

COLECCIÓN TEMAS EN DERECHOS HUMANOS
TOMO V

INTELIGENCIA ARTIFICIAL Y DERECHOS HUMANOS

PONENCIAS DE LAS JORNADAS 2024



unesco

Centro
Bajo los auspicios
de UNESCO



Centro Internacional
para la Promoción de
los Derechos Humanos

COLECCIÓN TEMAS EN DERECHOS HUMANOS

TOMO V

**INTELIGENCIA ARTIFICIAL
Y DERECHOS HUMANOS**

JORNADAS



unesco

Centro
Bajo los auspicios
de UNESCO



Centro Internacional
para la Promoción de
los Derechos Humanos

Directora Ejecutiva: Mg. Fernanda Gil Lozano
Director de Desarrollo de las Capacidades Institucionales: Dr. Rodrigo Karasik (a cargo)
Directora de Estudios e Investigación: Lic. Carolina Tellería (a cargo)
Director Técnico Administrativo: CPN Guido Napolitano
Coordinadora de Asuntos Jurídicos: Dra. Gabriela Pantuso
Responsable General de Gestión: María Tatiana Souza Korolkov
Responsable de Comunicación Institucional: Lic. Diana Martí
Responsable de Comunicación de Proyectos: Lucero Vega
Coordinación Proyecto Inteligencia Artificial: Lic. Elena Hanono
Compiladoras: Lic. Elena Hanono, María Tatiana Souza Korolkov, Dra. Gabriela Pantuso
Diseño Editorial: Lic. Marcela Braccelarghe
Corrección y Edición Gráfica: Anna Souza
Asistente de Diseño: Mercedes Araujo
Coordinadora Colección: Mg. Marina Guimpel

Publicado por:

**CENTRO INTERNACIONAL PARA LA PROMOCIÓN DE LOS
DERECHOS HUMANOS bajo los auspicios de UNESCO**

Av. España 2591, Buenos Aires, Argentina

Tel: +5411 5300 4000 extensión 78300

Las ideas y opiniones expresadas en esta obra son las de las y los autores; no reflejan necesariamente el punto de vista del CIPDH ni comprometen a la organización. Los gráficos expuestos han sido elaborados en base a la información proporcionada por cada autor

Cómo citar esta publicación: Centro Internacional para la Promoción de los Derechos Humanos bajo los auspicios de UNESCO (CIPDH). (2024). "Colección Temas en Derechos Humanos. Tomo V – Inteligencia artificial y derechos humanos", Buenos Aires, Argentina.

ISBN 978-631-90305-9-4

ISBN 978-631-90305-9-4



ÍNDICE

Prólogo	5
Introducción.....	7
Introducción a la inteligencia artificial y derechos humanos María Dolores Souza (Chile)	13
Cómo entiendo y cómo me entiende la inteligencia artificial Sebastián Waisbrot (Argentina)	29
Desórdenes informativos y narrativas hostiles María José Labrador (México)	41
Transformación digital y derechos humanos. Un diálogo a construir Gabriel Levy (Colombia)	55
Agencia humana a prueba de futuro. Cómo proteger los derechos humanos frente a las máquinas Cordel Green (Jamaica)	69
Inteligencia artificial para proteger a las personas en estado de vulnerabilidad Samridhi Arora (India)	79
Democratización del conocimiento a través de la tecnología Matías Karlsson (Argentina)	91
Educación e inteligencia artificial: un caso argentino Alicia Penayo (Argentina)	101
Ética de la inteligencia artificial: equidad de género Natalia González Alarcón (Colombia)	111

PRÓLOGO

El avance de la inteligencia artificial en los últimos años ha sido vertiginoso y está modificando múltiples aspectos de nuestras vidas. Desde la educación hasta la medicina, la IA ofrece oportunidades sin precedentes para mejorar el bienestar humano. Sin embargo, con esta transformación también surgen desafíos críticos, especialmente en el ámbito de los derechos humanos.

La IA, al depender de gran cantidad de datos, plantea serias preocupaciones: sus sistemas pueden poner en peligro nuestra autonomía y su uso indebido puede, además, perpetuar las desigualdades y vulnerabilidades existentes y generar nuevas formas de marginación, afectando gravemente el derecho a la equidad.

Un tema especialmente preocupante es la falta de transparencia y rendición de cuentas en la toma de decisiones automatizadas. En sectores como la Justicia, la salud y las finanzas, las determinaciones impulsadas por algoritmos de IA pueden ser difíciles de comprender o rechazar. Esto nos deja en una posición frágil frente a resultados abusivos, y pone en peligro derechos fundamentales como la igualdad ante la ley.

Además, el uso creciente de la IA en el control social puede amplificar prácticas de intolerancia. En efecto, la capacidad de predecir comportamientos a través de algoritmos puede generar el riesgo de que se adopten medidas arbitrarias y desproporcionadas que afecten las libertades individuales y perpetúen injusticias estructurales.

En este contexto, es imprescindible el debate sobre la ética de la inteligencia artificial. Los gobiernos, las empresas y la sociedad civil deben interactuar para desarrollar marcos normativos que garanticen que el uso de la IA esté alineado con los principios de los derechos humanos. Esto incluye diseñar sistemas éticos que atenúen riesgos, sensibilizarnos sobre las implicancias de esta tecnología y fomentar la cooperación internacional para establecer normas.

La inteligencia artificial está alterando la forma en que interactuamos con el mundo y la naturaleza del poder y, en última instancia, lo que significa ser humano. A medida que avanza, deberemos prestar atención a quienes controlan su desarrollo y su uso y qué responsabilidades y obligaciones recaen sobre ellos. En este sentido, los derechos humanos deberán orientar la regulación y observancia permanente de esta disciplina.

Este libro nos invita a reflexionar sobre los desafíos y las oportunidades que plantea la IA respecto de los derechos humanos para comprender sus aspectos técnicos y considerar en profundidad cómo podemos asegurarnos de que su desarrollo y aplicación promuevan una sociedad más libre.

La tarea es ardua, pero es fundamental para garantizar que el progreso tecnológico no debilite los derechos por los que tanto hemos luchado por promover y proteger.

Mi agradecimiento sincero a cada especialista que participó en este primer webinar sobre IA y DD.HH. y al equipo del CIPDH que asumió, con dedicación y compromiso, la realización del evento, llevándolo adelante con éxito y profesionalismo.

Mg. Fernanda Gil Lozano
Directora ejecutiva
Centro Internacional para la
Promoción de los Derechos Humanos

INTRODUCCIÓN

El Centro Internacional para la Promoción de los Derechos Humanos (CIPDH) bajo los auspicios de UNESCO procura poner en diálogo a la comunidad académica respecto de los principales desafíos globales en materia de derechos humanos que tiendan puentes entre diferentes campos de estudio.

Despliega distintos enfoques y miradas que aportan múltiples disciplinas que se ven interpeladas por los derechos humanos y por los temas que tratamos en cada uno de los tomos de esta colección al combinar artículos de tratamiento general con otros que, a través de casos concretos, exponen una problemática y su abordaje

La COLECCIÓN TEMAS EN DERECHOS HUMANOS comprende una serie de publicaciones periódicas donde se plantean y profundizan temas específicos que dan cuenta de los debates actuales y las perspectivas más novedosas de cada tópico.

El tema que trata esta obra, la inteligencia artificial (IA), se está consolidando rápidamente para transformar nuestras sociedades. Más allá de sus aplicaciones técnicas, plantea interrogantes fundamentales sobre la ética, la justicia y la responsabilidad social, especialmente en relación con los derechos humanos. Reúne perspectivas de destacadas autoras y autores que, desde sus áreas de especialización, aportan una reflexión profunda sobre el rol de la IA abordando temas críticos como el acceso al conocimiento, la libertad, la privacidad y la equidad.

El 17 y el 18 de julio de 2024, organizadas por el CIPDH y con el apoyo de la Oficina Regional de UNESCO Montevideo, se llevaron a cabo las primeras Jornadas sobre Inteligencia Artificial y Derechos Humanos. Estos encuentros tuvieron como principal objetivo promover un espacio de reflexión sobre las dimensiones éticas, los riesgos y los desafíos inherentes al impacto de la inteligencia artificial en la sociedad contemporánea.

Desde una visión integral, cada texto de esta compilación contribuye a comprender cómo se puede utilizar la IA como he-

ramienta para mejorar las condiciones de vida y promover los derechos humanos. A la vez, enfatiza en que este objetivo puede alcanzarse solo si la IA es puesta al servicio de la humanidad, con regulaciones sólidas y supervisión social activa. En este sentido surge la preocupación de un nuevo modo de feudalismo informativo con poder de manipulación y control de la información. A través de cada exposición se plantea la importancia de integrar principios que orienten el desarrollo y la implementación de la IA considerando sus beneficios y riesgos inherentes.

Los trabajos que presentamos aquí ofrecen un análisis crítico de la temática, teniendo en cuenta una variedad de dimensiones, perspectivas geográficas, de género y de trayectorias académicas de sus autores y autoras. Aunque se busca capturar un fenómeno global y complejo, es necesario aclarar que este enfoque hace necesariamente un recorte que limitará la inclusión de algunos conceptos relevantes para una comprensión integral del problema.

María Dolores Souza aporta una mirada introductoria sobre la IA y su interacción con la vida cotidiana. Señala que la IA puede contribuir a mejorar vidas, aunque también genera la necesidad de establecer políticas, principios y regulaciones nacionales e internacionales a fin de aumentar su gobernanza y minimizar los riesgos. En su exposición desarrolla el problema de los sesgos generados a partir de análisis estadísticos realizados por modelos de IA que tienden a subrepresentar a las minorías y a las opiniones divergentes, lo que puede generar la reproducción de desigualdades estructurales. Souza subraya, asimismo, la diversidad de iniciativas regulatorias de distintas regiones y organismos internacionales a nivel mundial.

Sebastián Waisbrot desarrolla en su ponencia los conceptos de algoritmo, modelo, inteligencia artificial y aprendizaje automático, y proporciona definiciones y ejemplos que facilitan su comprensión. Muestra su preocupación por la posibilidad de sesgos algorítmicos generados a partir de muestras sesgadas, contaminación de datos o características sustitutivas. También aborda la noción de daño algorítmico con sus características específicas de opacidad, escala y daño. En resumen, señala que la IA enfrenta desafíos considerables pero, a la vez, brinda numerosas ventajas en diversas áreas y sostiene la necesidad de fomentar espacios de debate y reflexión conjunta.

María José Labrador estructura su presentación en tres grandes ejes. El primero se centra en el fenómeno de la desinformación, donde prioriza definir su concepto, importancia y clasificación. En el segundo eje examina las vulnerabilidades y los aspectos éticos relacionados con la inteligencia artificial en el contexto de la desinformación. Para finalizar, ofrece un resumen de las investigaciones y estudios que lleva a cabo el Observatorio Desinformación Chile, del cual es fundadora.

Gabriel Levy afirma que, en la actualidad, la humanidad se enfrenta a un contexto complejo en el que la tecnología avanza a una velocidad vertiginosa reconfigurando las dinámicas de vida y relación entre las personas. Este artículo ofrece una visión sobre cómo la IA y otras tecnologías emergentes influyen en la vida cotidiana y en la estructura de la sociedad. Expone, de modo claro, la aceleración de los tiempos de desarrollo y la llegada a millones de personas de tecnologías como el telégrafo, el teléfono, la radio y la web. Señala también que este progreso acelerado conlleva importantes desafíos: la ampliación de la brecha por las disparidades en el acceso a la tecnología y a la información que, necesariamente, origina desigualdades económicas y suscita interrogantes acerca de los derechos humanos en un entorno digital.

Gabriel Levy y María José Labrador abordan el impacto de la desinformación basada en la utilización de la IA sobre la estructura social. Ambos plantean preocupaciones sobre los efectos de la manipulación de la información y las vulnerabilidades éticas que emergen de su uso.

Cordell Green ofrece una mirada crítica sobre los efectos de la IA en las libertades personales y la privacidad. Su planteo se centra en cómo impactan las nuevas tecnologías en la autonomía humana y señala que sin una regulación ética la IA podría erosionar derechos fundamentales como los mencionados. Expone, asimismo, cómo el control de la información digital ha dado lugar a una “nueva forma de feudalismo informativo” donde las grandes corporaciones tecnológicas asumen el rol de actores con poder en la manipulación y el control de la información. Aboga por la creación de un marco global de derechos digitales que contemple una “alfabetización mediática y digital” como derecho humano para asegurar que la ciudadanía sea consciente de cómo interviene la IA en su vida cotidiana.

Tanto Cordel Green como Gabriel Levy subrayan el modo en que la aceleración de tecnologías emergentes como la IA afecta las dinámicas sociales y plantea desafíos para los derechos humanos tales como la privacidad y la equidad en el acceso a la información.

Samridhi Arora explora el poder de la IA como herramienta de inclusión social. Su enfoque se centra en su uso para mejorar la vida de personas en situación de vulnerabilidad, desde aquellas con discapacidad hasta quienes no tienen acceso a servicios básicos como Justicia o salud. Reclama establecer marcos éticos claros que prioricen a las personas por encima de los beneficios comerciales, asegurando que la IA sirva para humanizar, no para mecanizar. Desde su óptica, la inteligencia artificial bien usada puede democratizar derechos y proporcionar oportunidades a quienes han estado históricamente marginados siempre que se implante desde un compromiso genuino con la Justicia.

Coincide con los especialistas Levy, Green y González Alarcón en la necesidad de que la IA actúe como un mecanismo de promoción de la inclusión social y abordaje de las desigualdades, especialmente en relación con los sectores más vulnerables, y priorice el bienestar de los seres humanos.

Desde la perspectiva de Matías Karlsson, la IA representa una oportunidad única para democratizar el conocimiento y subraya que, aunque el acceso a la información se ha ampliado con la era digital, la capacidad crítica para discernir su valor es fundamental para su aprovechamiento. Analiza cómo la IA puede transformar el sistema educativo y el acceso al conocimiento, pero advierte que esto solo puede lograrse si hay un cambio significativo en la educación y en la formación de una mentalidad crítica. En este sentido, la IA es un recurso que debe estar al servicio de la humanidad, promoviendo una distribución justa del conocimiento para superar las barreras tradicionales de acceso y fomentar una ciudadanía informada.

Alicia Penayo destaca la transformación de la educación y el acceso al conocimiento a partir de la incorporación de la IA en Misiones, Argentina. Señala que el reciente proyecto de ley provincial sobre su integración en las aulas propone un enfoque progresivo que incluye la capacitación docente y la mejora de la infraestructura tecnológica. Su objetivo es asegurar que la IA

funcione como una herramienta pedagógica que fomente habilidades futuras y respete los derechos humanos. Aboga por políticas inclusivas y equitativas para crear un modelo educativo en el que la IA y el aprendizaje colaboren a formar ciudadanos digitales informados y empoderados capaces de enfrentar los desafíos de una sociedad en constante cambio.

Al igual que Green y Karlsson, Penayo resalta el papel de la alfabetización digital, fundamental en la formación de una ciudadanía crítica que potencie el acceso equitativo al conocimiento.

Natalia González Alarcón centra su presentación en la capacidad de la IA para profundizar o reducir las desigualdades y brechas de género. Para la autora, la IA debe ser no solo un insumo técnico sino, también, un agente de cambio que promueva la inclusión. Desde su trabajo en UNESCO, impulsa la implementación de principios éticos que permitan una representación justa y diversa en datos y algoritmos. Su propuesta destaca la importancia de integrar a mujeres y a grupos vulnerables en el desarrollo de la IA para evitar que se amplíen las brechas sociales existentes.

González Alarcón, Gabriel Levy y Samridhi Arora destacan la necesidad de que la IA actúe como un mecanismo de promoción de la inclusión social y abordaje de las desigualdades, especialmente en relación con los sectores más vulnerables.

Cada especialista reconoce tanto los beneficios como los riesgos y desafíos que presenta la IA. Consideran que puede constituir una herramienta que, puesta al servicio de la humanidad, permita mejorar la vida cotidiana y promover la inclusión social, aunque advierten sobre los sesgos algorítmicos y las desigualdades que puede generar. Un tema recurrente es la necesidad de políticas y marcos regulatorios para minimizar los riesgos y proteger los derechos fundamentales así como sostener enfoques éticos y humanos.

En resumen, nuestra intención ha sido presentar una reflexión amplia y exhaustiva y abrir nuevos interrogantes sobre los retos sociales y legales de la inteligencia artificial en el contexto de los derechos humanos. Cada autor, cada autora, ha planteado, a su manera, los usos positivos de la IA y las respon-

sabilidades relacionadas con su desarrollo. Sobre la base de este enfoque interdisciplinario, los invitamos a reflexionar acerca del papel de la inteligencia artificial en un contexto donde la tecnología debe constituirse en un medio al servicio de la humanidad.

Lic. Elena Hanono

Lic. Diana Martí

INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL Y DERECHOS HUMANOS

MARÍA DOLORES SOUZA (CHILE)

Psicóloga, Universidad Real de Leiden, Holanda. Se desempeñó como jefa del Departamento de Estudios del Consejo Nacional de Televisión (regulador de la televisión) en Chile durante treinta y un años (hasta mayo de 2023). Creó el proyecto de televisión educativa Novasur en 1999 (hoy CNTV Infantil) y la plataforma de reguladores del audiovisual iberoamericano (PRAI) en 2010. Actualmente conforma el grupo de trabajo UNESCO en acceso a la información (IFAP, por su sigla en inglés).

RESUMEN

La aplicación de la inteligencia artificial cubre esferas técnicas y del ámbito social. En el campo de la interacción humana, de sus instituciones y del funcionamiento de la sociedad, la IA genera valor, sentido, confianza, y por ello estas herramientas deben utilizarse con seguridad y sin posibles riesgos. La IA puede contribuir a mejorar vidas y brinda soluciones basadas en la ciencia, aunque también genera la necesidad de establecer políticas, principios y regulaciones nacionales e internacionales a fin de aumentar su gobernanza y minimizar los riesgos.

Gran parte de los procesos de toma de decisiones se realizan mediante algoritmos, que están programados a partir de datos. Para sistemas de modelado de lenguaje a gran escala, como ChatGPT y similares, los investigadores pueden utilizar una importante variedad de contenidos ya codificados en grandes bases de datos de libre acceso. La calidad de los resultados de un modelo dependerá en gran medida de los datos recopilados. La mayoría de los análisis estadísticos realizados por modelos de IA clasifican los datos codificados y ejecutan estadísticas que tienden a representar en exceso los promedios y a jerarquizar los datos en base a frecuencias. Con ello se tiende a subrepresentar minorías u opiniones divergentes, por lo que estos modelos pueden reproducir desigualdades estructurales y sesgos.

Existen variadas iniciativas regulatorias de países y organismos internacionales, como la Recomendación UNESCO sobre la Ética de la Inteligencia Artificial (noviembre 2021) y el marco europeo sobre IA, que es la primera legislación del mundo en estas materias (noviembre 2023). En Estados Unidos hay iniciativas de política general, pero no planes específicos para regular la IA generativa. En América Latina todavía queda mucho por hacer en términos de regulación y gobernanza. Las regulaciones existentes en nuestra región se relacionan principalmente con la ciberseguridad y la protección de datos en general, pero no abordan las especificidades de la inteligencia artificial.

Se requiere responsabilidad de parte de los usuarios de tecnologías, no solo de los gobiernos o entidades supranacionales y empresas tecnológicas, para hacer oír su voz y pedir por estándares de DD.HH., seguridad y privacidad.

Palabras clave

inteligencia artificial generativa, derechos humanos, algoritmos, sesgos, regulaciones, gobernanza algorítmica

ABSTRACT

The application of artificial intelligence covers technical and social spheres. In the field of human interaction, its institutions, and the performance of society, AI generates value, meaning, trust and that is the reason why these tools must be used safely and without possible risks. AI can contribute to improving lives and provide science-based solutions, but it also generates the demand for national and international policies, principles and regulations in order to increase its governance and minimize risks.

Much of the decision-making processes are performed by algorithms, which are programmed from data. For large-scale language modeling systems, such as Chat GPT and similars, researchers can use a significant variety of content already encoded in large, freely accessible databases. The quality of the results of a model will depend to a large extent on the data collected. Most statistical analyses carried out by AI models classify the coded data and run statistics that tend to over-represent averages and rank the data based on frequencies. This tends to under-represent minorities or differing opinions, so these models can reproduce structural inequalities and biases.

There are various regulatory initiatives undertaken by countries and international organizations, such as the UNESCO Recommendation on the Ethics of Artificial Intelligence (November 2021) and the European Union's Framework for Artificial Intelligence (November 2023) which is the first legislation of the world on AI. In the United States, there are general policy initiatives, but no specific plans to regulate generative AI. In Latin America, there is still much to be done in terms of regulation and governance. Existing regulations in our region are mainly related to overall cybersecurity and data protection, but do not address the specificities of Artificial Intelligence.

Responsibility is required from technology users and not only from governments or supranational entities and technology companies to make their voices heard and ask for human rights, security and privacy standards.

Keywords

artificial generative intelligence, human rights, algorithms, algorithmic governance, biases, regulations, algorithmic governance

Introducción

Cuando hablamos de inteligencia artificial nos referimos a ciencia y tecnología, especialmente el desarrollo de la ingeniería, la física, las matemáticas, la robótica y otros. Pero la aplicación de la inteligencia artificial no solo cubre esferas técnicas, sino también las del ámbito social. Hoy en día se utiliza en las salas de redacción de medios, en campañas políticas, redes sociales y otros ámbitos relevantes de la esfera pública.

Con esta tecnología, por un lado, se resuelven tareas específicas, pero en el amplio campo de la interacción humana, de sus instituciones y del funcionamiento de la sociedad tenemos que pensar que con la IA existe creación de valor, de sentido, de confianza, y por ello se deben utilizar estas herramientas con seguridad y sin posibles riesgos.

Hacer que las máquinas sean inteligentes es permitirles aprender de los datos y “funcionar apropiadamente y con previsión en su entorno”,¹ como afirma Nils Nilsson, definición que también se utiliza en el glosario de derecho práctico del Reino Unido.²

Según la Organización de los Estados Americanos, “la inteligencia artificial generativa y las tecnologías detrás impactan los cuatro pilares fundamentales de la OEA”³:

DESARROLLO - DEMOCRACIA - DERECHOS HUMANOS - SEGURIDAD

Lo expresado anteriormente está en línea con los planteos de la OCDE (OCDE, marzo de 2023).⁴

El desarrollo de la inteligencia artificial es muy relevante: puede actualizar grandes cantidades de datos para orientar programas sociales, ayuda a mejorar vidas y brinda soluciones basadas en la ciencia para el acceso a alimentos, agua y energía y salud, según la OEA.

