Artificial puede desempeñar un papel clave como referencia para que los países fortalezcan la transparencia, la rendición de cuentas, la protección de datos y la cooperación internacional en materia de IA.

En el caso de Costa Rica, la reciente publicación de la Estrategia Nacional de Inteligencia Artificial (ENIA) representa un avance significativo, evidenciando un alto grado de alineamiento con los principios establecidos en la Carta. Desde esta perspectiva, el país se encuentra en una posición privilegiada para ejercer un rol de liderazgo en la región, impulsando un marco normativo integral que garantice la transparencia, la ética y la protección de los derechos humanos. No obstante, la ENIA también revela la persistencia de desafíos relacionados con el fortalecimiento de los mecanismos de gobernanza, la asignación adecuada de recursos y la definición clara de las responsabilidades institucionales para su implementación.

En este sentido, garantizar un desarrollo responsable de la IA en el sector público requerirá robustecer el marco normativo nacional, aumentar la inversión en la creación de capacidades técnicas e institucionales y fomentar una innovación responsable, socialmente inclusiva y centrada en las personas. La colaboración internacional, la adopción de buenas prácticas y el monitoreo continuo de los impactos de la IA serán factores esenciales para lograr una integración sostenible, ética y justa de estas tecnologías en la gestión pública.

16

El futuro es hoy: La IA como catalizador de oportunidades para Costa Rica

Marlon Ávalos Elizondo

Marlon Ávalos Elizondo. Director de Investigación, Desarrollo e Innovación del Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (Micitt).

Máster en Gestión de la Innovación Tecnológica de la Universidad Nacional de Costa Rica, especialidad en Gestión de Proyectos de Ciudades Inteligentes de la Universidad de Salamanca (España) y del Gobierno de Singapur, así como en Comunicación de Mercadeo y Periodismo. Actualmente se desempeña como Director de Investigación, Desarrollo e Innovación del Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (MICITT). Durante 15 años, ha trabajado como asesor en comunicación política, gestión de proyectos de innovación, transformación digital, planificación estratégica para ciudades inteligentes, diseño y estrategia digital, así como en la colaboración con medios de comunicación, agencias de cooperación y organismos internacionales.

Marlon Ávalos Elizondo

Resumen

La inteligencia artificial (IA) se ha consolidado como una tecnología clave para el desarrollo económico y social a nivel global. Costa Rica, consciente de este potencial transformador, ha diseñado la Estrategia Nacional de Inteligencia Artificial (ENIA) como un instrumento de política pública que busca promover la adopción ética, segura y centrada en las personas de la IA. Desde este enfoque la ponencia describe el contexto internacional de la IA, los avances nacionales en su implementación y los retos que enfrenta el país en términos de regulación, capacidades humanas e infraestructura tecnológica. Asimismo, se destacan las líneas de acción que la ENIA propone para fomentar el aprovechamiento de la IA en distintos sectores, impulsar el desarrollo territorial y mejorar la calidad de vida de la población.

Palabras clave: inteligencia artificial, despliegue de la inteligencia artificial, estrategia nacional, innovación tecnológica, ética digital.

Introducción

Actualmente, la inteligencia artificial (IA) es considerada como una de las tecnologías más disruptivas de nuestra era, pues está modificando la manera como ocurren diversos procesos económicos, sociales y políticos. Su acelerado desarrollo a nivel mundial ha demostrado el potencial de la IA para impulsar mejoras significativas en áreas como la salud, la educación, la producción agrícola y la eficiencia de los servicios públicos. Esto no solo ha promovido mayores inversiones para fortalecer las infraestructuras y la investigación y desarrollo (I+D), sino que también ha develado la existencia de retos éticos, normativos y laborales de gran relevancia.

Costa Rica no ha sido ajena a esta transformación y es por eso que, durante los últimos años, nuestro país ha mostrado un creciente interés por integrar la IA de una forma estratégica y responsable. Para ello, ha optado por desarrollar la Estrategia Nacional de Inteligencia Artificial (ENIA), un instrumento de política pública con el que se busca aprovechar las oportunidades que ofrece la IA y mitigar sus riesgos mediante el uso ético, seguro y responsable de la IA, centrado en las personas y alineado con los valores democráticos que caracterizan al país.