Dicho esto, la IA nos genera al mismo tiempo la necesidad de establecer políticas, principios y regulaciones nacionales e internacionales a fin de aumentar su gobernanza de acuerdo con nuestros objetivos y minimizar los riesgos.

IA como tecnología potente y algoritmos como sistemas de toma de decisiones

En nuestra vida cotidiana, gran parte de los procesos de toma de decisiones se realizan mediante algoritmos, que están programados a partir de una serie de datos. Pueden servir para sugerirnos qué serie ver en Netflix, o qué periódico leer en línea, en función de nuestras elecciones pasadas. En las organizaciones se utilizan algoritmos para la selección de personal, para conceder un préstamo bancario y para reconocimiento facial. En el sistema judicial, por ejemplo, los algoritmos de evaluación de riesgos se pueden utilizar para predecir quién tiene probabilidades de cometer un delito o convertirse en víctima. Y así, un sinnúmero de utilidades.

Según el nivel de importancia, la resolución de problemas y la toma de decisiones en base a inteligencia artificial conlleva diferentes tipos de riesgos, con distintos niveles de daño. Algunos de ellos son producto de errores técnicos o humanos, como en cualquier otro ámbito de la vida. Pero otros riesgos son intencionales, como falsedades alimentadas en los sistemas para generar información errónea y noticias falsas, o la suplantación de identidad.

De ahí que es importante conocer en parte cómo funcionan estos sistemas, por ejemplo los Modelos de Lenguaje Masivos o a Gran Escala (LLM en inglés).⁵

¿De dónde provienen los datos?

La inteligencia artificial trabaja con datos que deben tomarse de alguna parte. Para sistemas de modelado de lenguaje a gran escala, como ChatGPT y similares, los investigadores pueden usar conjuntos de datos disponibles. Se trata de una gran variedad de contenido, por ejemplo, de historia, literatura, información académica, de las ciencias médicas, de procedimientos judiciales, e incluso contenidos de e-mails públicos ya codificados en grandes bases de datos de libre acceso.

La calidad de los resultados de un modelo dependerá en gran medida de los datos recopilados. Según los científicos, es necesario mejorar los grandes conjuntos de datos, ya que no se trata solo de grandes volúmenes, sino también de diversidad de datos. Así se pueden añadir bases de datos más pequeñas, diversas y de alta calidad. Pero los conjuntos de datos son frag-

mentos de información ya existente, proveniente de diferentes fuentes —como se ha dicho— y, en su mayoría, en inglés. Así, están insertas cultural e históricamente en un contexto de tradiciones, valores y visiones de mundo.

La IA generativa y sus sistemas “aprenden” de los datos existentes que fueron creados por humanos específicos en momentos y lugares específicos. Por lo tanto, la data puede tener sesgos de género, étnicos o religiosos o contenido peyorativo que a veces resultan invisibles, porque están naturalizados en nuestra vida cotidiana.

Desde las ciencias sociales sabemos que la cultura no es estática, sino que evoluciona con el tiempo, al igual que las definiciones de derechos humanos. Así, lo que parece normal en un lugar o tiempo, puede considerarse inadecuado en otro tiempo o lugar. Por ello se han desarrollado una serie de estrategias —entre los científicos y codificadores— para abordar el problema de modo que los usuarios puedan detectar contenido potencialmente inapropiado.⁶

Existen softwares que incluyen modelos de toxicidad para identificar ciertos tipos de contenido —por ejemplo, contenidos sexualmente explícitos—. También se han creado sistemas que pueden ejecutar pruebas de co-ocurrencia para descubrir si hay prejuicios hacia algunos grupos de personas: por ejemplo, asociando estos grupos con palabras que provocan un sentimiento negativo. Aun así, algunos investigadores⁷ han señalado que, para todos los datos demográficos, el sentimiento promedio tiende a ser negativo, y la hipótesis es que esto se debe al contexto real de los datos, por ejemplo, artículos, análisis de hechos actuales y contenido de noticias, que suelen tener connotaciones negativas. Este es un ejemplo de cuán ligados están nuestros datos a la cultura.

Se han corregido estos problemas alineando las máquinas con ciertos valores, por ejemplo, de diversidad social. Pero en algunos casos esto ha llevado a que en los resultados de una búsqueda la máquina “corrige” el resultado. Algunas búsquedas históricas han arrojado mujeres papisas —corrigiendo la tradición de que los sumos pontífices de la Iglesia Católica son todos hombres—, o que entre los soldados nazis hubiera imágenes de soldados de color.

Este y otros ejemplos ya han ocurrido.

¿Y qué sucede con los modelos matemáticos y la estadística?

Hemos visto que la IA trabaja con datos, pero ¿cómo funciona? ¿Cómo analiza los datos? En esta etapa también podemos enfrentarnos a una serie de problemas. A continuación se describen dos de ellos.

1. El primero es que los modelos se basan en información y datos ya procesados —tales como artículos académicos, libros y noticias— y, por tanto, no podrán prever situaciones emergentes o nuevas formas de razonamiento. Y aunque el análisis puede tener en cuenta una gran cantidad de datos, no puede transmitir significado a situaciones sociales complejas ni analizar contenidos de forma crítica, algo propio de la interpretación humana. Un ejemplo de esto es la serie audiovisual generada por IA “Nada para siempre”, estrenada en Twitch y X (Twitter), que debió ser retirada por contenido homofóbico. De hecho, un artículo de Santiago López de la Cruz, de la Universidad de la República del Uruguay, llama “pastiche” a estas narrativas digitales —copiadas de contenidos preexistentes— y afirma que, de este modo, “cancelan” el futuro.⁸

La mayoría de los análisis estadísticos realizados por modelos de IA clasifican los datos codificados y ejecutan estadísticas, tales como análisis de conglomerados y correlaciones, que tienden a representar en exceso los promedios y a jerarquizar los datos en base a frecuencias, por ejemplo, con respecto a grupos de población o tipos de opiniones.

Con ello se tiende a subrepresentar minorías u opiniones divergentes, por lo que estos modelos pueden reproducir desigualdades estructurales y sesgos. Esto se puede ver en el documental *Coded Bias*, que contó con la contribución de varios investigadores y el trabajo de Joy Buolamwini, del Instituto de Tecnología de Massachusetts (MIT).⁹ Esta investigadora descubrió cómo el aprendizaje automático y los algoritmos pueden conducir a ciertos resultados asociados a cierto tipo de personas. Sus principales ejemplos se centran en los softwares de reconocimiento facial y

demuestra el sesgo de estos en relación con personas no blancas.

Estos sesgos pueden corregirse a nivel técnico, y de hecho se hace entre los investigadores y técnicos expertos. Pero para ello estas personas deben tener un propósito científico y/o ético explícito. Por ejemplo, velar por la no discriminación y el respeto a valores sociales como la pluralidad, la inclusión y la participación. Dichos valores pueden servir como estándares para una inteligencia artificial ética, en términos de promover buenas prácticas a nivel técnico, lo que tendrá incidencia en los resultados.

Es por eso que, frente a análisis complejos en materia social, el uso de datos y sus instrumentos de análisis estadístico requiere de un enfoque multidisciplinario y de una política, a fin de considerar lineamientos inspirados en principios de derechos humanos.

2. Un segundo desafío que presentan estos modelos tiene relación con que la mayoría de las personas no sabe cómo funcionan los algoritmos, y muchas veces los directivos de entidades públicas o privadas que hacen uso de estas herramientas tampoco lo saben. En ese sentido, existe una falta de transparencia respecto de cómo funcionan dichos instrumentos y de la forma como se toman decisiones en base a sus resultados.

Para un individuo, las decisiones algorítmicas son muy relevantes, porque podrían resolver un problema apremiante, por ejemplo ante una necesidad financiera de salud o vivienda por lo cual ha acudido a un banco, un ministerio público o una municipalidad.

Existe falta de “gobernanza algorítmica” o cualquier otra forma propia del ámbito legal. Tal como se mencionó, si la persona no es experta no puede entender todos los procesos detrás de la codificación y clasificación de datos, tampoco puede identificar un error, ni puede saber si en algún momento el proceso de codificación o de análisis fue supervisado por humanos. Finalmente, tampoco conocemos las razones para utilizar un algoritmo en particular y no otro.

Las tecnologías surgen de un medio ambiente científico y cultural —como se ha dicho—. Por lo tanto es importante que estas tecnologías que sirven a la toma de decisiones no se desarrollen en un único ambiente, porque la IA generativa utiliza ciertos mecanismos y reglas basadas en un punto de vista o visión de mundo. Es como crear las reglas del juego. Entonces, para el desarrollo de la IA hay que preguntarse cuáles son los valores de ese ambiente científico y técnico, del sistema político, incluso del país en el que surge. Debido a esto, la investigación y el desarrollo no deberían estar en manos de unas pocas empresas de innovación tecnológica, para evitar la concentración. Tampoco de un país en particular.¹⁰

Es decir, la tecnología es hoy una preocupación tan importante como la capacidad nuclear o el desarrollo de misiles.¹¹

En el desarrollo tecnológico hay distintos niveles de agencia. A un nivel está la proveniencia de los fondos para un modelo. Luego, el equipo humano —los desarrolladores— que crea un cierto modelo para ciertos fines. También está la plataforma que lo difunde; y el sistema de IA en sí mismo, que crea datos nuevos, o lleva a ciertos resultados. Esto es importante para la gobernanza y para determinar responsabilidades.

El aprendizaje de las máquinas no es transparente, no siempre tiene trazabilidad.

Por otro lado, la IA ha traspasado las fronteras de realidad y ficción, algo que la sociedad ya ha experimentado con noticias falsas y otros fenómenos comunicacionales, con efecto desestabilizador a nivel personal y social.

De ahí que en el mundo han surgido iniciativas regulatorias y/o de buenas prácticas desde hace un tiempo.

Iniciativas regulatorias

Los países se han enfrentado al problema de los riesgos de la IA desde hace un par de años. Por ello, han surgido variadas iniciativas de SOFT LAW o DERECHO BLANDO.

UNESCO

La Recomendación Sobre la Ética de la Inteligencia Artificial de UNESCO (noviembre de 2021) se enfoca en el diseño de los sistemas de IA que pueden funcionar con distinto grado de autonomía, como el aprendizaje de las máquinas y el razonamiento, que puede incluir sistemas de planificación y otros. En este contexto, la UNESCO sostiene que las cuestiones éticas atañen a todas las etapas de ciclo de creación de estos softwares, así como a todos los actores involucrados en este ciclo, partiendo por el financiamiento y terminando con el uso de sistemas de IA.

Propone que se eviten algoritmos que refuerzan sesgos, que se respete la dignidad humana, que se aborden inquietudes culturales y ambientales, entre otros. Para ello, se debe educar en nuevas competencias frente a la digitalización, involucrar a las ciencias exactas y sociales, respetar la identidad y diversidad cultural y respetar la igualdad, promoviendo también la participación, la comunicación y la información.



Esta Recomendación está dirigida tanto a las autoridades responsables y a los reguladores como a las empresas tecnológicas, de medios y hasta el usuario final.

En este contexto, por ejemplo, la UNESCO hizo un llamado (junio de 2024) para que se tomen acciones urgentes frente a la desinformación y el discurso de odio, temáticas que aborda la Recomendación antes citada.

EE.UU. DE AMÉRICA

En Estados Unidos existen iniciativas de política general, pero no planes específicos para regular la IA generativa, si bien la orden ejecutiva del presidente Biden de octubre de 2023 contiene ocho principios guía que el gobierno federal debe considerar. Por un lado se entregan lineamientos para el desarrollo de la IA, y por el otro se sugiere la supervisión permanente de investigación y pruebas. También se hace un llamado a convocar a distintos actores, incluyendo la sociedad civil



Luego existen otros dos cuerpos regulatorios sobre derechos y otros sobre manejo de riesgos, para avanzar en equidad y derechos civiles. Sin embargo, la Corte Suprema de ese país ha solicitado recientemente más especificidad en este sentido. Por ejemplo, el Acta de este año (mayo de 2024) sobre seguridad de la IA está hoy en proceso de revisión en el Senado. Si se transforma en ley, se podrán trazar los incidentes relacionados con la IA que han violado la seguridad de sistemas, personas o comunidades.

LATINOAMÉRICA

En cuanto a nuestra región, el Índice Latinoamericano de Inteligencia Artificial¹² afirma que es importante que los países tengan una estrategia sólida sobre IA, y que dicha estrategia debe institucionalizarse con una participación mínima de las partes interesadas, para garantizar su implementación y gobernanza. Además, el índice sugiere que la estrategia debe cumplir con los criterios de bien común acordados por la UNESCO. Chile tiene hoy una política en IA en base a Metodología Unesco de Evaluación de la Preparación¹³ de cada país para introducir esta tecnología.

Desafortunadamente, todavía queda mucho por hacer en la región latinoamericana en términos de regulación y gobernanza —de acuerdo con este índice—. Las regulaciones se relacionan principalmente con la ciberseguridad y la protección de datos en general, pero no abordan las especificidades de la inteligencia artificial (con algunas excepciones).¹⁴

En cuanto al sector privado, algunas empresas en el mundo también han transparentado su compromiso voluntario con ciertas reglas valóricas y de seguridad, como por ejemplo la empresa de IA Open AI, la que creó ChatGPT. Elaboró diez reglas (2023), las que se han revisado y adaptado para presentarlas internacionalmente.¹⁵

EUROPA. REGULACIÓN VINCULANTE

Europa ya contaba con una legislación general para las plataformas en línea: la DSA¹⁶. A esto se suman otras dos piezas legislativas, el acta europea sobre Libertad de Medios, que atañe a los contenidos, y la regulación de transparencia, y en particular en relación con la propaganda electoral.

En dicho contexto surge el Marco sobre Inteligencia Artificial (fines de noviembre de 2023), que se transforma así en la primera regulación del mundo sobre inteligencia artificial¹⁷ y se abrirá para la firma en septiembre de este año. Fue un trabajo de dos años y participaron cuarenta y siete Estados, la UE y once no miembros, como Argentina, México, Costa Rica, Perú y Uruguay, más observadores del sector privado.



Es un cuerpo legal con principios y reglas basado en los instrumentos de DD.HH. Su enfoque es de riesgos, e identifica contenidos generados por sistemas de IA, clasificando sus riesgos en cuatro categorías, de acuerdo con los usos de la IA.

Contiene requerimientos de transparencia y de supervisión, hechos a medida para contenidos específicos. Asimismo, medidas para enmendar o corregir errores o faltas.

En mayo recién pasado se realizó una convención del Consejo de Europa sobre AI, DD.HH. y el imperio de la ley, en la que se hizo un llamado a generar un marco legal con principios y reglas aplicable internacionalmente.



Conclusión

Como se ha visto, existe una discusión robusta para velar por un sistema de IA que promueva el bien común y evite posibles daños a nivel individual o social, si bien la legislación ha ido en general detrás del desarrollo tecnológico, como suele suceder en este ámbito.

Es por ello que existe una responsabilidad de parte de los usuarios de tecnologías y no solo de los gobiernos o entidades supranacionales y empresas tecnológicas, para hacer oír su voz y pedir por estándares de DD.HH., seguridad y privacidad.

NOTAS

¹ Nils Nilsson (2010), en Andreas Jungherr, “Inteligencia artificial y democracia: marco conceptual”. <https://journals.sagepub.com/doi/10.1177/20563051231186353>

² Una definición actual generalmente aceptada de IA se basa en dos pasos, abordando la inteligencia de las máquinas y luego las cualidades de la inteligencia: “La inteligencia artificial es la actividad dedicada a hacer que las máquinas sean inteligentes, y la inteligencia es la cualidad que permite a una entidad funcionar apropiadamente y con previsión en su entorno”. Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge University Press, 2010). El informe *AI and Life* también proporciona una definición más empírica de la IA como una rama operativa de la informática, afirmando que “el campo de la IA es un esfuerzo continuo para impulsar la frontera de la inteligencia de las máquinas” y señala la paradoja de que cuando la IA deja de estar en la frontera o en la vanguardia, deja de ser considerada IA. De manera más prosaica, “es solo IA hasta que sabes lo que hace, entonces es solo software”. <https://uk.practicallaw.thomsonreuters.com/Glossary/UKPracticalLaw/l25ebae628e5f11e9adfea82903531a62?-transitionType=Default&contextData=%28sc.Default%29>.

³ OEA (4 de marzo de 2023), Washington DC, Mesa Redonda de Alto Nivel de Diálogo Político. “IA: imperativos de políticas públicas para las Américas”, borrador de palabras de apertura.

⁴ Organización intergubernamental de 38 países.

⁵ LLM: Large Language Models.

⁶ “The Pile”, texto en inglés de 825 GB, en su mayoría de fuentes académicas y profesionales. “Common Crawl”, corpus con 46 % de texto en inglés y el resto en otros seis idiomas.

⁷ Leo Gao *et al.* (diciembre de 2020). “The Pile”. <https://www.eleuther.ai/>.

⁸ Santiago López de la Cruz. “Un vínculo paradójico”. Universidad de la República de Uruguay. https://www.researchgate.net/publication/371148491_Un_vinculo_paradojico_narrativas_audiovisuales_generadas_por_inteligencia_artificial_entre_el_pastiche_y_la_cancelacion_del_futuro

⁹ Ver el documental de Netflix *Coded Bias* (2020).

¹⁰ En Estados Unidos, por ejemplo, se han tomado medidas para que

Rusia o China no tengan acceso a los microchips norteamericanos.

¹¹ En este contexto se ubica la propuesta de la ONU (2019) de prohibir las armas autónomas, es decir, ataques sin intervención humana.

¹² CENIA (Centro Nacional de inteligencia artificial) (2023). Chile. <https://indicelatam.cl/>

¹³ Metodología RAM (Readiness Assessment Methodology). Chile es hasta ahora el primer país en adoptarla.

¹⁴ Brasil y Colombia y en segundo lugar Chile y Perú.

¹⁵ Cumbre de Seúl, presencial y virtual, llevada a cabo en mayo de este año (2024), sobre inteligencia artificial, seguridad, innovación e inclusión.

¹⁶ DSA: Digital Services Act.

¹⁷ <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>

CÓMO ENTIENDO Y CÓMO ME ENTIENDE LA INTELIGENCIA ARTIFICIAL

SEBASTIÁN WAISBROT (ARGENTINA)

Programador. Docente de Derecho Informático en la Cátedra de Justicia Digital de la Facultad de Derecho (UBA). Trabajó en grandes empresas tecnológicas internacionales como Twitter, y dictó numerosas charlas en conferencias de tecnologías informáticas en los últimos años. Se especializó en aprendizaje automático. Es miembro del Observatorio de Derecho Informático Argentino (ODIA).

RESUMEN

En esta exposición se desarrollan los conceptos de “algoritmo”, “modelo”, “inteligencia artificial” y “aprendizaje automático”, con definiciones de ellos y con ejemplos que permiten comprenderlos.

“Algoritmo” puede definirse como un conjunto de operaciones que permite hallar la solución de un problema. Cuando se habla de “modelo” se trata de un esquema teórico que ayuda a facilitar la comprensión y el estudio de un comportamiento. El concepto de “inteligencia artificial” se refiere a la comparación con la mente humana. En cuanto al “aprendizaje automático”, es una rama dentro de la inteligencia artificial que toma los datos, aplica un algoritmo y en base a eso genera un modelo, que puede servir para predecir y explicar comportamientos.

También se aborda la noción de “daño algorítmico”, que es comparable a un arma de destrucción masiva y que se caracteriza por contar con tres componentes: la opacidad, la escala y el daño.

Otro concepto a desarrollar es el sesgo. Se puede pensar un sesgo como un error sistemático. Los modelos que fueron creados por aprendizaje automático pueden tener sesgos algorítmicos, que consisten en una desviación que se va a repetir cada vez que el modelo haga una predicción. Estos sesgos en general son muy difíciles de evitar y tienen diversos orígenes, aunque principalmente se generan por los datos utilizados: puede haber muestras sesgadas, contaminación de los datos y, por último, las características sustitutivas o proxy.

El último tema abordado es el marco jurídico, en el que se desarrollan las distintas modalidades del daño.

A modo de conclusión, se puede decir que el aprendizaje automático actualmente presenta desafíos significativos, pero ofrece diversas ventajas en diferentes áreas.

Son importantes estos espacios de debate, de reflexión compartida y poder hacerlo en comunidad.

Palabras clave

algoritmo, modelo, inteligencia artificial, aprendizaje automático, excedente conductual, daño algorítmico, sesgo, marco jurídico

ABSTRACT

The concepts of algorithm, model, artificial intelligence and machine learning are developed in this presentation, with definitions of the concepts and examples that allow understanding them.

Algorithm can be defined as a set of operations that allows finding the solution to a problem. A model is a theoretical scheme that helps to facilitate the understanding and study of a behavior. The concept of artificial intelligence refers to the comparison with the human mind. As for machine learning, it is a branch within artificial intelligence, which takes data, applies an algorithm and based on that generates a model, which can be used to predict and explain behaviors.

It also addresses the notion of algorithmic damage, which is comparable to a weapon of mass destruction and is characterized by three components: opacity, scale and damage.

Another concept to be developed is bias. A bias can be thought of as a systematic error. Models that were created by machine learning can have algorithmic biases. It is a deviation that is going to be repeated every time the model makes a prediction. These biases in general are very difficult to avoid, they have various origins although mainly the bias is generated by the data used: there may be biased samples, contamination of the data and finally, proxy characteristics.

The last topic addressed is the legal framework, in which the different types of damage are developed.

By way of conclusion, it can be said that machine learning currently presents significant challenges, but offers several advantages in different areas

It is important to have these spaces for debate, for shared reflection and to be able to do so in community.

Keywords

algorithm, artificial intelligence, algorithmic harm, algorithmic model, bias, machine learning, legal framework, behavioral surplus, algorithmic damage

Algoritmos

“Algoritmo” es una palabra muy utilizada, aunque no todo el mundo está familiarizado con lo que significa. La Real Academia Española dice en su definición que es un conjunto infinito de operaciones que permite hallar la solución de un problema. Simplificando, puede asemejarse a una receta de cocina. Es decir: “tengo hambre y quiero comer una tortilla, ¿qué tengo que hacer? Tengo que agarrar un huevo, partirlo, batirlo, etc.”.

Toda esa serie de pasos que ayudan a llegar a la solución de un problema es un algoritmo.

Otra definición necesaria es el significado de “modelo”. Siguiendo a la Real Academia Española, es un esquema teórico. Pero el concepto más importante es que ayuda a facilitar la comprensión y el estudio de un comportamiento. Por ejemplo, en economía existe el clásico modelo de oferta y demanda que plantea que cuando aumenta el precio la cantidad ofertada va a ser mayor, mientras que la cantidad demandada va a ser menor. Esta es una simplificación de la realidad, es una abstracción, es una forma que ayuda a comprender la realidad. Pero no es necesariamente algo real, algo tangible, algo visible.

El matemático George Box, en 1976, lo puso en palabras: “Todos los modelos son incorrectos, pero algunos son útiles”. O sea, necesitamos las dos cosas. Por un lado, el modelo, necesariamente, tiene que ser, de alguna manera, incorrecto porque un modelo que sea correcto es tan grande que no sirve. Es como un mapa de escala uno a uno, como el cuento de Borges “Del rigor en la ciencia”.¹

Para poder comprender un fenómeno se requiere simplificarlo, y eso va a llevar a ciertos errores, pero se requiere que resulte útil para predecir, comprender el problema. Un ejemplo en un fenómeno cultural son los ejes cartesianos de orientación política, que tienen ejes de derecha, izquierda, autoritario, libertario, y es posible ubicar a distintas personas o países en estos ejes. Se entiende que esta es una simplificación de la realidad, que no se puede medir la distancia entre Gandhi y Mandela, pero ayuda a entender con mayor o menor precisión dónde están paradas respecto de ciertos ejes, ciertas opiniones políticas o algunas personalidades.

Por último, la definición de “inteligencia artificial” de la Real Academia Española habla también de la comparación con la mente humana. Cuando se hace una analogía, con un concepto que es complicado o desconocido, o que no se conoce lo suficientemente bien, la analogía pierde fuerza. La idea de una analogía es poder entender mejor algo a través de una comparación.

El aprendizaje automático es una rama dentro de la inteligencia artificial. Se lo define como la creación de modelos usando algoritmos sobre datos. Los datos son fundamentales, y en el caso del aprendizaje automático lo que se hace es tomar los datos, aplicar un algoritmo y en base a eso generar un modelo. Un modelo que puede servir para predecir y explicar comportamientos.

La ciencia de la computación trae los algoritmos, las estadísticas traen los modelos y la sociedad lo interpreta de alguna manera como una inteligencia artificial.

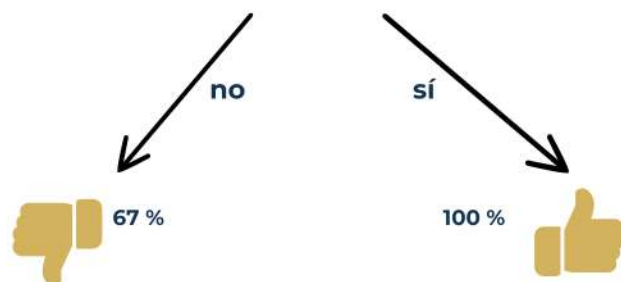
Un ejemplo muy simplificado de cómo funciona un algoritmo de aprendizaje automático es el de un recomendador de películas. Se pueden elegir o no películas como *Interestelar* o *Piratas del Caribe*, es decir, cuál es del gusto de una persona determinada y cuál no, y tratar de crear un modelo a través del aprendizaje automático. Es decir, que pueda predecir si la película le va a gustar o no.

Con estas películas se puede armar un modelito muy sencillo, obviamente incorrecto. Quizás sea útil, pero definitivamente incorrecto, en cuando a qué inferencia puede hacer sobre los gustos de una persona a través de estos datos. Por ejemplo, si se selecciona una película dramática, el modelo va a ofrecer otro drama y considerar que hay un 100 % de chances de que sea una película que le guste, y si no es un drama hay un 67 % de chances de que no le guste. Esa es la inferencia que va a sacar el modelo de aprendizaje a través de estas películas.

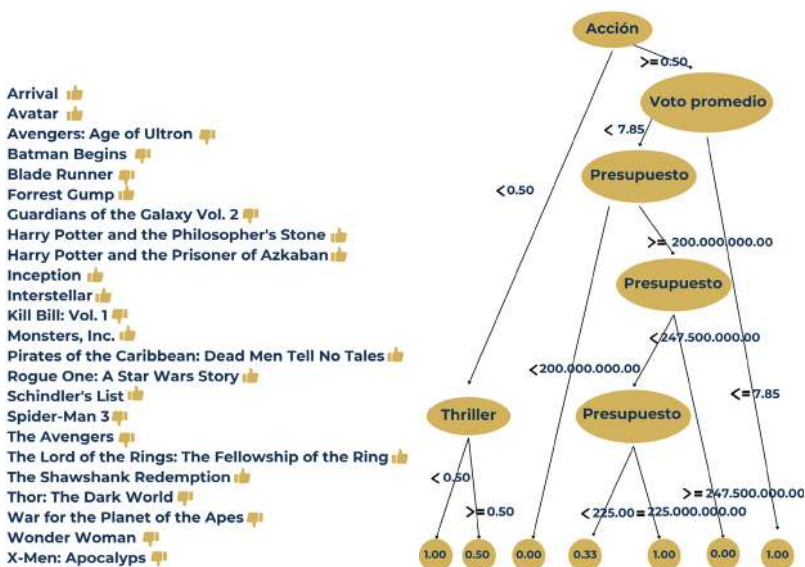
Esto se puede seguir complicando agregando cada vez más datos. Por un lado, existen el volumen de datos y la calidad de los datos. En este caso se consideran todas estas películas con sus respectivos géneros, el presupuesto con el que se armaron, el voto que les dieron distintas personas respecto de cuánto les gustaron, etc.

Ejemplo de aprendizaje automático

¿Es un drama?



Ejemplo de aprendizaje automático



En base a eso, se puede generar un modelo de aprendizaje automático a través del aprendizaje automático que sirva para predecir si una película va a gustarle o no a alguien. En este caso, si es una película de acción o si no lo es, o si es un thriller, hay un 50 % de posibilidades que sea del gusto de esta persona, y ahí se puede empezar a recorrer este árbol, este modelito

que se creó a través de un algoritmo. Este sería un ejemplo muy sencillo de un modelo de aprendizaje automático que sirve para entender cómo funcionan este tipo de modelos.

Excedente conductual

Shoshana Zuboff propone el concepto de “excedente conductual”, que plantea que cuando las personas hacen cosas, particularmente en internet, a través del software van dejando trazas de lo que van haciendo. Todo ese comportamiento y todos estos datos que van dejando generan información que después puede explotarse para hacer predicciones de comportamiento, crear nuevos productos o mejorar los productos existentes.

Un ejemplo muy común dentro de la industria del software son las pruebas AB. Estas pruebas consisten en que una empresa genere dos versiones de un mismo producto con pequeñas diferencias entre ellas y determine cuál de las dos tiene mejores resultados según las estadísticas y la función que quiere maximizar o minimizar. Esto sería el equivalente médico de los ensayos clínicos, las pruebas controladas aleatorizadas, donde a un grupo de control le dan solución salina, no le dan el componente activo, a otro sí se lo dan, y de esa manera determinan cuál es la diferencia entre ambos grupos. Las empresas de software hacen esto constantemente para mejorar sus productos e ir experimentando de manera continua cuál de ellos es el mejor.

Daño algorítmico

A continuación se aborda la noción de daño algorítmico y de cómo se puede llegar a generar un daño. El daño algorítmico se basa en el concepto de “armas de destrucción matemática”, de Cathy O’Neil, que en su libro que lleva ese nombre define estas armas con un juego de palabras con arma de destrucción masiva, y las caracteriza por tres componentes, tres elementos.² Plantea que cuando se está frente a un algoritmo que tiene estas tres propiedades nos encontramos frente a un arma de destrucción matemática.

El primero de ellos es la opacidad, que es el no poder entender cómo funciona.

Esto funciona en realidad a dos escalas. Primero, como usuarios de un servicio, de un producto, no se tiene acceso a la información de la empresa: qué decidió, cuál fue su motivación o qué algoritmo utilizó. E incluso aunque se tuviese acceso, suponiendo que existiera un procedimiento judicial que obligara a abrir las computadoras y se pudiera ver todo el código fuente, surgiría el problema de que estos algoritmos tienen una complejidad intrínseca tan grande que una persona no puede comprenderlos. Por eso, en general, se habla del concepto de “caja negra”, con gran oscuridad u opacidad. Como dice la autora, una persona puede proporcionar datos y el modelo puede hacer una predicción, pero es muy difícil leer el modelo, es muy difícil interpretar qué es lo que el modelo pensó, qué razonamiento realizó para llegar a esa predicción. Entonces, con esos datos puede hacer una predicción, pero lo que pasa por dentro es demasiado complicado para que una persona pueda comprenderlo.

Otra de las propiedades que reconoce la autora es la escala. Si un modelo que recomienda películas es utilizado por una sola persona, no habría ningún problema. Aunque pueda ser opaco, no afectará la vida de miles o millones de personas, no llegará a un volumen importante de posibles afectados. En ese caso no sería un problema particularmente grave porque la escala no es considerable.

El último concepto, que es fundamental, es el daño.

Si existe un modelo que recomienda películas y llega a ser escalable, el daño de recomendar una película mala no es muy significativo para las pobres víctimas que van a tener que ver una mala película: simplemente esa película no les va a gustar y van a haber desperdiciado dos horas de su vida.

Entonces, hay que ponderar qué podría pasar si hubiera vidas en juego, o si se estuviera hablando de despedir personas en un trabajo, que sí es algo significativo (los que estudian abogacía o los que son abogados están acostumbrados a pensar el derecho de daños).

Hay que considerar estas tres cualidades para definir si se está frente a un arma de destrucción matemática, para determinar la importancia que se le asignará a este modelo.

Sesgos algorítmicos³

Otro concepto a desarrollar es el tema del sesgo.

En cuanto a la definición de sesgo, es un término que en realidad viene de la estadística y luego se popularizó. Se puede pensar un sesgo como un error sistemático. Es decir, si a todas las reuniones a las que asiste una persona llega diez minutos tarde tiene un sesgo, tiene un error sistemático, porque siempre se equivoca de la misma manera. Pero si un día llega tarde y otro día temprano no se trata de un sesgo. Simplemente es mala estimando el tiempo. Pero si siempre llega tarde sí es un sesgo, un error sistemático.

En particular, estos modelos que fueron creados por aprendizaje automático tienen un sesgo algorítmico. Es un sesgo que se va a repetir una y otra vez, cada vez que haga una predicción. Estos sesgos en general son muy difíciles de evitar, ya que tienen diversos orígenes, aunque principalmente se generan por los datos utilizados.

Un ejemplo de esto son las muestras sesgadas. Si los datos que se van a utilizar para que el modelo aprenda no son representativos de la población, es decir que fueron tomados de un subconjunto de personas específico y que tienen ciertas cualidades distintas del resto, cuando se los quiera aplicar sobre otro conjunto humano no van a dar buenas predicciones. Va a tener un error sistemático en ese otro conjunto humano simplemente porque quienes lo componen no fueron representados en el momento de recolectar los datos.

Otro posible problema es la contaminación de los datos. Un ejemplo es la contaminación por subjetividades, como ocurre en el caso de las noticias como fuente de datos, que pueden tener un sesgo negativo. En este caso se trata de un sesgo subjetivo que le da el periodista, el reportero que escribe, el que brinda la información que se va a publicar y el que decide presentar lo que tienen. Se trata del que contamina los datos con su propia perspectiva. Entonces, cuando se utilice el modelo no va a ser representativo neutral de la realidad, sino que va a estar contaminado por la perspectiva de las personas que le pusieron su propia impronta al interpretar los datos recibidos.

Esta no es una enumeración taxativa ya que la cantidad de posibles problemas con los datos es vasta e inabarcable.

Por último, es necesario mencionar las características sustitutivas, que en inglés se llaman proxis. Suponiendo que se quiera hacer un modelo que sea neutro en cuanto a género, es decir, no se quiere discriminar entre mujeres y hombres y se busca que tengan las mismas respuestas frente los mismos datos. La primera decisión que se puede tomar en este caso es no incluir la pertenencia o no de la persona a un género. Simplemente no se le dice al modelo si esta persona es hombre o mujer, entonces se va a tratar a todas las personas por igual.

El problema que tiene esto es que, en el mundo real, hombres y mujeres fueron tomando distintas decisiones a lo largo de su vida, y eso está correlacionado con otras variables que son, a primera vista, neutrales. Por ejemplo, el uso de redes sociales. Si se dice que una persona usa Pinterest no se va a estar especificando si es hombre o mujer, pero sí habrá ciertas probabilidades de que sea uno u otro. Porque simplemente hay más mujeres que usan Pinterest que hombres. Esto es algo que se da en la práctica. No es algo que se pueda modificar.

Se trata de cómo funciona el mundo que conocemos, que es un mundo bastante imperfecto. Simplemente existen distintas correlaciones entre las variables. Entonces, aunque el modelo use todos estos datos que son neutros en cuanto a género, puede llegar a tener ciertas inferencias en lo relativo a género.

Marco jurídico

El último tema a abordar es el marco jurídico con el que se cuenta.

Para los que no son abogados, hay distintas áreas del derecho que tienen distintas normas jurídicas. Por ejemplo, cuando se habla de daño (la materia tradicional de daño es el derecho civil), donde la persona que lo ocasiona tiene el deber de reparar a la persona afectada, se evaluará la conducta en el sentido de si actuó con prudencia o no, si fue negligente o si fue doloso. Todo esto entra dentro del análisis del derecho civil. Esto aplica-

ría tanto a un daño tradicional como chocar con el auto como a un daño ocasionado a través de un algoritmo creado con aprendizaje automático.

De la misma manera, el derecho penal sanciona la actividad o el resultado y no el medio con el que se realizó. Es decir, no hay un tipo penal que sea asesinato con un arma, asesinato con otra arma, asesinato con una enumeración de armas. Si se produce un asesinato y si se usa un modelo va a caer dentro del tipo penal correspondiente al homicidio. De la misma forma, el Código Nacional Electoral, en el caso de Argentina, tiene incluida la creación de datos fraudulentos para modificar el voto de una persona. Entonces, si se utiliza un modelo de inteligencia artificial generativo para mostrar un candidato diciendo algo que no dijo, eso ya estaría de alguna manera contemplado dentro del Código Electoral Argentino.

Por último, el derecho administrativo, que trata sobre qué es lo que puede hacer el Estado, es mucho más complicado. Es bastante más difícil saber qué es lo que puede hacer o no. Hay elementos para dar un gran debate sobre cuál es el límite y qué criterios se deben adoptar para permitirle o no al Estado usar este tipo de modelos, si es la forma menos restrictiva de los derechos de las personas.

Conclusión

A modo de despedida, el aprendizaje automático presenta desafíos significativos en la actualidad, pero ofrece diversas ventajas en diferentes áreas. No se puede hacer que sea realmente perfecto; ponerle esa expectativa es una barrera que podría destinarlo al fracaso. No se puede ser perfecto cuando se intenta simplificar la realidad, como decía George Box en la frase citada. De alguna forma hay que entenderlo, hay que saber cómo controlarlo, y es realmente un problema difícil.

Son importantes estos espacios de debate, de reflexión compartida y poder hacerlo en comunidad. Van a ser necesarios muchos más intercambios y debates para llegar a una conclusión, para clarificar las ideas que podamos tener al respecto.

NOTAS

¹ “Del rigor en la ciencia” es un cuento corto de solo un párrafo escrito por Jorge Luis Borges, acerca de la relación mapa-territorio, escrito en la forma de una falsificación literaria.

² Cathy O’Neil (2017). *Armas de destrucción matemática. Cómo el big data aumenta la desigualdad y amenaza la democracia*. Madrid. Capitán Swing.

³ El autor y Valeria Edelsztejn publicaron un paper que se dedica a la descomposición de la brecha salarial entre hombres y mujeres en tecnología en Argentina.

DESÓRDENES INFORMATIVOS Y NARRATIVAS HOSTILES

MARÍA JOSÉ LABRADOR (MÉXICO)

Doctora en Ciencias de la Información por la Universidad Complutense de Madrid. Licenciada en Ciencias y Técnicas de la Comunicación por la Universidad Anáhuac México Norte. Directora de Investigación en la Universidad del Desarrollo de Chile. Miembro del Grupo AI4AI¹ del proyecto IFAP² UNESCO. Conferencista de trayectoria internacional.

RESUMEN

Esta presentación se estructura en tres grandes ejes. El primero aborda el fenómeno de la desinformación, para lo cual es prioritario hacer énfasis sobre su concepto y significación, así como su clasificación y caracterización. En el segundo eje se analizan las vulnerabilidades y las perspectivas éticas frente a la inteligencia artificial en este mismo marco de la desinformación. Para concluir, se efectúa una breve reseña sobre las investigaciones y estudios que se están realizando en el Observatorio Desinformación Chile, del cual la autora es fundadora.

Palabras clave

inteligencia artificial, desinformación, clickbait, vulnerabilidades, desiertos informativos, troll, perspectivas éticas, bot

ABSTRACT

This presentation is structured around three main axes. The first addresses the phenomenon of disinformation, emphasizing its concept and significance, as well as its classification and characterization. The second axis analyzes vulnerabilities and ethical perspectives regarding artificial intelligence within the same framework of disinformation. To conclude, a brief overview is provided of the research and studies being conducted at the Disinformation Observatory in Chile, of which the author is a founder.

Keywords

artificial intelligence, disinformation, clickbait, vulnerabilities, information deserts, troll, ethical perspectives, bot

El fenómeno de la desinformación, clasificación y caracterización

En un contexto de sobreabundancia informativa derivada de la democratización y el amplio acceso a internet (Vizoso y Vázquez-Herrero, 2019) surge la desinformación como uno de los fenómenos globales de mayor trascendencia y todavía impredecible (Wardle y Derakhshan, 2017), expresado mediante desórdenes informativos que inciden en todo tipo de procesos sociales y democráticos.

Este cambio de paradigma ha traído un aumento muy grande en la producción de la circulación de información y, junto con ello, una desinformación sin precedentes cuyas características están ligadas a sesgos cognitivos acoplados y a la llamada economía de la atención.

La desinformación es una de las principales amenazas a las sociedades libres y a las democracias latinoamericanas y del mundo. Así lo han hecho ver instituciones globales como el Fondo Económico Mundial desde 2013, la UNESCO desde 2020 a la fecha y la Unión Europea con su plan de acción 2018-2024 que finaliza con las leyes de Servicios Digitales³ y la de Mercados Digitales⁴ (Unión Europea, 2018). Estas tienen por objeto crear un espacio digital más seguro en el que se protejan los derechos fundamentales de los usuarios y establecer condiciones de competencia equitativas para las empresas.

El plan de acción de lucha contra la desinformación, de la Unión Europea (2018), y a la vez la Comisión Europea, a partir del informe de resultados de un enfoque multidisciplinario a la desinformación, elaboran una distinción basada en desinformación (*disinformation*), información errónea (*misinformation*) e información maliciosa (*malinformation*) (Wardle & Derakhshan, 2017).

Los desórdenes informativos incluirían formatos, productos informativos y presentación de la información caracterizada como sátira y parodia, conexión falsa, contenido engañoso, contenido impostor, contenido fabricado, contexto falso y contenido manipulado.



En el marco de los procesos electorales estos desórdenes informativos se han convertido en una amplia preocupación tanto en la región como en Chile.

En el análisis del impacto de lo que algunos han llamado la “crisis de la información” se plantea una idea de una “crisis epistémica” de la democracia (Bimber y Gil de Zúñiga, 2020), ya que la desinformación en contextos electorales supone una gran amenaza al sistema democrático. En este entorno, la cuestión central es cómo los ciudadanos están filtrando, verificando, discerniendo y circulando la información y, en definitiva, quién define lo verdadero y prioriza lo importante en el espacio público (Dahlgren, 2018) y de qué forma se generan capacidades en la ciudadanía para detectar la desinformación (UNESCO, 2023).

Ante esta necesidad surge el chequeo de datos como un movimiento transnacional (Graves, 2016), conformado por profesionales de diversos sistemas de medios y culturas periodísticas, quienes bajo procedimientos específicos (Moreno Gil y Salgado de Dios, 2023) y conscientes de lo perjudicial de los desórdenes informativos (Mena, 2018), se dan la tarea de verificar datos. Como explican López y Rodríguez (2019), el actual movimiento de plataformas de verificación no se concreta en departamentos asignados a una redacción que corrigen los errores antes de publicar, sino que su objetivo es recopilar afir-

maciones sobre hechos, encontrar pruebas que los confirmen o cuestionen e informar sobre las conclusiones, desacreditando así contenidos viralizados en redes sociales, que circulan con rapidez y en cascadas informativas.

¿Cuál es la categorización que tiene la desinformación?

Existen diferentes categorías de la desinformación. Una de ellas es el *clickbait*, que es un cebo, una técnica de redacción que consiste en crear encabezados y descripciones sensacionalistas para llevar a la opinión pública a determinados tipos de enlace. En los años pasados, junto con ellos se han generado y se han abierto camino estas granjas de *clickbait*. Es decir, una cantidad enorme de *clickbait* con una funcionalidad que en Facebook era Instant Articles⁵, y que permite consumir mucho contenido periodístico.

También Ad Breaks, a partir de sus pausas publicitarias, que permiten además monetizar el video, ha fomentado una cantidad enorme de estas granjas. En un momento dado, Allen y otros investigadores del MIT⁶ registraron que hasta un 60 % de estos dominios o estas tácticas se estaban utilizando en Facebook. De hecho, frente al caso Cambridge Analítica⁷, Facebook tiene que comparecer ante el Congreso de Estados Unidos y, a partir de ello, cambia su nombre a Meta. Cambia también entonces la lógica de su dueño, Mark Zuckerberg, quien en el lanzamiento de la red declara que la privacidad había terminado y, a partir de estos hechos, Meta explica que la privacidad es un tema fundamental para la empresa, lo que admite un análisis en materia de investigación.

Otra categorización son los bots, que son cuentas automatizadas de redes sociales administradas por un algoritmo y no por una persona real. Esto se ve habitualmente sobre todo en el anonimato, que es uno de los problemas más graves de la desinformación, por el tema de la opacidad. Es decir, no se puede saber con claridad el origen de la desinformación.