Bajo esta perspectiva, la ponencia presenta un panorama general del contexto global y nacional de la IA y ante esto, se destacan los principales desafíos que Costa Rica enfrenta en términos de infraestructura, gobernanza, regulación y talento humano; así como las intervenciones estratégicas que se pretenden implementar en el marco de la ENIA.

Panorama Global de la Inteligencia Artificial

La capacidad de la IA para procesar grandes volúmenes de datos, aprender de ellos y tomar decisiones informadas ha permitido avances significativos en áreas como la salud, la educación, la agricultura y el sector público. Esto indica oportunidades económicas y la posibilidad de que la IA se convierta en una herramienta para mejorar los servicios y transformar positivamente a nuestra sociedad. Por eso no es de extrañar que durante los últimos años se haya dado un incremento en el número de Estados y empresas que están invirtiendo más recursos en esta tecnología.

De hecho, según Goldman Sachs, se proyecta que la inversión global en IA se acerque a los USD$200.000 millones de dólares para 2025 y alcance hasta los USD$600.000 millones de dólares al 2030 (cifra que representa el doble del Producto Interno Bruto (PIB) combinado de los países centroamericanos). Asimismo, se estima que la inversión relacionada con IA podría alcanzar entre el 2,5% y el 4% del PIB en Estados Unidos, y entre el 1,5% y el 2,5% del PIB en otros líderes en IA.

Se cree que el gasto empresarial para la adopción de la IA tendrá un impacto económico acumulado a nivel mundial de USD$20 billones hasta 2030, año en el que esta tecnología generará el 3,5% del PIB del planeta. Costa Rica no ha sido ajena a esta tendencia y por el contrario datos recientes de Microsoft indican que más del 70% de las Pymes del país están invirtiendo o planean invertir en IA; mientras que el 14% ya destina más del 25% de su presupuesto a esta tecnología.

Además, según el Centro de Investigación y Estudios Políticos (CIEP) el 70% de las personas jóvenes entre los 10 y los 35 años conocen sobre la IA; lo que refuerza la idea de que la IA se ha posicionado a nivel nacional, no sólo a nivel empresarial sino también en la ciudadanía en general.

¿Cómo queremos maximizar la IA en Costa Rica?

En este contexto de crecimiento acelerado de la IA, el país ha identificado una oportunidad para maximizar esta tecnología a partir del diseño de un instrumento de política pública con el que se promueva el uso, la adopción y se impulse el desarrollo de la IA de forma ética, segura y responsable; procurando beneficios para la ciudadanía y evitando que su uso cause algún daño a las personas. Con este propósito nació la Estrategia Nacional de Inteligencia Artificial (ENIA) la cual se diseñó sobre cuatro pilares fundamentales:

1. Diseño de marcos éticos y normativos
2. Generación de capacidades y habilidades en la población
3. Fomento de la adopción de IA
4. Mejora de los servicios públicos

La construcción de la ENIA fue liderada por el Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (MICITT) y adoptó un enfoque participativo e intersectorial en el que se analizaron instrumentos nacionales e internacionales y estrategias que están implementando otros países, con el fin de analizar tendencias y entender hacia donde se puede orientar el desarrollo de la IA de acuerdo a la realidad nacional. A partir de esto se busca posicionar a Costa Rica como un referente en la adopción responsable y sostenible de la IA a nivel mundial.

Uno de los principios clave que interesa promover desde la ENIA es que los sistemas de inteligencia artificial, especialmente aquellos de alto riesgo, deben estar sujetos a supervisión humana. Esto es particularmente importante en áreas sensibles como la salud o la selección de personal. Desde este punto de vista se pretende distinguir entre los sistemas de IA en los que se busca mayor automatización de los servicios e interesa que tengan una intervención humana más reducida; de los sistemas críticos de alto riesgo que pueden afectar los derechos de las personas y que por tanto, requieren de supervisión humana.

La estrategia también contempla la prevención de sesgos en los sistemas de IA que implementen la Administración Pública, el sector privado y la Academia. Para esto, la ENIA contempla el establecimiento de lineamientos obligatorios para las instituciones del gobierno central y recomienda que otros sectores promuevan la investigación, desarrollo e innovación (I+D+i) y generen un marco de gestión de riesgos.