Los trolls son usuarios reales que amplifican algunos mensajes con bots, y han irrumpido con fuerza los llamados *deepfake*, videos que en la actualidad tienen un análisis y una investigación muy profundos porque son un tipo particular

de medios sintéticos. Es decir, una persona, una imagen, un video existente es reemplazado por la imagen de otra persona. Se utilizan técnicas muy poderosas de aprendizaje automático y de inteligencia artificial y se manipula y se genera el contenido visual y el contenido de audio. Esto tiene un alto potencial de engaño y está siendo material de análisis para los investigadores, precisamente por los graves riesgos a los que conlleva.

Finalmente, el Brain Hacking o las teorías conspirativas, que tienen un efecto exponencial en los tipos de desinformación y en las categorías y los discursos de odio o las narrativas hostiles, las encontramos vinculadas, generalmente, a grupos con problemáticas específicas, por ejemplo, migrantes o grupos de enfermos o con algún tipo de discapacidad.

Vulnerabilidades y perspectivas éticas frente a la IA en el marco de la desinformación y las narrativas hostiles

El segundo eje de análisis de esta presentación tiene que ver con las vulnerabilidades y las perspectivas éticas frente a la inteligencia artificial en el marco de la desinformación.

¿CUÁLES SON ESTAS VULNERABILIDADES?

Por ejemplo, en Chile existen grupos de mujeres, grupos de adultos mayores y personas con necesidades, que son vulnerables, a los que se presta especial atención.

Un aspecto que incide en tales vulnerabilidades tiene que ver con los llamados desiertos informativos, con las brechas de género y los paradigmas de liderazgo y, junto con ello, se exponen algunas perspectivas éticas al respecto. Un reciente estudio que hizo la Asociación de Mujeres Geógrafas en Chile da cuenta de que cuatro de cada diez mujeres de Chile viven en zonas de nivel alto o muy alto de vulnerabilidad.

En regiones como Santiago, Antofagasta, Valparaíso y Concepción hay un alto nivel de vulnerabilidad de las mujeres. En 2021 la Subsecretaría de Telecomunicaciones de Chile (SUBTEL) refería que el 77 % de las mujeres presentan mayor dificultad y vulnerabilidad en el momento de acceder a servicios digitales.

Es decir, existe un problema de raíz que tiene que ver primero con la brecha digital y, por tanto, aumenta esta vulnerabilidad respecto de los avances relacionados con inteligencia artificial o la sociedad tecnológica.

Los desiertos informativos representan un problema que se ha expuesto a nivel regional y que por tanto plantean múltiples desafíos. El primero sería que millones de latinoamericanos viven en zonas rurales, en zonas urbanas pobres, con lo cual habitan auténticos desiertos de noticias. Es decir, no existen ecosistemas mediáticos robustos y, por tanto, enfrentamos un problema anterior: no tener información, y que la información que llega está distorsionada o manipulada. Además, hay problemas posteriores que tienen que ver con el pluralismo informativo. Son comunidades que carecen de estos medios locales y que proporcionan, además, contenidos creíbles y completos. Es decir, muchas veces tienen algo de las noticias, pero con precarización.

Existe una precarización del periodismo y de otras fuentes locales, y estas comunidades no tienen acceso a información esencial, ni siquiera sobre eventos locales, y mucho menos de gobierno o de elecciones y otros temas importantes que, particularmente en este 2024, son de mucha relevancia.

La Inter Press Service⁸ y otras organizaciones, así como el *Time Journalist Review*, han dado cuenta del caso interesantísimo de Brasil, que tiene prácticamente un 48,7 % de municipios o de comunidades que carecen de medios de comunicación. Argentina bordea también el 47 %. El caso Colombia, todos ellos dignos de estudio y a pesar de que surgen cada año centenares de medios en línea, muchos no son independientes o están sesgados porque dependen, muchas veces, de autoridades locales, de alcaldías u otro tipo de organizaciones. En Venezuela siete millones de personas viven en municipios donde no hay medios de comunicación, y la cifra sube prácticamente al doble si se consideran como semidesiertos los municipios que tienen uno o dos medios de comunicación. Por lo tanto, los desafíos están en la sostenibilidad de estos medios de comunicación y en la viabilidad como empresas digitales.

Este año se acaban de ejecutar iniciativas del Consejo de la Unión Europea como “Medios locales para la democracia”⁹. Tam-

bién un estudio da cuenta de la importancia de la movilización de los agentes públicos y privados para contribuir a reactivar el panorama de los medios de comunicación locales, y la importancia de reforzar la independencia, la sostenibilidad y la capacidad de innovación de los medios de comunicación.

A la vez, desde las perspectivas de la educación es fundamental concretar planes de acción para la alfabetización digital. En esta línea se está trabajando con diversos investigadores en el ámbito nacional e internacional.

Finalmente, ¿Cuáles son las perspectivas éticas? ¿Cuáles son los principios éticos o la ética que habremos de poner en marcha frente a la inteligencia artificial?

Uno de los grandes desafíos éticos que se presentan al respecto se propone a partir de un análisis sobre las características fundamentales que se extraen de la teoría de la ética del cuidado, de Carol Gilligan: cuidado, responsabilidad, comunidad, atención e interdependencia. La fuerza que las mueve es la cooperación mutua que se lleva a cabo por medio de aptitudes como la empatía y la capacidad para sostener relaciones interpersonales. Esto está planteado así, a diferencia de la ética de justicia, cuyos valores iban más encaminados hacia la individualidad, la independencia, la objetividad, la libertad, la igualdad y la justicia, y son impulsadas por la razón y el cumplimiento de las normas estipuladas. Las mujeres, según Gilligan, se centran menos en la justicia abstracta y la equidad, y piensan más en las responsabilidades que tienen hacia personas específicas de su entorno. En su modelo, Gilligan postula que las preferencias de las mujeres hacia el cuidado de los demás como responsabilidad moral se fundamenta en el hecho de ser responsables, en mayor medida, del cuidado de los recién nacidos. Las mujeres piensan más en términos relacionados con el cuidado, en un plano contextual, y por ello se requiere plantear los principales desafíos para la política social y de desarrollo de la sociedad tecnológica en esa dirección.

Por lo tanto, valores como equidad, justicia, intimidad y privacidad, desde sus abordajes, así como la privacidad del norteamericano y la intimidad en el enfoque europeo, deben de ser base inmutable para que las decisiones algorítmicas tengan

que pasar necesariamente por una supervisión humana de los sistemas de inteligencia artificial.

Reviste especial importancia aportar un enfoque similar en la región que se relaciona con el desarrollo ético de la inteligencia artificial. Existe la necesidad de avanzar y profundizar en garantizar este desarrollo ético. ¿Qué significa eso? ¿Qué significa la transparencia algorítmica? ¿Qué significa que los chatbots no produzcan contenidos ilegales? ¿Qué significa velar por derechos humanos tan significativos como la libertad de expresión y el derecho a la información?

Observatorio Desinformación Chile. Investigaciones y estudios

El Observatorio Desinformación Chile¹⁰ es una iniciativa académica que nace a partir del proyecto “Garantías institucionales y regulatorias, autoridades electorales y de supervisión digital ante interferencias y narrativas hostiles”, de publicidad segmentada y polarización, que llevan a cabo 69 investigadores de todo el mundo, con un fondo adjudicado por el Ministerio de Ciencia e Innovación de España a la Facultad de Ciencias de Información de la Universidad Complutense de Madrid.

El Observatorio Desinformación Chile está anclado al Observatorio de la Desinformación Complutense. Las investigaciones que se abordan se centran en la regulación de la desinformación y en estudios sobre la desinformación en redes sociales. Un aspecto central de Estudios sobre la Regulación de la Desinformación es la investigación de todas las regulaciones que aborden tanto los aspectos procedimentales e institucionales (quién debe conocer y resolver los conflictos) como la regulación de las plataformas tecnológicas. Se ha puesto especial atención en el estudio de las estructuras de red y el flujo informacional de hashtags desinformativos. Asimismo, se ha realizado seguimiento y monitoreo de cuentas que realizan desinformación durante los procesos electorales.

Chile focaliza en el proceso electoral del plebiscito constitucional de 2023, que fue interesantísimo para el análisis en la temática de las desinformaciones. Algunos resultados de interés son los siguientes: un 68,8 % de las desinformaciones provienen de la red social X, un 12,5 % de la TV y un 6,3 % de otras plata-

formas como TikTok, Instagram y Facebook. El estudio concluyó también que el autor causante de la desinformación, en un 12,5 % de los casos, es atribuible a partidos políticos, mientras que en un 87,5 % a un actor de origen desconocido. Respecto de la temática de las desinformaciones, un 58,8 % de ellas están vinculadas con deficiencias en el sistema electoral, seguidas de un 11,8 % sobre seguridad y otro 11,8 % sobre legislación y funciones del Congreso.

Cabe destacar que el análisis contempló los desórdenes informativos encontrados en los meses de noviembre y diciembre (antes de la fecha del plebiscito de 2023), para los cuales se aplicaron criterios de selección relacionados con el rango de interés público, el potencial de daño que generan, el nivel de viralización y las solicitudes de la audiencia. En total, se investigaron 17 desórdenes informativos.

Adicionalmente, el Observatorio elaboró un análisis comparado de la cobertura mediática de diversos procesos políticos y/o electorales tanto en medios tradicionales como independientes y extrasistémicos o los llamados pseudomedios. El objetivo es caracterizar con detalle los enfoques y encuadres noticiosos de cada uno de estos soportes, para determinar el grado de injerencia que tienen en la validación de información por parte de las audiencias y su posterior asimilación y difusión. Esto implica indagar tanto en el consumo informativo en diversos grupos etarios como en las rutinas y los métodos de los periodistas e informadores empíricos para concebir información, la responsabilidad profesional que emana de dicho proceso y los mecanismos de media accountability que utilizan los medios de comunicación en Chile.

NOTAS

¹ Artificial Intelligence for Information Accessibility.

² Programa Información para Todos (PIPT) o Information for All Programme (IFAP) es un programa intergubernamental único de la UNESCO que apoya a los Estados miembros en el fomento de sociedades equitativas mediante la promoción del acceso universal a la información y al conocimiento para el desarrollo sostenible. <https://www.unesco.org/en/ifap>

³ Esta ley entró en vigor el 16 de noviembre de 2022. Su objetivo es crear un entorno digital más seguro y abierto y proteger los derechos fundamentales de los usuarios en línea. La DSA establece normas para com-

batir la oferta de contenidos ilícitos en línea, empoderar a los usuarios y a la sociedad civil, evaluar y paliar los riesgos y supervisar y controlar el cumplimiento de las plataformas en línea.

⁴ Esta ley tiene como objetivo garantizar la equidad y la competencia en los mercados digitales. La DMA impone normas a las plataformas que actúan como “guardianes” en el sector digital. Su objetivo es proteger la privacidad de los usuarios y fomentar la competencia leal, la transparencia y la innovación en el mercado digital.

⁵ Facebook Instant Articles ofrece a los editores una manera rápida y atractiva de presentar sus contenidos a una gran audiencia.

⁶ Massachusetts Institute of Technology.

⁷ Facebook se halla sumida en una tormenta política global por la fuga de datos de 50 millones de usuarios estadounidenses, que habría utilizado la consultora Cambridge Analytics para afinar con perfiles psicológicos las estrategias de atracción de votos de la campaña de Donald Trump en 2016.

⁸ Inter Press Service (IPS) es una agencia mundial de noticias comprometida con el ejercicio del periodismo independiente y especializada en reportajes analíticos sobre procesos y acontecimientos económicos, políticos, sociales, artísticos y culturales. Su cobertura se caracteriza por incluir a las organizaciones que conforman la sociedad civil como fuente informativa, y los procesos de la era de la globalización como temas de especial interés.

⁹ El Consejo de la Unión Europea señala la importancia de la libertad de los medios de comunicación en la UE en un momento en el que la libertad de los medios de comunicación se está cuestionando en muchos países del mundo. Se han presentado nuevas propuestas con el fin de seguir reforzando la libertad de los medios de comunicación en la UE. <https://www.consilium.europa.eu/es/policies/media-freedom-eu/>

¹⁰ <https://observatoriodesinformacion.cl/>

REFERENCIAS

- Dahlgren, P. (2018). La participación en línea en la esfera pública. Las ambigüedades del afecto. *In Mediaciones de la Comunicación*, 13. <https://doi.org/10.18861/ic.2018.13.1.2824>
- Graves, L. (2016). *Deciding what's True: The Rise of Political Fact-checking in American Journalism*. Columbia University Press.
- López, M. y Rodríguez, J. (2019). El papel de las plataformas de verificación en la lucha contra la desinformación: un análisis de su impacto en redes sociales. *Revista internacional de Comunicación Digital* Nro. 23, 87-110.

- Mena, P. (2018). Los desórdenes informativos y su impacto en el proceso democrático: Herramientas de lucha contra la desinformación. *Revista de Estudios Políticos y Sociales*, 15(2), 45-65.
- Moreno Gil, V. y Salgado de Dios, F. (2023). El cumplimiento del código de principios de la International Fact-Checking Network en las plataformas de verificación españolas. Un análisis cualitativo. *Revista de Comunicación*, 293-307.
- UNESCO (2023). *The challenge of disinformation in electoral processes*. Obtenido de <https://www.unesco.org/es/articles/el-desafio-de-la-desinformacion-en-procesos-electorales>.
- Vizoso, A. y Vázquez-Herrero, J. (2019). Plataformas de fact-checking en español. Características, organización y método. *Communication & Society*.
- Wardle, C., y Derakhshan, H. (2017). *Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking*. Council of Europe. <https://rm.coe.int/information-disorder-report-november-2017/1680764666>

TRANSFORMACIÓN DIGITAL Y DD.HH.: UN DIÁLOGO A CONSTRUIR

GABRIEL LEVY (COLOMBIA)

Conferencista y moderador de trayectoria internacional, docente y asesor en comunicación digital, transmedia, nuevos medios, transformación digital y TIC. Investigador de las universidades de Antioquia, Externado y Javeriana. Coordinador académico de Andina-link. Investigador asociado del Observatorio Latinoamericano de Regulación, Medios y Convergencia.

RESUMEN

Se vive un momento desafiante en el que la tecnología avanza a una velocidad asombrosa, transformando la forma en que las personas viven y se relacionan. Este artículo ayuda a comprender cómo la inteligencia artificial (IA) y otras innovaciones están moldeando la vida cotidiana y la sociedad en general.

A través de la evolución de tecnologías como el telégrafo, el teléfono, la radio y la web, se explora cómo estas herramientas han alcanzado millones de usuarios en tiempos cada vez más cortos. Sin embargo, este rápido progreso presenta importantes desafíos: las brechas en el acceso a la tecnología y la información se están ampliando, y van dejando a muchas personas atrás, lo que no solo genera desigualdades económicas, sino que también plantea preguntas cruciales sobre los derechos humanos en un mundo digital.

Palabras clave

teoría de la aceleración, desigualdades, oportunidades, brecha digital, YouTube, Chatgpt, TikTok, evolución

ABSTRACT

We are living in a challenging moment where technology is advancing at an astonishing speed, transforming the way people live and interact. This article helps to understand how artificial intelligence (AI) and other innovations are shaping daily life and society as a whole.

Through the evolution of technologies such as the telegraph, telephone, radio, and the web, it explores how these tools have reached millions of users in increasingly shorter periods. However, this rapid progress presents significant challenges: the gaps in access to technology and information are widening, leaving many people behind. This not only generates economic inequalities but also raises crucial questions about human rights in a digital world.

Keywords

theory of acceleration, inequalities, opportunities, digital divide, YouTube, ChatGPT, TikTok, evolution

Aceleración tecnológica y sus indicadores

Con el desarrollo de la tecnología, nos enfrentamos a un fenómeno acumulativo de aceleración. La Teoría de la Aceleración es una teoría que varios autores han analizado durante los últimos años. Básicamente, plantea que el impacto de la tecnología en nuestras vidas ocurre de manera cada vez más rápida y acelerada.

Existen dos indicadores comúnmente utilizados para determinar cuándo una tecnología se convierte en masiva: uno es alcanzar los 100 millones de usuarios, y el otro es llegar a 50 millones de usuarios. Aunque no existe un estándar global al respecto, el indicador de 50 millones es el más citado por los autores.

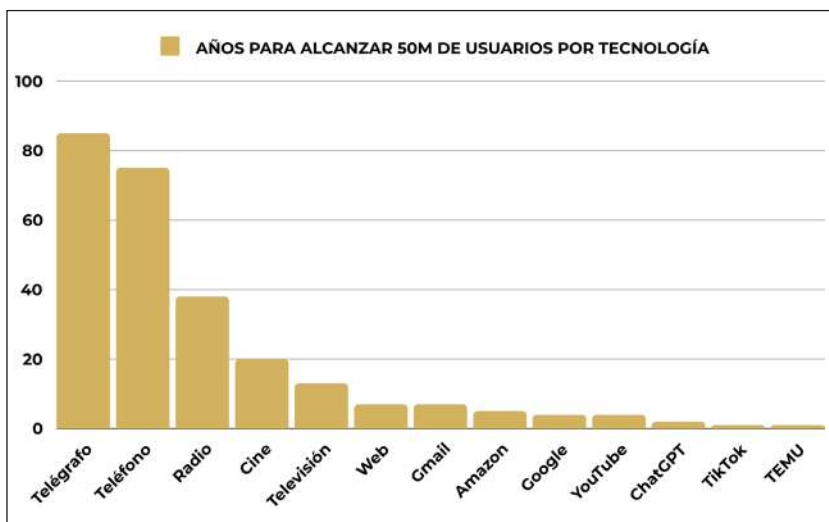
Por ejemplo, al telégrafo, la primera tecnología posterior al desarrollo de la electricidad y el inicio de la electrónica, le tomó 85 años alcanzar los 50 millones de usuarios. Esto equivale a unas ocho generaciones y media, tiempo suficiente para que los académicos pudieran investigar su funcionamiento, los reguladores pudieran establecer normas para su uso, y las personas pudieran apropiarse de la tecnología y entenderla plenamente.

El teléfono requirió una década menos, 75 años, para llegar a los 50 millones de usuarios.

Aunque los medios impresos, como la prensa y los libros, también habían alcanzado una gran audiencia, no lo hacían en tiempo real.

Con la radio, la primera tecnología de medios masivos en tiempo real, este proceso se aceleró. A la radio le tomó solo 38 años alcanzar los 50 millones de usuarios, la mitad del tiempo que necesitó el teléfono. Aun así, este periodo fue suficiente para que los académicos estudiaran la tecnología, los reguladores la normaran y los usuarios la adoptaran.

El cine, por su parte, fue la primera tecnología en alcanzar esta cifra en dos décadas. Aunque surgió antes que la radio, tardó más en ser adoptado, pero logró llegar a los 50 millones de usuarios en veinte años.



Posteriormente, la televisión alcanzó esta cifra en poco más de una década, 13 años. La web fue la primera tecnología en llegar a los 50 millones de usuarios en menos de una década: logró esta cifra en solo siete años. Este es el mismo tiempo que le tomó a Gmail alcanzar los 50 millones de usuarios.

Brechas digitales y desigualdades

A medida que avanzamos hacia tecnologías que se popularizan en menos de una década, surgen dificultades en los procesos de apropiación social.

Por un lado, la academia no tiene tiempo suficiente para realizar estudios exhaustivos ni para generar conocimiento *ex post* sobre la tecnología. Por otro, los reguladores no logran mantener el ritmo, y quedan por debajo de la línea ideal de una década para establecer regulaciones. Amazon lo logró en cinco años, y Google, en solo cuatro años. YouTube alcanzó los 50 millones de usuarios en cuatro años, ChatGPT en dos años y TikTok en apenas un año.

Actualmente, todas las tecnologías emergentes, especialmente aquellas derivadas de la inteligencia artificial, están alcanzando esta cifra en menos de un año. Esto presenta numerosas aristas y múltiples implicaciones de análisis, cada una de gran relevancia y digna de interpretación.

Por una parte están las brechas que se acumulan de manera constante. ¿Qué significa esto? Aún no hemos superado las brechas del analfabetismo básico: todavía existen personas en el mundo que no saben leer y para quienes el acceso al conocimiento representa una enorme barrera.

A esto se suma la primera brecha digital, es decir, el acceso a una computadora y a internet. Esta brecha se acumula con las demás.

Luego surge la brecha de acceso a dispositivos móviles y, más recientemente, la brecha entre quienes tienen o no acceso a un smartphone y a internet en el mismo.

Todas estas brechas se acumulan a medida que las nuevas tecnologías avanzan más allá de las líneas de apropiación.

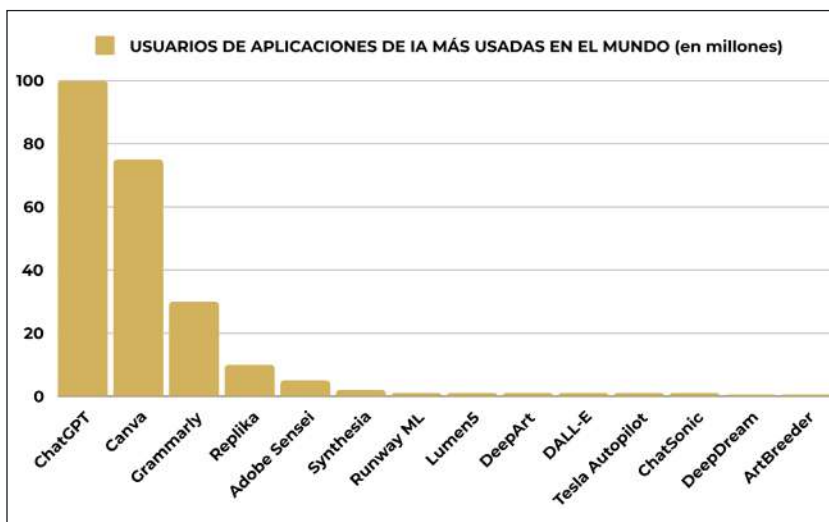
Esto, inevitablemente, incrementa las desigualdades. A medida que estas brechas se profundizan, enfrentamos un serio problema de derechos humanos, que es el tema que nos convoca.

La transformación digital no ocurre de forma simétrica en todo el planeta; no se da de la misma manera entre todas las poblaciones, países o culturas, lo cual tiene un profundo impacto en términos de oportunidades.

Por supuesto, una persona que ya domina ChatGPT tiene muchas más oportunidades laborales que alguien que ni siquiera tiene una computadora, o una persona que no sabe leer y escribir, o quien ni siquiera conoce qué es ChatGPT.

Esto se aplica a diferentes niveles y demuestra que las brechas son acumulativas. Por ejemplo, alguien podría no tener electricidad en su hogar, otra persona podría no tener computadora ni teléfono, o puede que simplemente no tenga acceso a internet, o que desconozca por completo la existencia de ChatGPT.

En cualquiera de estos niveles ya existe una barrera significativa.



Estas barreras están relacionadas con las posibilidades de acceso de las personas que cuentan con mayores oportunidades educativas, mayor capacidad adquisitiva o un mayor nivel de escolaridad.

De esta forma, mientras la humanidad no resuelva los problemas previos, las brechas seguirán acumulándose, lo que se traduce en una mayor distancia entre ricos y pobres, una mayor desigualdad en el acceso a la sociedad de la información y una disparidad creciente en la capacidad de usar estas herramientas, entre otros problemas.

Otro factor sumamente problemático es el uso de las aplicaciones derivadas de este fenómeno. Es decir, incluso en un escenario ideal en el que todas las personas tuvieran acceso a internet y a un dispositivo móvil, el problema radica en que, en periodos de tiempo cada vez más cortos, las aplicaciones están alcanzando millones de usuarios rápidamente. ChatGPT, por ejemplo, logró niveles extraordinarios de uso en solo dos meses, un ritmo impresionante en comparación con Google Translate en sus inicios.

Este fenómeno se explica mediante la Ley de los Rendimientos Acelerados, propuesta por Raymond Kurzweil, director de Ingeniería en Google.

Basándose en la Teoría de Moore¹, Kurzweil² sugiere que, en los últimos cien años, la humanidad ha experimentado más desarrollo y progreso que en los 20.000 años previos.

Esto se debe, en gran medida, a la aceleración exponencial de la velocidad y potencia de los chips, lo que incrementa drásticamente la capacidad tanto de las máquinas como de las personas para utilizarlas.

De la revolución agrícola al ChatGpt

Nos encontramos ante una gran paradoja. Aunque las tecnologías actuales alcanzan a una gran cantidad de personas en tiempos récord, al mismo tiempo están ocurriendo dos fenómenos: no logramos apropiarnos o conocer plenamente estas tecnologías, y dejamos a muchas personas fuera de este avance.

Esto se traduce en múltiples brechas, incluyendo brechas de acceso al conocimiento, diferencias en la capacidad adquisitiva y, en última instancia, mayores desigualdades entre riqueza y pobreza.

Para ilustrar esta aceleración, consideremos la evolución de las computadoras. En solo cuatro décadas, pasamos de las primeras computadoras a las supercomputadoras modernas, cuya capacidad de procesamiento ha aumentado de manera exponencial.

En 2011 y 2012 celebramos que una computadora había alcanzado la capacidad de procesamiento equivalente al cerebro de un ratón, algo que se confirmó en 2015. Menos de una década después, en 2023, logramos que ChatGPT superará la Prueba de Turing.

Esta prueba consiste en que un ser humano, al interactuar, no sea capaz de discernir si está conversando con otra persona o con una máquina.

Hasta 2023, no habíamos superado esta prueba establecida por el matemático Alan Turing. Sin embargo, ahora, gracias a los avances en hardware, software y la inteligencia artificial generativa, hemos logrado que los humanos no puedan distinguir si están chateando con una persona o con una máquina.

En menos de ocho años, pasamos de una computadora con capacidad de procesamiento equivalente a un ratón a una capaz de procesar información al nivel de un cerebro humano.

Si continuamos con esta línea de aceleración, que es exponencial y acumulativa, se prevé que para el año 2045 una sola computadora tendrá la capacidad de procesamiento equivalente a todos los cerebros humanos en el planeta, es decir, unos diez mil millones de personas.

Esto muestra cómo se han ido reduciendo las ventanas de tiempo: entre la Revolución Agrícola y la Revolución Industrial transcurrieron 8.000 años; entre la Revolución Industrial y la invención de la bombilla, solo 120 años; entre la bombilla y el primer viaje a la Luna, 90 años; entre el viaje a la Luna y la creación de la web, 22 años; y entre la web y la secuenciación del genoma humano, apenas 9 años.

El desafío de seguir el ritmo

Thomas Friedman aborda esta aceleración en su libro *Gracias por llegar tarde*³. En este texto, el periodista norteamericano analiza el avance y la transformación digital.

El título proviene de la práctica de Friedman de citar a expertos en una cafetería en Nueva York, donde muchos llegaban tarde debido al tráfico. Friedman solía responder a sus disculpas diciendo: “No, gracias a ti por llegar tarde”, ya que durante esos minutos podía disfrutar de su café, revisar sus notas y observar a su alrededor, desconectándose de las pantallas y experimentando una pausa en su vida. Él sugiere que esta capacidad de hacer una pausa en un mundo que se mueve a toda velocidad es quizás uno de nuestros activos más valiosos.

Al analizar la Ley de los Rendimientos Acelerados⁴ y contrastar con el conocimiento de neurocientíficos, Friedman descubrió que hasta el año 2015 la capacidad humana para procesar información mantenía el ritmo con el avance tecnológico. Sin embargo, desde entonces, la evolución de la tecnología, impulsada principalmente por la inteligencia artificial, ha superado con creces nuestra capacidad de adaptación y comprensión.

Aunque se modificaran los modelos educativos y se ampliara la manera en que entrenamos nuestro cerebro, no se alcanzaría la línea de evolución tecnológica actual.

Este desfase implica que la tecnología avanza mucho más rápido de lo que nuestro cerebro puede procesar. Cada vez más personas quedan excluidas de la incorporación a estos avances tecnológicos, y, aunque algunos sean la élite en investigación, corremos el riesgo de quedarnos rezagados en un laboratorio, buscando alcanzar una tecnología que, a este paso, continuará alejándose rápidamente.

La inteligencia artificial nos sobrepasará, tanto en cantidad como en profundidad, como ya se observa en el acelerado crecimiento de las plataformas. Ni siquiera alcanzamos a comprender la cantidad de tecnologías emergentes en los últimos meses; ya no hablamos de años. Nos superan tanto en posibilidades como en potencial.

Reflexiones sobre inteligencia artificial para la humanidad

Está sucediendo la génesis de la IA y poco a poco se va comprendiendo lo que realmente implica. Se puede asegurar que la IA será el mayor generador de brechas que la humanidad haya presenciado.

La IA se considera la nueva electricidad, y representa un desarrollo de dimensiones e impacto comparables a la electricidad. Si en los últimos cien años la electricidad revolucionó el conocimiento humano más que en los 20.000 años previos, la IA generará un fenómeno igual, pero de manera exponencialmente mayor.

En la actualidad, la cantidad de desarrollos tecnológicos en períodos tan cortos resulta abrumadora, y existe temor de quedar excluidos; sin embargo, aún no se ha visto nada de lo que sucederá en los próximos años. Este es solo el inicio de algo mucho mayor.

Por supuesto, este fenómeno genera respuestas como el miedo, la resistencia e incluso el rechazo, tal como ocurrió con la electricidad. La electricidad fue la tecnología que más resistencia produjo.

Había dos tipos de resistencia: la ideológica-conceptual y la objetiva. Por ejemplo, algunas comunidades amish en Estados Unidos aún se niegan a usar la electricidad, ya que la consideran un elemento demoníaco por ser invisible pero existente, lo que responde a un tema religioso y cultural. También había razones prácticas para rechazarla, pues en sus inicios la electricidad provocaba numerosos incendios al no estar protegida adecuadamente. Sin embargo, con el tiempo se resolvieron estos problemas, y hoy en día los incendios causados por la electricidad son más controlados.

La IA genera una reacción similar. Paul Daugherty⁵ dice que, aunque no tenemos evidencia de que la IA vaya a reemplazar a los seres humanos, sí sabemos con certeza que aquellos que usen la IA reemplazarán a quienes no la usen. La diferencia estará en quién incorpore la IA en sus actividades diarias y en sus proyectos. Este es el núcleo de las nuevas brechas: la brecha de conocimiento, la brecha en el acceso a la información y la creciente desigualdad entre ricos y pobres.

Por su parte, Oren Etzioni⁶ afirma que la IA es una herramienta, y el modo de usarla depende de cada ser humano. Habrá quienes la empleen para desarrollos que beneficien a la humanidad, y otros que la usen con fines destructivos. Para ilustrar esto, se usará el ejemplo de una silla. Las sillas fueron creadas para sentarse; ese es su propósito. Sin embargo, una persona malintencionada podría usar una silla para agredir a alguien, convirtiéndola en un arma peligrosa. Pero esto no es culpa de la silla ni de quien la creó; es el uso que se le da. Lo mismo ocurre con la IA.

Actualmente, en el Reino Unido se ha desarrollado un prototipo de antibiótico capaz de combatir a las bacterias más resistentes, gracias a datos procesados por un sistema de IA. Según el equipo científico responsable, sin la IA este desarrollo habría tomado más de cien años. En este caso, la herramienta ha contribuido a enfrentar uno de nuestros mayores desafíos: las bacterias resistentes a los antibióticos. Es un uso adecuado y seguramente uno de los objetivos para los que se creó la IA. Sin embargo, es probable que existan personas con intenciones destructivas que intenten desarrollar el veneno más letal utilizando inteligencia artificial. Esto no sería culpa de Open AI, ni de la IA, ni de los avances tecnológicos; sería un problema inherente a la humanidad, a su capacidad de construir o destruir.

Por ello, la regulación de la IA es fundamental. No obstante la velocidad a la que avanza, esta tecnología plantea desafíos, ya que en cuestión de meses los reguladores no alcanzan a adaptarse, por eso la nueva Directiva Europea de Regulación para IA⁷ es prematura. Se observa cómo afecta la implementación del nuevo sistema operativo de Apple en Europa, debido a la falta de tiempo para comprender y adaptar la normativa. Como dijo Stephen Hawking⁸ antes de conocer ChatGPT, “la inteligencia artificial puede ser lo mejor o lo peor que le haya ocurrido a la humanidad”. Dependerá de nosotros y de cómo la usemos; la IA no tiene la culpa.

Naveen Rao⁹ señala: “La inteligencia artificial no es una industria; es una ola que transformará a todas las demás”. No podemos considerar la IA como opcional. Es una tecnología transformadora, y negarse a interactuar con ella sería como negar el impacto de la electricidad. La IA cambiará todas las industrias y afectará todos los aspectos de la vida humana.

Resulta imposible considerar el área de la inteligencia artificial (IA) como si se dijera: “A mí no me interesa el fútbol” o “la botánica no es lo mío” o “no tengo relación con la manufactura”. No se trata de un tema de interés o elección; se trata de un desarrollo que será transversal a cualquier actividad que se realice. No importa cuál sea la profesión de cada persona, la IA tendrá un impacto en su vida. No hay excepciones posibles.

La IA impactará en todo lo que las personas hagan, tanto en el ámbito profesional como en la vida cotidiana. Cuanto más se resista su avance, más se quedarán al margen, y eso es precisamente lo que se debe evitar: que millones de personas en el mundo queden excluidas.

El problema más grave es que la IA será el mayor generador de brechas que la humanidad haya visto. Ni la electricidad, ni la computación, ni los smartphones, ni el acceso a internet pueden producir un nivel de desigualdad como el que desencadenará la IA.

Por lo tanto, se presenta como la mayor amenaza de división entre ricos y pobres, entre quienes tienen acceso a sistemas educativos y a la información y aquellos que no cuentan con esa posibilidad, especialmente en lo que respecta a los pro-

cesos de transformación digital. Si se permite que esto ocurra, podría convertirse en la mayor amenaza a los derechos humanos que se haya presenciado.

Sin embargo, esta situación podría revertirse, tal vez usando la IA como una herramienta para resolver el problema, o simplemente enfrentando con seriedad lo que no hemos logrado solucionar hasta ahora: el acceso a internet.

Afortunadamente, hay personas en todo el mundo que están trabajando para mejorar el acceso a internet, y se espera que encuentren la sabiduría necesaria para asegurar que la IA tenga un impacto positivo en todos. Este es un gran desafío. Para quienes ya están en contacto con la IA, aquellos que son curiosos y cuentan con oportunidades como el acceso a internet, la educación, los libros y los dispositivos, el camino es difícil. Es aún más preocupante imaginar lo que enfrentarán aquellos que carecen de estos recursos.

NOTAS

¹ En 1965, el cofundador de Intel, Gordon Moore, predijo que el número de transistores en un chip se duplicaría aproximadamente cada dos años, con un aumento mínimo en el costo. G. E. Moore (1965). "Cramming more components on integrated circuits", *Electronics*, vol. 38, n° 8, pp. 114 ff., en *IEEE Solid-State Circuits Society Newsletter*, vol. 11, n° 3, pp. 33-35, sept. 2006, doi: 10.1109/N-SSC.2006.4785860.

² Raymond Kurzweil (Queens, Nueva York, 12 de febrero de 1948) es un inventor estadounidense, además de músico, empresario, escritor y científico especializado en ciencias de la computación e inteligencia artificial. Desde 2012 es director de Ingeniería en Google.

³ *Gracias por llegar tarde. Cómo la tecnología, la globalización y el cambio climático van a transformar el mundo los próximos años* https://www.google.com.ar/books/edition/Gracias_por_llegar_tarde/K9BHDwAAQBAJ?hl=es&gbpv=0

⁴ La Ley de Rendimiento Acelerado de Ray Kurzweil plantea que el progreso tecnológico no avanza de manera lineal, sino de forma exponencial. Esto significa que cada nueva tecnología no solo se suma al desarrollo existente, sino que también impulsa la creación de tecnologías aún más avanzadas, generando un ciclo continuo de innovación. Aunque a menudo se piensa en el avance tecnológico como algo gradual, Kurzweil señala que la velocidad de cambio se duplica en intervalos regulares. Así, lo que podríamos considerar un siglo de desarrollo

tecnológico podría equivaler a miles de años de progreso, incluyendo la creación de máquinas que superen la inteligencia humana y que incluso puedan mostrar emociones.

⁵ Director ejecutivo de tecnología del grupo Accenture y su director de Tecnología (CTO).

⁶ Oren Etzioni es profesor emérito de Informática y director ejecutivo fundador del Instituto Allen de Inteligencia Artificial. Etzioni es el fundador y director ejecutivo de TrueMedia.org, una organización sin fines de lucro dedicada a luchar contra los *deepfakes* políticos, que se lanzó en abril de 2024.

⁷ https://www.europarl.europa.eu/news/es/press-room/20240308IP_R19015/la-eurocamara-aprueba-una-ley-historica-para-regular-la-inteligencia-artificial

⁸ Stephen Hawking fue un físico teórico, astrofísico, cosmólogo y divulgador científico británico. Sus trabajos más importantes consistieron en aportar, junto con Roger Penrose, teoremas respecto de las singularidades espaciotemporales en el marco de la relatividad general y la predicción teórica de que los agujeros negros emitirían radiación, lo que se conoce hoy en día como radiación de Hawking (o a veces radiación Bekenstein-Hawking).

⁹ Ex vicepresidente y actual gerente general del Grupo de Productos de Inteligencia Artificial de Intel Corporation.

AGENCIA HUMANA A PRUEBA DE FUTURO. CÓMO PROTEGER LOS DERECHOS HUMANOS FRENTE A LAS MÁQUINAS

CORDEL GREEN (JAMAICA)

Abogado. Ex fiscal general adjunto de Jamaica. Director ejecutivo de la Comisión de Radiodifusión. Miembro de la Comisión Nacional de Jamaica para la UNESCO y representante en el Consejo Intergubernamental del Programa Información para Todos (IFAP) de la UNESCO. Encabezó la Conferencia Mundial Anual sobre Inteligencia Artificial para la Accesibilidad a la Información AI4IA de UNESCO.

RESUMEN

Este artículo analiza de manera profunda y filosófica la nueva realidad a la que los seres humanos se enfrentan y que pone en jaque los derechos humanos fundamentales como la libertad de expresión, la privacidad, la libertad de pensamiento. Reflexiona sobre cómo las nuevas tecnologías invaden no solo los mercados sino las mentes, introduciendo mensajes y conceptos que rompen con el criterio personal y la capacidad de reflexión a través de la manipulación de la información. No hay nuevos derechos humanos para la virtualidad, y por esa razón se ponen en riesgo los principios más elementales concebidos en la Declaración Universal. Los límites entre lo humano y las máquinas se están desdibujando a pasos agigantados y la sociedad debe protegerse de una nueva forma de feudalismo y control de las vidas humanas.

Palabras clave

los límites entre lo humano y lo cibernético, protección de identidad, protección de la privacidad, alfabetización digital, Nueva Carta Magna, feudalismo informativo, determinismo tecnológico, patrones de consumo, violación de datos personales, control masivo

ABSTRACT

This article delves deeply and philosophically into the new reality that human beings are facing and putting in danger fundamental human rights such as freedom of expression, privacy, freedom of thought. Reflects on how new technologies invade not only markets but also minds, introducing messages and concepts that break with personal criteria and the capacity for reflection through the manipulation of information. There are no new human rights for virtuality and for that reason the most elementary principles conceived in the Universal Declaration are put at risk. The boundaries between humans and machines are fading by leaps and bounds and society must protect itself from a new form of feudalism and control of human lives.

Keywords

the limits between the human and the cybernetic, identity protection, privacy protection, digital literacy, New Magna Carta, information feudalism, technological determinism, consumption patterns, personal data breach, mass control -

La distancia que separa a la Argentina de Jamaica es de 6.441 km y aproximadamente once horas de vuelo. Pero esas limitaciones de tiempo y espacio no tienen importancia, no tienen significancia en la medida que nuestra separación se ha reducido a la velocidad del pensamiento.

Un conocido líder religioso opina que estamos experimentando una verdadera revolución cognitiva e industrial.

Desde la pandemia, la capacidad de la computación, la manipulación de la información y la generación de datos se han convertido en una de las formas de energía más significativas, uniéndose en importancia a la energía nuclear, las energías fósiles y las renovables, lo cual ha sido producto de la capacidad humana de innovación.

Las posibilidades de acceso al conocimiento para miles de millones de personas conectadas por dispositivos inteligentes con un poder de procesamiento sin precedentes están al borde de ser infinitas. Al mismo tiempo, nuestro concepto de lo que significa ser “humano” está siendo desafiado.

Es cierto que somos seres de información. Todo sobre nosotros está codificado en genes. ¿Esto debería significar que compartimos una aproximación con las máquinas? ¿El valor humano debe reducirse a la recolección arbitraria de puntos de datos y de estadísticas?

La respuesta, para algunos, es una clara equivalencia entre lo que podría ser una computadora y lo que podría ser un cerebro humano (la aproximación de lo humano y las máquinas) que si no se controla va a borrar las líneas, los límites, entre lo humano y lo cibernético.

No se necesita mucha inteligencia para poder discernir la tensión entre ese ecosistema y los impactos sociales sobre las libertades individuales. Esto significa que se debe considerar la inteligencia artificial como un factor especial y con ese alto grado de particularidad. De lo contrario, el rediseño y la gobernanza de la sociedad no van a ser óptimos.

Se puede suponer que la inteligencia artificial y la actividad interactiva, que es la próxima fase, van a presentar una antro-

pomorfización¹ de las máquinas que, junto con su capacidad neurotecnológica para interpretar ondas cerebrales, inferirán en las preferencias mentales de las personas y podrán preparar, imprimir o incluso desencadenar esas preferencias diseñadas intencionalmente para eludir nuestras defensas racionales.

El Foro Económico Mundial lo expresa así: “La cuarta Revolución Industrial² finalmente va a cambiar no solamente lo que hacemos sino también quiénes somos. Es decir, va a afectar nuestra identidad y todos los temas asociados con esa identidad, nuestro sentido de la privacidad, nuestras nociones de la propiedad, los patrones de consumo, el tiempo que le dedicamos al trabajo y al ocio o cómo desarrollamos nuestras carreras, cómo cultivamos nuestras habilidades, cómo conocemos a las personas, cómo fomentamos nuestras relaciones. Ya está cambiando la salud y llevando a un *yo cuantificado* y, antes de lo que se piense, podría llevar a una medición del valor de lo humano. La lista es interminable porque lo único que lo limita es nuestra imaginación”.

Un regulador, como por ejemplo la Comisión de Radiodifusión de Jamaica³, que es el nuestro, podría estar entonces menos preocupado por la capacidad compartida de la inteligencia artificial generativa para crear texto, imágenes, video, música. Sin embargo, sí va a estar más preocupado por el papel de la inteligencia artificial en la selección de la información y las noticias que la gente lee, la música que la gente escucha, o las decisiones que toma la gente, así como también su interacción y compromiso político.

Subyacente a ese interés estaría la preocupación de que los sistemas de inteligencia artificial son cajas negras que abren una especie de abismo entre aquellos que crean la inteligencia artificial y todos los demás. Entonces, así se concibe, se acumula y se distribuye información en una escala sin precedentes.

La mayoría de las personas desconocen la naturaleza o el alcance en que la información sobre ellas está siendo almacenada, accedida o compartida, o incluso utilizada para manipular su comportamiento y sus decisiones. Esto significa que las suposiciones y expectativas que tenemos son cada vez más irrealistas. El diseño para la protección de la privacidad, así como se la conoce, se basa en la expectativa de que deberíamos tener una

capacidad inviolable para determinarnos a nosotros mismos sobre cuándo, cómo y en qué medida se comunica a otros la información sobre nosotros.

En base a la observación y a la experiencia, se puede afirmar que muy pocas decisiones se toman a nivel racional. De hecho, muchas personas quizás ni siquiera consideran un gran problema que sus derechos personales sean violados o comprometidos. Cada vez más, se ve la devaluación de la privacidad como algo común, muy común, en la sociedad digital. Es el precio que se paga por estar conectados.

Entonces, la transparencia y el consentimiento informado, que eran baluartes contra la violación de la privacidad, ya dejaron de ser eficaces. Este es el mundo nuevo que se nos presenta. En el mundo antiguo, el ciudadano podía confiar en la Declaración de los Derechos Humanos de la ONU⁴ como una base cierta para poder exigir que sus derechos a la libertad, a la libertad de expresión, a la conciencia y a la privacidad fueran respetados. Pero esos derechos nunca se contemplaron para la persona virtual ni para los nuevos sistemas operativos de la sociedad.