Riesgos de la IA, lo que debemos gestionar

Como parte del ejercicio reflexivo realizado en el marco de la ENIA, se han identificado diversos riesgos que pueden desencadenarse con la implementación de la IA en el país, si esta no llega a ser direccionada de una manera adecuada. En ese sentido, el primer desafío que se identifica es la **sostenibilidad de la IA**, pues el uso de esta tecnología tiene un impacto directo en el ambiente por la cantidad de energía y recursos que consume, así como por el calor que emiten los centros de datos en los que se entrenan los sistemas de IA. Para solventar esto, la ENIA propone acciones que desde la política pública mitiguen el consumo energético de las instituciones públicas al adoptar sistemas de IA.

Otro riesgo importante es el **desplazamiento de los puestos de trabajo**, para lo cual la ENIA dispone un fuerte componente para promover la generación de capacidades, impulsar el upskilling y el reskilling, considerando el impacto que la IA generará en el trabajo y la formación académica de las personas. En ello, se considera fundamental la vinculación con instituciones como el Ministerio de Educación Pública (MEP), el Instituto Nacional de Aprendizaje (INA) y las universidades.

También se identifican riesgos como la **desinformación, la ciberseguridad y la protección de datos personales**. Esto último implica impulsar reformas legales desde la Asamblea Legislativa y fortalecer las capacidades de ciberseguridad del país

Oportunidades para sectores clave

La IA está produciendo una transformación clave en sectores estratégicos como la salud, la educación y la agricultura. Es por esto, que desde el MICITT se ha venido trabajando en distintos proyectos piloto entre los que cabe mencionar: 1) el uso de la IA, el IoT y la analítica de datos para mejorar la producción agrícola en tres cantones de la zona de Los Santos y 2) aplicaciones de IA para la detección temprana de enfermedades como en cáncer en colaboración con la Caja Costarricense de Seguro Social (CCSS) y el sector privado[1].

Figura 16.1. Transformación de sectores clave

Fuente: Elaboración propia.

1 Este tipo de iniciativas muestra el potencial para promover alianzas público-privadas que generen valor público.

Proyectos como los mencionados previamente muestran que la IA ofrece grandes oportunidades que deben ser aprovechadas por nuestro país, para lo cual resulta necesario que:

- Se desarrolle la infraestructura tecnológica para ampliar la capacidad de servicio y la potencia computacional.

- Impulsar y mejorar el talento especializado que se forma, ya que si bien se cuenta con capital humano de calidad, persiste una brecha de profesionales en IA que no permite cumplir con las demandas que el mercado exige en este ámbito.

- Se requiere de un marco regulatorio que no limite el despliegue la inteligencia artificial. En este momento, se encuentran tres proyectos de ley en corriente legislativos, de los cuales el Micitt no comparte criterio pues considera que los tres son restrictivos para la IA.

- Se requiere más inversión pública y privada en I+D+i. En este momento el país no cuenta con la suficiente capacidad para financiar proyectos de I+D+i y por ello, es esencial generar mecanismos de financiamiento diferenciados para instituciones públicas y el sector privado.

Adicionalmente, Costa Rica debe aprovechar su compromiso con la sostenibilidad, la estabilidad política y la calidad del talento humano para continuar posicionándose a nivel internacional y aprovechar estas fortalezas para atraer inversión extranjera directa (IED) en IA.

Líneas de Acción de la Estrategia

En línea con las necesidades identificadas, la ENIA pretende contribuir al despliegue de la IA mediante las siguientes líneas de acción:

1. **Inteligencia artificial ética y responsable**. Esto debe estar presente en el diseño de mecanismos regulatorios y la normativa técnica que se genere.
2. **Articulación territorial y el desarrollo económico**. Esto es de especial relevancia, pues la innovación está muy concentrada en la Gran Área Metropolitana (GAM). Esto supone promover este tipo de procesos en las provincias periféricas y las zonas fronterizas.
3. **Promoción de la IA en el sector público.** Con esto se pretende que el Micitt fomente la integración de la IA en instituciones del gobierno central y descentralizadas y que emita un marco de recomendación para instituciones autónomas, las municipalidades, el Poder Judicial, la Asamblea Legislativa y el Tribunal Supremo de Elecciones (TSE), entre otras.
4. **Formación de talento.** En este ámbito se busca capacitar y formar talento humano con el apoyo de instancias como el Ministerio de Educación Pública (MEP) y el Instituto Nacional de Aprendizaje (INA).