Por consiguiente, como afirma el Instituto de Ingenieros Eléctricos y Electrónicos (IEEE)⁵ en su diseño éticamente alineado, los derechos humanos en la esfera digital no existen hasta que las personas a nivel mundial o global estén empoderadas con los medios, incluyendo las herramientas que puedan asegurar su dignidad a través de alguna forma de soberanía, agencia simétrica o control respecto de su identidad y sus datos personales.

Si se parte de un lugar de equidad y justicia, ninguno de nosotros podría argumentar en contra de la premisa de que las personas informadas tienen que ser o deberían ser un precursor de las máquinas inteligentes, y no al revés.

Las inversiones en chips y en sensores se deben priorizar en igual o menor medida que las prioridades de la UNESCO y la preservación de la información, accesibilidad a la información, ética de la información o información para el desarrollo y multilingüismo en el espacio cibernético.

Frente a ese marco, a continuación presentamos algunas recomendaciones específicas sobre cómo podríamos avanzar

teniendo plena conciencia de la necesidad de evitar la vulneración de los derechos.

Primero: La alfabetización digital mediática e informativa tendría que promoverse como un derecho humano específico, no solamente como uno derivado del derecho a la libertad de expresión o de información, y así debería quedar plasmado en el Artículo 19 de la Declaración Universal.

Segundo: Se debería considerar la posibilidad de crear un nuevo conjunto de derechos neuroespecíficos, que abarquen el derecho a la libertad cognitiva o a la autodeterminación mental. Esto se debe a que, como observa la profesora Nita Farahany⁶, no existen protecciones legales contra el hecho de que nuestra mente sea leída de forma involuntaria.

La ley, según lo que existe actualmente, contempla las externalidades en el sentido de que no se debe restringir nada para poder formular pensamientos y para expresarlos libremente.

Pero en realidad de lo que se está hablando acá es de la jurisprudencia de la mente. Tiene que ver con proteger el cerebro como sistema operativo, que debería ser inviolable a la intrusión mediante la computación neural.

Es decir que los terceros no puedan acceder a la información mental de ninguna persona ni hacer inferencias a partir de sus datos mentales, ya sean neurológicos, cognitivos y/o afectivos, ni tampoco recopilar esos datos sin el consentimiento del dueño de estos, el sujeto de los datos.

Como solución de la información a la asimetría y el desequilibrio de poder entre las grandes tecnológicas, los gobiernos y la gente común tenemos la obligación de proteger los intereses y los derechos de las partes vulnerables.

Tercero: La ley debería imponer responsabilidades fiduciarias a las empresas tecnológicas como solución a la asimetría de la información y el desequilibrio de poder entre las grandes tecnológicas, los gobiernos y los ciudadanos de a pie. Podemos tomar como modelo otras relaciones de poder y confianza, como abogado/cliente, médico/paciente, donde el fiduciario (en este

caso las grandes tecnológicas) tiene la obligación de proteger los intereses y derechos de la parte vulnerable.

Cuarto: Tiene que haber una potente defensa respecto de un marco global de realismo humano construido sobre principios, como la recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial⁷.

Conclusión

Es necesario decir también que las expectativas del multilateralismo deben ser moderadas porque hay muchas organizaciones multilaterales que parecen estar más debilitadas en el momento en el que más se las necesita.

Por lo tanto, la sociedad civil debe ser un jugador fundamental en una forma de cohesión global de fuerzas progresistas que involucren actores a través de todos los vectores de la sociedad y la cultura. En otras palabras, es fundamental que se construya un nuevo multilateralismo de abajo hacia arriba.

Entonces, no se puede negar que las grandes tecnológicas son los nuevos oligarcas, son la raza moderna de conquistadores. Para ellos la conquista no tiene que ver con el territorio sino con la mente. Por eso necesitamos una nueva Carta Magna⁸ mediante la cual la alfabetización digital, mediática e informativa esté garantizada como un derecho humano básico.

Lo que propongo es que si eso se materializa se podría evitar el determinismo tecnológico y el feudalismo informativo. De esa forma podríamos preservar un mundo mejorado a partir de la inteligencia artificial pero que todavía se centre en lo humano, con los valores, con la dignidad humana, la privacidad mental, física y digital y la libertad de expresión. Estos valores deben preservarse y mantenerse intactos.

NOTAS

¹ Conceder forma o cualidades humanas a una cosa o a un ser sobrenatural. Se utiliza también como pronominal: “La fantasía antropomorfiza lo inanimado”.

² Foro Económico Mundial. “La Cuarta Revolución Industrial genera un mundo en el que los sistemas de fabricación virtuales y físicos cooperan entre sí de una manera flexible a nivel global. Sin embargo, no consiste solo en sistemas inteligentes y conectados. Su alcance es más amplio y va desde la secuenciación genética hasta la nanotecnología y de las energías renovables a la computación cuántica. Es la fusión de estas tecnologías y su interacción a través de los dominios físicos, digitales y biológicos lo que hace que la Cuarta Revolución Industrial sea diferente a las anteriores”. <https://es.weforum.org/>

³ <https://www.broadcastingcommission.org/>

⁴ <https://www.un.org/es/about-us/universal-declaration-of-human-rights>

⁵ <https://www.ieee.org/>

⁶ Nita Farahany es una autora iraní-estadounidense y distinguida profesora y académica sobre las ramificaciones de las nuevas tecnologías en la sociedad, el derecho y la ética. <https://www.nitafarahany.com/>

⁷ <https://www.unesco.org/es/artificial-intelligence/recommendation-ethics>

⁸ <https://www.unesco.org/es/memory-world/magna-carta-issued-1215>

INTELIGENCIA ARTIFICIAL PARA PROTEGER A LAS PERSONAS EN ESTADO DE VULNERABILIDAD

SAMRIDHI ARORA (INDIA)

Abogada de la Suprema Corte de India, donde ejerce en el Tribunal Supremo de Justicia. Miembro del grupo de trabajo UNESCO-IFAP sobre accesibilidad a la información. Presidenta de la Conferencia AI4IA de UNESCO. Se dedica a hacer accesible la Justicia a personas desfavorecidas, en estado de vulnerabilidad y personas con discapacidad. Posgrado en la Universidad de Oxford y en Harvard Kennedy School de Estados Unidos.

RESUMEN

Este artículo pone de manifiesto cómo a través de distintos programas de inteligencia artificial creados por empresas y organismos en todo el mundo es posible brindar servicios a personas en estado de vulnerabilidad, con discapacidades, sin acceso a la Justicia y/o que desconocen sus derechos. Salud, justicia, educación, acceso a programas de agricultura y a diferentes servicios están disponibles con la inteligencia artificial para acceder a beneficios que puedan ayudar a las personas con escasos recursos. Muestra la importancia de estar bien informados en los nuevos mecanismos, programas y proyectos que funcionan con IA a nivel mundial para saber defender y proteger nuestros derechos más básicos.

Palabras clave

Herramientas digitales, chatbots, discapacidad, vulnerabilidad, salud, agricultura, desastres naturales, defensa legal, inclusión financiera

ABSTRACT

This article highlights how through different Artificial Intelligence programs created by companies and organizations around the world, provides a service to people in vulnerable situations, with disabilities, without access to justice and who do not know their rights. Health, Justice, Education, Access to agricultural programs and different services are available with Artificial Intelligence to access benefits that can help people with few resources. Talk about the importance of being well informed on the new mechanisms of AI to know how to defend and protect our most basic rights.

Keywords

Digital tools, chatbots, disability, vulnerability, health, agriculture, natural disasters, legal defense, financial inclusion

La inteligencia artificial (IA) está transformando rápidamente nuestra forma de vivir, trabajar e interactuar con el mundo. Más allá de sus aplicaciones en la vida cotidiana, la IA posee un inmenso potencial para abordar los retos críticos a los que se enfrentan las poblaciones vulnerables. La IA aprovecha el poder del análisis de datos, el aprendizaje automático y la automatización para empoderar a las personas y a las comunidades, lo que conduce a soluciones más equitativas y sostenibles.

IA y acceso a la salud

Los sistemas impulsados por inteligencia artificial pueden analizar los datos de las personas enfermas, colaborar con la medicina preventiva a partir de incluir en el historial médico la información genética y los factores del estilo de vida, de riesgos, para crear planes de tratamiento personalizados adaptados a las necesidades individuales. Esto puede dar lugar a tratamientos más eficaces y con mejores resultados para los pacientes.

Los algoritmos de inteligencia artificial pueden analizar imágenes médicas, como radiografías y escáneres, e identificar posibles signos de enfermedad en una fase temprana. Esto permite intervenir a tiempo y puede mejorar significativamente las posibilidades de éxito del tratamiento.

Los chatbots y asistentes virtuales basados en la inteligencia artificial pueden proporcionar consejos médicos básicos y clasificar a los pacientes, lo que ayuda a aliviar la carga de los profesionales de la salud y hace que la asistencia sanitaria sea más accesible para las comunidades remotas o desatendidas.

La inteligencia artificial acelera el descubrimiento y el desarrollo de fármacos analizando grandes conjuntos de datos para identificar posibles candidatos y predecir su eficacia. Esto puede conducir al desarrollo de nuevos tratamientos para enfermedades que actualmente carecen de terapias eficaces.

Hay algunos ejemplos de la inteligencia artificial en la salud con analíticas predictivas para la gestión de enfermedades crónicas, donde los modelos desarrollados se vuelven más precisos y pueden promover el avance de la prevención de enfermedades, así como las imágenes diagnósticas ayudan a llegar a diagnósticos oportunos y a tiempo, sobre todo en enfermedades como el cáncer, posibilitando una mejora en los tratamientos.

La IA capta el estado de vulnerabilidad en las siguientes áreas



Estabilidad económica

CALIFICACIÓN CREDITICIA BASADA EN IA WORD

Algunos modelos de IA precisos pueden evaluar la solvencia crediticia con mayor eficacia, permitiendo la inclusión financiera de personas sin historial de crédito tradicional y de escasos recursos. Por ejemplo, Zest AI¹ utiliza el aprendizaje automático para evaluar el riesgo crediticio y poder brindar créditos a personas marginadas que no pueden aplicar a los sistemas bancarios tradicionales. De esa forma también ayuda a mejorar la calidad económica de esos grupos y el acceso a servicios sociales. Sus modelos incorporan miles de puntos de datos, y logran una mejora del 15 % en la precisión con respecto a los métodos tradicionales de calificación crediticia, lo que ayuda a reducir las tasas de morosidad y ampliar el crédito a

poblaciones desatendidas. Esto se traduce en mejores tasas de reembolso de los préstamos y estabilidad financiera para los prestatarios, lo que hace que se convierta en una herramienta única para acceder a ayuda financiera, cualquiera sea el destino, y permite la inclusión a personas que de otro modo quedarían marginadas de nuevas posibilidades y de ayuda para sus emprendimientos.

Como los avances en prótesis de la IA mejoran la funcionalidad y la calidad de vida en personas con discapacidades

El caso del señor Shivranshu

Se trata de un joven que perdió los brazos por una electrocución y pudo generar una mejora en su vida a través de la prótesis basada en inteligencia artificial. Todo un proceso de pruebas reiteradas resultaron en que él pudiera contar con sus dos brazos amputados y llevar una vida normal. Este joven fue asistido legalmente para obtener estos beneficios a través de asistencia legal gratuita. La vida de Shivranshu se transformó gracias a las prótesis con inteligencia artificial. Esta avanzada tecnología le permitió recuperar el control motor fino y realizar tareas como escribir o tocar instrumentos musicales. Estos dispositivos ofrecen esperanza e independencia a las personas con pérdida de extremidades y les permiten llevar una vida plena.

La inteligencia artificial en los derechos de la discapacidad se integra con innovaciones que mejoran la movilidad y la independencia de las personas con discapacidad motora, con amputaciones y problemas de la marcha. Hay varios ejemplos que nos muestran estos sistemas y las metodologías de uso.

Por ejemplo, para explorar las prótesis existen varias plataformas que permiten analizar, utilizar las instrucciones y generar comandos para controlar los movimientos, como por ejemplo las de EEG² o prótesis inteligentes, donde el brazo artificial está equipado con una serie de sensores adecuados que le dan al paciente una devolución inmediata sobre el entorno. Otro ejemplo es el guante táctil, desarrollado por el MIT³. El sistema utiliza sensores para poder captar señales de presión que son procesadas por una red neural que clasifica objetos y para predecir su peso, en base al toque solamente, sin un estímulo visual.

Futuro de las prótesis con IA

Mejorar el diseño de las prótesis

1 Los avances en la ciencia de los materiales y las técnicas de fabricación permitirán obtener prótesis más ligeras, resistentes y estéticas.

Mejores interfaces cerebro-ordenador

2 Las interfaces cerebro-ordenador más sofisticadas permitirán un control aún más intuitivo y sensible, lo que podría posibilitar la manipulación basada en el pensamiento y una mayor variedad de movimientos.

Soluciones protésicas personalizadas

3 Los algoritmos de IA personalizarán el diseño y el funcionamiento de prótesis en función de las necesidades, preferencias y estilo de vida de cada persona.



Otro desarrollo es el Angus 3, como sistema de prótesis retinal. También existe una plataforma con IA para personas con problemas auditivos y sordera como Widex's Moment⁴. Se utiliza para generar programas auditivos personalizados, analizando los entornos típicos en los que el usuario está presente. A su vez, para individuos con discapacidad auditiva se desarrolló Oticon More, primer equipo de audición equipado con una red neural profunda incorporada. También existe el Tongue Drive System⁵, que hace seguimiento de los movimientos de la lengua dentro de la boca e interpreta los comandos conforme a pautas predefinidas.

Esto tiene que ver con empresas como, por ejemplo, Atom Limbs⁶, que utiliza sensores avanzados y aprendizaje automático para interpretar señales eléctricas del cerebro al cerebro y utilizadas para moverse y para mover el miembro incorporado a la persona. La evolución desde las prótesis del pie de madera a

las prótesis actuales cambió hasta brindar soluciones que mejoran la funcionalidad y la calidad de vida del usuario.

Prótesis con energía de inteligencia artificial, una nueva era para la movilidad que mejora los movimientos con sensores y con el aprendizaje automático, y que, a su vez, ofrecen avances en las funcionalidades y en la experiencia del usuario. Combina los umbrales perceptuales con mediciones de imperancia de electrodos para crear tablas de búsqueda individualizadas para cada persona.

Asistencia social frente a la violencia. Monitoreo

En relación con la protección de las infancias existen modelos precisos de evaluación de peligros que pueden identificar a los niños en situación de riesgo y dar prioridad a los casos de intervención, reduciendo así las situaciones de negligencia y maltrato. La herramienta predijo casos de maltrato infantil con una precisión del 89 %, significativamente superior a la de los métodos tradicionales. El impacto que eso tuvo condujo a una asignación más eficiente de los servicios sociales y a intervenciones oportunas, mejorando la seguridad y el bienestar de los niños.

La inteligencia artificial actúa en el monitoreo de situaciones de vulnerabilidad en muchas partes del mundo. En India, tenemos lo que se llama Safecity⁷, que genera datos sobre abuso o acoso sexual. En Estados Unidos está la RAINN⁸, que opera a nivel nacional con una línea de teléfono destinada a intervenir en ataques sexuales para brindar ayuda a sobrevivientes.

En Australia existe Daisy App, una aplicación desarrollada por 1800RESPECT que ofrece información y recursos a las personas que sufren violencia doméstica y familiar.

En el Reino Unido CHAYN⁹ es una organización internacional que utiliza inteligencia artificial y herramientas digitales para proporcionar recursos y apoyo a las supervivientes del maltrato doméstico y la violencia de género en todo el mundo.

La inteligencia artificial ayuda también a la resolución de casos en el cumplimiento de la ley y mejora la equidad y la

transparencia con el uso de las plataformas de la policía y de investigaciones.

Estas plataformas generan un destino más óptimo de los recursos y colaboran a la resolución de casos. Las herramientas de análisis de patrones digitales del delito, y también el software que detecta de evidencia digital, procesan muchos datos y brindan asistencia virtual. Los Chatbot¹⁰ también brindan respuestas automáticas frente a los pedidos de ayuda a los agentes de la ley.

IA y el manejo de desastres

La inteligencia artificial en el manejo de los desastres ecológicos, huracanes, temporales e incendios mejora la respuesta y los esfuerzos para una rápida recuperación. Con la predicción y el monitoreo, la inteligencia artificial puede analizar datos para estar más preparados en tiempo real y predecir catástrofes. También la información compartida en las plataformas de inteligencia artificial facilita la disponibilidad de información durante los desastres, lo que mejora la coordinación, el destino de los recursos y la gestión de los mismos.

La inteligencia artificial optimiza el destino de los recursos durante el desastre, y en cuanto a la recuperación posdesastre, puede brindar asistencia para ayudar a evaluar los daños y lograr una recuperación más rápida. Hay algunos ejemplos de inteligencia artificial que han dado respuestas eficaces, por ejemplo el sistema IBM® Watson,¹¹ que ha sido utilizado para predecir el camino y el impacto de huracanes permitiendo organizar las evacuaciones y minimizando la pérdida de vidas. Otro ejemplo es un modelo de predicción de inundaciones en el sudeste asiático, donde se mejoró la preparación y la respuesta con una ventana de 72 horas.

IA y vulnerabilidad legal

Otro aspecto importante es el uso de la inteligencia artificial para proteger a la gente en situación de vulnerabilidad legal. La IA puede analizar gran cantidad de datos para poder brindar apoyo a través de herramientas automatizadas. Por ejemplo, DoNotPay Chatbot, que brinda asistencia legal gratis. Los sistemas de IA recopilan y analizan grandes volúmenes de datos

procedentes de diversas fuentes, como expedientes jurídicos, datos de clientes, sentencias judiciales y actualizaciones de la legislación. Sacan provecho de algoritmos avanzados para procesar y analizar los datos recopilados, como el procesamiento del lenguaje natural (PLN), el aprendizaje automático (AM) y el análisis predictivo. Puede evaluar los riesgos jurídicos prediciendo los resultados de los casos, analizando contratos, realizando comprobaciones de cumplimiento legal y detectando fraudes. El análisis predictivo en la Justicia Criminal ayuda a mitigar los prejuicios y garantiza sentencias más equitativas, lo que aumenta la confianza en el sistema de justicia penal.

El Kira Systems¹² y el Luminance¹³ reducen significativamente el tiempo y los recursos necesarios para la revisión de contratos y la diligencia debida, mejorando la precisión y el cumplimiento así como CourtNav en el Reino Unido capacita a los usuarios proporcionándoles información y apoyo jurídico claros y accesibles, con lo cual se mejora el acceso a la Justicia de quienes pueden encontrar obstáculos por su coste o complejidad.

IA y la agricultura frente a los puntos vulnerables

La inteligencia artificial también está mejorando y revolucionando la agricultura, dando soluciones de gestión para los productores y agricultores, desde los campos que están cultivados y otros tipos de desafíos. Hay que proteger, darle poder y al mismo tiempo permitirle a la población el acceso a todas estas herramientas aplicadas a varios sectores de la economía y de los derechos humanos básicos de las personas. Continuando con la implementación de más y más soluciones, haciéndolas accesibles a las poblaciones vulnerables, podemos crear un mundo más equitativo e inclusivo.

Los drones con inteligencia artificial proporcionan a los agricultores información en tiempo real sobre sus cultivos, lo que les posibilita identificar problemas, aplicar soluciones específicas y maximizar el rendimiento. Permiten supervisar el uso del agua, identificar las zonas que requieren riego y minimizar el despilfarro, lo que se traduce en prácticas sostenibles y ahorro de costos. Los drones equipados con sensores pueden detectar precozmente brotes de enfermedades, deficiencias de nutrientes y otros problemas de salud de los cultivos, lo que permite intervenir a tiempo.

Iniciativas agrícolas mundiales con IA

- **Plataforma de agricultura de precisión**
Utiliza IA para ofrecer a agricultores/as recomendaciones personalizadas sobre aplicación de fertilizantes, programas de riego y estrategias de control de plagas.
- **Sistema automatizado de cosecha**
Se están desarrollando robots impulsados por IA para automatizar el proceso de cosecha, reduciendo costos de mano de obra y mejorando la eficiencia.
- **Mejora de cultivos asistida por IA**
Científicos/as utilizan IA para analizar datos genéticos y acelerar el proceso de obtención de nuevas variedades de cultivos con mayor rendimiento y resistencia.
- **Predicción meteorológica asistida por IA**
Agricultores/as acceden a previsiones meteorológicas precisas para prepararse ante posibles tormentas, sequías y otros fenómenos meteorológicos.



La IA tiene el potencial de ser una fuerza transformadora para el bien, abordando los retos críticos a los que se enfrentan las poblaciones vulnerables en diversos sectores. Aprovechando sus capacidades en los ámbitos sanitarios, de gestión de catástrofes, jurídico y económico, la IA puede contribuir a un futuro más equitativo y sostenible para todos. Es esencial desarrollar e implantar sistemas de IA de forma responsable, garantizando que sean accesibles, éticos, sin sesgos discriminatorios y diseñados para beneficiar a todos los miembros de la sociedad por igual.

NOTAS

¹ <https://www.zest.ai/>

² <https://www.clinicbarcelona.org/asistencia/pruebas-y-procedimientos/electroencefalograma>

- ³ <https://www.mit.edu/>
- ⁴ <https://www.widex.com.ar/audifonos/moment>
- ⁵ <https://mobilitymgmt.com/the-tongue-drive-system/>
- ⁶ <https://atomlimbs.com/>
- ⁷ <https://webapp.safecity.in/>
- ⁸ <https://rainn.org/>
- ⁹ <https://www.chayn.com>
- ¹⁰ <https://www.chatbot.tap.ai>
- ¹¹ [https://es.wikipedia.org/wiki/Watson_\(inteligencia_artificial\)](https://es.wikipedia.org/wiki/Watson_(inteligencia_artificial))
- ¹² <https://kirasystems.com/>
- ¹³ <https://www.luminance.com>

DEMOCRATIZACIÓN DEL CONOCIMIENTO A TRAVÉS DE LA TECNOLOGÍA

MATÍAS KARLSSON (ARGENTINA)

Ingeniero. Experto en inteligencia artificial con más de veinte años de experiencia. Profesor en Le Wagon y en la Universidad Austral en Argentina, así como en otros países de América Latina. Chief Technology Officer de Traditum y de otras dos empresas ubicadas en Estados Unidos y Nueva Zelanda. Ha sido ponente en diferentes países.

RESUMEN

La llegada de la inteligencia artificial y de todas las nuevas formas de interacción y aprendizaje a través de las nuevas tecnologías ha modificado la manera de acceder al conocimiento y se vuelve necesario democratizarlo. Se tiene mucha información, hay gran cantidad de datos, y sin un cambio en la educación se pierde la distancia crítica necesaria para saber qué es correcto o no de lo que nos dicen las plataformas y los sistemas. La forma de aprender, de incorporar los nuevos conocimientos y ponerlos al servicio positivo de la humanidad depende en gran parte de mantener una actitud crítica. Este artículo describe cómo ha cambiado el modo de aprender desde la antigüedad hasta nuestros tiempos de la industria 4.0 y la inteligencia artificial.

Palabras clave

democratización, enseñanza, ChatGPT, aprendizaje, industria 4.0, historia del aprendizaje, información digital

ABSTRACT

The development of artificial intelligence and all the new forms of interaction and learning through new technologies has changed the way of accessing knowledge and it has become necessary to democratize it. We have a lot of information, a lot of data and if we not change the way of educating and teaching we may lose the critical distance necessary to know if what the platforms and systems tell us is correct or not. The way to learn, to incorporate new knowledge and put it in a positive service of humanity, depends largely on maintaining a critical attitude. This article describes how the way of learning has changed from ancient times to our times of Industry 4.0 and Artificial Intelligence

Keywords

democratization, knowledge, ChatGPT, learning, industry 4.0, history of learning, digital information

Es interesante reflexionar sobre la historia del conocimiento de la humanidad: cómo ha ido desplazándose y agrupándose, y cómo el revisionismo permite cuestionarnos sobre su origen. Uno de los grandes hitos de conocimiento en la antigüedad fue la biblioteca de Alejandría, en el siglo III a. de C. En esa época requisaban los barcos, y si llevaban papiros, los copiaban y los guardaban en la biblioteca. Solo unas pocas personas tenían acceso a estos conocimientos, ya que apenas algunos sabían leer y escribir y podían acceder a esa base de información. Aunque era el repositorio más grande del mundo en ese momento, en comparación con la información que tenemos hoy no llegaría ni a lo que cabe en un pendrive.

Con el avance de los siglos, el conocimiento comenzó a democratizarse un poco más, especialmente con la invención de la imprenta en el siglo XV. Se introdujo tecnología para evitar la copia manual, y el conocimiento empezó a expandirse más ampliamente. En ese momento, la humanidad pudo acceder a algunos materiales, aunque no a todos, pues aún se necesitaba contar con el libro adecuado para obtener ciertos conocimientos específicos.

Internet es otra herramienta que ha facilitado el acceso al conocimiento global, ya que ha permitido que esté al alcance de una computadora o un celular. Esto es significativo porque, mientras antes se requería memorizar datos debido a la limitada disponibilidad de información, ahora contamos con esta disponibilidad. Esto nos lleva a preguntarnos si nuestra forma de aprender debe adaptarse. Antes teníamos pocos datos y mucho tiempo; hoy tenemos abundancia de datos y menos tiempo. Por lo tanto, cambia mucho nuestra forma de aprender y de pensar, y más aún con la llegada de la inteligencia artificial.

Este avance representa un cambio de paradigma, en el que no solo se tiene acceso a todo el conocimiento de la humanidad, sino también a herramientas que pueden ayudar a interpretarlo. Esto plantea una cuestión interesante: ¿por qué se necesita un agente que nos ayude a interpretar la información? ¿Por qué es necesario tener este agente que ayude a interpretar las cosas?

En 2010 comienzan a producirse datos en zettabytes, que eran mínimos en comparación con la escala actual. Antes de ese año, la cantidad de información digital en zettabytes era casi insignificante. Para 2024, ya se ha superado la marca de los 120 zettabytes, y se espera que para 2025 la cifra alcance los 175 zettabytes de información digital disponible. Un zettabyte es un uno seguido de veintiún ceros en bytes, una cantidad abrumadora de información.¹

Anteriormente había tiempo suficiente para estudiar veinticinco manuscritos; en cambio, hoy en día la cantidad de información es tal que resulta abrumadora. El problema ya no es acceder a la información, sino encontrar lo que realmente se busca en medio de tanta abundancia de datos.

Es aquí donde entra la tecnología para facilitar esta tarea, ofreciendo ayuda para encontrar la información deseada. Aunque quienes tienen conocimientos sobre inteligencia artificial entienden cómo funciona en su interior, sigue siendo algo similar a una “caja negra”, difícil de explicar en cuanto a por qué da una respuesta en lugar de otra. La IA se basa en el aprendizaje, que puede estar sesgado o contener errores, dada la inmensa cantidad de datos disponibles. Por eso, aunque se recurra a la inteligencia artificial para encontrar información, es importante mantener una actitud crítica, ya que también puede equivocarse y ofrecer datos incorrectos.

Para adaptarse a esta nueva forma de educación es necesario desarrollar una mentalidad crítica. Hoy en día no es suficiente estudiar para aprender y memorizar; ahora es fundamental aprender a pensar y a desarrollar un pensamiento crítico. Esto es lo que permite progresar en un momento donde la tecnología ya facilita el acceso al conocimiento.

El valor añadido humano radica en el pensamiento crítico. A lo largo de la historia, la humanidad ha pasado por diferentes niveles de acceso a la información, y aunque antes era fundamental memorizar y asimilar toda la información disponible, hoy el panorama ha cambiado. En la era de la industrialización, se buscó industrializar incluso el conocimiento, para simplificar la tarea docente, lo que llevó a algunos métodos extremos.

Por ejemplo, para reducir la carga de trabajo del docente en los exámenes, se utilizó la modalidad de *multiple choice* (opción múltiple), donde solo hacía falta una plantilla con las respuestas correctas para determinar si una respuesta era verdadera o falsa y si el examen estaba aprobado. Hoy en día, con el uso de computadoras, un examen de opción múltiple puede calificarse instantáneamente.



La pregunta es: ¿Este es el tipo de educación que se necesita hoy en día? ¿Qué ocurre cuando se ofrece este tipo de educación a un estudiante? ¿Qué pasa cuando se le dice a un alumno que no puede usar calculadora, internet o herramientas como ChatGPT? ¿Qué sucede cuando ese estudiante termina su carrera y se enfrenta al mundo laboral? Empiezan a surgir casos como el de un abogado que usó ChatGPT para redactar una demanda y obtuvo antecedentes legales inventados.

¿Por qué ocurre esto? Porque, en un entorno donde podría haberse formado para desarrollar una mentalidad crítica, no se le enseñó a cuestionar si esos antecedentes realmente existían. La universidad y el colegio deberían ser espacios seguros para aprender y cometer errores; lugares donde, si se usan herramientas como ChatGPT de forma incorrecta, se puede recibir la orientación necesaria para corregir ese error.

Como ejemplo, un estudiante que elabora un resumen en tamaño reducido para esconderlo y poder copiarse ha aprendi-

do en cierto sentido a hacer una síntesis, incluso si no se le enseñó explícitamente. Esto demuestra cómo el estudiante busca el conocimiento activamente, lo cual es interesante en esta era tecnológica. Debemos replantearnos la idea de que el propósito de la educación es solo evaluar, sin dar espacio para pensar y analizar.

Un excelente ejercicio sería plantear en un examen la pregunta: “¿Qué opina sobre el uso de ChatGPT por un abogado en la redacción de demandas? ¿Por qué considera que es correcto o incorrecto? ¿Debería ChatGPT asumir alguna responsabilidad por los antecedentes que inventó?”



Cuando se comienza a realizar este tipo de actividades, se observa cómo se estimulan las neuronas. Coloca tanto al profesor como al alumno en una situación incómoda, pero ambos aprenden. De la otra manera, ninguno de los dos realmente aprende, y es ahí cuando surgen problemas.

Hoy en día, la única forma de acceder al conocimiento y gestionarlo de manera efectiva es a través del pensamiento crítico. Antes de la Revolución Industrial, la economía global era bastante uniforme y, en general, muy pobre. La situación mundial no era favorable, ya que debía enfrentarse a constantes hambrunas y epidemias. Esto no quiere decir que actualmente no existan desigualdades ni hambrunas, pero en ciertos aspectos al mundo le va mejor que antes. Es importante recordar también el impacto del cambio climático, que es una realidad ac-

tual. Aun así, en términos de economía y tecnología, el mundo ha avanzado significativamente, y esto es evidente.

Están por un lado las naciones de Europa occidental, Australia, Canadá, Estados Unidos, Nueva Zelanda y Japón, y por otro la “periferia” o países en desarrollo, o emergentes. Con el inicio de la Revolución Industrial y la incorporación de tecnologías, los países del Primer Mundo que pudieron industrializarse experimentaron un crecimiento exponencial. Por otro lado, aquellos que no consiguieron integrarse en esa ola —la mayoría de los países del Tercer Mundo, como Argentina, que tardó en industrializarse— quedaron rezagados en comparación con las naciones que lograron ese avance.

Desde la era de las tecnologías de la información, y especialmente con la llegada de la industria 4.0² y la inteligencia artificial, estos países han comenzado a experimentar un salto algo mayor en su desarrollo. La clave aquí es que, aunque un país no haya aprovechado la ola de industrialización anterior, ahora tiene la oportunidad de hacerlo a través de la economía del conocimiento en esta nueva Revolución Industrial. La revolución tecnológica es aún más grande que la anterior, y países como Argentina tienen la posibilidad de posicionarse como líderes en conocimiento e inteligencia artificial.

Para lograr esto, es fundamental centrarse en democratizar el conocimiento y entender las nuevas formas de aprendizaje para poder guiar a los futuros estudiantes que liderarán esta nueva era industrial inteligente y transformarán la sociedad. Si un país decide alejarse de estos avances tecnológicos, limitando el uso de computadoras, herramientas como ChatGPT, o la inteligencia artificial en las aulas (no de manera indiscriminada, sino organizada y segura para el aprendizaje), se corre el riesgo de ampliar aún más la brecha entre las naciones desarrolladas y en desarrollo, incluso más que durante la era industrial.

Una frase inspiradora de María Montessori, destacada pensadora italiana en educación, refleja esta visión: “Los conocimientos no deben ser introducidos en la mente de los niños, sino que ellos deben tener la libertad de percibirlos y razonarlos”. En la actual era del conocimiento, donde la información está al alcance de todos, el verdadero desafío es captar el interés

de los estudiantes, desarrollar su capacidad crítica y fomentar su crecimiento, no empujando conocimiento, sino invitándolos a explorarlo.

El modelo tradicional de aulas con un profesor impartiendo contenido a una clase pasiva está siendo reemplazado. En esta nueva era, la educación tiende más hacia una interacción personalizada, donde el estudiante toma la iniciativa para consultar con el profesor sobre temas específicos. Esta es una de las transformaciones que la inteligencia artificial está generando en la educación.

NOTAS

¹ Según el informe Data Age 2025 de IDC, para el año 2025 la cantidad de datos podría alcanzar los 175 zettabytes. Eso es cuarenta veces más que el volumen de datos de 2013. <https://www.seagate.com>

² La industria 4.0 se puede definir como la integración de tecnologías digitales inteligentes en la fabricación y los procesos industriales. Abarca un conjunto de tecnologías que incluyen redes industriales de IoT, IA, big data, robótica y automatización. <http://www.energia.gob.ar>

EDUCACIÓN E INTELIGENCIA ARTIFICIAL. UN CASO ARGENTINO

ALICIA PENAYO (ARGENTINA)

Presidenta en Silicon Misiones¹. Máster en Business Intelligence de la Universidad de Barcelona. Coordinadora general en el POLO TIC Misiones. Licenciada en Administración de Empresas por la Universidad Argentina de la Empresa (UADE).

RESUMEN

Este artículo examina la intersección entre educación e inteligencia artificial (IA) en Misiones, Argentina, y cómo la IA transforma este sector a través de un ecosistema educativo disruptivo.

Con un 27 % de su población en el rango de 0 a 14 años, la provincia enfrenta el reto de preparar a los jóvenes para un futuro en constante cambio. Para eso la provincia ha desarrollado una escuela de robótica para niños a partir de los cuatro años, y también la Escuela Secundaria de Innovación; el Silicon Misiones está en el último tramo de este ecosistema.

Misiones está trabajando en un proyecto de ley que busca integrar la IA en la educación, con el propósito de promover modelos educativos innovadores, mejorar la alfabetización digital y garantizar un uso ético de la IA. Este proyecto propone la incorporación gradual de la IA evaluando su impacto y protegiendo los derechos de los estudiantes. Subraya la importancia de centrar la IA en las personas, asegurando la inclusión y que su implementación beneficie a toda la comunidad.

La provincia de Misiones ya está desarrollando políticas y programas para transmitir conocimientos sobre alfabetización digital y abordar problemáticas emergentes en el entorno digital, como la ludopatía y el ciberbullying.

Palabras clave

educación disruptiva, alfabetización digital, brecha digital, innovación educativa, perspectiva crítica, ciberbullying

ABSTRACT

This article examines the intersection between education and artificial intelligence (AI) in Misiones, Argentina, and how AI transforms this sector through a disruptive educational ecosystem.

With 27% of its population in the 0-14 age range, the province faces the challenge of preparing young people for an ever-changing future. For this purpose the province has developed a robotics school for children from the age of 4 and , also a Secondary School of Innovation, the Silicon Misiones is in the last leg of this ecosystem.

Misiones is working on a bill that seeks to integrate AI in education, with the purpose of promoting innovative educational models, improving digital literacy and ensuring an ethical use of AI. This project proposes the gradual incorporation of AI, evaluating its impact and protecting the rights of students. It stresses the importance of focusing AI on people, ensuring inclusion and that its implementation benefits the entire community.

The province of Misiones is already developing policies and programs to transmit digital literacy skills and address emerging issues in the digital environment, such as gamification and cyberbullying.

Keywords

disruptive education, digital literacy, digital divide, educational innovation, critical perspective, ciberbullying

La inteligencia artificial está revolucionando muchos aspectos y muchos sectores de la sociedad. El ámbito de la educación no es la excepción, y representa uno de los sectores más críticos en cuanto a su influencia.

La provincia de Misiones, en Argentina, tiene una historia marcada por la Ley de Educación Disruptiva², que vino a cambiar el paradigma tradicional de la manera de aprender a través de la implementación de las nuevas tecnologías de la información y la comunicación (TIC), así como también las tecnologías del aprendizaje y el conocimiento (TAC)³.

Misiones se encuentra en el noreste del país. El 90 % de su frontera es compartida con Paraguay y Brasil, y solo el 10 % con otra provincia argentina, lo que implica una situación compleja. El ecosistema de educación disruptiva se centra en innovar en la educación de su población, en especial si se considera que la provincia cuenta con la mayor proporción de jóvenes del país, con un 27 % de su población entre 0 y 14 años.

Esto plantea un gran desafío en la formación y el desarrollo del talento necesario para el futuro en un entorno totalmente desafiante y cambiante. Para enfrentar este reto, Misiones ha establecido un ecosistema de educación disruptiva que construye estas bases pedagógicas para el futuro; el Silicon Misiones está en el último tramo de este ecosistema, consolidando así su compromiso con la educación del futuro.

Los alumnos pueden formarse a partir de los cuatro años en la escuela de robótica, donde adquieren habilidades en programación y robótica y atraviesan diversos trayectos educativos. La provincia también cuenta con la Escuela Secundaria de Innovación,⁴ que presenta un modelo educativo innovador en comparación con la educación tradicional. En esta institución, las materias se imparten de manera interdisciplinaria, combinando asignaturas como inglés con matemáticas y con biología, y ciencias naturales con historia, y se mezcla la currícula aplicando el enfoque de *flipped learning*⁵, en el que el conocimiento se construye de manera colaborativa entre el estudiantado y los docentes, quienes actúan como facilitadores del aprendizaje.

En 2024 se creó la Secretaría de Inteligencia Artificial en la Cámara de Representantes de Misiones⁶, lo que refleja el enfoque de la provincia en abordar nuevos desafíos.

En particular, se está tratando en comisión un proyecto de Ley de Integración de Inteligencia Artificial en la Educación⁷.

Proyecto de Ley Integración de la IA en la educación

(En tratamiento en comisión - Cámara de
Representantes de la provincia de Misiones)

- DESARROLLO DE MODELOS EDUCATIVOS
INNOVADORES**
Incorporar IA para mejorar la calidad educativa, de manera transversal en diferentes niveles y modalidades, de forma gradual.
- OPORTUNIDADES DE APRENDIZAJE Y
ALFABETIZACIÓN DIGITAL**
Garantizar acceso a experiencias educativas potenciadas por IA para todos y usar IA para mejorar la alfabetización y competencias tecnológicas.
- REGULACIÓN Y ÉTICA**
Establecer marcos que prioricen la seguridad y derechos fundamentales del educando.

La provincia considera que la regulación es fundamental para asegurar que la implementación de la inteligencia artificial en el aula se realice de manera ética y responsable, siempre con una mirada crítica.

Para este proyecto de ley se adopta la definición de inteligencia artificial propuesta por UNESCO en 2022: “Son aquellos sistemas con capacidad para procesar datos de forma similar a un comportamiento inteligente”. Esta definición es importan-

te, ya que entiende el “comportamiento inteligente” en relación con capacidades como el aprendizaje, la percepción y el razonamiento lógico, y es determinante en un contexto de rápidos cambios tecnológicos.

También aborda la inteligencia artificial generativa como un tipo de IA que crea contenido nuevo a partir de grandes modelos de aprendizaje, no solo automatizando tareas.

La incorporación se plantea de forma gradual, iniciando una fase de transición para la experimentación sistemática con inteligencia artificial generativa. Se prevén instancias de revisión y seguimiento, analizando de manera crítica el impacto de la inteligencia artificial, siempre enfocados en el educando.

Este proyecto tiene como objetivo preservar los derechos de todas las personas y favorecer el desarrollo de modelos educativos innovadores, inclusivos y de calidad.

La provincia ha avanzado en este ámbito a través de diversas normas aprobadas por la Cámara de Representantes, con el objetivo de crear experiencias educativas innovadoras que promuevan el uso de la inteligencia artificial. Este enfoque busca ofrecer oportunidades de aprendizaje para la totalidad de las personas a lo largo de toda la vida, y abarca también contextos más allá del entorno escolar.

A la vez, se busca facilitar la integración de la inteligencia artificial en el ámbito educativo mediante espacios de cuestionamiento que permitan establecer regulaciones para su uso.

Se propone también aprovechar la inteligencia artificial para abordar problemáticas relacionadas con la alfabetización, fomentando el acceso al conocimiento y al manejo de tecnologías, claves para el desarrollo de una ciudadanía crítica. Esto incluye fortalecer los procesos de enseñanza y aprendizaje a través de estrategias didácticas que incorporen tecnología o inteligencia artificial generativa, con el objetivo de generar aprendizajes significativos.

La norma se propone favorecer el empoderamiento de los estudiantes y busca integrar la inteligencia artificial de manera inclusiva y acortar la brecha digital. Para lograr esto, se pretende

garantizar el acceso a internet en todo el territorio de la provincia con la instalación de antenas satelitales para promover la conectividad en las áreas más remotas.

Los principios guía del proyecto de ley hacen foco en el uso de la inteligencia artificial centrado en el ser humano, así como en las necesidades, capacidades y experiencias de las personas, para que sean las que diseñen, implementen y utilicen la inteligencia artificial.

Se busca que las experiencias educativas sean diseñadas para mejorar el aprendizaje de los estudiantes y servir como herramienta de apoyo para los docentes.

Se trata de fomentar un cuestionamiento crítico sobre el impacto del uso de la inteligencia artificial, identificando los riesgos potenciales y regulando su uso en el aula y en las experiencias educativas. Esto debe realizarse reconociendo la necesidad de impulsar la innovación educativa, teniendo como eje al ser humano y garantizando la seguridad y el respeto de sus derechos.

La ley prevé el acompañamiento y la capacitación continua de los docentes ya que existe la necesidad de impulsar la adaptación y actualización didáctico-pedagógica de estos para que se conviertan en guías y facilitadores del proceso de aprendizaje de los estudiantes.

La provincia ya cuenta con una Ley de Educación Disruptiva, que marca el inicio de este camino. Se pone énfasis en la centralidad del estudiante y en el desarrollo del pensamiento computacional, una habilidad esencial para los desafíos actuales.

El gobierno provincial ya está implementando políticas y programas que contribuyen a alcanzar estos objetivos. Un ejemplo es la formación de formadores a nivel provincial, que puedan transmitir conocimientos sobre alfabetización digital y abordar problemáticas emergentes en el entorno digital, como la ludopatía y el ciberbullying.

Igualmente, se está trabajando en la interacción con la inteligencia artificial, educando sobre los desafíos y riesgos asocia-

dos. Se aspira a integrar la inteligencia artificial en las herramientas digitales utilizadas a diario, para abordar fenómenos como la polarización y los sesgos existentes. Todo esto se enmarca en un esfuerzo por concientizar y transmitir este conocimiento a las comunidades.

En esta nueva ley se identifican los beneficios de implementar la inteligencia artificial en el aula, tales como la personalización del aprendizaje, la optimización de la gestión educativa como herramienta para los docentes y la mejora en la accesibilidad, todo ello acompañado de políticas concretas para reducir la brecha digital.

Se pretende crear espacios que permitan escuchar al estudiantado y aprender de su experiencia.

En esta norma se consideran los desafíos éticos y pedagógicos, por lo que este proyecto de ley se apoya en un trabajo de investigación en el territorio. Esto incluye la continuidad en la creación de un marco práctico y legal que regule el uso de la inteligencia artificial en el aula.

Para concluir, la misión de esta ley en el ámbito educativo es reafirmar la relevancia de la inteligencia artificial para el desarrollo de competencias futuras y la mejora continua del sistema educativo, garantizando su uso ético, el desarrollo de habilidades necesarias y la equidad en el acceso a la tecnología.

NOTAS

¹ <https://siliconmisiones.gob.ar/>

² Ley VI N° 212 https://digestomisiones.gob.ar/archivospdf/1688646267_LEY%20VI%20%E2%80%933%20N%2012%20Texto%20Definitivo.pdf

³ Tecnologías de la Información y las Comunicaciones.