5. **Desarrollo de infraestructura digital.** Para esto se pretende alinear las acciones de la ENIA con instrumentos que ya están en ejecución como el Plan Nacional de Desarrollo de Telecomunicaciones (PNDT) 2022-2027.
6. **Liderazgo internacional.** Implica aprovechar los espacios de diálogo regional e internacional en los que está Costa Rica como el Sistema Centroamericano de Integración (SICA), el Partnership on AI, la Alianza Global para la Inteligencia Artificial (GPAI) y el Comité de Gobernanza de la Inteligencia Artificial de la Organización para la Cooperación y Desarrollo Económicos (OCDE). Esto permite que el país tenga acceso al conocimiento y experiencias de otros países, así como posicionar al país en el tema de la IA.

Además, el país está colaborando con el SICA en el desarrollo de la Estrategia Regional de Inteligencia Artificial para la Región Centroamericana y se espera que durante el 2025 de a conocer, de manera conjunta con la OCDE, una caja de herramientas para guiar en la implementación de los principios de Inteligencia Artificial de la OCDE. Finalmente, debe mencionarse que, en la ejecución de estas líneas de acción, se adoptará un enfoque de articulación intersectorial de modo que ello permita integrar los aportes del sector público, la academia, el sector privado, la sociedad civil y los organismos internacionales presentes en el país, entre otros.

Conclusiones

Costa Rica ha comprendido que la inteligencia artificial no es solo una herramienta tecnológica, sino un motor estratégico para impulsar su desarrollo, así como para transformar los campos de la salud, la educación, la producción agrícola y la gestión pública, entre otros. En ese sentido, la formulación de la ENIA representa un paso muy importante para integrar esta tecnología en distintos sectores socio productivos del país, así como un claro reconocimiento del potencial transformador y de los riesgos que puede generar la implementación de un salto tecnológico con dirección.

Esta estrategia, enraizada en valores democráticos y centrada en las personas, busca garantizar que la adopción de la IA contribuya al bienestar colectivo. Además, el diseño de la estrategia muestra la puesta en práctica de un enfoque inclusivo y multisectorial en el que la participación de una diversidad de actores ha permitido la identificación de necesidades estratégicas como: el fortalecimiento de la infraestructura tecnológica, la creación y fomento de habilidades especializadas y de contar con una regulación ética de la IA, pero que no frene su desarrollo.

No obstante, persisten desafíos estructurales que deben ser abordados para catalizar los beneficios de la IA en nuestro país. La brecha digital, la desigualdad territorial en el acceso e infraestructuras tecnológicas y la falta de inversiones en I+D+i, son solo algunas de las barreras que requieren atención

urgente para establecer condiciones favorables al despliegue de la IA. Por otro lado, no debe olvidarse que la sostenibilidad y efectividad de la ENIA dependerán de la voluntad política, la obtención de financiamiento, la articulación interinstitucional y la continuidad de los esfuerzos institucionales a largo plazo.

S

Síntesis analítica

Las intervenciones contenidas en esta Memoria evidencian las complejidades que se intersecan y surgen con la inteligencia artificial (IA). Desde este punto de vista la IA no sólo debe ser vista como una poderosa herramienta con capacidad transformadora, sino también como un fenómeno sociotécnico, multidimensional y con una amplia gama de impactos que continuaremos experimentando durante los próximos años y décadas. El potencial disruptor de esta tecnología se avizora no sólo como una oportunidad para modernizar a las Administraciones Públicas, hacer más eficientes a las empresas y mejorar todo tipo de procesos; sino también como un medio para que los países impulsen estrategias de desarrollo que les permitan posicionarse en los mercados y negocios emergentes que puede fomentar la IA.

Ante un contexto de competencia global, situar la mirada hacia la IA se ha vuelto un asunto estratégico y por ello, su integración y desarrollo ha pasado a ser una cuestión de atención prioritaria. Sin embargo, esta carrera tecnológica no debe limitarse al mero ejercicio de integrar una nueva tecnología, sino que esta debe considerar aspectos que puntualicen el objetivo, el alcance y los resultados buscados mediante estos procesos de transición tecnológica. En esto resulta esencial impulsar condiciones habilitantes para el desarrollo de la IA, así como encontrar un balance entre la innovación y el respeto de la dignidad humana en el que se aborden las tensiones y los nuevos dilemas éticos surgidos con esta tecnología emergente.

La diversidad de enfoques plasmadas en la Memoria, permiten identificar al menos 7 consideraciones principales que deben tenerse en cuenta a la hora de estructurar y definir acciones que estén destinadas a promover el desarrollo y utilización de la inteligencia. Estas se sintetizan en los apartados que se presentan a continuación.

1. La IA, una herramienta con grandes potencialidades

Los avances en materia de automatización son vistos como un aspecto central que ayudará a revolucionar a distintos sectores productivos (como la agricultura, las telecomunicaciones y la educación, entre otros) gracias a la vinculación de la IA con otras tecnologías disruptivas como las redes de quinta generación (redes 5G), el Internet de las Cosas (IoT) y los gemelos digitales. Los beneficios asociados con esta integración tecnológica contemplan la optimización de recursos, la posibilidad de anticipar necesidades, fortalecer la toma de decisiones, promover la innovación, modernizar la gestión pública, mejorar la transparencia y la calidad de los servicios e inclusive, contribuir a reducir el impacto ambiental.

Estas potencialidades podrían ayudar a las organizaciones para que generen mayor valor público, sin embargo, más allá de que la IA puede llegar a constituirse en una herramienta que propicie el desarrollo económico y social, su implementación no deja de enfrentar retos, así como sugerir la presencia de riesgos que deben ser examinados cuidadosamente.

2. Grandes capacidades, pero también riesgos importantes

Ningún desarrollo tecnológico es neutral o completamente autónomo, ya que en estos procesos se cuenta con una par-

ticipación humana que influye en las decisiones que determinan el diseño, la intencionalidad y el propósito con el que son fabricadas las tecnologías. La IA no es la excepción, pues en la construcción de los sistemas de IA, así como la selección de los datos de entrenamiento suelen intervienen creadores humanos que plasman sus visiones, preferencias e inclusive prejuicios sobre los mismos. Esto puede llevar a que los sistemas de IA exhiban sesgos algorítmicos que favorecen desigualdades y producen discriminación hacia determinadas poblaciones, lo que conlleva el riesgo de que estas tecnologías refuercen dinámicas de exclusión social y una concentración del poder, en lugar propiciar la democratización de oportunidades.

La falta de transparencia es una fuente de preocupación constante, puesto que la opacidad de los sistemas de IA dificulta entender la manera como estas tecnologías toman decisiones o generan determinados resultados. De igual modo, los avances de la inteligencia artificial generativa (IA generativa) han provocado retos no sólo con respecto a la autoría de un contenido generado y la veracidad del mismo, sino también porque determinada información puede generar manipulación, desinformación, discursos de odio y polarización social en nuestras democracias.

Otro riesgo es la posible exacerbación de la brecha digital entre países y a lo interno de los mismos. Debido a que el desarrollo de la IA (al igual que otras tecnologías) sigue siendo dominado por países del Norte Global, se plantea la importancia de impulsar acciones que fortalezcan las capacidades de aquellos países con menor ventaja tecnológica. A ello puede contribuir el desarrollo de tecnologías de IA que reflejen las prioridades y necesidades de los contextos socioculturales locales, la protección de derechos y la reducción de brechas históricas, así como la creación de visiones conjuntas que permitan la participación de diversos actores. También puede ayudar la construcción de apoyos y esfuerzos regionales que permitan el posicionamiento en los espacios de debate global sobre la IA.

Por otro lado, a nivel doméstico debe evitarse que la utilización de la IA ocurra de manera desigual ampliando las brechas existentes en áreas, segmentos de la población o entre los sectores productivos de los países, lo que quiere decir que es necesario democratizar su desarrollo. Tampoco debe olvidarse el impacto ambiental que trae aparejado la masificación y el uso intensivo de la IA. Cuestiones como el consumo de energético, la extracción de minerales requeridos para las infraestructuras que soportan a los sistemas de IA, la huella de carbono y hasta la generación futura de residuos de aparatos eléctricos y electrónicos (RAEE) ejemplifican solo algunas de las repercusiones ambientales que deben ser tenidas en los procesos de transformación tecnológica.

Por último, la IA también representa riesgos a nivel ético. Entre estos se encuentran la vigilancia masiva, la dependencia tecnológica, la pérdida de la privacidad y la deshumanización, así como otras dinámicas que aún no han sido explorados y que conforme avance la tecnología, es de esperar que aparezcan otras nuevas.