⁴ El término TAC en este caso se refiere a las Tecnologías para el Aprendizaje y el Conocimiento que incluyen a las TIC más un componente metodológico, necesario para generar un aprendizaje significativo. Este tipo de tecnologías apuestan a un entorno de aprendizaje más personalizado, o lo que se conoce como PLE (por sus siglas en inglés, *Personal Learning Environment*) donde el alumno es protagonista de la incorporación de su conocimiento. https://es.wikipedia.org/wiki/Tecnolog%C3%ADas_del_aprendizaje_y_el_conocimiento

⁵ <https://esim.edu.ar/web/>

⁶ El “*flipped learning*” o “aula invertida” es un enfoque pedagógico que ha ganado popularidad en la educación superior al invertir el modelo tradicional de enseñanza. En este modelo, los alumnos estudian los contenidos teóricos antes de la clase a través de recursos digitales como videos y lecturas. El tiempo en el aula se dedica a la aplicación práctica de esos conceptos, lo que transforma el rol del profesor de transmisor de información a guía y facilitador del aprendizaje, lo que permite diseñar actividades prácticas, discusiones y debates que ayudan a los estudiantes a aplicar lo aprendido y profundizar en su comprensión. <https://campus.iesalc.unesco.org/inicio/blocks/coursefilter/course.php?id=200>

⁷ <http://www.diputadosmisiones.gov.ar/nuevo/secretarias>

⁸ <http://www.diputadosmisiones.gov.ar/nuevo/archivos/proyectos/P64132.pdf>

ÉTICA DE LA INTELIGENCIA ARTIFICIAL: EQUIDAD DE GÉNERO

NATALIA GONZÁLEZ ALARCÓN (COLOMBIA)

Economista y especialista en política tecnológica. Cuenta con más de diez 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 en América Latina y el Caribe en la UNESCO. Máster en Economía de la Universidad de los Andes, Colombia. Máster en Políticas Públicas de la Universidad de Georgetown.

RESUMEN

A través de la Recomendación sobre la Ética de la IA de la UNESCO (UNESCO, 2022) se presenta un enfoque global para guiar el desarrollo y el uso responsable de esta tecnología, con énfasis en la equidad de género y la inclusión de poblaciones vulnerables.

Se destacan tres pilares fundamentales: la protección de los derechos humanos, la creación de sociedades justas y la garantía de diversidad e inclusión. Además, se analizan los principios éticos cruciales para asegurar que la IA beneficie a todos, evitando la ampliación de brechas sociales.

La publicación explora la disparidad de género en el acceso a la tecnología, las habilidades digitales y el ámbito de la investigación. También se analiza cómo estas desigualdades impactan en el mercado laboral, donde las mujeres enfrentan barreras para acceder a empleos mejor remunerados en la economía digital.

Finalmente, iniciativas como Women4EthicalAI buscan fomentar la participación de mujeres en el desarrollo tecnológico y promover algoritmos sin sesgos de género. Esta publicación presenta estrategias concretas para avanzar hacia una IA más equitativa y accesible.

Palabras clave

inteligencia artificial, UNESCO, ética, equidad de género, inclusión, datos

ABSTRACT

Through UNESCO's Recommendation on the Ethics of AI, a global framework is presented to guide the responsible development and use of this technology, with a focus on gender equity and the inclusion of vulnerable populations.

Three fundamental pillars are highlighted: the protection of human rights, the creation of just societies, and the guarantee of diversity and inclusion. Additionally, the publication examines the crucial ethical principles necessary to ensure that AI benefits everyone, preventing the widening of social gaps.

The publication explores gender disparities in access to technology, digital skills, and the research field. It also analyzes how these inequalities affect the labor market, where women face barriers to accessing higher-paid jobs in the digital economy.

Finally, initiatives like Women4EthicalAI aim to promote the participation of women in technological development and foster algorithms free from gender bias. This publication presents concrete strategies for advancing towards a more equitable and accessible AI.

KEYWORDS

artificial intelligence, UNESCO, ethics, gender equity, inclusion, data

Recomendaciones sobre la Ética de la Inteligencia Artificial: ¿Por qué son importantes?

Se ha hablado del potencial de la inteligencia artificial, también de los potenciales riesgos y desafíos que presenta. La Recomendación sobre la Ética de la Inteligencia Artificial de UNESCO se considera el primer instrumento global que pretende guiar ese desarrollo, uso y/o despliegue de esta tecnología emergente.

Cuando se habla de equidad de género, se habla de un enfoque respecto de la inclusión, de cómo incorporar a esas poblaciones vulnerables en la conversación y cómo hacer que estas tecnologías usen ese potencial que tienen sin generar un incremento de la desigualdad. Es decir, que no amplíen brechas que ya existen en diferentes aspectos sociales. Ese es el objetivo detrás de tener este tipo de recomendaciones. Los 193 países miembros de la UNESCO adoptaron unánimemente este instrumento, por lo que este documento se convierte en la base para llevar la conversación a otro nivel.

La Recomendación tiene tres grandes momentos en los que describe y establece subfocos:

Un primer momento relacionado con los valores que propone gira en torno al respeto de los derechos humanos y la dignidad humana. El segundo es el de tener sociedades justas y pacíficas. Un tercer momento es garantizar la diversidad y la inclusión.

Hay además un cuarto momento que, en su primera instancia, fue el enfoque novedoso de la Recomendación, que es el del efecto y el impacto sobre el medio ambiente y los ecosistemas.

En una segunda parte, la Recomendación presenta los principios que deberían guiar el uso y el desarrollo de esta tecnología emergente con diez principios básicos:

1. Proporcionalidad e inocuidad (se enfatiza la necesidad de evitar riesgos y prevenir posibles daños al utilizar la IA, basándose en fundamentos científicos rigurosos).
2. Seguridad y protección.
3. Equidad y no discriminación.

4. Sostenibilidad.
5. Derecho a la intimidad y protección de datos.
6. Supervisión y decisión humanas.
7. Transparencia y explicabilidad.
8. Responsabilidad y rendición de cuentas.
9. Sensibilización y educación.
10. Gobernanza y colaboración adaptativas de múltiples partes interesadas.

Uno de estos principios es particularmente importante y relevante para el tema que nos compete: el de equidad y no discriminación.

Este principio plantea que se precisa un enfoque inclusivo para garantizar que los beneficios de las tecnologías estén disponibles y accesibles para todos. Todos no son solamente hombres y mujeres. En “todos” se incluyen necesidades específicas de diferentes grupos poblacionales, muchos de ellos en situación de vulnerabilidad. Considerando diferentes variables, se tienen en cuenta el género, la edad, diferentes sistemas culturales y/o grupos lingüísticos, personas con discapacidad, niños, niñas y adolescentes y personas que están en situación de vulnerabilidad y al margen de beneficios sociales.

Esto se traduce en temas de representatividad. ¿Por qué es tan importante que estas tecnologías estén alimentadas? Porque al final los resultados o los impactos que puedan tener las tecnologías de inteligencia artificial están 100 % basados en los datos que las alimentan. Estos datos, al final, tienen que garantizar la representatividad de los diferentes grupos poblacionales.

Otro aspecto de la Recomendación es el aterrizaje en la práctica de aplicación de estos principios en el día a día, en diferentes sectores y aspectos de nuestras vidas y en diferentes ámbitos de acción política¹. Estos ámbitos son los siguientes:

1. Evaluación de impacto ético.
2. Política de datos.
3. Género.
4. Gobernanza.

5. Medio ambiente y ecosistemas.
6. Cultura.
7. Educación e investigación.
8. Economía y trabajo.
9. Comunicación e información.
10. Salud y bienestar social.
11. Desarrollo y cooperación internacional.

La inclusión de género en la IA: retos en acceso, habilidades e investigación y desarrollo

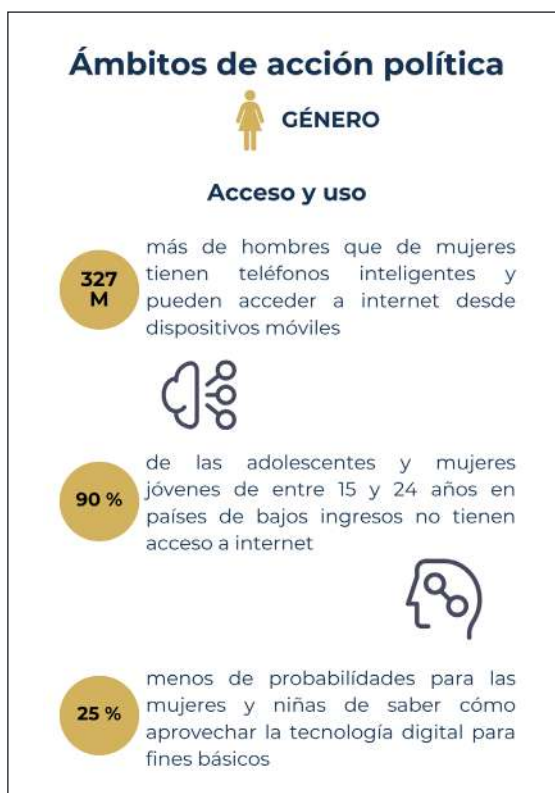
En género, la pregunta es: ¿por qué es importante? ¿Por qué es relevante hablar sobre temas de género cuando hablamos del desarrollo y del uso ético de la inteligencia artificial? Cuando hablamos de ética de la inteligencia artificial, y en general de la inteligencia artificial, para todos ha sido evidente que la inteligencia artificial deja de ser un tema netamente técnico. Se convierte en una cuestión profundamente social, porque tiene muchas implicancias en diferentes aspectos del día a día que nos afectan como sociedad y como personas, a todos y cada uno de nosotros.

El género se presenta como una de las áreas claves que se deberían tener en cuenta en la Recomendación y en los foros de debate.

Hay distintos aspectos en los que el género se ve afectado ya que existen diferentes brechas: en el acceso, habilidades e investigación y desarrollo. Hay muchas más, pero estas tres son las abordadas en cuanto a inequidad de género y en el uso de la tecnología.

En temas de acceso y uso, en 2023 se registró que más de 327 millones de hombres que de mujeres acceden a teléfonos inteligentes y por ende pueden tener internet y los servicios extra que se obtienen a través de sus dispositivos móviles.

A su vez, el 90 % de las adolescentes y mujeres jóvenes de entre 15 y 24 años que viven en países de ingresos bajos todavía no tienen acceso a internet, mientras que los pares masculinos adolescentes, hombres de entre 15 y 24 años, tienen el doble de probabilidad de tener este acceso.



Además, las mujeres y las niñas tienen un 25 % menos de probabilidades que los hombres de saber cómo aprovechar la tecnología digital para fines básicos. Esto no solamente tiene que ver con el acceso, sino con el tipo de usos que, como usuarias, las mujeres le dan a la tecnología. Más allá del acceso a la conectividad, más allá del acceso a dispositivos móviles con conexión a internet, también la utilización de la tecnología por parte de las mujeres difiere de la de los hombres.

Lo que muestran diferentes estudios es que los hombres utilizan la tecnología para acceder a temas más complejos, por ejemplo a servicios financieros, a temas de salud, a búsquedas de servicios o de empleo. Mientras que las mujeres se quedan en un uso de las tecnologías para situaciones más básicas: hacer llamadas, temas de comunicación y para estar en contacto con las personas que las rodean, más allá de acceder a diferentes servicios de mayor valor agregado.

En cuanto a habilidades, en América Latina y el Caribe menos del 40 % de las y los graduados de carrera STEM (ciencia, tecnología e ingeniería y matemáticas) son mujeres. Es una cifra muy baja en tecnologías de información y comunicación; hay países, como Brasil y Chile, en los que las mujeres solo representan el 15 % y el 13 % de los graduados y graduadas, respectivamente. Otros países están un poco mejor, en algunos la proporción es mayor, pero en ningún caso se supera el 50 % de graduados en estas carreras.

Lo que muestra un estudio de UNESCO es que la combinación de estos factores genera que los hombres tengan cuatro veces más probabilidades que las mujeres de poseer habilidades avanzadas en cuanto al uso y el manejo de tecnologías de información y comunicación. Una de ellas es saber programar. Las mujeres en promedio tienen 4 % menos de probabilidades de saber cómo programar y cómo usar los beneficios de la inteligencia artificial y de la generación de algoritmos para generar soluciones.

Impacto de las disparidades de género en el mercado laboral

En el contexto actual, estas disparidades están teniendo una implicación directa en temas como el mercado laboral. Actualmente, el 90 % de los empleos demandan competencias básicas digitales. Una de ellas, programar, pero en el conjunto de competencias básicas digitales las mujeres siguen por detrás.

Esto limita no solamente las oportunidades laborales a las cuales se puede acceder, sino también la capacidad de participar en la economía digital que presenta diferentes oportunidades. Sin duda esto tiene efecto en la remuneración y en los ingresos. Viendo en detalle las habilidades digitales y el uso de las mismas para diferentes cargos y oportunidades laborales, se observa que las mujeres tienen menos probabilidades de llegar a puestos de mayor nivel y mejor remunerados.

Mientras las mujeres se especializan más en temas de preparación y exploración de datos, los hombres se concentran en temas de aprendizaje automático, big data y temas más sofisticados de ciencia de datos que son más valorados por el mercado y, por ende, mejor remunerados.

Ámbitos de acción política



GÉNERO

Investigación y desarrollo

22 %

de los/las profesionales de la IA son mujeres

13,8 %

de los/las autores/as de productos de investigación en el campo de la IA son mujeres

18 %

de ponentes principales en conferencias sobre IA son mujeres

10/15 %

de los/las investigadores/as en aprendizaje automático en empresas líderes de tecnología son mujeres

Vemos el impacto negativo de esta conversión en cuanto a habilidades y en cómo aprovechar esta tecnología de inteligencia artificial.

Un último aspecto es el tema de investigación y desarrollo. Acá también la situación y los datos son muy críticos y alarmantes. A nivel mundial, solamente el 22 % de los profesionales en inteligencia artificial son mujeres. Solamente el 13,8 % de autores de productos y documentos de investigación en el campo de la inteligencia artificial son mujeres. Y apenas el 18 % de las ponentes en conferencias de inteligencia artificial son mujeres.

Aquí se comienza a ver el tema de la representatividad, de cómo esto constituye un problema y que todavía, a pesar de los avances, falta mucho.

En el informe del Alan Turing Institute titulado “Dónde están las mujeres” se destaca que entre el 10 y el 15 % de los investigadores y las investigadoras, en cuanto a aprendizaje automático dentro del sector privado en empresas líderes de tecnología, son mujeres.

Esta falta de representatividad se visualiza tanto en la academia como en el sector privado y en la industria, que es la que está liderando el desarrollo y la innovación en este aspecto. Esto tiene un impacto directo en temas económicos.

Los reclutadores o las empresas que buscan capital humano, por ejemplo en Silicon Valley, que es un “hub”² de innovación y de tecnología, reportan que para los puestos más técnicos de inteligencia artificial y ciencia de datos, dentro de todo el grupo de aplicantes que buscan este tipo de empleos, en promedio solo menos del 1 % de los y las aplicantes son mujeres.

Hay un tema de representatividad de habilidades que está teniendo impacto en el mercado laboral. Además tiene impacto en cómo se está desarrollando y qué se está teniendo en cuenta en el momento de programar, de incluir o no los datos para alimentar algoritmos de la inteligencia artificial.

Todo esto tiene una repercusión que puede ser negativa en cuanto a brechas, sesgos y discriminación.

Este es el diagnóstico, el porqué se está trabajando en esta agenda y el cómo, o por qué, la recomendación sobre la inteligencia artificial de la UNESCO presenta un área política dedicada a temas de igualdad de género (Collet, Neff y Gomes, 2022). Esto es para garantizar que en este momento nuestra conversación no solamente sea con gobiernos, sino que para que a través de sus iniciativas de políticas públicas tengan este tema en el radar y se vuelva una prioridad para garantizar que estos derechos humanos y estas libertades fundamentales para niños y niñas no se violen en ninguna etapa del ciclo de vida del desarrollo de sistemas de inteligencia artificial.

Esto es un poco el por qué. Ahora veremos cómo se intenta poner en práctica desde la UNESCO, qué se está haciendo, cómo se están abordando esos principios, esas recomendaciones, ese diagnóstico de la situación actual de la región en la práctica.

Trabajo de UNESCO con Estados miembros y el sector privado

Desde la UNESCO se está trabajando en dos grandes áreas. Una con los Estados miembros, con los gobiernos, y otra, con el sector privado, que es parte indispensable de la conversación.

En la conversación con el sector público, con los gobiernos, la recomendación plantea y propone una herramienta que se llama Metodología de Evaluación del Estado de Preparación (UNESCO, 2023). Es una herramienta diagnóstica para ayudar a los países a entender cuál es su situación actual alrededor de diferentes dimensiones. Una de ellas es el tema social y cultural, donde también se considera el aspecto de impacto social, de discriminación, de sesgos, de cuál es la calidad de los datos, de qué tan representadas están las poblaciones más vulnerables dentro de los datos que tiene el gobierno a su disposición, o con lo que está usando para desarrollar tecnologías, y qué tan justa, qué tan representativos o qué tan alta calidad tiene en tema científico y educativo.

De nuevo, hay que ver la participación de las mujeres en temas no solamente de las carreras STEM, sino también en temas de investigación, de cuántas mujeres se están dedicando a investigar y a desarrollar, por ejemplo, patentes. Patentes es otro tema crítico. Si seguimos en el mismo ritmo de cuántas mujeres están registrando patentes relacionadas con inteligencia artificial, recién en 2080 nosotras, las mujeres, podríamos llegar a un nivel comparable con las patentes que registran los hombres.

Entonces todas estas variables nos informan que la política pública busca mostrar prioridades, necesidades para que los países utilicen esa información y este instrumento diagnóstico para comprender cómo o para tomar decisiones más informadas en cuanto a política pública, y uno de esos aspectos de este RAM, como se le llama en inglés (Readiness Assessment Methodology), está relacionado con el tema de inclusión y de igualdad de género.

Se está haciendo esto en más de cincuenta países a nivel mundial y más de doce países en América Latina y el Caribe, donde parte de este ejercicio con muchos países es el diseño de sus estrategias nacionales de inteligencia artificial, y parte de la agenda de esas estrategias, sin dudas, en la mayoría de los países, en casi todos, es incluir una agenda que sea transversal a la estrategia en cuanto a temas de género e inclusión.

Por otro lado, se trabaja con gobiernos, pero además se está empezando a trabajar con el sector privado a través de una iniciativa que se llama el Consejo Empresarial para la Ética de la Inteligencia Artificial. Es una plataforma donde trabajamos con empresas actualmente lideradas por Microsoft y Telefónica, grandes compañías que están liderando la conversación sobre el uso ético de la inteligencia artificial.

Con ellos, intentamos entender esas prácticas que, dentro de su desarrollo de tecnología interno y corporativo, están aplicando para diseñar y tener soluciones de inteligencia artificial más inclusivas, que respeten y promuevan la equidad de género, y poder servir de puente entre el sector privado y lo público. Porque al final hay un montón de oportunidades de aprendizajes, de lecciones aprendidas entre un sector y el otro.

Una de las herramientas que propone también la Recomendación es la Evaluación de Impacto Ético (UNESCO, 2023), la cual estamos ajustando con el sector privado para estar seguros de que en el momento en que las empresas, y no solamente las grandes compañías sino las startups, las medianas y pequeñas empresas que están empezando a implementar o incluso a desarrollar temas de inteligencia artificial, puedan saber cuáles son las preguntas que se tienen que hacer, para que entiendan cuáles son los requisitos que tienen que cumplir y qué aspectos deberían tener en cuenta desde el momento del diseño y la conceptualización de las herramientas de inteligencia artificial para evitar impactos negativos en contra de equidades de género.

Iniciativa para promover la igualdad de género en la IA

Una de las iniciativas de UNESCO para resaltar esta agenda de igualdad de género y dar visibilidad es una plataforma que se llama Mujeres por la Ética de la Inteligencia Artificial, Women for Ethical AI (UNESCO, 2024).

Tiene como objetivo contribuir a esa identificación de buenas prácticas para evitar sesgos de género en el desarrollo de la tecnología, impulsar la participación de niñas, mujeres y grupos no tan representados en el desarrollo de la inteligencia artificial

y promover que estos algoritmos y los datos no tengan sesgos discriminatorios contra las mujeres.

Entre las acciones concretas, en el diálogo con países y en el estudio del diagnóstico actual en la región, hay tres acciones que quisiera destacar para cerrar mi intervención.

Existen tres prioridades: una primera prioridad es el desarrollo de competencias desde educación primaria y secundaria para garantizar que las niñas y los niños y las mujeres tengan un mayor acceso al mercado laboral, o al menos despierten el interés por carreras como las STEM. Una segunda es la conformación de equipos de inteligencia artificial y de ciencia de datos diversos, donde haya representación no solo de mujeres sino de todas las poblaciones vulnerables. Y una tercera prioridad es la participación desde la academia que tiene muchos diagnósticos de implicaciones a largo plazo.

NOTAS

¹ La principal acción consiste en que los Estados miembros establezcan medidas eficaces, como marcos normativos, y promuevan que actores del sector privado, instituciones universitarias y la sociedad civil se adhieran a ellas, controlando la evaluación del impacto en los derechos humanos, el Estado de derecho, la democracia y la ética.

² Hub es un dispositivo de red que conecta diferentes nodos de, por ejemplo, una red Ethernet con configuración en estrella.

REFERENCIAS

Collet, C., Neff, G. y Gomes, L. G. (2022). *Los efectos de la IA en la vida laboral de las mujeres*. UNESCO, Organización de Cooperación y Desarrollo Económico, y Banco Interamericano de Desarrollo, Edits.

UNESCO (01 de 01 de 2022). <https://www.unesco.org/es/articles/recomendacion-sobre-la-etica-de-la-inteligencia-artificial>. Obtenido de <https://www.unesco.org>: <https://www.unesco.org/es/articles/recomendacion-sobre-la-etica-de-la-inteligencia-artificial>

UNESCO (01 de 11 de 2023). *Ethical impact assessment: a tool of the Recommendation on the Ethics of Artificial Intelligence*. Obtenido de <https://unesdoc.unesco.org>: <https://unesdoc.unesco.org/ark:/48223/pf0000386276>

UNESCO (01 de 01 de 2023). *Metodología de evaluación del estadio de preparación: una herramienta de la Recomendación sobre la Éti-*

ca de la Inteligencia Artificial. Obtenido de <https://unesdoc.unesco.org>: https://unesdoc.unesco.org/ark:/48223/pf0000385198_spa

UNESCO (23 de 10 de 2024). *Women4Ethical AI*. Obtenido de <https://www.unesco.org>: <https://www.unesco.org/en/artificial-intelligence/women4ethical-ai>