3. **Un fenómeno complejo, un abordaje integral**

La IA conforma un terreno complejo con aristas diversas, lo que la convierte en algo más que un campo de la informática. Esto quiere decir que no se acaba en aspectos técnicos como los datos, los algoritmos o los sistemas, sino que por el contrario, sus implicaciones sociales, culturales, políticas e incluso, antropológicas, evidencian interrelaciones entre la ciencia, el derecho, la filosofía y las ciencias sociales. Por esto requiere de abordajes multi e interdisciplinarios que abonen a su adecuada comprensión. Lo anterior es fundamental para atender a retos que en primera instancia pueden parecer paradójicos, pero que cristalizan desafíos a los que nos enfrentamos ante la IA (por ejemplo, promover la innovación sin dejar de proteger los derechos fundamentales).

La inter y multidisciplinariedad es vital para fomentar una colaboración efectiva y posible gracias a la participación de actores diversos (academia, gobierno, sector privado, organizaciones de sociedad civil y ciudadanía) en las discusiones sobre IA.

4. **Resignificar lo humano**

La inteligencia artificial no es un fin en sí misma y por el contrario, su desarrollo debe colocar a la persona humana en el centro. Esto nos permite cuestionarnos a fondo los impactos que la IA puede ocasionar en nuestras vidas y reflexionar a profundidad sobre sus repercusiones éticas, al tiempo que nos ayuda a preguntarnos por el sentido de la transición tecnológica en nuestras sociedades. En consecuencia, la integración de la IA debería ser realizada con la intención de mejorar nuestra calidad de vida, contribuir en la resolución de problemas públicos o potenciar nuestras capacidades humanas, entre otros propósitos que pueden estar orientados al bienestar colectivo.

A su vez, retomar lo humano implica incluir una gestión de riesgos consciente y rigurosa de la IA en la que puedan prevenirse situaciones indeseables y riesgos potenciales. Más que identificar riesgos y amenazas o aplicar estrategias para prevenirlos, se debe reconocer que cada decisión técnica trae aparejada una dimensión moral que repercute en el diseño e implementación de los sistemas de IA. No tomar en cuenta el aspecto moral puede alejarnos de la posibilidad de proteger a las personas y de fortalecer su confianza en la tecnología, evitando estas situaciones que socaven los derechos humanos, la privacidad o la seguridad de la información.

Asimismo, debido a que los datos son un activo clave que tiene un valor ético, político y humano, en su gestión se debe velar por la protección de derechos fundamentales y evitar que se reproduzcan desigualdades estructurales. Por eso es necesario que la ética sea integrada en el diseño de los sistemas de IA como un componente central de todo el ciclo de vida de la tecnología. Para esto, se puede recurrir a la aplicación de marcos internacionales que faciliten la evaluación y la mitigación de riesgos, como MITRE, ATLAS y OWASP AI/ML.

Las indagaciones sobre la relación humano-máquina están lejos de acabar y probablemente continúen con los avances científi-

ficos venideros que no harán más que complejizar lo humano, resignificando las intersecciones entre el componente biológico, los elementos tecnológicos, la psicología y las relaciones humanas; surgiendo con ello nuevos interrogantes y dilemas.

5. ¿Es posible regular la IA?

El surgimiento de nuevos ecosistemas digitales dominados por la presencia de tecnologías emergentes como la IA, demanda mecanismos para potenciar sus beneficios y mitigar riesgos. Por tal motivo, se ha considerado oportuno regularla bajo estándares éticos que propicien el uso seguro, sostenible e informado de la IA. En consecuencia, la regulación no puede quedarse en cuestiones éticas superficiales, sino abogar por un abordaje que profundice en el impacto e implicaciones de esta tecnología.

Sin embargo, esto es un reto no sólo porque esta tecnología cambia con suma rapidez, sino también porque existe una tendencia a crear marcos normativos rígidos que no favorecen la flexibilidad ni la adaptación, los cuales son condiciones muy necesarias cuando se está frente a una tecnología como la IA. En ese sentido, la regulación no puede continuar siendo estática o aplicar modelos tradicionales, sino que debe permitir la innovación y mitigar riesgos potenciales de la IA. Al mismo tiempo, deben adoptarse marcos éticos y regulatorios sustentados en principios éticos, derechos humanos y normas claras que prevengan riesgos como la discriminación algorítmica, la opacidad, la tecnovigilancia y/o la concentración de poder, entre otros.

Estas tensiones entre la innovación y los marcos regulatorios tradicionales sugieren pensar en las maneras como el Estado puede potenciar el desarrollo tecnológico, anticiparse a sus impactos y promover una gobernanza digital, inclusiva, adaptativa y enfocada en la persona humana. Para eso, se sugiere evitar la creación de legislación que pueda quedar obsoleta en poco tiempo y en su lugar, adoptar enfoques regulatorios basados en la evidencia y la experimentación. Una práctica muy recomendada es la aplicación de sandbox regulatorios, previo a la emisión de cualquier regulación.

Por otro lado, la regulación de la IA no solo contempla marcos o legislaciones específicas que puedan emitirse, sino también normativas supletorias que rigen y/o pueden afectar áreas específicas de la IA como la protección de datos personales, la ciberseguridad, la propiedad intelectual, el desarrollo tecnológico y la promoción de la innovación, entre otros. En consecuencia, la regulación en IA tiene un carácter evolutivo, inter intersectorial y dinámico.

6. La necesidad de un abordaje estratégico

Maximizar las potencialidades que ofrece la IA requiere de una toma de decisiones responsable, en la que el progreso tecnológico sea promovido desde un enfoque que considere sus distintas implicancias e incorpore cuestiones como la transparencia, la sostenibilidad y los derechos humanos. En ese sentido, debe evitarse que los desarrollos y el uso de estas tecnologías respondan únicamente a intereses corporativos o económicos y por el contrario, deben incluirse activamente a los sectores históricamente marginados en estos procesos.

Las acciones de política pública juegan un papel central para orientar la integración de la IA hacia objetivos claros, identificando prioridades y aquellas condiciones necesarias para impulsar el aprovechamiento efectivo de la IA. Además, estas herramientas pueden alinear la tecnología con las necesidades reales de las personas, para lo que resulta necesario que se cuestione que se puede hacer con la IA, para quién y con qué propósito.

Para definir estos aspectos, es indispensable contar con datos y evidencias que guíen en la construcción de las herramientas de políticas públicas. Por eso, es fundamental continuar generando mediciones, índices o herramientas similares que identifiquen los progresos, las falencias y vacíos presentes en los países y en distintos sectores; lo que resulta útil a la hora de crear intervenciones públicas. Además, permiten mapear desigualdades estructurales que pueden condicionar el aprovechamiento de la IA, y a partir de ello, redireccionar la atención hacia dichos aspectos.

En estos procesos los Estados tienen un rol central, más no son el único actor relevante pues ante la multiplicidad de retos planteados por la IA, se requiere de una respuesta estratégica e inclusiva que solo puede ser alcanzada con la participación de distintos sectores que contribuyan a la construcción de marcos de acción comunes y visiones conjuntas sobre la IA. En la práctica, este tipo de abordajes ha propiciado una cooperación internacional enfocada en fomentar el intercambio de experiencias, la armonización de estándares e incluso, promover posiciones regionales con respecto a la IA.

7. ¿Cómo catalizar su desarrollo?

Apoyar el desarrollo de la IA requiere impulsar acciones que faciliten la construcción de entornos habilitantes en los que se fortalezcan las infraestructuras digitales, se aceleren las capacidades tecnológicas (como la potencia computacional) y robustezcan los ecosistemas de investigación y desarrollo (I+D) desde una perspectiva local. Aunado a ello, es necesario invertir en la formación de talento humano especializado, promover el desarrollo de startups tecnológicas y actualizar la regulación, entre otros aspectos.

UNIVERSIDAD DE COSTA RICA

PROSIC

Programa Institucional
Sociedad de la Informacion
y el Conocimiento

Jornadas de Investigación y Análisis

Inteligencia Artificial, ¿Moda pasajera o cambio permanente de paradigma?

Con la publicación de la Memoria de las Jornadas de Investigación y Análisis IA , ¿moda pasajera o cambio permanente del paradigma? el Programa Sociedad de la Información y el Conocimiento (Prosic) pone a disposición un producto del conocimiento que, desde una mirada crítica y multidisciplinaria, busca aportar a las discusiones sobre el desarrollo de la IA en Costa Rica, resaltando los retos y oportunidades que el país enfrenta ante un mundo cada vez más digitalizado, interconectado e incierto.

ISBN: 978-9968-510-31-8

9 789968 510318